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

George-Levi Gayle and Limor Golan

Tepper School of Business

Carnegie Mellon University

August 11, 2006

Abstract

This paper investigates the role of labor-market attachment, on-the-job human-capital accumulation, occupational sorting, and discrimination in the narrowing gender wage gap over the past three decades. This paper contributes in three ways: First, we formulate a dynamic adverse-selection in which self-fulfilling beliefs about future employment spells arise endogenously in equilibrium, affecting gender differences in labor-market experience, and occupational sorting. Second, we develop a new three-step estimation technique that allows us to estimate the model without solving it. This is particularly important with this class of model because it may exhibit multiple equilibria. We estimate the model using the PSID. Third, we decompose the changes in the gender earnings gap into the different components and quantify the effect of statistical discrimination on the changes in labor-market experience and the earnings gap. Increase in overall productivity, demographic changes and statistical discrimination patterns account for a large percentage of the decline in the gender earnings gap and the increase in female labor market experience. Whereas, relative increase in productivity in professional occupations raise representation of women in professional occupations.

KEYWORDS: Adverse selection, dynamic general equilibrium, occupation sorting, labor market experience, gender earning gap, discrimination, semi-paramtric estimation, structural estimation, estimation of dynamic games, kernel density, discrete choice.

*We thank Zvi Eckstein, John Kennan, Bob Miller, Steve Spear, Yoram Weiss, the participants of SOLE 2005, Tel-Aviv summer workshop 2005, SITE 2006, and the the seminar participants at Purdue University, University of Minnesota, Carnegie Mellon, and Northwestern for comments and suggestions.

1 Introduction

One of the most striking changes in the U.S. labor market over the past three decades has been the significant reduction in the gender wage gap during the 1980s. In 1968 the unconditional median gender wage differential was about 40%; this gap was reduced to around 28% by 1992.[1] While the wage gap was declining, there were significant changes in labor-market attachment and women's occupational composition. According to figures from the Michigan Panel Study of Income Dynamics (PSID), the participation rate of women 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. While these trends were taking place in women's labor-market attachment, there were little or no changes in the figures for men. Furthermore, occupation composition has changed significantly over the period with women entering in greater numbers the traditionally male occupations. For example, according to Lewis (1996), in 1976 42% of women and 49% of men held federal jobs in which 95% of their coworkers were of the same sex. By 1993, this had fallen to 12% and 3% respectively. Lewis (1996) also found that the portion of women holding a professional or administrative job went from 18% to 45% between 1976 and 1992. In the PSID, the percentage women holding professional jobs went from 28% in 1968 to 43% in 1992. The *unexplained* portion of the earnings gap, which is sometimes attributed to discrimination, has declined as well.[2] These significant changes prompt the question: What are the main driving forces of these changes in gender patterns of labor-market outcomes?

This paper investigates the roles of labor-market attachment, on-the-job human-capital accumulation, occupational sorting, and discrimination in the narrowing gender wage gap. The main challenge in quantifying these effects is to account for the endogeneity of labor supply, discrimination, and earnings. This paper contributes in three ways: First, we formulate a dynamic adverse-selection model of labor supply and human-capital accumulation, in which gender discrimination and the earnings gap arise endogenously. Second, we develop a three-step estimation technique and estimate the model using the PSID. Third, we decompose the changes in the gender earnings gap into the different components and quantify the effect of the incomplete information and other factors on the changes in experience and the earnings gap.

There are two broad types of employers' discrimination in the literature. The first type is

[1] For example, according to Blau and Kahn (1997), the log male/female wage differential declined from 0.47 to 0.33 between 1979 and 1988.

[2] 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 (2005), among others.

taste-based discrimination, formulated by Gary Becker (1971); the second type results from incomplete information, pioneered by Kenneth Arrow (1972) and Edmund Phelps (1973). Discrimination of the first type may not persist in a competitive environment, but some frictions, such as search frictions, may lead to persistent group differentials in the long-run equilibrium (see Bowles and Eckstein, 2002). Such models, however, are more adequate to explain race-based discrimination because their predictions do not match patterns of gender-related labor-market differences.[3] Our model belongs to the class of incomplete-information discrimination models and, in particular, to the literature on statistical discrimination first analyzed by Coate and Loury (1993). Whereas the statistical discrimination literature focuses on the effect of beliefs about productivity differences across groups, in our model the uncertainty is about the turnover propensity of workers. In particular, employers beliefs that women are less attached to the labor force and have shorter employment spells than men may be self-confirming, leading to women accumulating less labor-market experience and hence begin paid less than men.[4]

We incorporate statistical discrimination based on beliefs about employment-spell length into a framework of a general equilibrium dynamic adverse-selection model.[5] In the model, workers are heterogeneous with respect to characteristics affecting disutility of working. These characteristics evolve according to a known Markov process. Every period, workers choose consumption, whether to participate in the labor market (extensive margins), and how many hours to work (intensive margin) in order to maximize lifetime utility. Utility is an increasing function of consumption and *nonmarket* hours (leisure time and hours spent producing home goods). *Nonmarket* hours is time inseparable in the utility function to allow for the possibility that the *stock* of past nonmarket hours affects the current disutility of hours worked. Every period there is a random utility shock to the utility of participating and not participating. We assume asset markets are complete, firms compete over workers, and there is free entry of firms into the market. The returns to experience, hours worked and cost of hiring new workers, differ across occupations—Eckstein and Wolpin (1989), Altug and Miller (1998), Buchinsky et al. (2005), Beffy et al. (2006), and Gayle and Miller (2004), among others, estimate models of labor supply in which returns to experience are

[3]For example, Such models predict that the discriminated group is less likely to work in professions in which they have extensive contact with costumers. Women, however, tend to work in service industries.

[4]Albanes and Olivetti (2005) develop a one-period model of household good production and labor supply in which gender discrimination also arises in equilibrium; however in their model the asymmetric information is due to unobserved effort. Thus, employers face moral hazard problem and not adverse selection.

[5]Pay differences can be generated in our model if the groups of male and female are ex-ante identical. However, in the estimation stage, we allow for differences across groups. Furthermore, characteristics that are exogenous in our model, such as education, can generate pay gaps. We are able, however, to estimate these effects separately.

endogenous. There are firm-specific costs of hiring new workers that are identical within occupations and different across occupations. Employers observe age, experience, education, and other skills, but there are characteristics affecting the disutility of working that are the worker's private information. In particular, because these characteristics evolve according to some known process, there is a correlation in the "worker's" type over time. Based on observable characteristics, employers form beliefs regarding the worker's future employment spell when they offer wage contracts. Wage contracts consist of hours and earning. That is, we solve for a contract posting equilibrium.

Because of the firm-specific cost of hiring new workers, employers make rent on workers a year after they are hired (since the market is competitive, when hiring workers firms make zero expected profits over the workers' employment spells). In equilibrium, beliefs about the worker's future participation in the firm enter the earnings equations. Therefore, if women face lower wages, since there is disutility of working, they will work less and sort into occupations with lower returns to experience and lower costs of hiring workers. Thus, on average, they will accumulate less experience than men. The closest papers to ours are Lee and Wolpin (2000, 2006). They develop and estimate a model with cost of switching sectors and include an endogenous formation of human capital; however, there is no private information in their model and they do not model discrimination.

Our model is a signaling model. Individual labor-supply decisions (participation and hours) may provide information on the worker's type. In equilibrium, information on workers is revealed gradually over time (this is a typical feature of dynamic adverse-selection models with correction in the *types* over time and incomplete contracts. See, for example, Tirole (1996). Over time, employers update beliefs based on individual labor-market history. Thus, working more today may affect the worker's potential earnings not only through accumulation of experience, but also because of the possible effect on the employers' beliefs. Therefore, the model predicts that the information employers have on experienced workers is more accurate than their information on young workers. These predictions are consistent with empirical regularities found in several papers that empirically examine the relationship between the pay gap and the length of employment spells. Light and Ureta (1992) found that, controlling for all observable characteristics, older women are more likely to stay in a job and earn as much as men. Light and Ureta (1992) found evidence that *stayers* are paid more than workers with short employment spells. They also found that women are more likely to move and that this difference becomes smaller as workers age. Altonji and Paxton (1992) found that job mobility is strongly linked to hour changes; women who face changes in family responsibility adjust their hours and this may lead to lower earnings. Baron et al. (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. The main theoretical contribution of the model is to formally show how self-fulfilling beliefs about future employment spells arise endogenously, affecting gender differences in labor-market experience, occupational sorting, and attachment to the labor force.

One of the goals of this paper is to account for changes in relative earnings, the wage gap, over time. The literature focuses on several factors that may have caused these changes, some of which are exogenous in our model; they drive changes in beliefs, earnings, the gender earnings gap, and labor supply. We explore which of these factors drove changes in the relative wages of men and women, and quantify their relative importance. The first factor is technological changes in the economy, which we model as occupation-specific shocks to productivity. They raise productivity for all workers within the occupation equally. Our model, however, predicts that if women's participation is lower, positive productivity shocks may increase female participation and employment spells more than males', driving a decline in the gender wage gap. The second possible source of changes in relative earnings is a decline in costs of producing home goods. In our model, there are fixed costs in the utility when an individual participates in the labor market. If costs of home-produced goods declined over time, these costs should decline as well. It may affect the beliefs about women's attachment to the labor force as well. The third factor is changes in education, marriage, and fertility over time. Such changes may cause changes in labor supply behavior because they affect the disutility of working.

Our model exhibits a multiplicity of equilibria. Different beliefs generate different equilibrium outcomes. This is standard in models of statistical discrimination (Moro, 2003; Antonovics, 2004). In the standard statistical discrimination model, an identification problem of the following form arises: Given the observed wage distribution and human-capital investment, an econometrician tries to recover the preference parameters and the beliefs about an unobserved variable that affects productivity. The source of the identification problem is that there may be more than one combination of preferences parameters and beliefs that could have generated the observed wage distribution. Our model, however, does not have that problem. This is because our beliefs are about future participation probability, conditional on observed characteristics. Panel data, however, allows us to observe the individual's next-period participation decision, which is correct across the population conditional on the observed characteristics. Therefore we can nonparametrically identify these beliefs under the assumption that for each cohort one equilibrium plays out.

Toward this end, we developed a new multistep semiparametric estimation strategy that allows us to estimate the model without solving it. This is particularly important in this class model which may exhibit multiple equilibria. Our model yields several moment conditions. First, there is a consumption Euler equation. From labor-supply decisions, we obtain the Euler equation for hours. In addition, workers face a discrete-choice problem each period of whether to participate in the labor market or not, yielding another moment condition. Finally, the free entry into the labor market yields another moment condition as the wage contracts are offered so employers make zero expected profit from hiring a worker. We use these conditions for identification and estimation. First, from the standard Euler equation, we identify the individual marginal utility from wealth used in the hours Euler equation. Second, the beliefs are identified nonparametrically from the equilibrium condition: Conditional on the current information available to the employer, the expected participation is correct. The zero-profit condition identifies the cost of hiring employers in each occupation. This is identified over variation in beliefs about future participation in the earnings equation. The production function parameters in each occupation are also identified through this equation. We then use the hours Euler equation and the choice probabilities to identify the rest of the utility parameters (Altug and Miller, 1998). The estimation then proceeds in three steps. The idea is to estimate the employer's problem in the first step, along with other inputs from the individual consumption Euler equation. The estimates above are used to nonparametrically estimate each individual's choice-specific probabilities and their derivatives in the second step. In the final step, these estimates are combined with a tractable alternative representation of the agents' choice-specific valuations to form moment conditions to estimate the structural parameters of the agents' utility function. The estimates of the structural parameters are $\sqrt{N}$ consistent (where N is the number of individuals in the sample) and asymptotically normal although the second step is estimated nonparametrically. Our estimate 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, 2004; 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.

We find that the cost of hiring workers is roughly 3.5 times higher in professional occupations compared to nonprofessional occupations. The returns to labor-market experience are also substantially higher in the professional occupations. These findings are consistent with the model's prediction that the earnings gap should be smaller in occupations with low costs of hiring workers and that women will sort into occupations in which the costs of hiring is smaller and returns to labor-market experience are lower. Our model predictions matched the raw data well. The predicted decline in the wage gap between the mid- to late 1970s and the mid- to late 1980s in professional occupations is 28% and the raw decline is 30%; for nonprofessional occupations, the predicted decline in the earnings gap is 22% and the raw decline is 25%. Further analysis shows that private information, demographic changes and aggregate shocks to market productivity account for most of the change in the accumulation of labor-market experience by women over the period relative to men.[6]

The decomposition of the change in the gender wage gap reveals that changes in statistical discrimination patterns account for 16% of the decline in the gender wage gap in professional occupations and 14% in nonprofessional occupations. We then computed the decline in the wage gap due only to factors unrelated to the friction (hiring costs). That is, we computed the decline in the wage gap in an economy with no hiring costs (in which wages are the marginal productivity). These changes are decomposed into changes in demographic characteristics, changes in costs of home production and occupation-specific productivity shocks. We find that changes in demographic characteristics account for 19% and 22% of the closing gender wage gap in professional and nonprofessional occupations, respectively. These changes include changes in education and family structure (such as fertility and marriage). Our production-function parameter estimates show that the positive trend in productivity shocks is larger in professional occupations than in nonprofessional occupations. Occupation-specific productivity shocks account for 12% of the change in the earnings gap in professional occupations and only 7% in nonprofessional occupations. The estimation results do not support the hypothesis that changes in home production technology explain the increase in women's labor-market participation (see similar findings in Jones, Manuelli, and McGrattan, 2003). Such changes should have caused a decrease in the fixed cost of participating in the labor market (estimated as part of the utility-function specification). Our estimation results suggest that fixed costs of participation account for only 1% in professional occupations and 2% in nonprofessional occupations.

This paper is organized as follows. Section 2 describes the model. Section 3 contains the equilibrium analysis and the theoretical model's predictions. The identification and estimation methods are presented in section 4. The data description is in section 5, and section 6 contains the estimation results, empirical analysis, and the gender earnings gap decomposition. Section 7 concludes. The appendices present proofs, the asymptic property of our estimator and a detailed data description.

2 Theoretical Model

This section describes the basic structure of the economy and the dynamic general-equilibrium model. In the labor market, firms compete over workers, and the gender wage gap and oc-

[6]Notice that although our model uses nonlinear labor pricing, we can compute earnings. From earnings, we can compute the implied wages.

cupation sorting are determined together with worker's labor supply. We then analyze the model's predictions and empirical properties.

2.1 Workers' problem

We assume that there exist a continuum of workers (men and women) on the unit interval $[0,1]$. Let $g = \{f, r\}$ denote the worker's gender (female and male respectively) and let a denote the worker's age–education cohort (all workers who were born in the same year, who have the same number of years of completed education). Workers are finitely lived (denote the worker's age by $\tau = \{18, \ldots, 65\}$). The calendar year is indexed by t ($t = 1975, \ldots, 1994$). Each individual has preferences about nonmarket hours (time in which the individual does not work) and consumption. The consumption allocated to individual n at date t is denoted by c_{nt}. The other choice variable is h_{nt}, the time spent working by agent n in period t. There is a fixed amount of time in each period available for working, which implies that the amount of time worked in each period can be normalized as $0 \leq h_{nt} \leq 1$. If $h_{nt} = 0$, the agent does not work at time t. Otherwise, the agent works the fraction of time $h_{nt} > 0$. For notational convenience, a participation indicator d_{nt} is defined, where $d_{nt} = 1$ if and only if $h_{nt} > 0$ and 0 otherwise.

Preferences are additive in consumption and leisure both contemporaneously and over time. It is assumed that there are both observed and unobserved (by the econometrician) exogenous, time-varying characteristics that determine the utility associated with alternative consumption and leisure allocations. Denote the former by the $K \times 1$ vector z_{nt} and the latter by the 3×1 vector $(\varepsilon_{0nt}, \varepsilon_{1nt}, \varepsilon_{2nt})'$. It is assumed that z_{nt} is independently distributed over the population with a known distribution function, $F_{0ga}(z_{nt+1} \mid z_{nt})$, the vector $(\varepsilon_{0nt}, \varepsilon_{1nt}, \varepsilon_{2nt})'$ is independent across (n,t) and drawn from the population with a distribution function $F_1(\varepsilon_{0nt}, \varepsilon_{1nt}, \varepsilon_{2nt})$.

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

$$U_{nt} \equiv d_{nt} u_{t0}(z_{nt}) + u_1(z_{nt}, 1 - h_{nt}) + u_2(z_{nt}, c_{nt}, \varepsilon_{2nt}) + (1 - d_{nt})\varepsilon_{0nt} + d_{nt}\varepsilon_{1nt}, \quad (1)$$

where u_{t0} represents the fixed utility costs of working, which depends on the observed individual-specific characteristics but may change from period to period. We allow this to change over time in order to capture the possible changing home production technology. It takes the following form:

$$u_{t0}(z_{nt}) = B_{ot} + z'_{nt} B_0.$$

u_2 represents the current consumption utility (which depends on consumption and observed characteristics of the individual and the idiosyncratic shocks). $u_2(z_{nt}, c_{nt}, \varepsilon_{2nt})$ is concave increasing in c_{nt} for any z_{nt} and ε_{2nt}. We assume the following functional form:

$$u_2(c_{nt}, z_{nt}, \varepsilon_{2nt}) = \exp(z'_{nt}B_4 + \varepsilon_{2nt})c^{\alpha}_{nt}/\alpha.$$

and $E(\varepsilon_{2nt} \mid z_{nt}) = 0$. $u_1(z_{nt}, 1-h_{nt})$ is the utility of nonmarket hours consumed. Intuitively, u_1 is the utility from consuming a greater fraction of leisure or spending a greater fraction of time on home production (which varies with current time spent at work as well as the observed individual-specific characteristics). The utility from non–labor-market hours is also a function of spouse characteristics such as education, and income. It depends on past nonmarket hours and there is interaction between past hours. This formulation allows us capture the element of experience accumulation in home production

$$u_1(z_{nt}) = z'_{nt}l_{nt}B_2 + \theta_{0g}l^2_{nt} + \sum_{s=1}^{\rho} \theta_{sg}l_{nt}l_{nt-s}$$

We allow men and women to have different distributions of z_{nt} and a different utility parameter θ_s. We do not, however, impose these differences.[7]

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

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

To provide a tractable solution to the model, we assume that asset markets are competitive and complete (CCM).[8] Here the word *competitive* is synonymous with price-taking behavior and "complete markets" means that there are no frictions in the markets for loans, a common interest rate facing borrowers and lenders, and a rich set of financial securities exists to hedge against uncertainty.[9]

[7] Wage gap, in our model can arise even if men and women are ex-ante identical in all respects.

[8] Whereas to some the assumption of complete markets might be controversial, it is empirically tractable and serves as a useful benchmark, which allows us to focus our analysis on the primary source of asymmetry in our model. A popular alternative to the complete-market assumption is to put wages directly into the utility function, this is an even stronger assumption than complete markets and can only be justified under two very strong assumptions; (1) wealth maximization or (2) no markets to borrow or save. Hence we feel that at least by assuming complete markets, we know the source of our restrictions on behavior.

[9] Altug and Miller (1990, 1998) have used this condition to estimate both males' and females' consumption and labor supply with aggregate shocks. Other papers that discuss complete markets and estimate frameworks based on this assumption include, Card (1990); Mace (1991); Townsend (1994); Altonji, Hayashi and Kotlikoff (1996); Miller and Sieg (1997); and Gayle and Miller (2004), among others.

The CCM assumption allows us to compactly write the lifetime individual budget constraint. A complete market implies that individuals can condition their choice 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. This assumption allows us to rewrite the workers' budget constraint in each period as

$$E_0 \left\{ \sum_{t=0}^{T_n} \beta^t \lambda_t \left[c_{nt} - \overline{S}_{nt} \right] \right\} \leq W_n, \tag{3}$$

where $\overline{S}_{nt}$ is the total household labor-market income (i.e., if the individual is single then it is only one income, but if the individual is married it the sum of two incomes), λ_t is the expected price of the contingent claim, and W_n is an exogenously determined quantity denoting bequests net of inheritances. In any state, ξ, the price of a contingent claim is

$$\overline{\lambda}_t(\widetilde{\xi}) = \int_{\widetilde{\xi}} \lambda(\xi_t) g(\xi_t) \, \mathrm{d}\widetilde{\xi}.$$

The states are determined by realizations of ε_{2nt}, z_{nt} (recall that the densities are $F_0(z_{nt+1} \mid z_{nt})$; $F_1(\varepsilon_{2nt})$); they are independent and $g(\tau_t)$ is the joint density.

The aggregate feasibility condition equates the sum of labor income by all households and the aggregate resource endowment W_t

$$\int_0^1 [c_{nt} - \overline{S}_{nt}] \, \mathrm{d}\mathsf{L}(n) \leq W_t, \quad t \in \{0, 1, \ldots\}. \tag{4}$$

In this expression, $\mathsf{L}(.)$ is the Lebesgue measure, which integrates over the population.

2.2 Firms' problem, labor market and employment contracts

Employers are infinitely lived. Firms (employers) compete over workers. There are M occupations, $m = 1, \ldots, M$. There are costs to the employer when a new worker is hired (for example, specific training). These costs are specific to the employer and vary across occupations. Within occupations, the costs are the same for all employers. We denote employer's cost of hiring a new employee (switching costs) in occupation m by γ_m. There is a free entry into the market. We assume that there is a homogeneous product with price normalized to 1.

Let d_{mnt} be an indicator function which takes the value 1 if worker n works in occupation m in period t, and zero otherwise. H_{nt-1} denotes the worker's labor-market experience at the beginning of period t (participation in each occupation, and hours worked in every period), i.e. $H_{nt-1} = [h_{n1}, h_{n2}, \ldots, h_{nt-1}, \{d_{mn1}, d_{mn2}, \ldots, d_{mnt-1}\}_{m=1}^{M}]$. Let z_{nt}^p denote a vector of worker characteristics that affect production. The production function and costs are identical within occupations, but vary across occupations. Output in period t in occupation m is denoted by: $y_{mt}(h_{nt}, H_{nt-1}, z_{nt}^p)$. Employers maximize lifetime expected discounted profits. Each occupation has a large number of identical firms that competing for workers.

Workers and firms can only commit to (noncontingent) spot contracts.[10] Each firm offers one contract each period and commits to hours (for the current period) before posting the contract. Output is a function of human capital, hours, and other characteristics, z_t^p. Occupations may vary in the hours they offer. We assume that a firm in occupation m only offers a contract with hours h for a worker with experience H_{t-1} and characteristics z_t^p if there exists no other occupation m' such that

$$f_{m't}(h_t \mid H_{t-1}, z_t^p) - \gamma_{m'} > f_m(h_t \mid H_{t-1}, z_t^p) - \gamma_m. \tag{5}$$

This is a *sorting* assumption that implies that occupations *specialize* in hours in which they have an advantage over other occupations.[11] This assumption implies that any choice of hours, h_{nt}, of any particular worker with characteristics H_{t-1}, z_t^p (education and skill) *maps* into a unique occupation choice in period t. That is, for any given characteristics that affect production, each occupation offers contracts of $\underline{h_{mt}}(H_{t-1}, z_t^p) \leq h_t(H_{t-1}, z_t^p) \leq \overline{h_{mt}}(H_{t-1}, z_t^p)$. Because each employer can offer an employee a contract with any number of hours in the occupation, turnover of workers who remain in the labor force only occurs across occupations.[12]

The profit function of the firm is simply the expected profits from hiring an individual worker. Thus, an optimal contract offer can be solved separately for each job (hours and salary). Denote by π_{tm} the value of a vacancy for each employer in occupation m in period t. Thus, participation in a firm in occupation m denoted by the indicator d_{nmt}, where $d_{nmt} = 1$ if and only if $\underline{h_{mt}}(H_{t-1}, z_t^p) \leq h_{nt}(H_{n,t-1}, z_{nt}^p) \leq \overline{h_{mt}}(H_{t-1}, z_t^p)$ and $d_{nmt} = 0$ otherwise.

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

[11] Although this assumption does not change qualitatively the theoretical prediction of the model, it simplifies the estimation stage substantially.

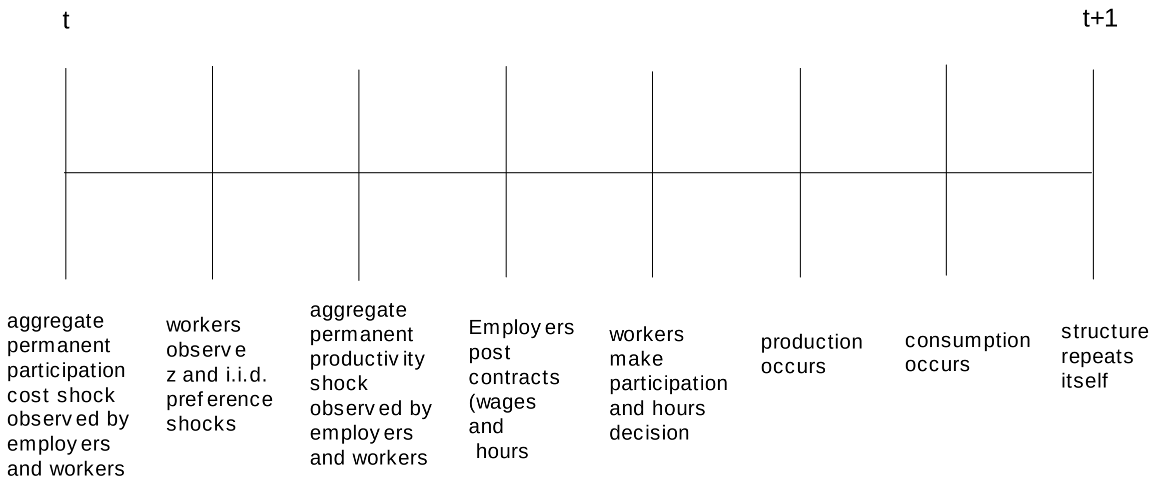
[12] The dataset we use does not have information on employers. Only on occupations. Therefore, we cannot detect job-to-job transitions within an occupation.

2.3 The structure of the game: Timing, information, and strategies

Labor-market experience, H_{nt-1}, is common knowledge. The utility function parameters and functional forms are common knowledge as well, but outside employers do not observe past wage history of a worker.[13] Some characteristics affecting labor supply, z_{nt}, are unobserved by potential employers in the labor market. The worker's private characteristics are wealth, spouse income, spouse characteristics (such as education), marital status, and number of kids.[14] The i.i.d. shock to utility are also the worker's private information. These characteristics affect labor-market–participation decisions and labor supply. Denote by z^*_{nt} the worker's publicly observable characteristics which are at time t except for the worker's history. Denote by z^p_{nt} a subset of z^*_{nt} that directly affects production (such as education, age, and individual *fixed-effect*). So the information structure is as follows: $z^p_{nt} \subset z^*_{nt} \subseteq z_{nt}$. Workers know the complete structure, but potential employers only know z^*_{nt}. Workers' consumption is not observed by the employer (knowledge of z_{nt} and ε_{0nt} is sufficient to recover consumption, but both are unknown to the employer). Notice that the worker's *type* here is characterized by z_{nt} and ε_{0nt}. Whereas ε_{0nt} is i.i.d., z_{nt} evolves according to a known process, $F_{ga}(z_{nt+1} \mid z_{nt})$; therefore, there is a correlation between types over time (for example, number of children, marital status, spouse characteristics). Thus, the worker has better information about the probability of remaining in the firm in the future (the employer knows $F_{ga}(z_{nt+1} \mid z^*_{nt})$). To be exact the timeline of the actions is as follows.

[13] This assumption is made to simplify the off-equilibrium path analysis and to reduce notation.

[14] These characteristics affect the marginal utility of consumption and the disutility of working. The criterion we used to determine what characteristics are unobserved by potential employers is whether they are unverifiable by potential employers. For example, legal restrictions on asking workers to provide this information and on firing them later if they do not report it truthfully. We also assume implicitly that no markets for this information exists. Furthermore, we conduct a counterfactual experiment in which the above characteristics are common knowledge. This provides an informal test of whether the characteristics are indeed private information. Notice, that the important assumption is that workers have information potential employers do not have, as opposed to their current employer. The reason is that wages are determined by the offers made by potential employers.



Strategies and beliefs

The worker's strategy is a choice of consumption, participation, and labor hours in each period for every possible state. We denote a strategy—the probability of participation and how many hours to work given the worker's type, labor-market experience, and observable characteristics—by $\sigma_g(h_{nt-1}, d_{nt-1} \mid z_{nt-1}, H_{nt-2})$ and the consumption by c_{nt}.

The firm's strategy is a choice of hours, $h_{mt} \in [\underline{h}_m(H_{nt-1}, z^p_{nt}), \overline{h}_m(H_{nt-1}, z^p_{nt})]$, and salary, $S_{mg}(h_{tm}, H_{nt-2}, z^*_{nt})$, in every period for any observable characteristics and history of a worker. At the beginning of each period, firms form a (common) set of prior beliefs on each individual worker's type, i.e. $\mu_{tg}(z_{nt} \mid H_{nt-1}, z^*_{nt})$. At the end of each period, firms update their beliefs about each worker's type by observing the worker's actions in the previous period. We denote *posterior* beliefs—which are formed at the end of period t, after observing the worker's decision—by $\widetilde{\mu_{tg}}(z_{nt} \mid H_{nt}, z^*_{nt})$. Notice that $\widetilde{\mu_{tg}}(z_{nt} \mid H_{nt}, z^*_{nt})$ is used to form the posterior beliefs in period $t+1$, as the types are correlated over time, and evolve according to the Markov process specified above. Therefore, the worker's behavior in time t may convey information about the worker's future type. We also define the firms' beliefs at the beginning of period t about *future participation* by $\widetilde{p}_{gt+j}$. The beliefs about future participation will be derived from the beliefs about the worker's type. We discuss belief formation further below.

3 Equilibrium Analysis

We first describe the equilibrium concept and then analyze the workers' optimal strategies and the optimal contracts. The rest of this analysis is for a given age–education cohort; for expositional ease, we drop the cohort subscript.

Definition 1 (Equilibrium) *A Perfect Bayesian Equilibrium consists of strategies*

$(c_n, \sigma^o_{gn}, S^0_{gm}(h, H, z^*), h_m)$, *where* σ^o_{gn} *is the probability of participation and hours worked in every period for every type in every state,* $S^0_{gm}(h, H, z^*)$*, are the contracts offered in each occupation in each period for any observable characteristics, history, and belief system, such that*

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

2. *The posterior beliefs,* $\widetilde{\mu}$*, satisfies Bayes' rule when possible.*[15]

$$\widetilde{\mu}_{tg}(z_{nt} \mid h_{nt}, d_{nmt}, z^*_{nt}, H_{nt-1}) = \frac{\mu_{tg}(z_{nt} \mid z^*_{nt}, H_{nt-1})\sigma(h_{nt}, d_{nmt} \mid z_{nt}, H_{nt-1})}{\int \mu_t(\widetilde{z}_{nt} \mid .)\sigma(h_{nt}, d_{mnt} \mid .)\, \mathrm{d}\widetilde{z}_{nt}} \tag{6}$$

3. *At the beginning of period* $t+1$ *firms form priors about the worker's type at that period based on past history (types changed exogenously)*

$$\mu_{t+1g}\ (z_{nt+1} \mid H_{nt}\ , z^*_{nt}\ ) = f_{gz}(z_{nt+1} \mid z_{nt})\widetilde{\mu}_{tg}(z_{nt} \mid .) \tag{7}$$

4. *The probability of participation in the proceeding period is*

$$\widetilde{p}_{gm,t+1}(H_{nt-1}, z^*_{nt}) = E_t[E_{t+1}(\sigma(h_{nt+1}, d_{mt+1} \mid z_{t+1}.)) \mid H_{nt-1}, z^*_n]$$

Next we characterize the equilibrium strategies of the workers and firms. We begin with the solution to the individual problem. Let η_n denote the Lagrangian multiplier associated with the budget constraint in Equation (3), then the optimal consumption satisfies the necessary conditions,

$$\frac{\partial u_2(c_{nt}, z_{nt}, \varepsilon_{2nt})}{\partial c_{nt}} = \eta_n \lambda_t, \tag{8}$$

for all $n \in [0, 1]$ and $t \in \{0, 1, 2 \ldots\}$.

With the contemporaneous separability of the consumption from the labor supply choices, the condition in (8) can be used to solve for the individual Frisch demand functions, which determine consumption in terms of the time-varying characteristics, z_{nt}, the idiosyncratic shocks to preferences, ε_{2nt}, and the shadow value of consumption, $\eta_n \lambda_t$.[16] Characterizing the optimal labor-market participation and hours-of-work decisions is more involved. Note that each person's labor supply contributes an infinitesimal addition to aggregate output. Define

[15] The restriction on the beliefs is stronger than usual as it applies to updating in histories that are reached with probability zero. See Definition 8.2 of Fudenberg and Tirole (1996) for the formal description of the conditions of equilibrium.

[16] This solution assumes complete information in the consumption contingencies market. One may worry about the unraveling of workers' private labor information that can happen if there is a market for information. In order to resolve this, we assume trade is anonymous in the consumption market.

the conditional valuation functions associated with the decisions to work and not work at time t as V_{1nt} and V_{0nt} respectively. The necessary condition for the optimal participation decision is

$$d_{nt}^o = \begin{cases} 1 & V_{1nt} + \varepsilon_{1nt} \geq V_{0nt} + \varepsilon_{0nt} \\ 0 & \text{otherwise.} \end{cases} \tag{9}$$

The expected value of d_{nt}^o, conditional on the observed (to the researcher) state variables, ω_{nt}, is the conditional choice probability of participating in the labor force and can be written as

$$p_{nt} \equiv E[d_{nt}^o \mid \omega_{nt}] = \int_{-\infty}^{V_{1nt}-V_{0nt}} (\varepsilon_{0nt} - \varepsilon_{1nt}) \, \mathrm{d}F(\varepsilon_{0nt}, \varepsilon_{1nt}, \varepsilon_{2nt}). \tag{10}$$

In our model $\omega_{nt} \equiv [H_{nt-1}, z_{nt}, \eta_n \lambda_t, \mu_t]$. This equation basically characterizes the participation decision.

The equations below are the value function for an individual who chooses not to participate in period t and to behave optimally thereafter.

$$\begin{aligned} V_{0nt} + \varepsilon_{0nt} \quad \equiv \quad & \max_{\{h_{nr}\}_{t=t}^{T}} E_t \left\{ \sum_{s=t}^{T} \beta^{s-t} \left[d_{ns} u_0(z_{ns}) + u_1(z_{ns}, 1 - h_{ns}) + d_{ns} \varepsilon_{1ns} \right.\right. \\ & \left.\left. + (1 - d_{ns}) \varepsilon_{0ns} + \eta_n \lambda_s \sum_{m=1}^{M} d_{nms} S_m(h_{ns}, z_{ns}^*, H_{ns-1}, \mu_s) \right] \mid h_{nt} = 0 \right\}. \end{aligned} \tag{11}$$

Using the Bellman principal, the value function for an individual who choose to participate in the labor force in period t and to behave optimally thereafter is

$$\begin{aligned} & V_1(\omega_{nt}) \\ = \quad & \max_{h_{nt} \in (0,1)} \left[u_1(l_{nt}, H_{nt-1}, z_{nt}) + \eta_n \lambda_t \sum_{m=1}^{M} d_{nmt} S_m(h_{nt}, z_{nt}^*, H_{nt-1}, \mu_t) \right. \\ & + \beta E_t \{ p_{nt+1} [V_1(\omega_{nt+1}) + \varepsilon_{1nt+1}] \mid \omega_{nt}, h_{nt} > 0 \} \\ & \left. + \beta E_t \{ (1 - p_{nt+1}) [V_0(\omega_{nt}) + \varepsilon_{1nt+1}] \mid \omega_{nt}, h_{nt} > 0 \} \right]. \end{aligned} \tag{12}$$

We next analyze the worker's labor-supply decision. We define by h_{nt}^o, the optimal labor-supply decision for individual n in period t and by $h_{nt}^* \in (0, 1)$, the optimal interior solution of the labor-supply decision for individual n in period t. Similarly, we define by d_{nt}^o the optimal participation decision for individual n in period t. Using the above formulation, the necessary condition for an optimal interior solution for labor supply is

$$0 = \frac{\partial u_1\left(l_{nt}, H_{nt-1}, z_{nt}, \{S^0_{mt}(h,\mu_t)\}_{m=1}^M\right)}{\partial h_{nt}} \tag{13}$$
$$+ \eta_n \lambda_t \sum_{m=1}^{M} d_{nmt} \frac{\partial S_{nt}(h_{nt}, z^*_{nt}, H_{nt-1}, \mu_t)}{\partial h_{nt}}$$
$$+ \beta E_t\{\partial[p_{nt+1}(V_1(\omega_{nt+1}) - V_0(\omega_{nt+1}))]/\partial h_{nt}$$
$$+ \partial V_0(\omega_{nt+1})\backslash \partial h_{nt} \mid \omega_{nt}, h_{nt}\}$$

There are several points to notice in the condition for the interior solution in the hour equation. The first component is the change in the disutility of working when hours change and the last two terms are the change in the future expected payoffs. In the model, the change in earnings that results when one changes hours is nonlinear.

$$\frac{\partial S_m(h_{nt}, z^*_{nt}, H_{nt-1}, \mu_t)}{\partial h_{nt}} = \frac{\partial f_m(h_{nt}, H_{Tnt-1}, z^*_{nt})}{\partial h_{nt}} + \gamma_m \beta \frac{\partial \widetilde{p}_{nmt}(h_{nts}, H_{mnt-1}, z^*_{nt})}{\partial h_{nt}}$$

It is worth noticing that the beliefs component in the salary is based on beliefs about the selection of workers with different unobserved characteristics in the specific contract. Furthermore, increasing hours worked today has two effects on future earnings. First, it affects the productivity in the future (as $f_{mt}(h_t, H_{t-1}, z^p_{nt})$ is a function of past experience) and therefore, affects future earnings. The second effect is due to the information asymmetry. Choice of hours affects not only current beliefs about type, and therefore, current salary, but also affects future beliefs. That is, work experience is a signal about the worker's type $\widetilde{p}_{gm,t+1}(H_{nt-1}, z^*_{nt})$. In the model, occupation sorting is endogenous: The assumption in equation 5, together with the assumption that all human capital is general, implies that by choosing hours, the worker selects into an occupation and a salary that is paid in this occupation for the given number of hours worked.

The participation decision is simply derived from a comparison between the value of working and the value conditional on not working,

$$d^o_{nt}(\omega_{nt}, \varepsilon_{0nt}, \varepsilon_{1nt}) = \begin{cases} 1 & \text{if } V_1(\omega_{nt}) + \varepsilon_{1nt} \geq V_0(\omega_{nt}) + \varepsilon_{0nt} \\ 0 & \text{otherwise.} \end{cases} \tag{14}$$

Next, we characterize the optimal contracts firms offer. Define

$$I_m(z_{nt+1}, H_{nt}) \equiv I\{\underline{h}_m(H_{nt}, z^p_{nt}) \leq h_{nt+1}(z_{nt+1}, H_{nt}) \leq \overline{h}_m(H_{nt}, z^p_{nt})\}$$

to be an indicator function which takes the value 1 if the worker's hours are offered in occupation m and zero otherwise, and denote the probability that type z_{nt}, with history H_{nt-1} and observable characteristics z^*_{nt}, works in that period by

$$Q_t(z_{nt}, .) = E[d_{nt} \mid z_{nt}, .].$$

Proposition 1 (Employers' optimal strategies)

1. *The optimal contracts of the firms are*

$$S_{gm}(h_{nt}, H_{nt-1}, z^*_{nt}) = y_{mt}(h_{nt}, H_{nt-1}, z^p_{nt}) - \gamma_m + \beta\gamma_m \widetilde{p}_{gm,t+1}(H_{nt-1}, z^*_{nt}) \quad (15)$$

for all $h_{nt} \in [\underline{h}_m(H_{nt-1}, z^*_{nt}), \overline{h}_m(H_{nt-1}, z^*_{nt})]$*, where* $\widetilde{p}_{gm,t+1}(H_{nt-1}, z^*_{nt})$ *is the belief that the worker will work in the firm in the proceeding period.*

2. *The beliefs about the participation of the worker in the proceeding period is*

$$\widetilde{p}_{gm,t+1} = \int Q_{t+1}(z_{nt+1}, H_{nt}, h_{t+1}) I_{t+1}(z_{nt+1}, H_{nt}, h_{t+1}) \mu_{gt+1}(z_{nt+1} \mid H_{nt}, z^*_{nt})\, \mathrm{d}z_{nt+1}$$

3. *Off-equilibrium path: if* $\forall z_{nt}(h^t)$, $\sigma_g(h_{nt}, d_{mnt} \mid H_{t-1}, z_{nt}) = 0$*, then* $\mu_{gt}(z_{nt} \mid H_{nt-1}, z^*_{nt})$ $= \underline{\mu}$ *if* $h_{nt} < \underline{h}_{nt}$*, and* $\mu_{gt}(z_{nt} \mid H_{nt-1}, z^*_{nt}) = \overline{\mu}$ *if* $h_{nt} > \overline{h}_{nt}$*. That is, if a worker's hours are above the highest hours a worker of his type works, he is believed to be the type who works the most hours. If his hours are below the lowest hours, he is believed to be the lowest type given the observable characteristics.*[17]

The above propositions characterize the employers' optimal contracts given their beliefs. These beliefs, in equilibrium, satisfy Bayes' law and are consistent with workers' labor supply. A formal proof is in Appendix A, but the intuition is as follows. Because of the free entry assumptions, all offers yield zero expected payoffs over time. Earnings are determined by competitive bidding; thus, each period an outside employer offers a contract, which determines the earnings. The solution is obtained by solving backwards. In the worker's final period, T, each outside firm offers $y_{mt}(h_{nt}, H_{nt-1}, z^p_{nt}) - \gamma_m$, which is the worker's marginal product in a new firm. Consider the period prior the final period $(T-1)$. Taking into account the earnings in the final period and the probability the worker will remain in the firm (leaving the current employer a rent of γ_m), and since the current productivity net of hiring cost is $y_{mt}(h_{nT-1}, H_{nt-2}, z^p_{nt}) - \gamma_m$, the salary that makes the expected profit over the

[17] Note that the prior belief $\widetilde{\mu}(z_{nt} \mid .)$ is always strictly positive. This is because given any possible history, $F(z_{nt+1} \mid z_{nt}) > 0\ \forall z_{nt+1}, z_{nt}$

two periods zero is as described in equation 15. Earnings is computed in the same fashion for all periods.

It is worth pointing out the difference between this contract and the optimal contract in the symmetric information case. In the symmetric information case, each worker earns the expected productivity over time net of hiring costs. In the asymmetric information model, each worker earns the expected marginal product (net of costs) over the expected employment spell with the employer of a worker with the same observable characteristics. That is, each period the earnings equation is given by

$$S_{mg}(h_{nt}, H_{nt-1}, z_{nt}) = y_{mt}(h_{nt}, H_{nt-1}, z_{nt}^{p}) - \gamma_m + \beta\gamma_m \widetilde{p}_{gm,t+1}(H_{nt-1}, z_{nt}). \tag{16}$$

Thus consider two workers with the same publicly observable characteristics, z_{nt}^{*}, who differ with respect to unobservable characteristics, z_{nt}, and choose the same number of hours worked in equilibrium (asymmetric information). The worker who has a lower expected employment spell earns on average more than the productivity (net of costs), and the worker who is more "attached" is paid less because $\widetilde{p}_{gm,t+1}$, the belief that the worker will remain in the firm in the proceeding period, is higher than the actual probability for the worker with the lower *attachment* and lower than the probability of the worker who is more *attached* to the firm.[18]

Next, we show that an equilibrium exists. Establishing existence requires that a solution to the worker's problem given current and future expected earnings exists. To show that, it is sufficient (in addition to some standard regularity conditions) to prove that there exist beliefs about future participation that are self-fulfilled. Self-fulfilling beliefs mean that the expected probability that workers who accept a contract (given information firms currently have) will remain in the firm in proceeding period is indeed the proportion of workers who remain in the firm. By construction, this will imply that the beliefs about workers' types satisfy Bayes' law. Notice that firms and workers are forward looking: workers consider the effect of current choices on future earnings, and firms predict the probability workers will remain in the firm in all future periods (until retirement). Therefore, there is a system of equations in which the self-fulfilling beliefs are solved simultaneously.

Proposition 2 (Existence)

1. *For any state variables, there exists a unique solution to the worker's problem.*

[18] In the symmetric equilibrium, these workers would have faced their true expected employment spell and, therefore, their individual expected productivity over the employment spell, net of the hiring costs. Therefore, they would have worked different numbers of hours.

2. *There exists* $\{\widetilde{p}_{mn,2}(h_1, z^*_{n2}), \ldots, \widetilde{p}_{mn,T}(z^*_{nT}, H_{nT-1})\}$ *which satisfies the equilibrium conditions.*

3.1 Discussion

Before we proceed with the characterization of the equilibrium contract, we demonstrate the nature of this repeated adverse-selection problem. First note that *types* are correlated over time and z evolves stochastically according to the density $F_g(z_{t+1} \mid z_t)$.[19] It is typical for this class of dynamic adverse-selection model that when only spot contracts are feasible, individuals do not truthfully report their future-participation probability.[20] More information is revealed over time and, in general, there is *semipooling.* The payoff from working more hours today in standard models (with symmetric information) in which workers accumulate human capital on the job, increases the future payoff through the accumulation of *productive* labor-market experience. Therefore, workers with more experience are more likely to participate because the opportunity cost of not working is higher (see Altug and Miller, 1998). Under asymmetric information, there may be additional benefit to accumulating labor-market experience because employers use experience as a noisy signal to workers' unobservable characteristics. This creates incentive for workers to *pool.* The current salary of a worker who has a long expected employment spell is lower than in the symmetric information case and therefore this worker's incentive to work long hours is lower. In contrast, the reward to a worker with a lower expected employment spell for working more hours increases (compared to the full-information case). In general, the equilibrium, however, is not one in which all workers supply the same hours. Workers for whom the disutility of working is very high still work less than workers who have a low disutility of working. The degree of separation may vary across occupations and over time.

The model exhibits *discriminatory equilibria.*[21] Although we allow men and women to be different, our model gives rise to discriminatory equilibria even if men and women are initially identical (same distribution of preferences and skills) in all aspects that affect labor supply, participation, and production. As a benchmark, we describe these equilibria below. Suppose employers have different beliefs about the likelihood of future employment spells, then women may face lower wages than men (conditional on all other relevant characteristics).

[19] Therefore, $\widetilde{p}_{gm,t+1}(H_{nt-1}, z^*_{nt})$ is a function of the beliefs of the worker's type. For example, the prior beliefs in period 3, $\mu_{3g}(z_{n3} \mid H_{n2}, z^*_{n3})$, depend on the beliefs of the type in period 2, $\mu_{2g}(z_{n2} \mid (H_{n1}, z^*_{n2})$, and are derived using Bayes' rule.

[20] If the firm can commit to long-term contracts and workers cannot, workers still do not report their type truthfully, although the adverse selection is weaker.

[21] See Tirole (1997) 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 between firms and workers is random. In our framework, workers select into *contracts* offering different hours and earnings.

These beliefs are self-fulfilling in equilibrium and induce women and men to make different participation and labor-supply decisions. One feature that generates persistence in labor supply (apart from past labor market experience) is the nonseparability of leisure in the utility function (captured by θ). Suppose $\theta > 0$.[22] That is, there is complementarity between nonmarket hours across time. Then, workers who worked less hours or did not work at all may tend to have lower employment spells because the cost of participation is higher. Therefore, for given random shocks to utility from participation and nonparticipation, they are less likely to work.

Our model is also consistent with Becker (1965) and a model of home-production division (the statistical discrimination mechanism is similar to Coate and Loury, 1993). In particular, if married women face lower wages, the efficient division of home-production hours is that women put in more hours at home, leading to the solution to the decision problem in which the worker decides whether to participate in the labor force and how many hours to work there. This decision depends on the returns to working in the labor market (current and future expected earnings). If, systematically, wages are lower for women, they may accumulate, on average, less labor-market experience and more *home-production* hours. Therefore, in equilibrium, employers beliefs on labor-market participation are correct.

Proposition 3 (Discrimination) *Suppose that the distributions of skills, characteristics, and preferences are ex-ante identical for men and women, but that employers believe that women's employment spells in the proceeding periods is lower than men's. That is, for some* z_{nt}*, and* m $p_{1f}(z^*_{nt}) < p_{1r}(z^*_{nt})$*, then women face lower wages than men. These wage differences are larger in occupations with large* γ_m*. Beliefs are self-fulfilling in equilibrium, and therefore women work less hours and are less likely to participate.*

The costs of hiring new employees are different across occupations. If the initial beliefs, p_{1m}, for men are lower than for women, then in equilibrium women face lower wages and expected future wages (over time and across the different states) than men face. In equilibrium, these beliefs are self-fulfilling and generate different labor-market histories for the two groups.

The Bayesian updating formula shows persistence of initial beliefs. Furthermore, these beliefs are self-fulfilling in equilibrium and, thus, create different experience and occupation choices in the two groups.

[22] We do not, however, impose this restriction on the empirical model. Instead, we estimate the utility function parameters and verify that there is complementarity in leisure. Whereas a gender wage gap can occur when the two groups are ex-ante similar, differences in the utility function parameter (higher θ_s), imply that for given wages, women's labor supply is smaller. Such utility differences may arise, for example, if women's productivity in home goods is higher.

In particular, the optimal hours are characterized by the following Euler equation:

$$\begin{aligned} h_{nt}^{*} &= \underset{h_{nt}^{*}\in(0,1)}{\arg\max}\Bigg\{u_0(z_{nt}) + u_1(z_{nt}, 1-h_{nt}) + \eta_n\lambda_t\sum_{m=1}^{M} d_{nmt}S_{gm}(h_{nt}, z_{nt}^{*}, H_{nt-1}, \mu_t) \\ &\quad + \beta E_t[p_{nt+1}V_{1nt+1} + (1-p_{nt+1})V_{0nt+1}]\Bigg\} \end{aligned} \tag{17}$$

Notice that for two *identical* workers with the same z_{nt}, who differ only by gender, the first two elements in the equation are similar, but the last two elements are different. The change in beliefs,

$$\frac{\partial\widetilde{p}_{gmt}(h_{nts}, H_{Tnt-1}, z_{nt}^{*})}{\partial h_{nt}},$$

may be gender specific and is part of the current earnings. Importantly, the last element, $E_t\left\{\frac{\partial V_{n,t+1}}{\partial h_{nt}} \mid z_{nt}, h_{nt} = h_{nt}^{o}\right\}$, which is the change in the continuation value from the marginal change in hours worked, may be different. If the returns to labor-market experience are sufficiently lower for women because they face a lower stream of wages across states and time, then in order for the FOC to hold, the current marginal disutility of working, $\frac{\partial u_1(z_{nr},1-h_{nt})}{\partial h_{nt}}$, is smaller.

This is intuitive: The current salary for any level of hours is higher for men, and working more hours more hours increases labor-market experience, which is rewarded when a worker is planning to work in the future. However, it is possible that for some workers in some states,

$$\frac{\partial\widetilde{p}_{gmt}(h_{nts}, H_{Tnt-1}, z_{nt}^{*})}{\partial h_{nt}}$$

is high and reduces the effect of the lower continuation value.

Whereas discrimination can be a result of pure coordination failure (in which case, if there exists a unique solution given fixed beliefs and there is no multiplicity, the two groups have the same outcomes), our model may exhibit discriminatory equilibrium due to cross-group (gender) effects. Moro and Norman (2004) analyze a statistical discrimination model in which a discriminatory equilibrium may arise as a result of complementarities in production. That is, if we have complementarities in the utility function, consumption depends on such household characteristics as spouse income. Once complementarity is established between the hours (participation) women and men work (via the utility function), a discriminatory equilibrium (asymmetric equilibrium) may establish itself, even if there is no coordination failure. If men are believed to participate more and earn more, women (married) have higher consumption and work less. It is important to note that although cross-gender complementarities exist through household consumption, this affects not only married women. The reason is that the model is dynamic, and single individuals take into account they may marry in the future, thus, household complementarities affect them through expectations.

3.2 Empirical Properties

Next, we discuss some implications of the model.

Corollary 1 *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 wages for men and women with the same observable characteristics; the gender wage gap in these occupations may be smaller. To sum, if employers believe that women are less likely to remain in the firm in the future, women may lower labor-market experience. Thereafter, they will sort into occupations that have lower costs of hiring workers and lower returns to experience.

Corollary 2 *The effect of initial beliefs on wages declines over time. Thus, for a given cohort, conditional on all observable characteristics (to the employer), the gender wage gap (for men and women with the same observable characteristics) declines with experience.*

This is an immediate implication of Bayesian learning. Over time, more information about labor-market participation and labor supply arrives. Therefore, the effect of the initial beliefs becomes smaller. The only wage component that generates difference in wages for equally productive men and women (with the same employment history) is the beliefs on future participation. It is interesting to notice that when workers are young, all workers with similar publicly observable characteristics face the same earnings, over time, their choices and histories reveal information, and therefore, earnings dispersion increases.

Next, we discuss the dynamic evolution of the gender wage gap and the factors that drive changes.

Corollary 3 *According to our model, the following exogenous (outside the model) changes could account for the narrowing in the observed gender wage gap over time.*

1. *Differences in education across the different cohorts*
2. *Occupation-specific aggregate productivity shocks*
3. *Demographic changes which affect the distribution,* $F(z_{nt+1} \mid z_{nt})$
4. *Changes in production costs of domestic goods*

Over time, women's educational attainment has increased and, therefore, beliefs about women's labor-market participation increase. Since education is constant for each individual, change in educational composition explains only earnings-gap differences across cohorts. The rest of the factors can account for changes in the earnings gap within cohorts. Suppose that there is an increase in overall productivity within an occupation. Such an increase affects the wages of all workers because $y_{mt}(h_{nt}, H_{nt-1}, z^*_{nt})$ increases, but if men's participation rate is high, beliefs about women's participation may increase women's wages relative to men's wages. This increase in wage will result in a bigger increase in labor supply and participation of women.

The third factor, changes in demographics (such as a decline in fertility or an increase in the divorce rate), affects beliefs about future participation. Lastly, if home production becomes cheaper over time (due to technological changes), the cost of participation in the labor market is reduced, possibly increasing participation. Since women are less likely to participate than men, changes in costs of participation may affect the relative wage because changes in beliefs about participation will raise women's wages more than they will raise men's wages.

It is important to notice that the equilibrium characterization is for each cohort separately. In the data, we observe several overlapping cohorts. A worker's cohort is an observable characteristic. Therefore, for workers who are identical in all observable characteristics except for the cohort they belong to, it is possible that employers' initial beliefs will be different. 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 can capture that.

4 Identification and Estimation

There are two types of multiple equilibria in this model that we have to consider before taking it to data. The first is the standard multiple equilibria that arise because it is a model of imperfect information; a given equilibrium is supported by specific off-equilibrium beliefs and different off-equilibrium beliefs may give rise to a different equilibrium. We never observed off-equilibrium behavior, but the earnings equation holds for any beliefs of histories that are on the off-equilibrium path. Hence we will be estimating a particular equilibrium outcome observed in the data.

The second type of multiple equilibria is more central to the estimation of our model. Statistical discrimination means that employers choose to believe that one group is less attached that other and these beliefs are then self-fulfilling in equilibrium. For any given belief, there is a different equilibrium. In the standard statistical discrimination model, an identification problem of the following form arises: Given the observed wage distribution and human-capital investment, an econometrician is trying to recover the preference parameters and beliefs about an unobserved variable that affects productivity. The source of the identification problem is that there may be more than one combination of preference parameters and beliefs that could have generated the observed wage distribution (see Moro, 2003, and Antonovics, 2004, for detail of this problem). Our model, however, does not have that problem. This is because our beliefs are about next-period participation probability conditional on observed characteristics. Panel data, however, allows us to observe the individual's next-period participation decision. This must be correct across the population conditional on the observed characteristics. Therefore we can nonparametrically identify these beliefs if the following assumption holds:

> **A1:(Equilibrium Selection)**: The data for each age–education cohort is generated by only one equilibrium.

This assumption is standard in the literature for estimating dynamic discrete games. It only rules out the possibility that for any given age–education cohort the data is generated by a mixture of two or more equilibria. It does not select any one equilibrium or restrict the type of equilibrium played across age–education cohorts.

This reduces our identification problem to recovering the preferences and technology parameters given the observed beliefs, conditional wage distribution, and labor-supply choices. As shown in Aguirregabiria (2005), these models are semiparametrically identified.

4.1 Outline of Estimation Strategy

Estimation proceeds in three steps. In the first step, we estimate the earnings and consumption equations using a panel data Generalized Method of Moment (GMM) estimation strategy. In the second step, we estimate the conditional choice probability and the firm's beliefs using a kernel-density–based nonparametric estimation strategy. In the third step, we use the estimates of the marginal utility of wealth, the production functions, and the switching cost from the first-step estimation, along with estimates of the conditional choice probabilities and the firms' beliefs in a set of moment conditions derived from the workers' optimization problem. We then minimize this GMM criterion function to obtain estimates of the utility function. We will describe the estimation strategy more in detail in what follows below.

Before we describe the details of the three estimation steps, we begin with a description of the basic preference estimating equations. First, note that the optimal participation decision in equation (14) implies that the equilibrium choice probabilities of working, $p(d^o_{nt} \mid \omega_{nt})$, must satisfy

$$p(d^o_{nt} \mid \omega_{nt}) = \Pr\{V_1(\omega_{nt}) + \varepsilon_{1nt} \geq V_0(\omega_{nt}) + \varepsilon_{0nt}\}, \tag{18}$$

where ω_{nt} are the econometrician's observed state variables, which also includes the firm's beliefs that the worker will remain in the firm next period. A very useful insight of 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 choice-specific value functions, $V_1(\omega_{nt}) - V_0(\omega_{nt})$.[23] We let $Q : R \rightarrow (0,1)$ denote the mapping from the choice-specific value function to the conditional choice probabilities. That is,

$$p(d^o_{nt} \mid \omega_{nt}) = Q(V_1(\omega_{nt}) - V_0(\omega_{nt})). \tag{19}$$

The inverse exist and gives

$$V_1(\omega_{nt}) - V_0(\omega_{nt}) = Q^{-1}(p(d^o_{nt} \mid \omega_{nt})). \tag{20}$$

This mapping, $Q(.)$, is only a function of the unobserved state variables, ε_{0nt} and ε_{1nt}. Proposition 1 of Hotz and Miller (1993) also states that there exists a mapping $\varphi_k : [0,1] \rightarrow R$, that measures the expected value of the unobservable in the current utility, conditional on action $k \in \{0,1\}$:

$$\varphi_k(p(\omega_{nt})) \equiv E[\varepsilon_{knt} \mid \omega_{nt}, d^o_{nt} = k]. \tag{21}$$

Given these two results, we can obtain the difference between the two conditional valuation functions once we know the distribution function of the unobserved state variables. However, we are interested in reformulating the problem as a function of only the primitives of our original problem, the utility functions and the equilibrium wage equation. Hence, we need, in addition, a representation of our conditional valuation functions as a function of only the model's primitives and the conditional probabilities. To do this, we need the following assumption, which states that the production and utility functions depend only on the most recent work history.

> **A2**: (Finite State Dependence): The production and utility functions depend only on a finite history of past work decisions, $H^p_{nt-1} = (h_{nt-\rho}, \ldots, h_{nt-1})'$.

This assumption states that the payoff-relevant history is finite. However, it still leaves

[23] This equation is central to estimation in a number of papers including, Hotz et al. (1993); 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 (2004); and Bajari and Hong (2005), among others.

the possibility that the information-relevant history includes the workers' complete work history. Note that the firm's set of beliefs is an aggregator of information, and hence helps to reduce the state space for the problem. Since the information-relevant history only enters into the firm's beliefs, and the beliefs are an aggregator that enters into the worker's state space, the worker's state space has finite dependence also.

First define vectors $\omega_{ont}^{(r)}$ and $\omega_{1nt}^{(r)}$, as

$$\omega_{ont}^{(r)} = (h_{nt-\rho+r}, \ldots, h_{nt-1}, 0, 0, \ldots, 0, z'_{nt+r}, \widetilde{p}_{1n,t+1+r}, \ldots, \widetilde{p}_{Mn,t+1+r})', \tag{22}$$

and

$$\omega_{1nt}^{(r)} = (h_{nt-\rho+r}, \ldots, h_{nt-1}, h_{nt}^{*}, 0, \ldots, 0, z'_{nt+r}, \widetilde{p}_{1n,t+1+r}, \ldots, \widetilde{p}_{Mn,t+1+r})' \tag{23}$$

for $r = 0, \ldots, \rho$. The vector $\omega_{ont}^{(r)}$ is the state at date $t + r$ for an individual who has accumulated the work history $(h_{nt-\rho+r}, \ldots, h_{nt-1})'$ up to period t, and then chooses not to participate at date t and for $r - 1$ periods following period t. Similarly, $\omega_{1nt}^{(r)}$ is the vector for an individual who has accumulated work history $(h_{nt-\rho+r}, \ldots, h_{nt-1})'$ up to period t and chooses to participate and supply hours h_{nt}^{*} at date t, but chooses not to participate for $r-1$ periods following period t.

A restatement of proposition 1 of Altug and Miller (1998) p. 64 in our context gives us the following representation. Suppose we define the primitives of our problem as follows:

$$U_k(\omega_{nt}) = \begin{cases} u_1(z_{nt}, 0) & \text{for } k = 0 \\ u_0(z_{nt}) + u_1(z_{nt}, 1 - h_{nt}^{*}) + \eta_n \lambda_t \sum\limits_{m=1}^{M} d_{nmt} S_m(z_{nt}^{*}, h_{nt}^{*}, \widetilde{p}_{mn,t+1}) & \text{for } k = 1. \end{cases} \tag{24}$$

and let $p_{knt}^{(r)} \equiv \Pr(d_{nt+1}^{o} = 1 \mid \omega_{knt}^{(r)} = \omega)$ for $k \in \{0, 1\}$ and $r = 0, \ldots, \rho$. Then for $k \in \{0, 1\}$, the conditional valuation functions can be expressed as

$$\begin{aligned} V_k(\omega_{nt}) \quad = \quad & U_k(\omega_{nt}) + E_t \Bigg\{ \sum_{r=1}^{\rho} \beta^r \left[U_0\left(\omega_{knt}^{(r)}\right) + \varphi_0\left(p_{knt}^{(r)}\right) \right. \\ & + \left. p_{knt}^{(r)} \left(Q^{-1}\left(p_{knt}^{(r)}\right) + \varphi_1\left(p_{knt}^{(r)}\right) - \varphi_0\left(p_{knt}^{(r)}\right) \right) \right] \\ & + \beta^{\rho+1} \left[V_0\left(\omega_{knt}^{(\rho+1)}\right) + \varphi_0\left(p_{knt}^{(\rho+1)}\right) + p_{knt}^{(\rho+1)} \right. \\ & \times \left. \left(Q^{-1}\left(p_{knt}^{(\rho+1)}\right) + \varphi_1\left(p_{knt}^{(\rho+1)}\right) - \varphi_0\left(p_{knt}^{(\rho+1)}\right) \right) \right] \Bigg\} \end{aligned} \tag{25}$$

The difference between the conditional valuation functions that characterize the participation decision does not depend on the unknown valuation function any more since $\omega_{nt}^{(\rho+1)}$ for

$k \in \{0,1\}$ are both equal to $(0,\ldots,0,z'_{nt+\rho},\widetilde{p}_{1n,t+1+\rho},\ldots,\widetilde{p}_{Mn,t+1+\rho})'$. Hence, equation (20) implies that the conditions for an optimal participation decision depend on state probabilities and payoffs for 2ρ dates into the future. Note that the valuation function now only depends on the utility function, the distribution of the utility shocks (through $Q^{-1}()$ and φ_k), and the conditional choice probabilities in a finite number of counterfactual states. We can now apply an estimation strategy that does not require the computation of the valuation functions and hence does not suffer from the problem of multiple equilibria.

4.2 First Step: Estimation of the Consumption and Earnings Equations

Suppose the econometrician has access to panel data of an age–education cohort of individuals ($n = 1,\ldots,N, t = 1,\ldots,T$). During each time period, the econometrician observes the actions and state variables except any idiosyncratic shocks to utility or any individual-specific, time-invariant state variables ($S_{nt}, d_{nt}, d_{nmt}, h_{nt}, c_{nt}, z_{nt}, z^*_{nt}$). Note that we are assuming that the econometrician observes the private information z^*_{nt} from survey data.

In the first step, we use the Euler equation for consumption (8) to form the moment condition:

$$E\left[\frac{\partial u_2(c_{nt},z_{nt},\varepsilon_{2nt},\theta_c)}{\partial c_{nt}} - \eta_n\lambda_t \;\middle|\; z_{nt}\right] = 0 \tag{26}$$

Here we are assuming that the functional form of $u_2()$ is known up to a finite-dimensional parameter vector, θ_c. Based on equation (26), the econometrician can estimate θ_c and $\eta_n\lambda_t$ up to a proportionality constant. In fact, we assume that

$$u_2(c_{nt},z_{nt},\varepsilon_{2nt},\theta_c) = \exp(z'_{nt}B_4 + \varepsilon_{2nt})c^{\alpha}_{nt}/\alpha.$$

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

$$(1-\alpha)^{-1}\triangle\varepsilon_{2nt} = \triangle\ln(c_{nt}) - (1-\alpha)^{-1}\triangle z'_{nt}B_4 + \triangle(1-\alpha)^{-1}\ln(\lambda_t) \tag{27}$$

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 (27), 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:

$$(1-\alpha)^{-1} \ln(\eta_n) \equiv \sum_{t \in T_1} \left[\ln(c_{nt}) - (1-\alpha)^{-1} z'_{nt} B_4 + (1-\alpha)^{-1} \ln(\lambda_t)\right] / T_1 \tag{28}$$

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 (28), 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.

In the first step, we also use the free-entry condition from *Proposition 1 (Optimal Contract)* to form the following orthogonality conditions.

$$E_t\{d_{nmt}[S_{nt} - y_{mt}(h_s, H_{s-1}, z_{ns}^p, \theta_e) + \gamma_m - \beta \gamma_m d_{n,mt+1}] \mid z_{nt}^*, H_{n,t-1}, h_{nt}^*\} = 0. \tag{29}$$

Again we are assuming that the functional form of $y_{mt}()$ is known up to a finite-dimensional parameter vector θ_e. Based on equation (29), the econometrician can estimate θ_e and $\beta\gamma_m$.

We assume the following functional form for our production function.

$$y_{mt}(h_s, H_{s-1}, z^p_{ns}, \theta_e) = \pi_{0mt} + \pi_{m1}h_{nt} + \pi_{m2}h^2_{nt} + \sum_{r=1}^{\rho} \pi_{m3r}h_{nt-r} + z^{*'}_{nt}B_{m5} + \nu_n$$

This production function allows for occupation-specific aggregate shocks and general human capital. However, it's rate of return is different across occupations. There is an unobserved individual-specific component, ν_n, which is completely general in nature. Since all the information set in equation (29) is public at period t, we have

$$E_t\{d_{nmt}d_{nmt-1}[\triangle S_{nt} - \triangle\pi_{0mt} - \pi_m \triangle HC_{nt} - \triangle z^{*'}_{nt}B_{m5} - \beta\gamma_m \triangle d_{n,mt+1}] \mid z^*_{nt}, H_{n,t}, h^*_{nt}\} = 0, \tag{30}$$

where $\triangle HC_{nt} = (\triangle h_{nt}, \triangle h^2_{nt}, \triangle h_{nt-1}, \ldots, \triangle h_{nt-\rho})'$ and $\pi_m = (\pi_{m1}, \pi_{m2}, \pi_{m31}, \ldots, \pi_{m3\rho})$.

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

$$\Theta_{em} = \begin{pmatrix} \pi_m \\ B_{m5} \\ \beta\gamma_m \\ \triangle\pi_{0m2} \\ \vdots \\ \triangle\pi_{0mT} \end{pmatrix}.$$

We also define $Y_{mn} = (d_{nm2}d_{nm1} \triangle S_{n2}, \ldots, d_{nmT}d_{nmT-1} \triangle S_{nT})'$ as a vector of endogenous variables and X_{mn}, the exogenous variables,

$$X_{mn} = \begin{bmatrix} \triangle x'_{m2} & D_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \triangle x'_{mT} & 0 & \ldots & D_T \end{bmatrix}$$

where $\triangle x'_{mnt} = d_{nmt}d_{nmt-1}(\triangle h_{nt}, \triangle h^2_{nt}, \triangle h_{nt-1}, \ldots, \triangle h_{nt-\rho}, \triangle z^{*'}_{nt}, \triangle d_{n,mt+1})$. Letting Z_n be the matrix of conditioning variables

$$Z_n = \begin{bmatrix} z^{*\prime}_{n2} & H_{n2} & h_{n2} \\ \vdots & \vdots & \vdots \\ z^{*\prime}_{nT} & H_{nT} & h_{n2T} \end{bmatrix}$$

and using equation (30), one can obtain a set of orthogonality conditions:

$$E\left[\left(Y_{mn} - X_{mn}\Theta_{em}\right)Z_n\right] = 0,$$

which can be exploited to estimate Θ_{em} using an optimal instrumental-variable technique. The aggregate effect and fixed effect in the earnings equation is estimated in a similar way to those in the consumption equation.

4.3 Second Step: Nonparametric Estimation of Choice Probabilities and Beliefs

The equilibrium earnings equation and the alternative representations for the Euler and participation conditions require estimates of the beliefs and the conditional choice probabilities. First, $\widetilde{p}_{mnt}$ are computed as a nonlinear regression of the next-period participation index d_{mnt+1} on today's public information variables, z^*_{nt}, work histories, H_{nt-1}, and hours worked, h_{nt}, conditional on working today in occupation m. Let $X_{nt} = (z^*_{nt}, H_{nt-1}, h_{nt}, \nu_n)$ and define $J_1[\delta_{1N}^{-1}(X_{nt} - X_{n's})]$ as the normal kernel. δ_N is the bandwidth associated with each argument. The nonparametric estimate of $\widetilde{p}_{mnt}$, denoted $\widetilde{p}^N_{mnt}$, is computed using the kernel estimator,

$$\widetilde{p}^N_{nmt} = \frac{\sum_{n'=1,n'\neq n}^{N}\sum_{s=1}^{T-1} d_{mn's}d_{mn's+1}J_1[\delta_{1N}^{-1}(X_{nt} - X_{n's})]}{\sum_{n'=1,n'\neq n}^{N}\sum_{s=1}^{T-1} d_{mn's}J_1[\delta_{1N}^{-1}(X_{nt} - X_{n's})]}. \tag{31}$$

The derivative is then estimated using the standard nonparametric derivative kernel estimator (see Pagan and Ullah, 1999). Theoretically, H_{nt-1} includes the complete working history of the worker, including number of hours worked and occupation. In practice, for older workers the history could be very large, and therefore, our estimator would suffer from the curse of dimensionality. In order to obtain, the optimal history length for information we use a cross validation procedure to find the optimal information history dependence.

The conditional choice probabilities, p_{nt}, are computed as nonlinear regressions of the participation index, d_{nt}, on the current state, $\omega^N_{nt} \equiv (z'_{nt,} H_{nt}, \widetilde{p}^N_{1nt}, \ldots, \widetilde{p}^N_{Mnt}, \eta^N_n \lambda^N_t)'$, where the N superscript denotes an estimated quantity. Define $J\left[\delta_N\left(\omega^N_{nt} - \omega^N_{n's}\right)\right]$ as the normal kernel and δ_N as the bandwidth associated with each argument. The nonparametric estimate of p_{nt}, denoted p^N_{nt}, is computed using the kernel estimator,

$$p^N_{nt} = \frac{\sum_{n'=1,n'\neq n}^{N}\sum_{s=1}^{T} d_{n's}J\left[\delta_N^{-1}\left(\omega^N_{nt} - \omega^N_{n's}\right)\right]}{\sum_{n'=1,n'\neq n}^{N}\sum_{s=1}^{T} J\left[\delta_N^{-1}\left(\omega^N_{nt} - \omega^N_{n's}\right)\right]}. \tag{32}$$

The conditional choice probabilities, $p_{knt}^{(r)}$, are also computed as nonlinear regressions of a participation index on the appropriate state variables. Define

$$d_{knt}^{(r)} = [1-(1-k)d_{nt-r} - k(1-d_{nt-r})] \prod_{l=1}^{r-1}(1-d_{nt-l}), \qquad k \in \{0,1\}.$$

Notice that $d_{1nt}^{(r)} = 1$ if the person participated at $t-r$, but did not participate for the preceeding $r-1$ periods. Similarly, $d_{0nt}^{(r)} = 1$ if the person did not participate between $t-r$ and $t-1$. Thus, $d_{knt}^{(r)}$ is an index variable that allows us to condition on the behavior of individuals with the labor-market histories defined by $H_{knt}^{(r)}$. The conditional choice probabilities, $p_{knt}^{(r,N)}$, are computed as

$$p_{knt}^{(r,N)} = \frac{\sum_{n'=1,n'\neq n}^{N} \sum_{s=1}^{T} d_{n's} d_{kn's}^{(r)} J\left[\delta_N^{-1}\left(\omega_{knt}^{(r,N)} - \omega_{n's}^{N}\right)\right]}{\sum_{n'=1,n'\neq n}^{N} \sum_{s=1}^{T} d_{kn's}^{(r)} J\left[\delta_N^{-1}\left(\omega_{knt}^{(r,N)} - \omega_{n's}^{N}\right)\right]}, \tag{33}$$

where $\omega_{knt}^{(r,N)} = (H_{knt}^{(r)}, z'_{nt+r}, \widetilde{p}_{n1t+r}^{N}, \ldots, \widetilde{p}_{nMt+r}^{N}, \eta_n \lambda_{t+r})'$ for $k \in \{0,1\}$ is the counterfactual state vector for individual n.

To evaluate the term $\partial p_{1nt}^{(r)}/\partial h_{nt}$, which appears in the Euler equation, define

$$f_{1nt}^{(r)} \equiv f_1\left(\omega_{1nt}^{(r)} \mid d_{nt+r} = 1\right) \tag{34}$$

as the probability density function for $\omega_{1nt}^{(r)}$, conditional on participating at date $t+r$. Likewise, let $f_{nt}^{(r)} \equiv f\left(\omega_{1nt}^{(r)}\right)$ be the related probability density that does not condition on participating in period $t+r$ for $r = 1, \ldots, \rho$. Denote their derivatives with respect to h_{nt}^{*} by $f_{1nt}^{'(r)}$ and $f_{nt}^{'(r)}$, respectively. It is straightforward to show that

$$\frac{\partial p_{1nt}^{(r)}}{\partial h} = \left[\frac{f_{1nt}^{'(r)}}{f_{1nt}^{(r)}} - \frac{f_{nt}^{'(r)}}{f_{nt}^{(r)}}\right] p_{1nt}^{(r)} \quad r = 1, \ldots, \rho. \tag{35}$$

We derive this expression using the fact that $p_{1nt}^{(r)}$ can be written as $p_{1nt}^{(r)} = \Pr\left(d_{nt+r} = 1 \mid \omega_{1nt}^{(r)}\right) = \Pr(d_{nt+r} = 1) f_{1nt}^{(r)}/f_{nt}^{(r)}$. Differentiating this expression with respect to h_{nt} yields the above expression. The nonparametric estimates of $f_{1nt}^{(r)}$ and $f_{nt}^{(r)}$ are defined, respectively, as the numerators and denominators of $p_{knt}^{(r,N)}$ in equation (33). The estimates of $f_{1nt}^{'(r)}$ and $f_{nt}^{'(r)}$ are obtained from the derivatives of the estimates, $f_{1nt}^{(r,N)}$ and $f_{nt}^{(r,N)}$, with respect to h_{nt} (Pagan and Ullah, 1999).

4.4 Third Step: Estimation of the Structural Parameter

Assume that $(\varepsilon_{0nt}, \varepsilon_{1nt})$ 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 $Q(p) = \sigma \ln[p/(1-p)]$, $\varphi_0(p) = \frac{\zeta}{\sigma} - \sigma \ln[(1-p)]$, and $\varphi_1(p) = \frac{\zeta}{\sigma} - \sigma \ln[p]$, where ζ is the Euler constant. Combining equations (20) and (25) along with the above expressions of $Q(p)$, $\varphi_0(p)$ and $\varphi_1(p)$, we obtain

$$\begin{aligned} \sigma \ln[p_{nt}/(1-p_{nt})] &= B_{0t} + z_{nt}' B_1 - z_{nt}' h_{nt} B_2 - \theta_0 \left(1 - l_{nt}^2\right) \\ &\quad + \eta_n \lambda_t \sum_{m=1}^{M} d_{nmt} \left[y_{mt}(h_{nt}, H_{nt-1}, z_{nt}^p, \theta_e) - \gamma_m + \beta \gamma_m \widetilde{p}_{mn,t+1} \right] \\ &\quad - \sum_{s=1}^{\rho} \theta_s h_{nt} \left(l_{nt-s} + \beta^s\right) + \sigma E_t \left[\sum_{s=1}^{\rho} \beta^s \ln \left(\frac{1 - p_{1nt}^{(s)}}{1 - p_{0nt}^{(s)}} \right) \right] \end{aligned} \tag{36}$$

Finally, combining the Euler equation for hours (13) and the alternative representation of value (25) along with the above expressions of $Q(p)$, $\varphi_0(p)$, and $\varphi_1(p)$, we obtain

$$\begin{aligned} & E_t \left\{ d_{nt} \left[\sigma \sum_{s=1}^{\rho} \beta^s \left(1 - p_{1nt}^{(s)}\right)^{-1} \nabla_{h_{nt}} p_{1nt}^{(s)} - z_{nt}' B_2 - 2\theta_0 l_{nt} - \sum_{s=1}^{\rho} \theta_s (l_{nt-s} + \beta^s) \right. \right. \\ & \left. \left. + \ \eta_n \lambda_t \sum_{m=1}^{M} d_{nmt} \left[\pi_{m1} + 2\pi_{m2} h_{nt} + \beta \gamma_m \nabla_{h_{nt}} \widetilde{p}_{mn,t+1} \right] \right] \right\} \\ = \ & 0, \end{aligned} \tag{37}$$

where $\nabla_{h_{nt}} \widetilde{p}_{mn,t+1}$ is the derivative of the beliefs with respect to current hours.

Note that from the first step, we have estimates of π_{m1}, π_{m2}, $\beta\gamma_m$, 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 second step yields estimates of p_{nt}, $p_{1nt}^{(s)}$, $\widetilde{p}_{mn,t+1}$, $\nabla_{h_{nt}} p_{1nt}^{(s)}$, and $\nabla_{h_{nt}} \widetilde{p}_{mn,t+1}$. We can form the moment conditions:

$$
\begin{aligned}
m_{1nt}\left(\Theta u, \Theta_c^{(N)}, \Theta_e^{(N)}, \psi^{(N)}\right) \;=\; & \sigma \ln\left[p_{nt}^{(N)} / \left(1 - p_{nt}^{(N)}\right)\right] - B_{0t} - z_{nt}' B_1 + z_{nt}' h_{nt} B_2 \\
& + \theta_0 \left(1 - l_{nt}^2\right) + \sum_{s=1}^{\rho} \theta_s h_{nt}(l_{nt-s} + \beta^s) \\
& - \sigma \sum_{s=1}^{\rho} \beta^s \ln\left(\frac{1 - p_{1nt}^{(s)(N)}}{1 - p_{0nt}^{(s)(N)}}\right) \\
& - \exp\left((1-\alpha)\phi_{nt}^{(N)}\right) \sum_{m=1}^{M} d_{nmt} \left[y_{mt}\left(h_{nt}, H_{nt-1}, z_{nt}^p, \theta_e^{(N)}\right)\right. \\
& \left. -\ \gamma_m^{(N)} + \beta \gamma_m^{(N)} \widetilde{p}_{mn,t+1}^{(N)}\right] \qquad (38)
\end{aligned}
$$

and

$$
\begin{aligned}
m_{2nt}\left(\Theta u, \Theta_c^{(N)}, \Theta_e^{(N)}, \psi^{(N)}\right) \;=\; & d_{nt} \left\{ \sigma \sum_{s=1}^{\rho} \beta^s \left(1 - p_{1nt}^{(s)(N)}\right)^{-1} \nabla_{h_{nt}} p_{1nt}^{(s)(N)} \right. \\
& - z_{nt}' B_2 - 2\theta_0 l_{nt} - \sum_{s=1}^{\rho} \theta_s (l_{nt-s} + \beta^s) \\
& + \exp\left((1-\alpha)\phi_{nt}^{(N)}\right) \sum_{m=1}^{M} d_{nmt} \left[\pi_{m1}^{(N)} + 2\pi_{m2}^{(N)} h_{nt}\right. \\
& \left.\left. +\ \beta \gamma_m^{(N)} \nabla_{h_{nt}} \widetilde{p}_{mn,t+1}^{(N)}\right] \right\},
\end{aligned}
$$

where $\psi^{(N)} = \left(p_{nt}^{(N)}, p_{0nt}^{(s)(N)}, p_{1nt}^{(s)(N)}, \widetilde{p}_{mn,t+1}^{(N)}\right)$ 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 (36) and (37). 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

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

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 estimator 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

$$\Theta_u^{(N)} \equiv \arg\min_{\Theta u} \frac{\sum_{n=1}^{N} m_{3n}^{(N)}\left(\Theta_u, \Theta_c^{(N)}, \Theta_e^{(N)}, \psi^{(N)}\right)\left(\Omega_n^{(N)}\right)^{-1} m_{3n}^{(N)}\left(\Theta_u, \Theta_c^{(N)}, \Theta_e^{(N)}, \psi^{(N)}\right)}{N}. \quad (39)$$

5 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. The variables used in the empirical study are h_{nt}, the annual hours work by individual n at date t; S_{nt}, the reported real annual earnings at t; c_{nt}, the real household food consumption expenditures; FAM_{nt}, the number of household members; $YKID_{nt}$, the number of children less than six years of age; $OKID_{nt}$, the number of children of ages between six and sixteen; AGE_{nt}, the age of the individual at date t; EDU_{nt}, the years of completed education of the individual at time t; NE_{nt}, W_{nt}, SO_{nt}, region dummies for northeast, west, and south, respectively; and MS_{nt}, denoting whether an individual is married or not. We also use variables related to the spouse of the individual if the individual is married. $SP.EDUC_{nt}$ and $SP.INCOME_{nt}$, the completed years of education and the labor-market income of the spouse of individual n in period t respectively. Individuals are classified into two different occupation categories, professional and nonprofessionals. We only keep white individuals between the ages of 25 and 65 in our sample. After eliminating missing variables 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 definition of the variables are described in greater detail in Appendix C.

Table 1 contains summary statistics of our main labor-market and human-capital variables. The participation rate for men is relatively constant over our sample period with a sight decline toward the end of the sample period. In contrast, the participation rate for women increased significantly over the sample period, starting around 54% in 1968 and increasing to 74% by 1992. The average annual hours worked by men is also relatively constant over our sample period, but the average annual hours worked by women increased by roughly 30% over the sample period, going from around 1,400 hours per year in 1968 to 1,800 per year in 1992. At the same time, the average annual earnings for men increased by roughly 18% over the period, going 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, going from US$16,200 in 1968 to US%24,100 in 1992. Note that the implied reduction in the gender earnings gap is less than that reported in the literature. This is because there are still significant differences between the hours worked by men and the hours worked by women. Figure 1 displays the median gender wage for the same period (wage is annual earnings divided by annual hours worked). There it is clear that the wage gap declined by around 30% over the period, going from around 40% in 1968 to around 28% in 1992. This is in line with other reports in the literature.[24] The percentage of women in the professional occupations increased by around 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. The gap in the average years of completed education between men and women has been almost completely erased by the early 1990s.

Table 2 contains summary statistics of our main demographic and wealth variables. The sample has aged and household size has declined, with the decline is most pronounced amongst young children. Roughly 80% of our sample is married through the sample. Household income has increased somewhat, but household consumption of food has declined. However, both food consumption and income per capita has increased over the sample period.

6 Empirical Results

The main purpose of estimating the consumption equation is to obtain estimates of the marginal utility of wealth for our main estimation equations. Therefore, we do not focus the discussion on these results. Table 3 contains the results from this estimation. These results

[24] It should be noted that because of the skewness of the distribution of earnings, the median earnings gap is less that the mean earnings gap. However, regardless of how it is measured, the earnings gap has declined less that the wage gap. In the results, we do our decomposition using the median wage gap to be comparable with the wage-gap literature.

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). Furthermore, we obtain *reasonable* estimates for our risk aversion parameter which is normally a problem in estimation of consumption equations (Altug and Miller, 1990; Gayle and Miller, 2004).

Figure 2 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. Our theoretical model implies that productivity shocks should have a more significant effect on female labor force participation than on men's participation, and therefore lead to a reduction in the gender earnings gap (see Corollary 3(ii)). These shocks should also lead to increased sorting of women into the professional occupations, as is documented in the literature (Lewis, 1996). We compute directly the effect of productivity shocks on the changes in the gender earnings gap below (Table 10).

The estimation results of the earnings equation is reported in Table 4. Consistent with Corollary 1, the professional occupations have significantly higher returns to labor-market experience than the nonprofessional occupations.[25] These results are consistent with the empirical fact that more women sort into the nonprofessional occupations than men. Additional evidence in support of Corollary 1 is that there is a larger cost of hiring a new worker, along with the higher returns to working less hours (part time) in the nonprofessional occupations. This can be seen by comparing the linear and quadratic terms in current hours.

Table 5 contains the estimates of the fixed cost of labor-force participation.[26] Here we find our first evidence in support of Proposition 4. The difference in the estimated coefficients for number of kids (both young and old kids) for men and women, is suggestive evidence that more women *specialize* in nonmarket 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 directly in our data. However, our results, using the number of kids as a proxy for home production hours, support this theory. There is no significant difference in the cost of participation for men and women with the same years of completed education. A larger number years of completed education raises the likelihood of working for men and women equally. The effect of marital status is highly nonlinear and depends on the education level of one's spouse. A married individual is more likely to participate in the labor force. A married women who is married to a more educated man, however, is less likely to participate. In contrast, a man who is married to a more educated women is more likely to participate in the labor force.

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

[26] In our model, unemployment is interpreted as a decision not to work. This is in keeping with the labor literature on female labor supply.

Table 6 contains the results of our estimates on the utility of leisure (or nonmarket production). Again, the estimates support the idea of specialization. However, the results are a little more subtle. For example, whereas women with kids are less likely to work in the labor market, conditional on working, they are more likely to work more hours; the opposite is true for men. The opposite is also true for education. That is, educated women are more likely to work in the labor market, but education has no significant effect on the likelihood 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 likelihood of participation, but it does increase men's likelihood of working more hours conditional on working in the labor market. Lastly, conditional on working in the labor market, marital status decreases the likelihood of working more hours. Conditional on working, having a more educated spouse increases the likelihood of working more hours for women but not for men. These results are also supportive of the idea of cross-group complementarities in the utility function (asymmetric equilibrium).

Table 7 contains the estimates for the time nonseparability in nonmarket hours. They show that there are significant complementarities between nonmarket hours across time for women. The results on complementarities between nonmarket hours across time for men are mixed. In particular, nonmarket hours for men are compliments one period back. However, they become substitutes two periods back. This further suggests that our model does not require a coordination failure in order to exhibit a "discriminatory equilibria." This could be the results of just an asymmetric equilibria that would generate self-fulfilling beliefs and different labor-market histories between men and women as discussed in section 3.1.

The fact that our structural estimates are consistent with our model's prediction does not mean that private information is quantitatively important or that the gender earnings gap is driven by discriminatory equilibria. This is even more problematic given that the estimated switching cost is not very high ($3,000 in professional and less, $1,000, in nonprofessionals). Although these numbers may be reasonable, they are still small relative to the earnings gap.

To investigate this, we decompose the earnings gap 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.[27] The wage gap for our sample is 87% and 76% for professionals and nonprofessionals, respectively. Our model predicts an earnings gap of 92% and 81% for professionals and nonprofessionals, respectively. Of the 92% predicted wage gap in the professional occupations, 60% is due to the difference in human capital, 12% is due to differences in firms' beliefs, 4% is due to differences in the estimated fixed effects,

[27] Where wage is compute, as earnings divided by hours worked.

and 10% is due to differences in education and age composition between men and women. In the nonprofessional occupation, of the 81% wage gap predicted by our model, 56% is due to differences in human capital, 9% is due to differences in employers' beliefs, 7% is due to differences in the estimated fixed effect, and 4% is due to differences in the education and age composition between men and women. Given how the fixed effect is estimated, one may be worried that it is capturing implicit discrimination which is not in our model. Given that it accounts, however, for only a small fraction of the predicted wage gap, we can safely ignore these other sources of possible discrimination.

Given that our model performs reasonable well in explaining the earnings gap, we are now in a position to look at 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 wage gap in both time periods and express the difference as a percentage of the median wage gap in the first period. The results are reported in Table 9. The raw changes in the wage gap over the period are 30% and 25% for professionals and nonprofessionals, respectively. Our model predicts changes of 28% and 22% for the two occupations, respectively. Therefore, our model is doing slightly better at predicting the change in the wage gap than the level. 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. Changes in firms' beliefs account for 8% and 6% of the changes in professional and nonprofessional, respectively, whereas education and age composition accounted for 25% and 29% of the changes in professionals and nonprofessionals, respectively.

Human capital is by far the most important factor in explaining both the wage gap and the changes in the wage 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. In order to disentangled these different effects on human-capital accumulation, we need 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 that the equilibrium we solve for is the one that actually played out in the data, the parameter values of our model may be consistent with many different equilibria. In order to get around this problem, we solve the model under two different situations. In both situations we know that there is a unique equilibrium. 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 situation is when the switching cost, γ_m, is zero. The second is when firms observe all the private information of the individual. Under both scenarios we can solve the model using backward induction.

We then calculate the effects of five sources of the changes in human-capital accumulationon on the earnings gap. These results are reported in Table 10. First, we calculate the effect of the hiring costs on the change in the gender wage gap. That is, we compute the change in the wage gap attributed to the friction of firm-specific hiring costs. Without costs of hiring new workers, earnings equal the worker's productivity, and the prospect of future participation does not affect them. To calculate the effect of hiring costs on the gender wage 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_m equal to zero.[28] Then we calculate the implied changes in the wage gap. Notice that we capture wage gap and changes in the earnings gap in a model in which earnings are the worker's productivity. In such a model, 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 as a percentage of the change attributed to human capital 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 the hiring costs. (That is, in a model without hiring costs, the change in human capital in professional occupation, would have been 38% smaller compared to the changes in human capital in our model.)

The second entry in Table 10 calculates the effect of the private information on the change in the wage gap. This can essentially be interpreted as the change in the earnings gap attributable to changes in patterns of discrimination. Recall that in our model, discrimination, as defined in the model, occurs only because employers do not have all the information workers have that is relevant to predicting employment spells. The effect of the asymmetric information is calculated as described above (the effect of the hiring costs calculation). However, instead of setting γ_m equal to zero, we solve the model backward, calculating the actual probability of working 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. This accounted for 12% and 13% of the changes in the professional and nonprofessional occupations, respectively. That is, the effect of changes in beliefs about next-period participation on human-capital accumulation is 12% and 13% smaller when information is symmetric. Note that the overall effect of the information friction (sta-

[28] Therefore, the earnings equation in the scenario is: $S_m(h_{nt}, H_{nt-1}, z^*_{nt}) = y_{mt}(h_{nt}, H_{nt-1}, z^*_{nt})$

tistical discrimination) is the effect on human capital plus the direct effect on earnings. In professional occupation it is 16% (0.12*0.67+0.08=0.16), and for non-professionals it is 14%.

To compute the effect of changes in demographic characteristics on the change in human capital, we again set γ_m equal to zero, holding all the demographic characteristics level—marital status, number of kids, years of completed education, and spouse education. We hold their transition probabilities at the 1974–1978 levels in both periods and repeat the calculations on the effect of hiring costs. By doing so, we shut down the effect of these characteristics on the earnings because they do not enter the production function. This accounts for 28% and 34% of the changes 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 (which we refer to as changes in home-production costs), we hold the aggregate shock to home production fixed at its 1974–1978 level in both time periods, while allowing demographic characteristics 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.

We again repeat the same exercise, holding the aggregate shock to market productivity 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.

From the decomposition, exercise we learn that private information, hiring cost, demographic changes, and changes in aggregate productivity in the market sector account for most of the changes in human capital. In turn, human capital accounts for most of the changes in the explained gender earnings gap over the period.

7 Conclusion

The focus of the paper is accounting for the changes in labor-market outcomes gap for males and females. Our estimates reveal that the increase in productivity (estimates of the year- and occupation-specific productivity shocks) over the years is larger in professional occupations than in nonprofessional occupations. Productivity increased by around 150% in professional occupations between the years 1975 and 1985, whereas it only increased by around 100% in nonprofessional occupations. Our model predicts that such an increase allows for relative gains for women causing an increase in female representation in professional occupations over time. We continue with a decomposition of the change in the gender gap. We find that 16% of the decline in the gender wage gap in professional occupations and 14% in nonprofessional occupations is due to changes in patterns of statistical discrimination (both the direct effect on wages and the indirect effect on labor supply and occupational sorting). We also find that Occupation-specific technological shocks and changes in family structure and education played an important role in the decline in the earnings gap. The estimation results do not support, however, the hypothesis that changes in home production technology explain the increase in women's labor-market participation.

Further extensions of our framework will include exploring the effect of changes in family structure on the gender wage gap. We find that changes in family structure are a significant factor driving the change in beliefs. In professional occupations, 28% of the change in human-capital accumulated (which accounts for 67% of the entire gap) is due to demographic changes (such as the decline in married women's fertility). For nonprofessionals, it is 34% (the human capital explains 65%). These changes drive belief changes about women's attachment to the labor force. In our model, these changes in family structure are assumed to be exogenous, and therefore, are identified as factors causing changes in beliefs, increases in the participation rate, and decline in the gender wage gap. Although our empirical findings suggest that changes in family structure may be important to further understanding the observed changes in the gender wage gap, these changes are endogenous to changes in earnings. Therefore, we cannot disentangle the causality relations. Inferring causality is beyond the scope of this paper.

8 Appendices

8.1 Appendix A: Theoretical Results

Proof of Proposition 1 (Employers' optimal strategies). This proof establishes that given the strategies and beliefs of the players, the contract in (15) is optimal. The free-entry condition implies that in equilibrium, the expected value of a vacancy in each occupation at any period, π_{tm}, is zero. Thus π_{tm} is the continuation value of hiring a new worker in occupation m in period t. Define π^e_{tm} by the continuation value of the current employer in occupation m in period t. That is, π^e_{tm} is the expected profits from employing a worker who was employed in the firm for more than one period. We use this to derive the optimal contract by solving backwards:

At time $t = T$ (the worker's final year), the free-entry condition that implies that for a new employer, expected profit from offering the worker a contract is zero. The expected profit from offering a contract, $s(h_t, H_{t-1,}z_t^*)$, is

$$\pi_T = y_{mT}(h_T, H_{T-1}, z_T^p) - S_T(h_T, H_{T-1}, z_{nT}^*) - \gamma_m = 0. \tag{40}$$

Therefore,

$$S_T(h_T, H_{T-1}, z_{nt}^*) = y_{mT}\,(h_T, H_{T-1}, z_T^p) - \gamma_m$$

The current employer's profit, substituting the earnings is

$$\pi_T^e = y_{mT}(h_T, H_{T-1}, z_T^p) - S_T(h_T, H_{T-1}, z_T^*) = \gamma_m \tag{41}$$

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

$$\begin{aligned}\pi_{T-1} = y_{mT-1}(h_{T-1}, H_{T-2}, z_{T-1}^p) \quad &- \quad \gamma_m - S(h_{T-1}, H_{T-2}, z_{T-1}^*) \\ &+ \quad \beta\widetilde{p}_{mTg}(z_{T-1}^*, H_{T-2}, h_{T-1})\pi_T^e = 0.\end{aligned}$$

Thus,

$$S(h_{T-1}, H_{T-2}, z_{nt}^*) = y_{mT-1}(h_{T-1}, H_{T-2}, z_{T-1}^p) - \gamma_m(1 - \beta\widetilde{p}_{mTg}(z_{T-1}^*, H_{T-2}, h_{T-1})).$$

A current employer's profit in period $T-1$ is therefore

$$\begin{aligned}\pi_{T-1}^e = y_{mT-1}(h_{T-1}, H_{T-2}, z_n^p) \quad &- \quad S(h_{T-1}, H_{T-2}, z_{T-1}^*) \\ &+ \quad \beta\widetilde{p}_{mTg}(z_{T-1}^*, H_{T-2}, h_{T-1})\pi_T^e = \gamma_m.\end{aligned}$$

Solving backwards, at any period $s < T$, the free-entry condition implies

$$\begin{aligned}\pi_{T-s} = y_{mT-s}(h_{T-s}, H_{T-s-1}, z_n^p) \quad &- \quad \gamma_m - S(h_{T-s}, H_{T-s-1}, z_{T-s}^*) \\ &+ \quad \beta\widetilde{p}_{gmT-s+1}(z_{T-s}^*, H_{T-s-1}, h_{T-s})\pi_{T-s+1}^e = 0\end{aligned}$$

and, therefore,

$$S(h_{T-s}, H_{T-s-1}, z_n^*) = y_{mT-s}(h_{T-s}, H_{T-s-1}, z_n^p) - \gamma_m(1 - \beta\widetilde{p}_{gmT}(z_{T-s}^*, H_{T-s-1}, h_{T-s})).$$

Therefore, the current employer's profit is

$$\begin{aligned}\pi_{T-s}^e = y_{mT-s}(h_{T-s}, H_{T-s-1}, z_n^p) \quad &- \quad S(h_{T-s}, H_{T-s-1}, z_{T-s}^*) \\ &+ \quad \beta\widetilde{p}_{gmT-s+1}(z_{T-s}^*, H_{T-s-1}, h_{T-s})\pi_{T-s+1}^e = \gamma_m\end{aligned}$$

Given the beliefs and worker's strategy to accept the highest offer, and other firm's strategies, the above contract is optimal.

Next, we show that no possible deviation from the competitive rate can be profitable to any player. Each occupation offers all hours in the occupation's range. No worker accepts a lower wage ($d_{mnt} = 0$ if $S_m^{'}(h_s, H_{s-1}, z_s^*) < S_{m'}(h_s, H_{s-1}, z_s^*)$). By construction, the inability of both firms and workers to commit implies that offering a higher wage for any given hours and observables leads to a negative expected profit. Notice that every contract for every hour in any occupation is offered. They all yield zero profit and firms are indifferent. Any hours contract that is not offered provides an opportunity for a new firm to enter and attract a worker at a wage that yields positive profits. The worker's choice of participation and optimal hours is described in (14) and (13).

Next, we argue that the worker's strategy, given any hours she chooses to work is to accept the highest salary. A higher salary increases utility today and does not affect beliefs or salaries tomorrow (recall salary is not part of the observed employment history). Thus, for any hours worked, if offered a higher salary by a firm in the occupation, the worker will take the highest offer (switching employers within occupation does not affect beliefs). Off the equilibrium path, workers who work fewer than the minimal hours receive lower salaries and there are no gains from deviation (beliefs will not increase), there is no profitable deviation for working more than the largest amount of hours, as beliefs are not adjusted to be higher than the beliefs for hours worked more than the highest optimal hours on the equilibrium-path. ■

Proof of Proposition 3 (Existence). Existence of a solution to the worker's consumption and hours problem follows immediately from continuity and strict concavity of the utility function the fact that there is a solution to the worker's problem for any contract offered.

First, notice that since choice of hours worked and participation depend on current salary, they are affected by the firm's beliefs. Consider the final period T. The salaries are independent of the firm's beliefs as all factors affecting production are observable. The following functions summarize the worker's strategy (participation decision and hours worked, respectively),

$$Q_T(z_{nT}, H_{nT-1}(S_{nT-1}(h_{nT-1}, \widetilde{p}_{mn,T}), \ldots, S_{n1}(h_{n1}, \widetilde{p}_{mn,2})), h_T^*)$$
$$I_T(z_{nT}, H_{nT-1}(S_{nT-1}(h_{nT-1}, \widetilde{p}_{mn,T}), \ldots, S_{n1}(h_{n1}, \widetilde{p}_{mn,2})), h_T^*)$$

and $\forall 1 < t < T$,

$$Q_t(z_{nt}, H_{nt-1}(S_{nt-1}(h_{nt-1}, \widetilde{p}_{mn,t}), \ldots, S_{n1}(h_{n1}, \widetilde{p}_{mn,2})), h_t^*)$$
$$I_T(z_{nt}, H_{nt-1}(S_{nt-1}(h_{nt-1}, \widetilde{p}_{mn,t}), \ldots, S_{n1}(h_{n1}, \widetilde{p}_{mn,2})), h_t^*).$$

In period $T-1$,

$$\begin{aligned}\widetilde{p}_{mn,T} \;=\; & \int Q_T(z_{nT}, H_{nT-1}(\widetilde{p}_{mn,T}, \ldots, \widetilde{p}_{mn,2}), h_T^*) I_T(z_{nT}, H_{nT-1}(\widetilde{p}_{mn,T}, \ldots, \widetilde{p}_{mn,2}), h_T^*) \\ & \mu_T(z_{nT} \mid H_{nT-1}, z_{nT}^*)\, \mathrm{d}z_{nT}.\end{aligned}$$

Next, in order to solve for $\widetilde{p}_{mn,t}$, notice that it appears in Q_T and I_T on the right hand side for any period t. In particular, $\widetilde{p}_{mn,t+1}$ is the solution to

$$\begin{aligned}\widetilde{p}_{mn,t+1} \;=\; & \int_{z_t} \Bigg\{ \int_{z_{t+1}} f(z_{nt+1} \mid z_{nt}) Q_{t+1}(z_{nt+1}, H_{nt}(\widetilde{p}_{mn,t+1}), h_{t+1}^*(\widetilde{p}_{mn,t+2}) \\ & I_{t+1}(z_{nt+1}, H_{nt}(\widetilde{p}_{mn_{nt+1},\ldots,}\widetilde{p}_{mn2}), h_{nt+1}^*(\widetilde{p}_{mn,t+2}), S_{nt+1}(h_{nt+1}^*) \underline{h}_{mt+1}(H_{nt}, z_{nt}^p), \\ & \overline{h}_{mt+1}(H_{nt}, z_n^p)\, \mathrm{d}z_{n_{nt+1}} \Bigg\} \mu_t(z_{nt} \mid H_{nt-1}, z_{nt}^*)\, \mathrm{d}z_{nt}\end{aligned}$$

A) Note that $\widetilde{p}_{mn,t+1} : [0,1]$, and that the left hand side, is also defined on the interval $[0,1]$. Thus, continuity of the RHS suffices to guarantee a solution to each one of the equations separately. We then show that a solution exists to the system of beliefs $\widetilde{p}_{mn,t+1}, \ldots \widetilde{p}_{mn,2}$ simultaneously.

B) To show continuity: Let $\underline{z}_{mt+1}$ be the *marginal type* for which $h^*_{mt+1}(H_{nt}, z^*_n, \underline{z}_{nt+1}) \equiv \underline{h}^*_{mt+1}(H_{nt}, z^*_{nt+1})$, and $\overline{z}_{mt+1}$ the type for which $h^*_{mt+1}(H_{nt}, z^*_n, \overline{z}_{mt+1}) \equiv \overline{h}^*_{mt+1}(H_{nt}, z^*_{nt+1})$. $h^*_{mt+1}(z_{nt+1}, .)$ is continuous and invertible in z_{nt+1} as the utility function is continuous. $\underline{z}_{mt+1}$ ($\overline{z}_{mt+1}$) is the lowest (highest) z, so a worker with history H_{nt} and characteristics z^*_{nt+1} chooses optimally to work $\underline{h}_{mt+1}$. Thus we can write, $\underline{h}^{*-1}_{mt+1}(H_{nt}, z^*_{nt+1}) = \underline{z}_{mt+1}$ and $\overline{z}_{mt+1} \equiv \overline{h}^{*-1}_{mt+1}(H_{nt}, z^*_{nt+1})$. Since $I_{t+1}(z_{nt+1}, H_{nt}(\widetilde{p}_{mn_{nt+1},\ldots,}\widetilde{p}_{mn2}), h^*_{nt+1}, S_{nt+1}(h^*_{nt+1}))$ is an indictor function that takes the value 1 when $h^*_{nt+1} \in [\underline{h}_{mt+1}(H_{nt}, z^*_{nt+1}), \overline{h}_{mt+1}(H_{nt}, z^*_{nt+1})]$ and 0 otherwise, we can rewrite the integral

$$\int_{z_{n_{t+1}}} f(z_{nt+1} \mid z_{nt}) Q_{t+1}(.) I_{t+1}(.)\, \mathrm{d}z_{n_{nt+1}}$$
$$= \int_{\underline{z}_{nt+1}(H_{nt}(\widetilde{p}_{mn,t+1}), z^*_n)}^{\overline{z}_{mt+1}(H_{nt}(\widetilde{p}_{mn,t+1}), z^*_n)} f(z_{nt+1} \mid z_{nt}) Q_{t+1}(h_t(\widetilde{p}_{mn,t+1}) I_{t+1}(.))\, \mathrm{d}z_{n_{nt+1}}$$

C) Since $h_t(\widetilde{p}_{mn,t+1})$ is continuous in $\widetilde{p}_{mn,t+1}$ and $Q_{t+1}(h_t(\widetilde{p}_{mn,t+1}).)$ is continuous in h_t, we only need to show that $\overline{z}_{mt+1}(H_{nt}(\widetilde{p}_{mn,t+1}), z^*_{nt+1})$ and $\underline{z}_{nt+1}(H_{nt}(\widetilde{p}_{mn,t+1}), z^*_{nt+1})$ are continuous in $\widetilde{p}_{mn,t+1}$. Let $m-1$ ($m+1$) denote the occupation of the worker if she chooses hours below (above) $\overline{h}^*_{mt+1}(H_{nt}, z^*_{nt+1})$ ($\underline{h}^*_{mt+1}(H_{nt}, z^*_{nt+1})$). These hours are determined by the following conditions,

$$\begin{aligned} f_m(\underline{h}^*_{mt+1}(H_{nt}, z^*_{nt-1})) - f_{m-1}(\underline{h}^*_{mt+1}(H_{nt}, z^*_{nt-1})) - \gamma_m + \gamma_{m-1} &= 0 \\ f_m(\overline{h}^*_{mt+1}(H_{nt}, z^*_{nt+1})) - f_{m+1}(\overline{h}^*_{mt+1}(H_{nt}, z^*_{nt+1})) - \gamma_m + \gamma_{m+1} &= 0. \end{aligned}$$

From the continuity of the production function in each occupation in all factors of production, $\underline{h}^*_{mt+1}(H_{nt}, z^*_n)$ and $\overline{h}^*_{mt+1}(H_{nt}, z^*_n)$ are continuous in h_{nt}. Therefore, there exists a solution to every period's beliefs separately.

Next we show that there exists a beliefs sequence that solves the following system of simultaneous equations (where g represent the RHS of the equations),

$$\begin{aligned} \widetilde{p}_{mnT} &= g(\widetilde{p}_{mnT}, \ldots, \widetilde{p}_{mn2}) \\ \widetilde{p}_{mnT-1} &= g(\widetilde{p}_{mnT-1}, \ldots, \widetilde{p}_{mn2}) \\ &\vdots \\ \widetilde{p}_{mn2} &= g(\widetilde{p}_{mn2}). \end{aligned}$$

The matrix is diagonal and nonsingular. Therefore, a solution exists. ■

8.2 Appendix B: Derivation of Asymptotic Variance

It is well known in the econometric literature that under certain regularity conditions, preestimation 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 preestimation. In order to conduct inference in this type of estimation, one has to correct the asymptotic variance for the preestimation. The method used for correcting the variance in the final step of estimation depends on whether the preestimation 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:

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

where $(Y_n - Z_n\Theta_c)\, Z_n^c$ is the orthogonality condition from the estimation of the consumption equation, $(Y_{mn} - X_{mn}\Theta_{em})Z_n$ 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 $\widetilde{d}_{nt}^j = [d_{mnt}, d_{mnt} d_{mnt+1}]'$ for the estimation of $\widetilde{p}_{mn,t+1}$. Therefore $\psi^{j(N)}(z^j) = \frac{1}{N} \sum_{n=1}^{N} \widetilde{d}_n^j K_{\delta_N}(z^j - z_n^j)$. 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 τ, greater than the dimension of z and $\psi_{10}(z) = f_z(z)$ is bounded away from 0.*

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

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

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) $E[Z'_{n'} Z_n] < \infty$, $E[X'_{mn} Z_n] < \infty$, $E[\exp((1-\alpha)\phi_{nt})] < \infty$, $E[z_{nt}] < \infty$, $E[y_{mt}(h_{nt}, H_{nt-1}, z^p_{nt}, \theta_e)] < \infty$, $\gamma_m < \infty$, $E[\nabla_{h_{nt}} \widetilde{p}_{mn,t+1}] < \infty$, $E[X_{mn}] < \infty$ for $m = 1, \ldots, M$;

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

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

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

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

where

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

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,

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

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

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

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

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

So that

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

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 $m(\omega,\psi^{(N)})$:

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

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. ■

8.3 Appendix C: Data Description

We used data from the Family, 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 only labor income, hours worked, and years of completed education 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 or 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 or 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 of the value of this variable for the year $t-1$ and 0.75 of 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 is 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. The age variable resulted in a loss of 462 individuals.

The race of the individual and the region where they are currently residing 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, Northcentral, 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 Oriental 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. 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 the 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 transfer 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 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.[29]

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 by the Chain-type price deflator for total personal consumption expenditures.

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

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

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

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

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

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

[7] 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.

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

[9] 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.

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

[11] Baron, J., D. Black and M. Lowenstein (1993), "Gender Differences in Training, Capital and Wages," *Journal of Human Resources 28(2):343-64.*

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

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

[14] Beffy, M., M. Buchinsky, D. Fougère, T. Kamionka, and F. Kramarz (2006): *The Returns to Seniority in France (And why are they Lower than in the United States?)*, IZA Discussion Paper No. 1935.

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

[16] Bowlus A. and Z. Eckstein (2002): "Discrimination and Skill Differences in an Equilibrium Search Model," *International Economic Review*, 43(4), 1309–1345.

[17] Buchinsky, M., D. Fougère, F. Kramarz, and R. Tchernis (2005): *Interfirm Mobility, Wages, and the Returns to Seniority and Experience in the U.S.*, IZA Discussion Paper No. 1521.

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

[19] Coate, S. and G. Loury (1993): "Will Affirmative Action Policies Eliminate Negative Stereotypes?" *American Economic Review*, 83(5), 1220–1240.

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

[21] Bowlus A. and Z. Eckstein (2002): "Discrimination and Skill Differences in an Equilibrium Search Model," *International Economic Review*, 43(4), 1309–1345..

[22] Eckstein, Z. and Nagypal (2004) "The Evolution of U.S. Earnings Inequality: 1961-2002, " *Federal Reserve Bank of Minneapolis Quarterly Review, Vol. 28, No. 2, December 2004, pp. 10-29.*

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

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

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

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

[27] Hotz, Miller, Sanders, and Smith (1994):"A Simulation Estimator for Dynamic Models of Discrete Choice," *Review of Economic Studies, Vol 64, 256-89*

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

[29] Laffont J., and J. Tirole (1988): "The dynamics of incentives contracts," *Econometrica, 56(5), pp. 1153-75*

[30] Lee, D. and K. I. Wolpin (2000): "Intersectoral Labor Mobility and the Growth of the Service Sector," *Econometrica*, 74, 1–46.

[31] Lee, D. and K. I. Wolpin (2006): *Accounting for Wage and Employment Changes in the U.S. from 1968–2000: A Dynamic Model of Labor Market Equilibrium*, PIER Working Paper 06-005.

[32] Lewis, G. (1996): "Gender Integration of Occupations in Federal Civil Service: Extent and Effects on Male–Female Earnings," *Industrial and Labor Relations Review*, 49(3), 472–483.

[33] Light and Ureta (1992),: "Panel Estimates of Male and Female Job Turnover Behavior: Can Female Nonquitters be Identified?" *Journal of Labor Economics 10(2), pp. 156-81*

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

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

[36] Moro, A. (2003): "The Effect of Statistical Discrimination on Black–White Wage Inequality: Estimating a Model With Multiple Equilibria," *International Economic Review*, 44(2), 467–500.

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

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

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

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

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

[42] Pakes, A., M. Ostrovsky, and S. Berry (2004): *Simple Estimators for the Parameters of Discrete Dynamic Games*, NBER Working Paper w10506.

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

[44] Phelps, E. (1972): "The statistical theory of racism and sexism," *American Economic Review 62: 659-61*

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

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

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 (0.25)	0.54 (0.50)	2,244 (631)	1,401 (731)	39.8 (24.9)	16.2 (10.8)	0.28	0.45	12.2 (3.4)	11.7 (2.6)
1969	0.96 (0.18)	0.60 (0.49)	2,240 (610)	1,371 (739)	41.2 (26.6)	16.2 (11.2)	0.29	0.48	12.1 (3.4)	11.7 (2.6)
1970	0.97 (0.18)	0.64 (0.48)	2,216 (593)	1,332 (758)	41.6 (26.5)	16.2 (11.0)	0.30	0.49	12.1 (3.4)	11.8 (2.6)
1971	0.96 (0.20)	0.63 (0.48)	2,175 (636)	1,382 (750)	41.7 (24.9)	17.0 (11.6)	0.32	0.48	12.2 (3.3)	11.8 (2.6)
1972	0.95 (0.21)	0.62 (0.49)	2,155 (636)	1,389 (728)	41.7 (26.2)	17.4 (11.8)	0.32	0.49	12.2 (3.3)	11.8 (2.6)
1973	0.96 (0.19)	0.60 (0.49)	2,188 (633)	1,411 (720)	43.2 (26.5)	17.8 (11.2)	0.30	0.47	12.2 (3.3)	11.9 (2.6)
1974	0.95 (0.23	0.62 (0.49)	2,130 (641)	1,424 (713)	43.2 (28.4)	18.1 (12.5)	0.32	0.47	12.2 (3.3)	11.9 (2.6)
1975	0.92 (0.27)	0.62 (0.49)	2,230 (641)	1,415 (726)	42.0 (30.6)	17.7 (11.9)	0.32	0.48	12.3 (3.2)	11.9 (2.6)
1976	0.92 (0.24)	0.62 (0.49)	2,092 (677)	1,395 (726)	40.9 (31.1)	17.8 (12.4)	0.35	0.49	12.3 (3.2)	12.0 (2.5)
1977	0.91 (0.27)	0.61 (0.49)	2,119 (668)	1,418 (706)	42.6 (31.0)	18.1 (12.4)	0.35	0.50	12.4 (3.1)	11.8 (2.5)
1978	0.87 (0.33)	0.62 (0.49)	2,115 (650)	1,454 (737)	44.2 (32.8)	18.6 (12.5)	0.33	0.46	12.4 (3.1)	12.0 (2.5)
1979	0.91 (0.29)	0.63 (0.48)	2,141 (675)	1,472 (711)	43.5 (30.3)	18.7 (12.7)	0.36	0.49	12.4 (3.1)	12.1 (2.5)
1980	0.91 (0.29)	0.65 (0.48)	2,112 (651)	1,450 (726)	42.4 (28.2)	18.5 (12.4)	0.38	0.50	12.4 (3.0)	12.1 (2.5)
1981	0.91 (0.28)	0.64 (0.48)	2,199 (578)	1,642 (607)	42.3 (28.3)	19.9 (13.7)	0.37	0.45	12.6 (2.8)	12.2 (2.4)
1982	0.91 (0.29)	0.64 (0.48)	2,166 (576)	1,630 (617)	41.3 (28.3)	19.7 (12.1)	0.36	0.46	12.6 (2.8)	12.3 (2.4)
1983	0.90 (0.30)	0.65 (0.48)	2,136 (600)	1,632 (628)	40.3 (31.2)	20.2 (13.7)	0.37	0.47	12.6 (2.8)	12.3 (2.3)
1984	0.90 (0.30)	0.67 (0.47)	2,142 (586)	1,635 (628)	40.7 (32.7)	20.4 (13.8)	0.38	0.47	12.6 (2.7)	12.3 (2.3)
1985	0.90 (0.30)	0.70 (0.45)	2,188 (615)	1,646 (680)	42.9 (39.9)	20.6 (13.1)	0.40	0.47	12.6 (2.7)	12.3 (2.3)
1986	0.90 (0.30)	0.70 (0.46)	2,192 (576)	1,665 (678)	44.0 (39.5)	21.6 (15.1)	0.39	0.48	12.7 (2.7)	12.3 (2.3)
1987	0.90 (0.30)	0.70 (0.46)	2,215 (612)	1,690 (662)	45.2 (41.5)	22.5 (15.1)	0.39	0.48	12.7 (2.6)	12.3 (2.3)
1988	0.90 (0.30)	0.71 (0.45)	2,230 (594)	1,691 (671)	46.7 (51.4)	23.2 (15.3)	0.41	0.48	12.7 (2.6)	12.4 (2.3)
1989	0.89 (0.31)	0.72 (0.45)	2,221 (610)	1,703 (676)	47.7 (54.0)	23.7 (16.6)	0.41	0.47	12.7 (2.6)	12.4 (2.3)
1990	0.88 (0.32)	0.72 (0.45)	2,251 (579)	1,683 (631)	48.0 (50.7)	23.8 (17.4)	0.41	0.48	12.7 (2.6)	12.4 (2.2)
1991	0.87 (0.33)	0.72 (0.49)	2,259 (576)	1,807 (641)	47.2 (41.5)	23.7 (18.7)	0.42	0.43	12.7 (2.6)	12.5 (2.3)
1992	0.87 (0.33)	0.74 (0.44)	2,221 (606)	1,815 (682)	47.2 (44.8)	24.1 (18.2)	0.43	0.50	12.8 (2.6)	12.6 (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	Number of Kids > 5 and < 17 years old	Marital Status
1968	46.7		4.0	37.8	0.56	0.94	0.85
	(27.7)		(1.9)	(10.7)	(0.82)	(1.3)	(0.35)
1969	49.9	7.7	4.0	38.5	0.53	0.93	0.86
	(31.1)	(3.7)	(1.9)	(10.9)	(0.83)	(1.3)	(0.35)
1970	50.3	7.7	3.8	38.6	0.49	0.87	0.85
	(30.0)	(3.6)	(1.8)	(11.3)	(0.79)	(1.3)	(0.36)
1971	51.0	7.5	3.7	39.0	0.44	0.85	0.83
	(31.4)	(3.5)	(1.8)	(11.6)	(0.76)	(1.2)	(0.37)
1972	51.4	7.4	3.7	39.3	0.42	0.80	0.82
	(31.5)	(3.5)	(1.8)	(11.8)	(0.72)	(1.2)	(0.38)
1973	53.3	7.4	3.6	39.5	0.39	0.77	0.82
	(35.2)	(3.4)	(1.8)	(12.2)	(0.69)	(1.1)	(0.38)
1974	54.3	7.3	3.4	39.8	0.37	0.71	0.82
	(35.2)	(3.4)	(1.8)	(12.5)	(0.68)	(1.1)	(0.39)
1975	52.9	6.9	3.3	39.8	0.35	0.66	0.81
	(34.6)	(3.2)	(1.7)	(12.6)	(0.67)	(1.0)	(0.39)
1976	53.4	6.8	3.3	39.7	0.35	0.62	0.80
	(35.6)	(3.2)	(1.7)	(12.6)	(0.68)	(1.0)	(0.39)
1977	52.1	6.7	3.2	39.7	0.34	0.60	0.79
	(35.6)	(3.3)	(1.6)	(12.6)	(0.68)	(0.96)	(0.40)
1978	52.1	6.5	3.2	38.9	0.41	0.51	0.77
	(35.6)	(3.5)	(1.6)	(12.7)	(0.72)	(0.87)	(0.42)
1979	55.5	6.7	3.1	39.8	0.34	0.53	0.77
	(55.2)	(3.3)	(1.5)	(12.5)	(0.66)	(0.88)	(0.42)
1980	55.1	6.6	3.1	39.9	0.35	0.50	0.78
	(39.1)	(3.3)	(1.5)	(12.5)	(0.69)	(0.84)	(0.42)
1981	56.2	6.4	3.1	38.8	0.39	0.50	0.80
	(68.0)	(3.1)	(1.4)	(11.9)	(0.69)	(0.82)	(0.39)
1982	54.1	6.3	3.1	38.9	0.38	0.50	0.80
	(40.3)	(3.1)	(1.4)	(11.8)	(0.69)	(0.82)	(0.40)
1983	53.1	6.3	3.1	39.0	0.38	0.51	0.80
	(39.4)	(3.1)	(1.4)	(11.8)	(0.67)	(0.84)	(0.40)
1984	54.8	6.3	3.1	39.1	0.38	0.51	0.80
	(43.2)	(3.1)	(1.4)	(11.7)	(0.70)	(0.84)	(0.40)
1985	57.9	6.5	3.1	39.6	0.37	0.53	0.80
	(51.3)	(3.8)	(1.4)	(11.5)	(0.68)	(0.85)	(0.40)
1986	59.2	6.4	3.1	40.2	0.37	0.56	0.81
	(48.8)	(3.2)	(1.4)	(11.2)	(0.69)	(0.87)	(0.40)
1987	61.9	6.5	3.1	40.6	0.36	0.57	0.81
	(51.4)	(3.1)	(1.3)	(10.8)	(0.67)	(0.87)	(0.40)
1988	64.1	6.6	3.1	41.4	0.34	0.58	0.81
	(63.6)	(3.0)	(1.3)	(10.6)	(0.66)	(0.87)	(0.40)
1989	65.9	6.5	3.1	42.1	0.33	0.59	0.81
	(69.7)	(2.9)	(1.3)	(10.3)	(0.63)	(0.87)	(0.39)
1990	66.3	6.6	3.1	42.7	0.30	0.61	0.81
	(63.7)	(3.2)	(1.3)	(10.1)	(0.62)	(0.89)	(0.39)
1991	61.4	6.4	3.0	43.7	0.29	0.58	0.77
	(52.3)	(3.1)	(1.4)	(10.2)	(0.60)	(0.87)	(0.43)
1992	65.6	6.7	3.1	44.0	0.29	0.62	0.82
	(62.1)	(3.8)	(1.3)	(9.6)	(0.61)	(0.89)	(0.39)

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

Figure 1: Gender Wage Gap

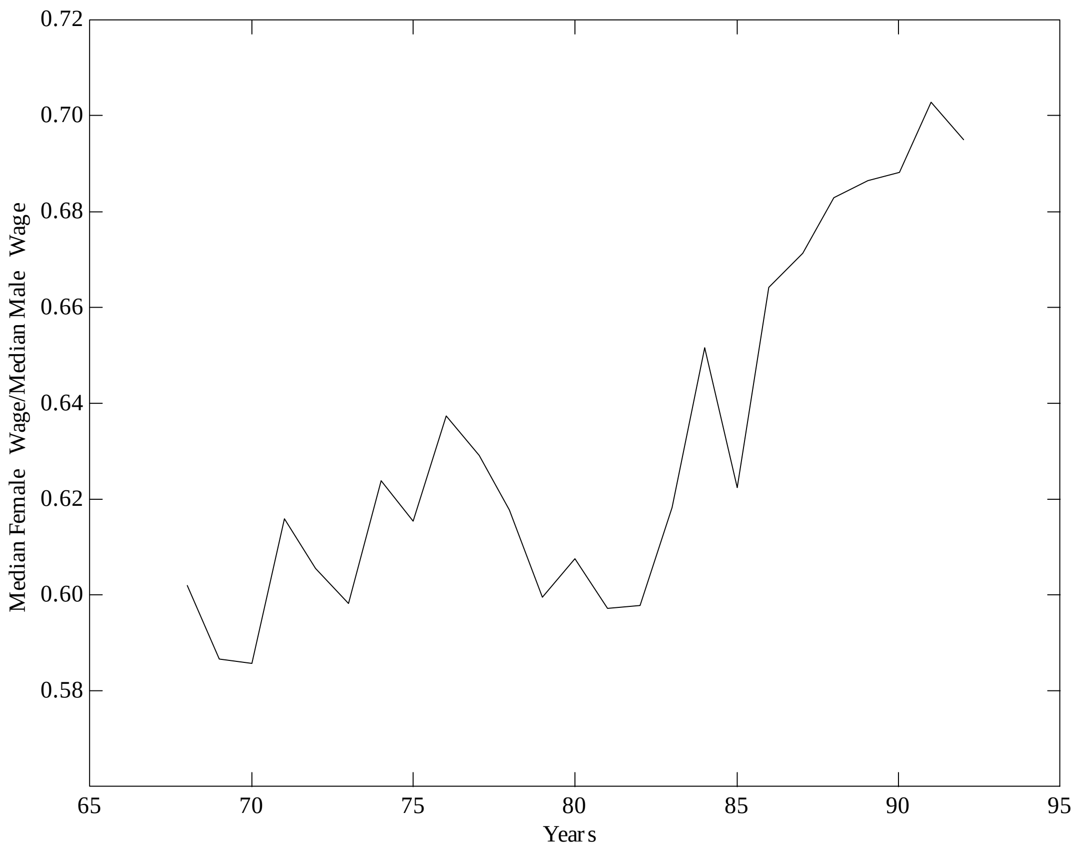


Table 3: Consumption Equation

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

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 2: Estimated Aggregate Productivity

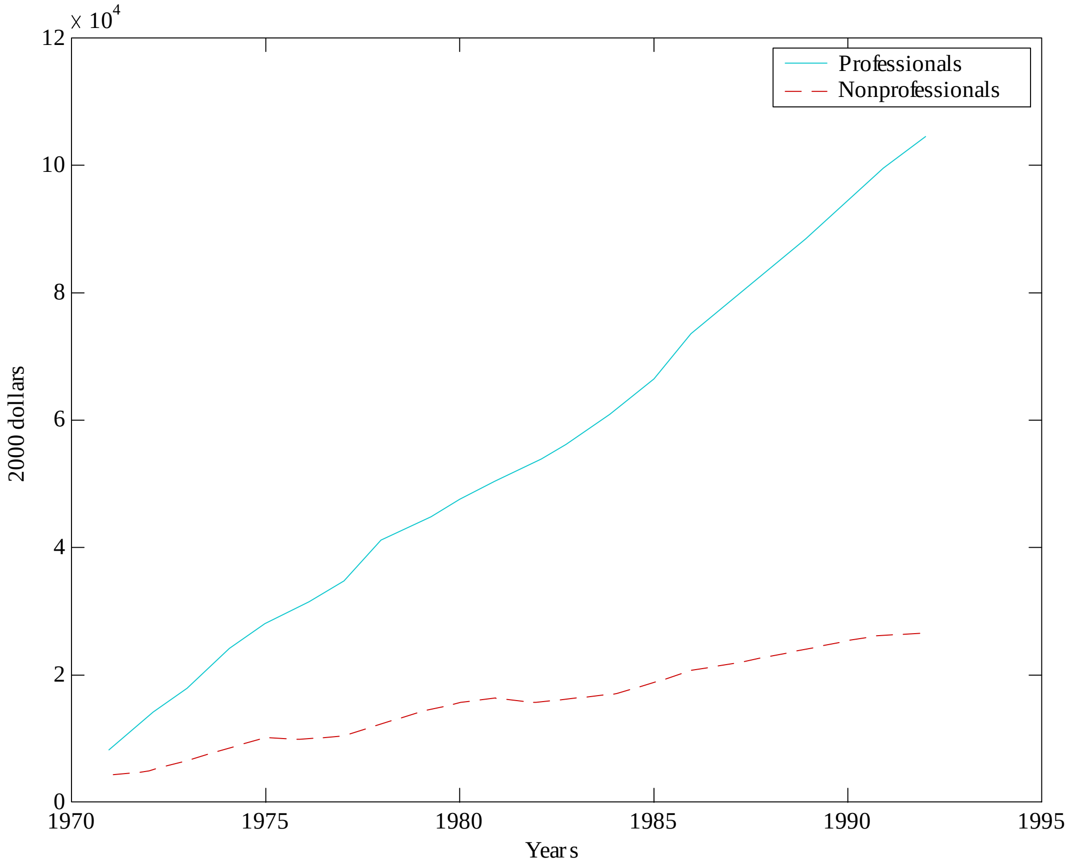


Table 4: Earning Equation

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}	6086 (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^2_{nt}	-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)
$AGE^2_{nt} \times l_{nt}$	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 Earnings Gap
(Median Women Wage over Median Men Wage(%))

Source	Professional	Nonprofessional
Raw	87	76
Predicted	92	81
Human Capital	60	56
Beliefs	12	9
Fixed Effect	5	7
Other	10	4

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

Source	Professional	Nonprofessional
Raw	30	25
Predicted	28	22
Human Capital	67	65
Beliefs	8	6
Other	25	29

Table 10: Decomposition of Change in Human Capital as a Source of Gender Wage 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