ESTIMATING A DYNAMIC ADVERSE-SELECTION MODEL: LABOR FORCE EXPERIENCE AND THE CHANGING GENDER EARNINGS GAP[1]

GEORGE-LEVI GAYLE
AND LIMOR GOLAN

This paper investigates the role of labor-market attachment, on-the-job human-capital accumulation, occupational sorting, and discrimination in the narrowing gender earnings gap over the past three decades. This paper contributes in three ways: First, it formulates an estimable dynamic model of labor supply occupational sorting and human capital accumulation, in which gender discrimination and earnings gap arise endogenously. Second, it demonstrates the identification and develops a three-step estimation technique. Third, it utilizes the framework to quantify the main driving forces behind the changes in gender patterns of labor-market outcome. The decomposition of the change in the gender earnings gap reveals that changes in private information and hiring cost accounts for over 33% of the change in professional occupation while increase aggregate labor market productivity accounted for 23% and demographic changes (mainly fertility decline) accounted for about 18%. Similar results were found for the Nonprofessional occupations. The estimation results do not support the hypothesis that changes in home production technology explain the increase in women's labor-market participation. Further analysis shows that market frictions significantly amplifies exogenous changes in our model. Without labor market frictions the earnings gap would have been small by at least 45%, women would have participation less in the labor force but the ones that participated would have worked more and earned income similar to men. We find that the there are significant higher return to professional occupation and that this has increased greatly over time, hence accounting for the increasing 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
.

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. 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. According to figures from the Michigan Panel Study of Income Dynamics (PSID), the

---

[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.

participation rate of women has increased from 54% in 1968 to 74% in 1992. The annual hours worked by women also increased over the period, from 1400 hours in 1968 to 1800 in 1992. 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, it formulates an estimable dynamic model of labor supply occupational sorting and human capital accumulation, in which gender discrimination and earnings gap arise endogenously. Second, it demonstrates the identification and develops a three-step estimation technique. Third, we utilize the framework to quantify the main driving forces behind the changes in gender patterns of labor-market outcome.

In the model, workers are heterogeneous with respect to characteristics affecting dis-utility 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 dis-utility 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. 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 dis-utility from 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. We solve for spot market contracts equilibrium.

Because of the firm-specific cost of hiring new workers, employers make rent on workers a

---

[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.

year after they are hired (since the market is competitive, when hiring workers firms make zero expected profits over the workers' expected employment spells). In equilibrium, beliefs about the worker's future participation in the firm enter the earnings equations. Our model incorporates several elements which may cause men and women to display different labor market outcomes. First, the utility parameters may vary across gender. Therefore, differences in dis-utility from working and differences in participation costs can induce men and women choices of participation, occupation, and hours worked to be different even in the absence asymmetric information. Secondly, there may be unobserved difference in skills across the gender; such differences can also give rise to earnings gap in the absence of asymmetric information. Thirdly, even if there are no differences in preferences and skills across gender groups, different outcomes may arise due to multiplicity of equilibria when there is asymmetric information.[3] That is, if employers have different beliefs about future participation of men and women, women may face lower wages; they therefore, may work less, and sort into occupations with lower returns to experience and lower costs of hiring workers. As a result, on average, women may accumulate less labor market experience than men. Because there may be differences in preferences and productivity across gender groups, we define discrimination as the difference between the labor markets outcome of men and women under symmetric information (which yields a unique equilibrium), and under asymmetric information. That is, the gender differences in earnings which arise due to observed "group affiliation" are refer to as discrimination, as opposed to differences that arise due to differences in preferences and productive skills.

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.

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 changes in productivity. They raise productivity for all workers within the occupation equally. Our model, however, predicts that if women's participation is lower, increase in productivity 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

---

[3]Thus, our framework is flexible enough to give rise to endogenous differences in labor market outcomes of men and women even if there are no initial differences, but does not impose the groups to be initially identical.

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 dis-utility from working.

There are two broad types of employers' discrimination in the literature. The first type is taste-based discrimination, formulated by Gary Becker (1971); the second type results from incomplete information (also called statistical discrimination), pioneered by Kenneth Arrow (1972) and Edmund Phelps (1973) and further analyzed by Coate and Loury (1993). 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). Our model belongs to the second class of models. Whereas the models in the typical statistical discrimination literature (see for example, Altonji(2005), Moro(2003), Antonovics(2004), among others) focuses on the effect of beliefs about productivity differences across groups, in our model the uncertainty is about the turnover propensity of workers. The main theoretical contribution of the model is to provide a theoretical framework for a dynamic model in which participation, accumulation of labor market experience and sorting into occupations are endogenous, when workers have private information. This allow us to demonstrate how beliefs about future employment spells arise endogenously, affecting gender differences in labor-market experience, occupational sorting, and attachment to the labor force. Another paper that has endogenous gender gap is Albanes and Olivetti (2005). They develop a one-period model of gender statistical discrimination in which effort in the labor market and hours worked at home are determined endogenously. The moral hazard problem in the labor market can generate endogenously gender discrimination. In equilibrium, women may have higher costs of effort as it is optimal to work more hours at home.[4]

One of the main problems in estimating games of multi-agent informational models is the possibility of multiple equilibria. Since the problem of multiple equilibria is not one of identification, we use a multi-stage estimation procedure to estimate the model; the procedure is based on necessary conditions for all equilibria. Thus assuming only one equilibrium played in the data per cohort (or other data partition), the data reveals which equilibrium is being played. Other papers taking this approach are Tamer (2003), Aguirregabiria(2004), Aguir-regabiria and Mira (2005), Penendorfer and Schmidt-Dengler (2003, 2006), Bajari, Benkard and Levin(2004), Pakes, Ostrovsky and Berry (2004), among others.

This paper then uses a constructive strategy, similar to Chesher (2003, 2005, 2007), Penendofer and Schimdt-Dengler(2006), to show that the model, of discrete and continuous choice, is semiparametrically identified. That is, conditional on the distribution of the unobserved preference shocks and the class of risk aversion, our model is nonparametrically identified up-to three additive normalizations. This results is stronger than what is obtain in the standard discrete choice model (see for example Magnac and Thesmar (2002) and Penendofer and Schimdt-Dengler(2006)) but have parallel in the literature on continuous choice models,(see Pesendorfer and Bonet (2004) for example).

---

[4]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 in non game theoretic setting,

Based on the necessary conditions for equilibrium, we developed a multi stage semiparametric estimation procedure. The estimation proceeds in three steps. The employer's problem, is estimated in the first step, along with other inputs from the individual consumption Euler equation. The estimates from the first stage are used to nonparametrically estimate each individual's choice-specific probabilities and their derivatives, and the employers beliefs about workers' future participation probabilities 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 estimator is akin to a number of estimators in the literature for the estimation of discrete games and single-agent models (Hotz and Miller, 1993; Altug and Miller, 1998; Pakes, Ostrovsky, and Berry, 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 without any simulation. To the best of our knowledge this is the first paper to estimate a structural dynamic signalling model.

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 earnings 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 market frictions significantly amplifies exogenous changes in our model. Without labor market frictions the earnings gap would have been small by at least 45%, women would have participation less in the labor force but the ones that participated would have worked more and earned income similar to men. We find that the there are significant higher return to professional occupation and that this has increased greatly over time, hence accounting for the increasing representation of women in professional occupations.[5]

The decomposition of the change in the gender earnings gap reveals that changes in private information and hiring cost accounts for over 33% of the change in professional occupation while increase aggregate labor market productivity accounted for 23% and demographic changes (mainly fertility decline) accounted for about 18%. Similar results were found for the Nonprofessional occupations. The estimation results do not support the hypothesis that changes in home production technology explain the increase in women's labor-market 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

---

[5]See Erosa, Fuster and Restuccia (2005) for a model of endogenous fertility, and human capital accumulation explaining the gender earnings gap.

of participation account for only 1% in professional occupations and 2% in nonprofessional occupations.

This paper is organized as follows. Section 2 gives an overview of the data and document the secular trends in the labor market. Section 3 describes the model. Section 4 contains the equilibrium analysis while Section 5 discusses the theoretical model's predictions for the earnings gap. The identification results are presented in section 7 while the empirical strategy are outline in Section 7. Section 8 contains the estimation results, empirical analysis, and the gender earnings gap decomposition. Section 9 concludes. The appendices present proofs, the implementation and asymptotic property of our estimator and a detailed data description.

2. DATA

The data for this study are taken from the Family File, the Childbirth and Adoption History File, the Retrospective Occupation file, and the Marriage History File of the PSID. The sample contains individuals who were either the *Head* or *Wife* of a household in the year of the interview. Individuals are classified into two occupation categories, professional and 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.

2.1. *Secular Trends*

We emphasize five important secular trends over the period that we intend to explain. First, earnings and wage gaps have narrowed significantly. Second, average hours worked by women increased significantly. Third, there has been a spectacular increase in labor force participation of women. Fourth, there has been a drastic change in the occupation composition with women increasingly entering traditionally male dominated occupations. Fifth, there has been significant changes in the demographic attributes of the population, especially the fertility rates.

2.1.1. *Wage and Earnings*

The average annual earnings for men increased by roughly 18% over the period, going from US$40,000 per year in 2000-constant dollars in 1968 to US$47,000 in 1992. Meanwhile, the average annual earnings for women increased by around 49% over the same period, going from US$16,200 in 1968 to US%24,100 in 1992. As seen in Figure (??) the gender earnings gap much greater that the wage gap. However, the declining trend in both are roughly the same. The wage gap declined by around 30% over the period, going from around 40% in 1968 to around 28% in 1992. At the same time the earnings gap declined by around 25%. The earning gap is 50 percent larger than the wage gap.

Therefore, in this paper we focus on the earnings gap instead of the traditional wage gap in order to understand the forces behind both.

2.1.2. *Participation, Hours Work, and Education*

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. Although the gap between the hours worked by women and men has narrowed significantly over the period, it however remain large. Therefore a larger fraction of the differences between the wage gap and the earnings gap is accounted for by the differences in the hours worked by men and women.

The gap in the average years of completed education between men and women has been almost completely erased by the early 1990s, yet still the participation and hours gap remains.

2.1.3. *Occupation*

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. Also the earning gap are significantly different in the two occupations. In professional occupation the earnings gap is much smaller than in the nonprofessional occupations. Also the earnings gap is narrowing at a faster rate in the professional than the nonprofessional occupations[6]. Therefore any theory that is going to attempt to explain the decline in the earnings gap and participation increases such account for that fact that over time women are entering occupations with high hours and lower earnings gap to begin with.

2.1.4. *Demographic*

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 final sample is married through the period. Therefore the major demographic change that occurs over the period has been declined in fertility rate over time. It is still an open question of what role does the decline in fertility plays in the labor trends documented above.

2.2. *Consumption and Other Variables*

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

---

[6]Infact with finer occupation classification Dey and Hill(2007) found that in some professional occupation , i.e., Engineering and Architecture, on average the women earns more than men.

per capita food consumption has increased over the period. Household income and income per capita has increased over the sample period.

3. THEORETICAL MODEL

In this section we describe the theoretical framework. The economy consists of infinitely lived firms and finitely workers. Labor markets are competitive with Υ occupations, $\tau = 1, \ldots, \Upsilon$. We describe below the preferences, technology, and the timing and information structure of the game-theoretical model.

3.1. *Workers' Preferences and Choice*

There exist a continuum of workers on the unit interval $[0, 1]$ in each age-education cohort (all workers who were born in the same year, and have the same number of years of completed education). These workers are divided into two observed gender groups , $i \in \{w, m\}$, women and men respectively. For notational ease we will denote age and calendar year for each cohort by t ($t = 0, \ldots, T$). The theoretical model is written for a given cohort. Workers have preferences about nonmarket hours, l_t (time in which the individual does not work) and consumption, c_t. Let h_t denote the time spent working at t. There is a fixed amount of time in each period available for working, which implies that the amount of time worked in each period can be normalized as $0 \le h_t \le 1$, thus $l_t = 1 - h_t$. If $h_t = 0$, the agent does not work at time t. Otherwise, the agent works the fraction of time $h_t > 0$. For notational convenience two additional indicators are defined: a work force participation indicator d_t, where $d_t = 1$ if and only if $h_t > 0$ and 0 otherwise, and an occupation participation indicator $I_{\tau t}$, where $I_{\tau t} = 1$ if and only if the worker is employed in occupation τ and 0 otherwise.

Let a_t be the set of labor market actions of the worker, i.e. $a_t \equiv (d_t, \{I_{\tau t}\}_{\tau=1}^{\Upsilon}, h_t)$. Define the employment history of a worker by $H_{t-1} = (a_0, a_1, ..., a_{t-1})$, therefore next employment history H_t, is $H_t = (H_{t-1}, a_t)$. Each worker preferences are additively separable in consumption and leisure contemporaneously. In addition, consumption is additive separate over time whereas leisure is not . It is assumed that there are two time-varying vectors of characteristics that determine the utility associated with alternative consumption and leisure allocations. Denote the first by the $K \times 1$ vector z_t and the second by the 3×1 vector $(\varepsilon_{0t}, \varepsilon_{1t}, \varepsilon_{2t})'$. It is assumed that z_t is independently distributed over the population (it includes number of household members, marital status, number of kids and spouse characteristics); over time it evolves according to a known group- specific transition distribution function, $F_{i0}(z_{t+1} \mid z_t, H_t)$. The vector $(\varepsilon_{0t}, \varepsilon_{1t}, \varepsilon_{2t})'$ is assumed to be independent across the population and time and drawn from a population with a common distribution function $F_1(\varepsilon_{0t}, \varepsilon_{1t}, \varepsilon_{2t})$. There are several things to notice: first, $F_{i0}(z_{t+1} \mid z_t, H_t)$ depends on H_{t-1}, a_t but not c_t, this property follows directly from the definition of time additive separability. Second, although we assume that characteristics in z_t, such as marriage and fertility are contemporaneous exogenous from work decision, $F_{i0}(z_{t+1} \mid z_t, H_t)$ is a function of the worker's actions a_t, and therefore are endogenous is a predetermined sense.

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

$$(3.1) \qquad U_{it} \equiv d_t u_{it0}(z_t) + u_{i1}(z_t, H_{t-1}) + u_2(z_t, c_t, \varepsilon_{2t}) + (1 - d_t)\varepsilon_{0t} + d_t \varepsilon_{1t},$$

where u_{it0} represents the fixed utility costs of working, which depends on the observed individual-specific characteristics but may change from period to period. We allow the costs to to change over time in order to capture the possible changes in home production technology. Particularly, if the cost of producing home good has changed over time, it should affect the costs of participating in the labor market. We allow for the possibility that the fixed cost of working and the utility from nonmarket hours to be group-specific (for example, it can differ by gender). Whether these utility components vary by gender is an empirical question which this paper hopes to answer. The utility from nonmarket hours is additive and non-separable over time. We assume the utility functions satisfies the standard regularity conditions: it is concave, continuous and twice differentiable. The distribution function $F_{i0}(z_{t+1} \mid z_t, H_t)$ is absolutely continuos with continuously differentiable $f_{i0}(z_{t+1} \mid z_t, H_t)$.

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

$$(3.2) \quad E_t \left[\sum_{r=t}^{T} \beta^{r-t} U_{it} \right]$$

To provide a tractable solution to the model, we assume that asset markets are competitive and complete (CCM).[7] 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.[8]

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

$$(3.3) \quad E_0 \left\{ \sum_{t=0}^{T} \beta^t \lambda_t \left[c_t - \overline{S}_t \right] \right\} \leq W,$$

where $\overline{S}_t$ 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 is the sum of two incomes), λ_t is the expected price of the contingent claim, and W is an exogenously determined quantity

---

[7]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.

[8]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.

denoting bequests net of inheritances. In any state, z^*, the price of a contingent claim is

$$\overline{\lambda}_t(z^*) = \int \lambda_t(z^*) g(z^*) \, dz^*.$$

The states are determined by realizations of ε_{2t}, and z_t, and they are independent with joint density $g(z_t^*)$. The aggregate feasibility condition equates the sum of labor income by all households and the aggregate resource endowment W_t

$$(3.4) \qquad \int_0^1 [c_t - \overline{S}] \, d\mathsf{L} \le W_t, \quad t \in \{0, 1, \ldots\}.$$

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

3.2. *Firms' Technology and Payoffs*

Firms offer jobs to maximize lifetime expected discounted profits. Each occupation has a continuum of identical firms competing for workers. Each firm offers one job each period. There is a homogeneous product with price normalized to 1. Let z_t^p be a vector of individual characteristics which affect productivity. It includes education, age, and an individual-specific skill which is constant over time. Output in period t in occupation τ is denoted by: $y_{\tau t} = y(h_t, H_{t-1}, z_t^p)$. The production function is identical within an occupation and varies across occupations. Output is a function of current hours input, the worker's past labor market experience, and other production relevant characteristics of the worker. Assume that gender does not affect output, that is, a woman and a man with the same labor market experience, and other production characteristics produces the same level of output if they supply the same amount hours. The production relevant characteristics, z_t^p, is meant to capture other production enhancing variables like education, ability (i.e. fixed effects) or potential experience (age less years of education). Past labor market experience augments the output produced per unit of hours input. Labor market experience is general but the returns to experience vary across occupations.

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

Workers and firms can only commit to (non-contingent) spot contracts.[9] Firms offers jobs (a job is equivalent to fraction of time worked) to a worker only if given the worker's characteristics, the output he/she produces net of the costs of hiring within the occupation is the highest among all occupation. This is formalized in the assumption below.

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

$$(3.5) \qquad y_{\tau' t}(h_t \mid H_{t-1}, z_t^p) - \gamma_{\tau'} > y_\tau(h_t \mid H_{t-1}, z_t^p) - \gamma_\tau.$$

---

[9]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.

If in addition to Assumption 3.1 the production function satisfies a single crossing condition then occupations would be segmented into jobs (hours) offered to workers, conditional on their characteristics affecting production. This property is formalized in the Assumption 3.2.[10]

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

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

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

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

Assumptions 3.1 and 3.2 imply that for every occupation τ, and worker with characteristics H_{t-1}, z_t^p, there is a range of hours in that occupation $h \subset (\underline{h}_\tau(H_{t-1}, z_t^p)$ $\overline{h}_\tau(H_{t-1}, z_t^p))$ where $\underline{h}_\tau(H_{t-1}, z_t^p) \leq \overline{h}_\tau(H_{t-1}, z_t^p)$ for which employers will offer jobs. While these assumptions are not necessary for the results derived in this paper, it simplifies the analysis significantly.

3.3. *The Structure of the Game: Information, Timing and Strategies*

Our model is a reputation game in which workers have private information, employer offers contracts which workers select into every period; observing workers' choices, employers update their beliefs. Below we describe the information, timing and strategies.

3.3.1. *Information*

All utility and production function parameters, hiring cost, labor market history, worker gender and production relevant characteristics, z_t^p, are common Knowledge. The worker's private information includes consumption, c_t, and time-varying characteristics that determine the utility associated with alternative consumption and leisure allocations, z_t, and ε_t. The distribution function of z_t and ε_t are common knowledge.

In the empirical section of this paper z_t will include variables such consumption, spouse characteristics, marital status, number of kids. The issue of which variables are private information to the worker is not is central to this paper, instead it is the existence of some private information allowing workers to better predict the likelihood of remaining in the firm in the future. In particular, it is crucial that this information is not observed by potential employers as opposed to the current employer. Assuming consumption is privately observed

---

[10]For example, a part time job for a worker with certain skill, experience and education is offered only by one occupation.

is natural provided trade in the consumption contingencies markets is anonymous. As to what we include in z_t is less obvious. Our decision of which variables are privately observed are based on two conditions: first, is the information private in nature and second, can a worker credibly report or convey this information. If this information is important, workers will have incentives to misrepresent this information. In our framework, employers cannot credibly commit to penalizing workers who misrepresented information after they are hired, because of the initial costs of hiring.[11] Empirical we will show that our model provides a set of natural over-identifying restrictions to guide in selection of what variables are private information or not.

Notice that the worker's *type* here is characterized by z_t and ε_t. Whereas ε_t is i.i.d., z_t evolves according to a known process, $F_{0i}(z_{t+1} \mid z_t, H_t)$; 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. This also means that an individual type is endogenously determined in the model because the transition probability of future type (i.e. $F_{i0}(z_{t+1} \mid z_t, H_t)$) depends current and past actions, H_t. Throughout the paper we denote the worker's *type by* $z_t^* = (z_t, \varepsilon_t)$.
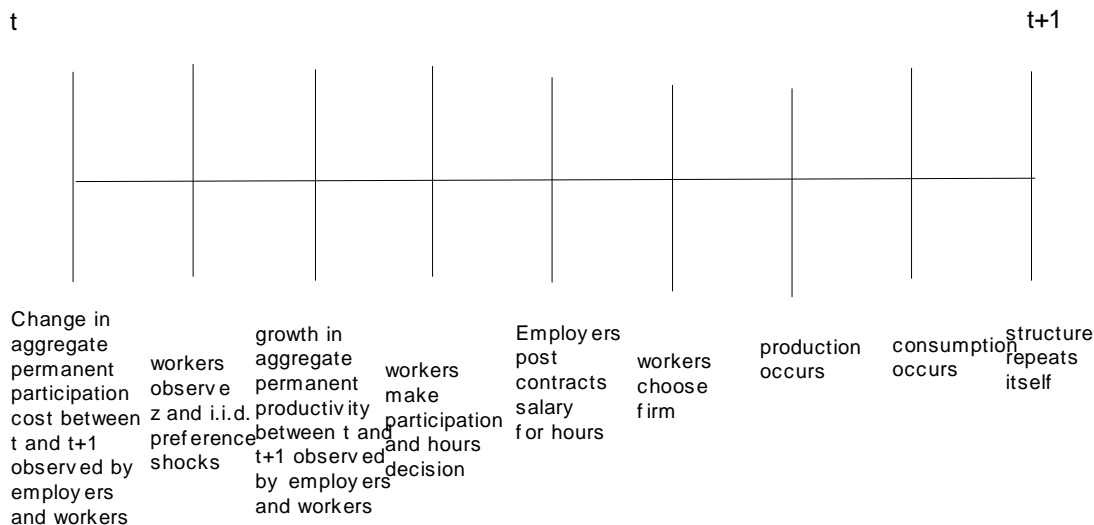
3.3.2. *Timeline and Strategies*

At the beginning of each period, the growth rates of aggregate utility costs of participation shocks, and aggregate permanent occupation specific productivity shocks between period t and $t + 1$ are observed by all agents in the economy. Workers privately observe their type, z_t^*. Given this information, workers make a participation, occupation and hours decisions, a_t. [12]Observing workers choices, firms offer simultaneously salaries for each worker. Workers observe the offers and chose a firm.

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

---

[11]If it is not costly to verify this imformation, then employers could conceivably obtain this information on a secondary market. This, however, is an empirical issue which we hope answer in the section.

[12]The choice of hours implies that a worker is committed to a job which is defined by hours and occupation and participation, for the current period. Clearly in equilibrium this choice will take into account the offers made in the later stage. We further discuss this choice of timing in section XX}

t   t+1

Change in aggregate permanent participation cost between t and t+1 observed by employers and workers

workers observe z and i.i.d. preference shocks

growth in aggregate permanent productivity between t and t+1 observed by employers and workers

workers make participation and hours decision

Employers post contracts salary for hours

workers choose firm

production occurs

consumption occurs

structure repeats itself

3.3.3. *Beliefs*

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

4. EQUILIBRIUM ANALYSIS

The equilibrium concept used in this paper is Perfect Bayesian. Below we provide a formal definition.

DEFINITION 4.1 (Equilibrium) *A Perfect Bayesian Equilibrium consists of strategies c_t, σ_{it}, and the worker's offer choice, $\{S_{i\tau t}\}_{\tau=1}^{\Upsilon}$, and a common belief system, such that*
1. *Each player's strategy is optimal given beliefs and other players' strategies*
2. *The posterior beliefs, $\widetilde{\mu}$, satisfy Bayes' rule when possible.*[13]

$$(4.1) \quad \widetilde{\mu}_{it}(z_t^* \mid z_t^p, H_{t-1}, a_t) = \frac{\mu_{it}(z_t^* \mid z_t^p, H_{t-1})\sigma_i(a_t \mid z_t^*, H_{t-1})}{\int \mu_t(z^* \mid .)\sigma_i(a_t \mid .)\,\mathrm{d}z^*}$$

and for all histories, types and actions

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

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 endogenously)*

$$(4.2) \quad \mu_{it+1}\left(z_{t+1}^* \mid H_t, z_t^p\right) = [f_{i0}(z_{t+1} \mid z_t, H_t), f(\varepsilon_{t+1} \mid \varepsilon_t)]\widetilde{\mu}_{it}(z_t^* \mid .)$$

---

[13]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.

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

The solution of our model proceeds as follows. For any period t and beliefs, derive optimal consumption choice, offer acceptance choice, optimal salary schedule, optimal participation, hours and occupation decisions (beginning with the final stage in that period). Then prove that the above strategies constitute an equilibrium given the beliefs. This is shown by first proving that there is no profitable unilateral deviation from the competitive salary schedule. Second, show that a necessary and sufficient condition for existence of equilibrium is that the implied probability that a worker participate in the firm next period, conditional on the employers information is correct and has a fixed point. Finally show that there exist such a fixed point.

Let η denote the Lagrangian multiplier associated with the budget constraint in equation (3.3), then the optimal consumption satisfies the necessary conditions,

$$(4.3) \qquad \frac{\partial u_2(c_t, z_t, \varepsilon_{2t})}{\partial c_t} = \eta \lambda_t,$$

for all $t \in \{0, 1, 2 \ldots\}$.

With the contemporaneous separability of the consumption from the labor supply choices, the condition in (4.3) can be used to solve for the individual Frisch demand functions, which determine consumption in terms of the time-varying characteristics, z_t, the idiosyncratic shocks to preferences, ε_{2t}, and the shadow value of consumption, $\eta \lambda_t$.[14]

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

LEMMA 4.1   *Given any choice of hours, a worker accepts the contract with the highest salary. Employed worker remains with the current employer if there is a tie. If not employed, the worker randomizes between identical offers.*

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

---

[14]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.

Proposition 4.1 states the optimal salary schedule.

PROPOSITION 4.1 (Competitive Salary Schedule)   *For any hours chosen by the worker, and firms' beliefs, the competitive salary is*

$$(4.4) \quad S_{i\tau t}(h_t, H_t, z_t^p) = y_{\tau t}(h_t, H_{t-1}, z_t^p) - \gamma_\tau + \beta \gamma_\tau \widetilde{p}_{i\tau t+1}(H_t, z_t^p)$$

for all $h_t \in (\underline{h}_\tau(H_{t-1}, z_t^p), \overline{h}_\tau(H_{t-1}, z_t^p))$, where $\widetilde{p}_{i\tau t+1}(H_t, z_t^p)$ is the probability that the worker will work in the firm the proceeding period.

The salary schedule is equal to the worker's productivity plus the function of the hiring costs and the probability the worker says with the firm next period. It is the highest outside offer a worker can receives. It is offered so an outside employer makes expected profits of zero over the worker's years in the labor market. This condition must hold for all workers in all occupations.

Under symmetric information, each worker earns the expected productivity over time net of hiring costs, however under asymmetric information each worker earns the expected productivity (net of costs) over the expected employment spell with the employer of all workers with the same observable characteristics. Therefore the only difference between the optimal salary under symmetric versus asymmetric information is the probability of future participation. Thus consider two workers with the same publicly observable characteristics, z_t^p, who differ with respect to unobservable characteristics, z_t^*, and choose the same number of hours worked in equilibrium under asymmetric information.. The worker whose z_t^* implies that she/he is less attached earns on average more than the productivity (net of costs), and the worker whose z_t^* implies she/he is more attached is paid less than the productivity (net of costs). This result is because $\widetilde{p}_{i\tau t+1}$ is higher than the actual probability for the worker with the lower *attachment* and lower than the actual probability for the worker who is more *attached* to the firm.

Conditional on the optimal salary schedule, we derive the workers participation, hours and occupation decisions. Let $\omega_t \equiv (H_{t-1}, z_t, \eta\lambda_t,)$ denotes the non-idiosyncratic state variables of the worker[15]. The complete state variables of the worker is $(\omega_t, \varepsilon_t)$. Define the ex-ante conditional valuation functions associated with the decisions to work and not work at time t as V_{i1t} and V_{i0t} respectively. The ex-ante conditional valuation function is the discounted sum of future payoffs before the choice specific idiosyncratic utility shock are realized and actions taken. Formally,

$$(4.5) \quad V_{ikt}(\omega_t) \equiv \max_{\{h_r\}_{t=t}^T} E_t \left\{ \sum_{s=t}^T \beta^{s-t} \left[d_s u_0(z_s) + u_1(z_s, h_s, H_{s-1}) + d_s \varepsilon_{1s} \right. \right.$$

$$\left. \left. + (1 - d_s)\varepsilon_{0s} + \eta\lambda_s \sum_{\tau=1}^\Upsilon I_{\tau s} S_{i\tau s}(h_s, z_s^p, H_s) \right] \mid d_{tt} = k \right\}.$$

The necessary condition for the optimal participation decision is then

$$(4.6) \quad d_{it}^o(\omega_t, \varepsilon_{0t}, \varepsilon_{1t}) = \begin{cases} 1 & \text{if } V_{i1t}(\omega_t) + \varepsilon_{1t} \geq V_{i0t}(\omega_t) + \varepsilon_{0t} \\ 0 & \text{otherwise.} \end{cases}$$

---

[15]Note that since the agents know the equilibrium functions/conjectures then a sufficient state statistic for $S_{it\tau}$ is H_{t-1} and z_t^p. If the worker is married then a sufficient statistic for his spouse work behavior is $\eta\lambda_t$ which enters the household budget constraint and equation(4.3).

That is, the worker chooses to participate if and only if the ex-ante conditional valuation function of working plus the utility shock of working is greater or equal to that of not working.

Define by h_{it}^o, the optimal labor-supply decision in period t and by $h_{it}^* \in (0,1)$, the optimal interior solution of the labor-supply decision in period t. Using the above notation the optimal worker strategies are summarized the the Proposition 4.2.

PROPOSITION 4.2 *Given the firms strategies, a worker's best response is characterized by,* p_{it}^o, *the conditional probability of participation,* $h_{it}^*(z_t, H_t)$, *the number of hours worker and* $I_{i\tau t}^0(z_t, H_t)$, *the optimal occupation choice, i.e.* $\sigma_{it} \equiv \{p_{it}^o, h_{it}^*(z_t, H_t), I_{i\tau t}^0(z_t, H_t)\}$:

$$(4.7) \quad p_{it}^o \equiv E[d_{it}^o \mid \omega_t] = \int_{-\infty}^{V_{i1}-V_{i0}} (\varepsilon_{0t} - \varepsilon_{1t}) \, \mathrm{d}F_1(\varepsilon_{0t}, \varepsilon_{1t}, \varepsilon_{2t}) = Q_i(V_{i1t}(\omega_t) - V_{i0t}(\omega_t)_t).$$

$$
\begin{aligned}
h_{it}^*(z_t, H_t) &= \arg\max_{h_t \in (0,1)} \left[u_1(l_t, H_{t-1}, z_t) + \eta\lambda_t \sum_{\tau=1}^{\Upsilon} I_{\tau t}^0 S_{i\tau t}(h_t, z_t^p, H_t) \right. \\
&\quad + \beta E_t\{p_{it+1}[V_{i1}(\omega_{t+1}) + \varepsilon_{1t+1}] \mid \omega_t, d_{it}^0 = 1\} \\
(4.8) &\quad \left. + \beta E_t\{(1 - p_{it+1})[V_{i0}(\omega_{t+1}) + \varepsilon_{1t+1}] \mid \omega_t, d_{it}^0 = 1\} \right].
\end{aligned}
$$

and

$$(4.9) \quad I_{i\tau}^0(z_t, H_t) \equiv I\{\underline{h}_\tau(H_{t-1}, z_t^p) < h_{it}^*(z_t, H_t) < \overline{h}_\tau(H_{t-1}, z_t^p)\}$$

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

Next we show that given Lemma 4.1, Propositions 4.1 and 4.2 no firm can profitably unilaterally deviate from the competitive salary schedule.

LEMMA 4.2 *Given firms are following the competitive salary schedule in Proposition 4.1, workers are following the strategies specified in Proposition 4.2 and the beliefs system are according to Definition 4.1 and Assumption B.1; no single firm has an incentive to unilateral deviate from the competitive salary schedule.*

Lemma 4.2 stands in contrast to the well unknown nonexistence result in the Rotschild-Stiglitz (1976) model. However, as pointed out in Hellwig (1987) this non-existence result is very sensitive to the timing of the players' moves. If the workers move first deciding how much to work , and then the firms make offers for these hours with the workers finally choosing which offer to accept then the non existence result of the original Rotschild-Stiglitz model is broken(see Hellwig (1987) for more details). Using the Walrasian Equilibrium concept instead of Perfect Bayesian would give the existence of equilibrium with the same optimal salary schedule[16]. For papers that demonstrated existence of equilibrium in Rotschild-Stiglitz

---

[16]For example Riley (1979), in a model with single dimensional type showed existence of equilbrium if under the assume that the distribution of type is every strictly concave and that there is a mass point at the lower end of the distribution of type. An equivalent set of assumptions can be derived for the multi-dimensional type model.

type models with Walrasian equilibrium, see Riley (1979), Prescott and Townsend (1984), Gale (1992), Bisin and Gottardi (2006), among other.

Next we show that there exist a fixed point in mutual best responses. The optimal hours worked and participation is function of firms beliefs about next period participation. To see this consider the beliefs above period T participation, i.e.

$$\widetilde{p}_{i\tau T} = \int Q(z_T^*, H_{T-1}(\widetilde{p}_{i\tau T}, \ldots, \widetilde{p}_{i\tau 2}), h_T^*) I_T(z_T^*, H_{T-1}(\widetilde{p}_{i\tau T}, \ldots, \widetilde{p}_{i\tau 2}), h_T^*)$$

(4.10) $$\mu_T(z_T^* \mid H_{T-1}, z_T^P) \, \mathrm{d}z_T^*.$$

Clearly $\widetilde{p}_{i\tau,T}$ is defined as a implicit function of itself. In fact there is a triangular system of implicit equilibrium beliefs of the following form:

$$\widetilde{p}_{i\tau T} = \Gamma_{iT}(\widetilde{p}_{i\tau T}, \ldots, \widetilde{p}_{i\tau 2})$$
$$\widetilde{p}_{i\tau T-1} = \Gamma_{iT-1}(\widetilde{p}_{i\tau T-1}, \ldots, \widetilde{p}_{i\tau 2})$$
$$\vdots$$

(4.11) $$\widetilde{p}_{i\tau 2} = \Gamma_{i2}(\widetilde{p}_{i\tau 2}),$$

where Γ_{it} is the RHS of equation 4.10.

COROLLARY 4.1   *A necessary and sufficient condition for existence of equilibrium in our model is that there exists a fixed point in $\{\widetilde{p}_{i\tau 2}, \ldots, \widetilde{p}_{i\tau T}\}_{i \in \{w,m\}}^{\tau \in \Upsilon}$ of the system equations in (4.11) for all τ and i*

Existence of equilibrium in our model established by showing there exist a fixed point to the system of equations in (4.11). Proposition 4.3 establishes that fact.

PROPOSITION 4.3   *There exists $\{\widetilde{p}_{i\tau 2}, \ldots, \widetilde{p}_{i\tau T}\}_{i \in \{w,m\}}^{\tau \in \Upsilon}$ which the fixed point of 4.11.*

Finally, we establish the main result of our model, self-fulfilling beliefs.

COROLLARY 4.2   *In equilibrium the perceived probability of the worker will work in the firm in the proceeding period is correct conditional on the what is observable to the firm.*

The fact that there exists a one to one mapping between the posterior beliefs and the implied participation probability in equation(4.10) comes directly from the requirement in Bayesian games that beliefs are consistent with player's strategies and Bayes' rule is satisfied (when possible), this condition holds by construction. Therefore, the expected profit condition on salary which are function of the firms beliefs, is also correct on-the-equilibrium path.

5. THE GENDER EARNINGS GAP

Next we discuss the different channels in which gender differences in the labor market arise in equilibrium. Under symmetric information the salary in 4.4 is different only with respect to the probability of future participation. It is conditioned on the information available to the worker, z_t^*, instead of z_t^p. If men and women were identical in all respects, then no wage gap arises in equilibrium, and men and women will have the same labor market participation, experience and occupation choice patterns. Suppose instead there are some z_t^* for which the cost of participation and the dis-utility from hours worked are larger for women. Then a

woman with the same characteristics as a man may earn lower wage because transitions into a state in which the costs of participation are higher may imply a lower probability of future participation. Facing states which can be reached and in which the likelihood of participation is lower, may cause current value of participation to be lower and the value of working long hours to be lower as well. Women would participate less, work less hours and sort into different occupations. Thus earnings gap and different labor market histories can be generated under symmetric information.

"Discrimination" in this paper refer to the difference between the earnings of men and women under symmetric versus asymmetric information. Under symmetric information individual worker's future participation probability is known to employers whereas under asymmetric information the gender is used to infer participation probability and the "group" affiliation affects earnings. Suppose that the distributions of skills, characteristics, and preferences are ex-ante identical for men and women then our model may gives rise to "discriminatory equilibria".[17] Suppose employers believes that women have a lower likelihood of future employment than men, then women would face lower earnings than men for identical characteristics. These beliefs are self-fulfilling in equilibrium and induce women and men to make different participation and labor-supply decisions.

Moreover, if women work less in equilibrium, they will sort into occupations with lower returns to labor-market experience and lower costs of hiring new workers. Occupations with lower costs of hiring new employees will have smaller differences in wages for men and women with the same observable characteristics; the gender wage gap in these occupations may be smaller.[18]

Whereas discrimination may 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 (see for example Moro and Norman (2004)). Because there are complementaries in the utility function, i.e. consumption depends on the household budget constraint. Then there is complementarity between the hours (participation) women and men work . A discriminatory equilibrium (asymmetric equilibrium) may then be established, even if there is no coordination failure. Then men would participate more and earn more, women (married) would have higher consumption and work less. Although cross-gender complementaries exist through household consumption, this affects not only married women. Since the model is dynamic, and single individuals take into account their probability of been married in the future, then household complementaries affect single women through their future expectations.

---

[17]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.

[18]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.

Finally, although the gender earnings gap, as well as differences in experience, occupation choice can be generated when preferences are different and the information is symmetric, the model of symmetric information and the model with asymmetric information are not observational equivalent. Under asymmetric information, labor market history is used by employers to update beliefs about an individual worker's type. If men participation and hours is high relative to women then the growth rate of earnings be higher for women than men. For example, a woman who works full time for a continuous period should have a larger increase in her earnings than for a similar man. Therefore, returns to experience for women would be larger than that of men with the same labor market history

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. .Thus, for a given cohort, conditional on all publicly observable characteristics, the gender earnings gap declines with experience.

5.1. *Changes in the Gender Earnings Gap Over Time*

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

COROLLARY 5.1   *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, $F_{i0}(z_{t+1} \mid z_t, H_t)$*
   4. *Changes in fixed costs of participating*
   5. *Changes in beliefs across cohort*

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 $y_{\tau t}(h_t, H_{t-1}, z_t^p)$ increases, but if men's participation rate is high, beliefs about women's participation may increase women's 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), affects beliefs about future participation. Forth, 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 earnings because changes in beliefs about participation will raise women's earnings more than they will raise men's earnings.

Note 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 would capture that.

6. MULTIPLICITY OF EQUILIBRIA AND IDENTIFICATION

One of the main problems in estimating games or multi-agent informational models is the possibility of multiple equilibria. Before discussing different solutions to this problem, it is important to make a clarification that equilibrium uniqueness is neither a *necessary* nor *sufficient* condition for the identification of a model. Put differently, the problem of multiple equilibria is not a problem of identification.

Formally, define the mapping $\mathcal{F}$ from the space of structural characteristic $\Theta(\mathcal{M})$ to the space of conditional distributions Π such that $\mathcal{F}(\theta(\mathcal{M}))$ contains all the conditional distributions predicted by the model when the structural characteristic is $\theta(\mathcal{M})$. Multiple equilibria means that $\mathcal{F}(\theta(\mathcal{M}))$ is not a function but a correspondence. While no identification means that the inverse mapping $\mathcal{F}^{-1}(.)$, evaluated at the population distribution $F_{Y|X}^o$, is a correspondence.[19]. Therefore, we will treat the solution to the problems of multiplicity and identification separately.

The solution to the multiple equilibrium problem is based on the following intuition. The players equilibrium strategies can be recovered from the data under the assumptions the data was generated by a single equilibrium (within observable groups of agents) and that the econometrician observes all the non-idiosyncratic state variables. Conditional on other players' equilibrium strategies, each player's decision becomes a single agent maximization problem(i.e. the best response function). Since this maximization problem is a necessary condition in all equilibria, an estimator of the structural parameters can be obtain from a criterion function based on these best response. This criterion function will be well defined once the model is identified.

The approach we take in this paper to show identification of the single agent problem is constructive in nature. We rely on Chesher (2007) that shows that if there exist a functional of the conditional distribution, $\mathcal{F}^{-1}(F_{Y|X})$, with the property that the functional returns the value θ^* in all the admissible structures of the model with $\theta(\mathcal{M}) = \theta^*$, then the value of a structural characteristic $\theta(\mathcal{M})$ is identified by the model when $\theta(\mathcal{M}) = \theta^*$. The advantage of this approach to identification is that there is a natural analog estimator of the value of the structural characteristic. Given a panel data with $\{a_{nt}, H_{nt-1}, z_{nt}, c_{nt}, S_{n\tau t}\}_{n=1,\tau=1,t=1}^{N_i, \Upsilon, T}$ where n index individual, t index the year, and (with some abuse of notation) z_{nt} contains an indicator for gender, we outline of our Identification strategy below:

1. First we show that under standard regularity condition on $u_2(z_t, c_t, \varepsilon_{2t})$, if it is multiplicative separate in ε_{2t}, then standard independence assumption between z_t and ε_{2t} allows us to identify $\eta\lambda_t$ up-to proportional constant.

---

[19]For more a detailed discussion of this problem see Jovanovic(1989)

2. Assuming that z_t^p is fully observed by the econometrician, we show that, under standard regularity conditions, the equilibrium salary schedule and β are identified.
3. Let $F_1(\varepsilon_{0t}, \varepsilon_{1t})$ be the marginal of $F_1(\varepsilon_{0t}, \varepsilon_{1t}, \varepsilon_{1t})$. Given that $S_{i\tau t}(h_t, H_{t-1}, z_t^p)$, β and $\eta\lambda_t$ are identified, we show that $u_0(z_t)$ and $u_1(z_t, l_t, H_{t-1})$ are identified up-to $F_1(\varepsilon_{0t}, \varepsilon_{1t})$ and two additive constants

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

6.1. *Identification of the Marginal Utility of Wealth*

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

ASSUMPTION 6.1   *Assume that the marginal utility of consumption has the following form*

$$(6.1) \qquad \frac{\partial u_2(c_{nt}, z_{nt}, \varepsilon_{2t})}{\partial c_{nt}} = u_{2c}(c_{nt})u_{2z}(z_{nt})\varepsilon_{2nt}$$

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

ASSUMPTION 6.2   *1) Assume that $E[\log \varepsilon_{2nt}|z_{nt}] = 0$ for all n,t; and 2) Assume that $E_n[\log(\eta_n)|z_{nt}] = 0$*

ASSUMPTION 6.3   *1) Assume that z_{nt} has a continuous element z_{cnt} with continuous variation on it support $[\underline{z_c}, \overline{z_c}]$; and 2) $u_{2z}(\underline{z_c}, .) = 0$*

Assumption (6.1) states that the marginal of consumption is multiplicatively separable, for example both the class of constant absolute risk aversion and constant relative risk aversion would satisfy this assumption. Assumption 6.2(1) formally states that the error is mean independent of z_{nt} with expectation zero. Assumption 6.2(2) is the standard normalization needed in panel data model in order recover the level of the time component. Finally, Assumption 6.3 (1) states that at least one variable with continuous variation on its support is required, and Assumption 6.3 (2) is a boundary condition. Assumption 6.3 can be replaced with a parametric assumption on the function $u_{2z}(z_{nt})$.

LEMMA 6.1   *Suppose $u_{2c}(c_{nt})$ is known, and assumptions 6.1-6.3 are satisfied, then $\eta_n\lambda_t$ is identified.*

6.2. *Identification of the equilibrium salary schedule*

This section establishes the identification of the Equilibrium salary schedule.

ASSUMPTION 6.4   *1. Suppose there exist observable characteristics, τ, H_{t-1}, z_t^p, on a set of positive measure, such that*

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

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

Assumption 6.4 (1) states that in each occupation there is a range of hours, labor experience and individual characteristic for which the employers hold different belief about the two gender, Assumption 6.4 (2) states that an input of zero hours produces zero output.

LEMMA 6.2   *Under Assumptions (3.1), (3.2) and (6.4)* $y_{\tau t}(h_t, H_{t-1}, z_t^p)$, β *and* γ_τ *are identified, and there are at least two over-identifying restrictions.*

Given the Lemma above, we are able to test the specification of our equilibrium salary schedule outside of the full solution of the model.

6.3. *Identification of the utility of Nonmarket Time*

In the literature on estimation of dynamic Markovian games, it is standard to use time-series data to estimate and identify models.[20]. We extend this approach to a panel data setting, considering age-education cohort partition of data generated by a single path of play, exploiting therefore, the information contained in the repeated observation of the same players in the cohort partition along the path of play. Because different cohorts may be playing different equilibrium, we also have variation across cohort partition. We formalize the equilibrium selection discussed in the previous paragraph below.

ASSUMPTION 6.5 (Equilibrium Selection)   *Conditional on the time invariant component of* z_t^p, *the data for each age-education cohort, is generated by only one equilibrium conditional.*

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

ASSUMPTION 6.6   *The econometrician observes all the private information of the worker except for the idiosyncratic components* ε_t.

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

Let's redefine the primitives of our problem as follows:

(6.2)
$$U_{itk}(\omega_{nt}) = \begin{cases} u_{i1}(z_{nt}, H_{nt-1}, 1) & \text{for } k = 0 \\ u_{it0}(z_{nt}) + u_{i1}(z_{nt}, H_{nt-1}, l_{nt}^*) + \eta_n \lambda_t \sum_{\tau=1}^{\Upsilon} I_{n\tau t} S_{n\tau t}(h_{nt}^*, z_{it}^p, H_{nt-1}) & \text{for } k = 1. \end{cases}$$

then we can write the ex-ante value functions as

$$(6.3) \quad V_{itk}(\omega_{nt}) \equiv \max_{\{h_r\}_{t=t}^T} E_t[\sum_{s=t}^T \beta^{s-t}\{d_s[U_{is1}(\omega_{ns}) + \varepsilon_{1s}] + (1 - d_s)[U_{is1}(\omega_{ns}) + \varepsilon_{0s}]\} \mid d_t = k]$$

---

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

There are two important dimensions our game differs from the games typically estimated. First, in our model, employers learn and update beliefs based on the complete history of workers' behavior. As a result, the Markov perfect equilibrium refinement, typically used in the literature on estimation of games, would assume away a central feature of our problem[21]. Second, the state variables do not have discrete support because labor market experience has continuous components[22].

In order to generalize results from the literature so we can apply them to our model, we borrow the concept of finite state dependence from Altug and Miller (1998).

DEFINITION 6.1   *Given any value of the initial state variable ω_0, there exists a finite integer $\rho(\omega_0)$, a value of the state variable ω_ρ , and a sequence of choices, over the next $\rho(\omega_0)$ periods, denoted by $d_{kt}^{\rho(\omega_0)}(\omega_t) = (d_{kt+1}^{\rho(\omega_0)}(\omega_{t+1}), ..., d_{kt+\rho(\omega_0)}^{\rho(\omega_0)}(\omega_{t+\rho(\omega_0)}))$, such that the state will be $\omega_{\rho(\omega_0)}$ at date $\rho(\omega_0)$ for $d_t = k$ for all k.*[23]

The finite dependence assumption is a generalization of the Markovian assumption to a larger class of model. Of all the possible types of state variables the finite dependence assumption only placed restriction on those that evolves stochastically but do depend on the action the individual takes. That is, finite dependence breaks the connection between these state variables and past choice. For example in our model the probability of marriages/or the number of kids today would only be allowed to depend on the decision to work in the past of a finite lag and a summary statistic.

ASSUMPTION 6.7 (Finite State Dependence)   *The non-idiosyncratic state variable ω_t has the property of finite state dependence.*

Next we characterize the necessary condition for equilibrium, namely condition 1) of Definition 4.1. First, note that the first result in Proposition 4.2 implies that the equilibrium choice probabilities of working, are given by

$$(6.4) \qquad p_{it}(\omega_{nt}) = E(d_{nt}^o \mid \omega_{nt}) = \Pr\{V_{it1}(\omega_{nt}) + \varepsilon_{1nt} \geq V_{it0}(\omega_{nt}) + \varepsilon_{0nt}\},$$

A 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 ex-ante value functions, $V_{it1}(\omega_{nt}) - V_{it0}(\omega_{nt})$.[24] Let $Q : R \to (0,1)$ denote the mapping from the choice-specific value function to the conditional choice probabilities. That is,

$$(6.5) \qquad p_{it}(\omega_{nt}) = Q(V_{it1}(\omega_{nt}) - V_{it0}(\omega_{nt})).$$

---

[21]See Maskin and Tirole (2001) for a discussion of the Markov Prefect Equilibrium.

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

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

[24]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 (forthcoming); and Bajari and Hong (2005), among others.

Lemma 1 of Hotz and Miller (1993) shows that the inverse exists:

$$(6.6) \quad V_{it1}(\omega_{nt}) - V_{it0}(\omega_{nt}) = Q^{-1}(p_{it}(\omega_{nt})).$$

Note that the 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 $\varphi_k : [0,1] \to R$, that measures the expected value of the unobservable in the current utility, conditional on action $k \in \{0,1\}$, i.e.

$$(6.7) \quad \varphi_k(p_{it}(\omega_{nt})) \equiv E[\varepsilon_{knt} \mid \omega_{nt}, d^o_{nt} = k].$$

To set some notation, let $\omega^{(s)}_{kt}$ denote state in period $t+s$ if at time t the kth decision is taken, i.e. $d_t = k$ and the sequence of decisions for the next s periods is $d^{\rho(\omega_0)}_{kt+1}(\omega_{t+1}), ...,$ $d^{\rho(\omega_0)}_{kt+s}(\omega_{t+s})$. Also denote by $p^{(s)}_{kit}$, the probability that $d_{t+s} = 1$ conditional on $\omega^{(s)}_{kt}$, i.e. $p^{(s)}_{kit} = E[d_{t+s}|\omega^{(s)}_{kt}]$.

Combining equations $(6.5),(6.6)$ and (6.7) with the ex-ante value function (6.3) allows us to write the ex-ante equilibrium value function for any initial state ω_0:

$$
\begin{aligned}
V_{itk}(\omega_0) = \ & U_{itk}(\omega_0) + E_t \Big\{ \sum_{s=t}^{\rho(\omega_0)} \beta^s [U_{is0}(\omega^{(s)}_{kt}) + \varphi_0(p^{(s)}_{kit}) \\
& + p^{(s)}_{kit}\{Q^{-1}(p^{(s)}_{kit}) + \varphi_1(p^{(s)}_{kit}) - \varphi_0(p^{(s)}_{kit})\}] \\
& + \beta^{\rho(\omega_0)+1}[V_{it+\rho(\omega_0)+10}(\omega_{\rho(\omega_0)+1}) + \varphi_0(p^{(\rho(\omega_0)+1)}_{it}) \\
(6.8) \quad & + p^{(\rho(\omega_0)+1)}_{it}\{Q^{-1}(p^{(\rho(\omega_0)+1)}_{it}) + \varphi_1(p^{(\rho(\omega_0)+1)}_{it}) - \varphi_0(p^{(\rho(\omega_0)+1)}_{it})\}]\Big\}
\end{aligned}
$$

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

Next we characterizes using 6.6 and 6.8 the necessary conditions for equilibrium in Proposition4.2. First we characterizes the equilibrium relationship from (4.7). Substituting (6.8) into (6.6) gives equilibrium

$$
\begin{aligned}
Q^{-1}(p_{it}(\omega_{nt})) = \ & U_{it1}(\omega_{nt}) - U_{it0}(\omega_{nt}) + E_t \Big\{ \sum_{s=t}^{\rho(\omega_{nt})} \beta^s [U_{is0}(\omega^{(s)}_{0t}) - U_{is0}(\omega^{(s)}_{1t}) \\
& \varphi_0(p^{(s)}_{0it}) - \varphi_0(p^{(s)}_{1it}) + p^{(s)}_{0it}[Q^{-1}(p^{(s)}_{0it}) + \varphi_1(p^{(s)}_{0it}) - \varphi_0(p^{(s)}_{0it})] \\
(6.9) \quad & - p^{(s)}_{1it}\{Q^{-1}(p^{(s)}_{1it}) + \varphi_1(p^{(s)}_{1it}) - \varphi_0(p^{(s)}_{1it})\}\Big\}
\end{aligned}
$$

Note that all the elements from period $\rho(\omega_{nt})+1$ onward are the same irrespective of whether action 1 or 0 are taken today by Assumption 6.7, hence they fall out of the above equation. Similarly the sufficient condition for equilibrium hours can be write as

$$
\begin{aligned}
-\frac{\partial U_{it1}(\omega_{nt})}{\partial h_t} = \ & E_t \Big\{ \sum_{s=t}^{\rho(\omega_{nt})} \beta^s [\frac{\partial U_{is0}(\omega^{(s)}_{1t})}{\partial h_t} + \frac{\partial \varphi_0(p^{(s)}_{1it})}{\partial h_t} \\
& + \frac{\partial p^{(s)}_{1it}}{\partial h_t}[Q^{-1}(p^{(s)}_{1it}) + \varphi_1(p^{(s)}_{1it}) - \varphi_0(p^{(s)}_{1it})] \\
(6.10) \quad & + p^{(s)}_{1it}\frac{\partial[Q^{-1}(p^{(s)}_{1it}) + \varphi_1(p^{(s)}_{1it}) - \varphi_0(p^{(s)}_{1it})]}{\partial h_t}\Big\}
\end{aligned}
$$

Note that all the elements from period $\rho(\omega_{nt}) + 1$ onward are independent of current participation choice, and choice of hours and by Assumption(6.7) and hence they fall out of the above equations.

By Proposition(4.2), equations (6.9) and (6.10) are necessary conditions and therefore must hold in all equilibria. $\left\{ p_{it}(\omega_{nt}), \left\{ p_{0it}^{(s)}, p_{1it}^{(s)} \right\}_{s=1}^{\rho(\omega_{nt})} \right\}_{i=w,m}$, and the distribution over which E_t is taken, are the only elements which differ, across the different equilibria. $\left\{ p_{it}(\omega_{nt}), \left\{ p_{0it}^{(s)}, p_{1it}^{(s)} \right\}_{s=1}^{\rho(\omega_{nt})} \right\}_{i=w,m}$, are conditional expectation functions, and given Assumption (6.5), can be recovered from the data, and therefore, are identified.[25] For the purpose of the identification of the utility of nonmarket hours, we can treat $\left\{ p_{it}(\omega_{nt}), \left\{ p_{0it}^{(s)}, p_{1it}^{(s)} \right\}_{s=1}^{\rho(\omega_{nt})} \right\}_{i=w,m}$ as known. Denote by $\eta_n^o \lambda_t^o$ and $S_{int\tau}^o(h_{nt}^*, z_{nt}^p, H_{nt-1})$ the shadow prices and salary schedule under the true equilibrium in the data respectively. The true probabilities under the true equilibrium in the data are denoted by $\left\{ p_{it}^o(\omega_{nt}), \left\{ p_{0it}^{o(s)}, p_{1it}^{o(s)} \right\}_{s=1}^{\rho(\omega_{nt})} \right\}_{i=w,m}$, and define

$$
\begin{aligned}
Y_{i1nt} \equiv\ & \eta_n^o \lambda_t^o \sum_{\tau=1}^{\Upsilon} I_{nt\tau} S_{in\tau t}^o(h_{nt}^*, z_{nt}^p, H_{nt-1}) + \sum_{s=t}^{\rho(\omega_{nt})} \beta^s \left\{ \varphi_0(p_{0it}^{o(s)}) - \varphi_0(p_{1it}^{o(s)}) \right. \\
& + p_{0it}^{o(s)}[Q^{-1}(p_{0it}^{o(s)}) + \varphi_1(p_{0it}^{o(s)}) - \varphi_0(p_{0it}^{o(s)})] \\
& \left. - p_{1it}^{o(s)}\{Q^{-1}(p_{1it}^{o(s)}) + \varphi_1(p_{1it}^{o(s)}) - \varphi_0(p_{1it}^{o(s)})\} \right\} - Q^{-1}(p_{it}^0(\omega_{nt}))
\end{aligned}
$$
(6.11)

and

$$
\begin{aligned}
Y_{i2nt} \equiv\ & \eta_n^o \lambda_t^o \sum_{\tau=1}^{\Upsilon} I_{int\tau} \frac{\partial S_{int\tau}^o(h_{nt}^*, z_{nt}^p, H_{nt-1})}{\partial h_t} - \sum_{s=t}^{\rho(\omega_{nt})} \beta^s \left\{ \frac{\partial \varphi_0(p_{1it}^{o(s)})}{\partial h_t} \right. \\
& + \frac{\partial p_{1it}^{o(s)}}{\partial h_t}[Q^{-1}(p_{1it}^{o(s)}) + \varphi_1(p_{1it}^{o(s)}) - \varphi_0(p_{1it}^{o(s)})] \\
& \left. + p_{1it}^{o(s)} \frac{\partial[Q^{-1}(p_{1it}^{o(s)}) + \varphi_1(p_{1it}^{o(s)}) - \varphi_0(p_{1it}^{o(s)})]}{\partial h_t} \right\}
\end{aligned}
$$
(6.12)

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

LEMMA 6.3   *In all equilibria the following system of equation holds :*

$$
\begin{aligned}
Y_{i1nt} =\ & u_{i1}(z_{nt}, H_{nt-1}, 1) - u_{it0}(z_{nt}) - u_{i1}(z_{nt}, H_{nt-1}, l_{nt}^*) \\
& + \sum_{s=t}^{\rho(\omega_{nt})} \beta^s [u_{is1}(z_{1nt}^{(s)}, H_{1ns-1}^{(s)}, 1) - u_{is1}(z_{0nt}^{(s)}, H_{0ns-1}^{(s)}, 1)] + \xi_{i1nt}
\end{aligned}
$$
(6.13)

$$
\text{(6.14)} \quad Y_{i2nt} = -\frac{\partial u_{i1}(z_{nt}, H_{nt-1}, l_{nt}^*)}{\partial h_t} - \sum_{s=t}^{\rho(\omega_{nt})} \beta^s \left[\frac{u_{i1s}(z_{1nt}^{(s)}, H_{1ns-1}^{(s)}, 1)}{\partial h_t} \right] + \xi_{i2nt}
$$

where $z_{knt}^{(s)}$ and $H_{kns-1}^{(s)}$ denote the worker's type and labor experience in period $t+s$, if at time t $d_t = k$, the sequence of decisions for the next s periods are $d_{kt+1}^{\rho(\omega_0)}(\omega_{t+1}), ..., d_{kt+s}^{\rho(\omega_0)}(\omega_{t+s})$, $E_t^o[\xi_{ijnt}|\omega_{nt}] = 0$ for $j = \{1, 2\}$, and $i = \{w, m\}$, and E_t^o is taken over the actual equilibrium played; formal definition of the residual ξ_{ijnt} is in the proof in the Appendix.

Note that the data is informative about E_t^o under Assumption(6.5). The identificability of

the structural functions depends on whether these functions can be deduced if $E_t^o[Y_{ijnt}|\omega_{nt}]$ is known. The following lemma establishes the main identification result.

LEMMA 6.4 (Nonmarket Utility Identification)   *Under Assumption(3.1)-Assumption(6.7)* $u_{it0}(z_{nt})$ *is identified up-to an additive constant, and* $u_{it1}(z_{nt}, H_{nt-1}, l_{nt})$ *is identified up-to an additive function of* z_{nt} *and* H_{nt-1}.

The above Lemma implies that, the levels of nonmarket hours utility are not identified, but the marginal utility of nonmarket hours are identified.[26] This result is stronger than identification results in discrete choice models with discrete state variables, in which only the difference between the utilities conditional on the choices is identified (see Magnac and Thesmar (2002) and Pesendorfer and Schimdt-Dengler (2006)). In this way our identification result is similar to that found in Bonet and Pesendorfer (2003). The continuous choice of hours in our model increases the identification power of the model: if the level of utility from not working is normalized, then the rest of to be fully identified.

7. EMPIRICAL STRATEGY

The estimation of the model follows the outline of the identification strategy. *First,* the marginal utility of wealth is estimated from the consumption Euler equation. *Second,* the earnings equation is estimated from the zero profit condition. *Third,* the conditional choice probabilities and the firms' equilibrium belief are estimated nonparametrically. *Lastly,* we form the empirical analogue of equations (6.13) -(6.14). Using method of moment estimation procedure, we recover the parameters of the utility from non-market work and the risk aversion parameter.

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

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

To address the first issue, we assume consumption is private information, and hence the marginal utility from wealth is private information[27]. All spouse related variables (such as education), marital status, the number of children and age distribution of children are assumed to be private. We assumed that employers observed the actual hours worked in each occupation for the most recent three years, and the total numbers years worked in each

---

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

[27]Note that the marginal utility from wealth is a sufficient statistic for spouse labor income and hence we are assuming that spouse labor income is private information as well.

occupation. This assumption is made in order to reduce the dimension of the conditioning variables in the nonparametric estimation of beliefs; this restriction is not imposed, however, by the finite state dependence assumption. Finally, given the functional form restrictions on the production function we can account for individual specific productivity differences. These differences are assumed to be observed by the firms and workers, but not by the econometrician. The following sections provide a brief overview of each stage of the estimation and a description of the functional form restrictions imposed at each stage. A detailed description of the estimation procedure can be found in the accompanying estimation appendix.

7.1. *Consumption Equation*

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

$$(7.1) \qquad u_2(c_{nt}, z_{nt}, \varepsilon_{2nt}) = \exp(z'_{nt}B_4 + \varepsilon_{2nt})c^{\alpha}_{nt}/\alpha.$$

The Euler equation of consumption becomes

$$(7.2) \qquad \exp(z'_{nt}B_4 + \varepsilon_{2nt})c^{\alpha-1}_{nt} = \eta_n\lambda_t$$

Under assumption(6.2), it is possible to estimate $(1-\alpha)^{-1}B_4$, $(1-\alpha)^{-1}\eta_n$, and $(1-\alpha)^{-1}\lambda_t$ using panel data on individual consumption, and characteristics z_{nt}; see discussion in Heckman and Macurdy(1980), MaCurdy (1981) and Altug and Miller (1990,1998)) among others. There are several issues to address when implementing this estimation strategy.

First, the only panel data set which includes reliable labor market variables and consumption related variable in the US is the PSID; the PSID, however, includes data on food consumption only[28]. Therefore, most papers estimating labor supply and consumption behavior jointly, use food consumption as the basis for estimating consumption equations (see Blundell and Macurdy (2007) for a survey of this literature). This can be justified under the assumption that food consumption is additively separable from non-food consumption with a quasi-linear utility function. Second, food consumption is only observed at the household level, whereas in the model consumption is chosen at the individual level. To address this issue, we divide consumption among the household members by allowing the data to implicitly adjust the weight given to each household member, accounting for the number of individual in the household and the age distribution of members of the household. Third, following standard practice we include regional dummies in z_{nt} to account for differences in consumption prices across regions.

Finally, $(1-\alpha)^{-1}\eta_n$ is estimated for each individual in our sample. Therefore, the traditional fixed effect estimations (as used in Heckman and MaCurdy(1980)) would be biased for small T. As shown in Macurdy(1981), $(1-\alpha)^{-1}\eta_n$ can be written as a function of individual specific variables. We follow the Altug and Miller(1998) implementation of this idea, and construct a nonparametric estimator which is consistent as the number individual goes to infinity; we use years of completed education, gender, marital status by age 30, age distribution of kids at age 35, home ownership at 35 and geographical location at age 30 in this estimation.

---

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

7.2. *Earning Equation*

We assume that the production function has the following functional form

$$(7.3) \qquad y_{\tau t}(h_{nt}, H_{nt-1}, z_{nt}^p) = b_{0\tau t} + b_{\tau 1} h_{nt} + b_{\tau 2} h_{nt}^2 + \sum_{r=1}^{\rho} b_{\tau 3r} h_{nt-r} + z_{nt}^{p'} B_{\tau 5} + \nu_n.$$

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

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

7.3. *Conditional Choice Probabilities and Equilibrium Beliefs*

There are five inputs of equations (6.11) and (6.12) that are needed to be estimated before we can form the empirical counterpart of Y_{i1nt} and Y_{i2nt}. First, Y_{i1nt} is a function of the equilibrium salary schedule, which is a function of the employers' beliefs (equation(4.1)). These beliefs will be estimated nonparametrically. Second, Y_{i2nt} is a function of the derivative of the equilibrium salary schedule with respect to current hours; we estimate these derivative nonparametrically. Third, Y_{i1nt} is a function of the current conditional choice probabilities, $p_{int}^0(\omega_{nt})$, which we will also estimate nonparametrically. Finally, Y_{i1nt} and Y_{i2nt} are both functions of $p_{kint}^{o(s)}$ and there derivative respectively. The following subsection discuss the estimation of all these elements.

7.3.1. *Estimation of Equilibrium Beliefs and Derivative*

The equilibrium beliefs for each occupation, $\widetilde{p}_{in\tau t}$, are computed as a nonlinear regression of the product of next-period participation and occupation choice index, $d_{nt+1} \times I_{n\tau t+1}$ on today's public information variables, z_{nt}^p, work histories, H_{nt-1}, and hours worked, h_{nt}, conditional on working today in occupation τ. Let $X_{nt} = (z_{nt}^p, H_{nt-1}, h_{nt}, \nu_n, Gender_n)$ and

$NY_{n\tau t-1}$ be the total number of years worked in occupation τ up to period $t-1$; since only two occupations are used in the estimation, $\tau \in \{1, 2\}$. The labor market history use in this paper is defined as

(7.4)
$$H_{nt-1} = (NY_{n1t-1}, NY_{n2t-1}, d_{nt-3}I_{n1t-3}, d_{nt-3}I_{n2t-3}, ..., d_{nt-1}I_{n1t-1}, d_{nt-1}I_{n3t-1}, h_{nt-3}, .., h_{nt-1})$$

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

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

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

7.3.2. *Estimation of Conditional Choice Probabilities*

The estimate of the conditional choice probabilities requires us to be more specific about the state variables. In contrast to the beliefs, the conditional choice probabilities are defined from the workers prospective and not the firms prospective. Recall, that the non-idiosyncratic state variable for the workers problem is $\omega_{nt} \equiv (H_{nt-1}, z_{nt}, \eta_n \lambda_t)$. From the estimation of the consumption equation $\eta_n \lambda_t$ is known up to a proportional constant. The elements included in z_{nt} are n*umber of individual in the family unit, number of children less than three, number of children between three and fourteen, age, years of completed education, marital status, the number of years of education of spouse if married, and gender.*

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

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

7.3.3. *Estimation of Finite State Path Probabilities and Derivative*

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

The hypothetical labor market history is defined as
$$\begin{aligned} H_{1nt}^{(s)} &= (NY_{n1t-1+s}, NY_{n1t-1+s}, d_{nt-3+s}I_{n1t-3+s}, d_{nt-3-s}I_{n2t-3-s}, ..., d_{nt-1}I_{n1t-11}, \\ &\quad d_{nt-1}I_{n2t-12}, I_{n1t}, I_{n2t,}, 0, ..., 0, h_{nt-3+s}, .., h_{nt-1}, h_{nt}^*, 0, .., 0) \end{aligned}$$
(7.7)
and
$$\begin{aligned} H_{0nt}^{(s)} &= (NY_{n1t-1+s}, NY_{n2t-1+s}, d_{nt-3+s}I_{n1t-3+s}, d_{nt-3-s}I_{n2t-3-s}, ..., d_{nt-1}I_{n1t-1}, \\ &\quad d_{nt-1}I_{n2t-1}, 0, 0, 1I_{n1t}, 1I_{n2t,}, 0, ..., 0, h_{nt-3+s}, .., h_{nt-1}, 0, h_{nt}^*, 0, .., 0) \end{aligned}$$
(7.8)

where

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

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

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

Thus, assuming a Markovian transition of the variables z_{nt}, we have found the sequence of decision satisfy the assumption of finite state dependence. Lets define the following participation indices that corresponds to the sequence of decisions that get one to $H_{knt}^{(s)}$;

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

and

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

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

$$(7.9) \quad p_{iknt}^{(s,N)} = \frac{\sum_{n'=1, n' \neq n}^{N} \sum_{r=1}^{T} d_{n'r} d_{kn'rs}^{(s)} J\left[\delta_N^{-1}\left(\omega_{knt}^{(s)N} - \omega_{n'r}^N\right)\right]}{\sum_{n'=1, n' \neq n}^{N} \sum_{r=1}^{T} d_{kn'r}^{(s)} J\left[\delta_N^{-1}\left(\omega_{knt}^{(s)N} - \omega_{n'r}^N\right)\right]},$$

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

$$(7.10) \quad f_{i1nt}^{(s)} \equiv f_{i1}\left(\omega_{1nt}^{(s)} \mid d_{nt+s} = 1\right)$$

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

$$(7.11) \quad \frac{\partial p_{i1nt}^{(s)}}{\partial h_{nt}} = \left[\frac{f_{i1nt}^{'(s)}}{f_{i1nt}^{(s)}} - \frac{f_{int}^{'(s)}}{f_{int}^{(s)}}\right] p_{1nt}^{(s)} \quad s = 1, \ldots, 3.$$

We derive this expression using the representation of $p_{i1nt}^{(s)}$ as $p_{i1nt}^{(s)} = \Pr\left(d_{nt+s} = 1 \mid \omega_{1nt}^{(s)}\right) = \Pr(d_{nt+s} = 1) f_{i1nt}^{(s)} / f_{int}^{(s)}$. Differentiating this expression with respect to h_{nt} yields the above expression. The nonparametric estimates of $f_{i1nt}^{(s)}$ and $f_{int}^{(s)}$ are defined, respectively, as the numerators and denominators of $p_{iknt}^{(s)N}$ in equation (7.11). The estimates of $f_{i1nt}^{'(s)}$ and $f_{int}^{'(s)}$ are obtained from the derivatives of the estimates, $f_{i1nt}^{(s)N}$ and $f_{int}^{(s),N}$, with respect to h_{nt} (Pagan and Ullah, 1999).

7.4. *Utility of Nonmarket Time*

The identification result in Proposition(6.4) assumes that the distribution of $(\varepsilon_{0nt}, \varepsilon_{1nt})$ is known. Because we are not estimating the utility from non-market labor nonparametrically, we can relax that assumption, and allow for the variance to be unknown; specifically, we assume $(\varepsilon_{0nt}, \varepsilon_{1nt})$ is distributed as a Type I extreme value with variance parameter σ^2 and mean zero. This distributional assumption for the preference shocks implies that $Q^{-1}(p) = \sigma \ln[p/(1-p)]$, $\varphi_0(p) = \frac{\zeta}{\sigma} - \sigma \ln[(1-p)]$, and $\varphi_1(p) = \frac{\zeta}{\sigma} - \sigma \ln[p]$, where ζ is the Euler constant. Under this distribution assumption Y_{i1nt} and Y_{i2nt} simplifies to

$$(7.12) \quad Y_{i1nt} \equiv \eta_n^o \lambda_t^o \sum_{\tau=1}^{\Upsilon} I_{in\tau t} S_{in\tau t}^o(h_{nt}^*, z_{nt}^p, H_{nt-1}) + \sigma \sum_{s=t}^{4} \beta^s \ln\left(\frac{1 - p_{i1nt}^{(s)}}{1 - p_{i0nt}^{(s)}}\right) - \sigma \ln[p_{int}/(1-p_{int})]$$

and

$$(7.13) \quad Y_{i2nt} \equiv \eta_n^o \lambda_t^o \sum_{\tau=1}^{\Upsilon} I_{in\tau t} \frac{\partial S_{in\tau t}^o(h_{nt}^*, z_{nt}^p, H_{nt-1})}{\partial h_{nt}} + \sigma \sum_{s=1}^{4} \beta^s \left(1 - p_{i1nt}^{(s)}\right)^{-1} \frac{\partial p_{i1nt}^{(s)}}{\partial h_{nt}}$$

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

Lastly we specify the form of the Non-market hours utility function. We allow the fixed cost of participating to change over time in order to capture the possible changing home production technology. It takes the form:

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

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

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

Under the above specifications, Lemma(6.3) yields:

$$Y_{i1nt} = -B_{0t} - z_{nt}' B_{i1} + z_{nt}' h_{nt} B_{i2} + \theta_{0i}\left(1 - l_{nt}^2\right) + \sum_{s=1}^{\rho} \theta_{si} h_{nt}(l_{nt-s} + \beta^s) + \xi_{i1nt}$$

$$(7.14) \quad Y_{i2nt} = z_{nt}' B_{i2} + 2\theta_{i0} l_{nt} + \sum_{s=1}^{\rho} \theta_{si}(l_{nt-s} + \beta^s) + \xi_{i2nt}$$

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

8. 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 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).

Next we discuss the results of the estimation of the salary schedule. The over-identifying restrictions can not be rejected at the 5% level of significance. First we will discuss the estimated productivity shock and the other component of the salary schedule.

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. or example, the estimated aggregate productivity increased by 80% from the mid 70s to the late 80s in the professional occupations and only by 33% in nonprofessional occupations over the same period.

The estimation results of the earnings equation are reported in Table 4. Coefficient on current hours worked is larger in professional occupation (183,392 versus 100,688).Consistent with Corollary 1, the professional occupations have significantly higher returns to labor-market experience than the nonprofessional occupations.[29] But, the returns to working less hours (part time) in the nonprofessional occupations is higher; this can be seen by comparing the linear and quadratic terms in current hours (COMPUTE NUMBER!). There is a larger cost of hiring a new worker (3032 versus 875).

These results are consistent with the empirical fact that women's representation in nonprofessional occupations is higher. Our model predicts, that women will sort into occupations with lower higher returns to labor-market experience, higher returns to working less hours and lower costs of hiring new workers if in equilibrium, they work less and are less attached to the labor market. The increase in productivity in professional occupations relative to nonprofessional occupation is consistent with increase in women's representation in professional occupation, as is documented in Table 1 and other literature (see Lewis, 1996)..[30] 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 5.1(2)).

Table 5 contains the estimates of the fixed cost of labor-force participation.[31] There is no significant difference in the cost of participation for men and women with the same years of completed education (notice that a positive coefficient implies higher dis-utility). 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

---

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

[30]This is in keeping with the empirical finding the increasing returns to skill.

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

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.

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

Table 7 contains the estimates for the time nonseparability in nonmarket hours. They show that there are significant complementaries between nonmarket hours across time for women. The results on complementaries 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.

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. These results are also supportive of the idea of cross-group complementaries in the utility function (asymmetric equilibrium). That is, 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. 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.[32] 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,

---

[32]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.

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.

First we assess the importance of labor market frictions, hiring costs, on labor supply, the earnings gap, occupation composition and the changes in those over these two disjoint time periods. Without costs of hiring new workers, earnings equal the worker's productivity, that is $S_{i\tau t}(h_{nt}, H_{nt-1}, z^p_{nt}) = y_{\tau t}(h_{nt}, H_{nt-1}, z^p_{nt})$. We then simulate the model with out hiring cost and computed the earnings gap, participation rates, hours worked and occupational composition for the two periods, respectively. The inputs into these simulations 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. The row labelled hiring cost in Tables 9 through 11 contains the result form from that simulation.

In Table 9 the predicted wage gap in professional occupation under this simulation, would have been 81% in 1974-1978 instead of 52% as predicted by bench mark model and 96% in 1984-1988 instead of 67%. Table 10 shows that with no labor market frictions women's' participation rate in the 70's and 80's (57%) would have been lower if there were no frictions (51% versus 62% and 57% versus 70%, respectively). Female representation in professional occupations would have been lower (30% instead of 34% and 35% instead of 40%) and that the increase in participation would have been smaller. Finally Table 11 shows that the hours worked by women who participate in an economy with no labor market frictions is larger than it is in an economy than in the benchmark model (1820 versus 1702 and 2050 versus 2100 respectively). In fact the 80's women who participate in an economy with no frictions are working almost as much as men (2050 versus 2100). A similar pattern is observed in Nonprofessional, except that the earnings gaps are lower that in professional but still higher than in the benchmark economy, there is a higher percentage of women in that nonprofessional than professional but still lower than in the benchmark economy and while the hours are higher than the benchmark economy, they are not as close to the hours worked by men as in the professional occupations.

The above results demonstrated two important things. First, without market friction the selection of the women than works is significantly different from the market friction economy. Women participates less and when the do the work more in the nonprofessional occupation but he ones who do participate works more hours and in professional occupations hours that are similar to men. This is because in an economy with private information the low attachment women pretend to be high attachment type and high attachment women work less. With no market friction there is a separation between the two types. Second, the change in the earnings gap over the two disjoint periods would be have been small by about 34% and 20% in professional and Nonprofessional Occupations respectively. That means that

market friction also plays an amplification role. For example, suppose that there is an increase in overall productivity within an occupation. Such an increase affects the wages of all workers because $y_{\tau t}(h_t, H_{t-1}, z_t^p)$ increases, but if men's participation rate is high, beliefs about women's participation may increase women's wages relative to men's wages. This increase in wage will result in a bigger increase in labor supply and participation of women. However, there would an additional increase that would come from the updating of employer's beliefs to reflect the higher participation of women.

The question still remains as to was the cause of the remaining 66% and 80% of the change in the gender earnings gap over the two period? In order to answer that question we simulate our model under that assumption of no market friction, holding each of our three remain exogenous source of variation, i,e, demographic change, aggregate participation cost change, and aggregate productivity increases, constant over two period. Tables 9 through 11 contain the results of these simulations under the headings, Demographics, participation Cost and Aggregate Production. Each is done by holding the other two components fixed at the 1970s level and only allowing the particular component to change in the 1980s. The result from that exercise show that of the remaining 66% change in the earning gap in professional occupation over the period; 35% is due to increase in aggregate productivity; 28% in to changes in demographic (mainly fertility decline) changes; and 3% is due in increase in the aggregate cost of participating in the labor force. In the Nonprofessional occupations, of the 80% change in the earnings gap 47% is due aggregate productivity increase; 29% is due demographic changes and 4% is due to changes in the aggregate cost of participation.

Finally we quantify the effect of the discrimination on the earnings gap and the change in the earnings gap. 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_τ 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. Table 9 shows that under symmetric information, the earnings gap would have been 71% instead of 52% in the 70's and 82% instead of 67% in the 80's. Table 10 demonstrates that participation would have been 62% instead of 70, percentage of women in professional occupations would have been 38 instead of 40 and hours 1820 instead of 1702 and 2090 instead of 1960. The intuition for the lower levels of participation and occupation composition is similar to the case with no frictions: there is no value of signalling. However, unlike no frictions, there is still larger earnings gaps due to lower attachment of women affecting hours supplied of women who work.

9. 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. Our model predicts that such an increase allows for relative gains for women causing an increase in female representation in professional

occupations over time. Further analysis shows that market frictions significantly amplifies exogenous changes in our model. Without labor market frictions the earnings gap would have been small by at least 45%, women would have participation less in the labor force but the ones that participated would have worked more and earned income similar to men. The decomposition of the change in the gender earnings gap reveals that changes in private information and hiring cost accounts for over 33% of the change in professional occupation while increase aggregate labor market productivity accounted for 23% and demographic changes (mainly fertility decline) accounted for about 18%. Similar results were found for the Non-professional occupations. The estimation results do not support 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. 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 earnings 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.

Tepper School of Business, Carnegie Mellen University
ggayle@andrew.cmu.edu, http://www.andrew.cmu.edu/user/ggayle/
                                and
Tepper School of Business, Carnegie Mellon University
limor@andrew.cmu.edu, http://www.andrew.cmu.edu/user/limor/web.html

APPENDIX A: 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.[33]

---

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

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.

APPENDIX B: THEORETICAL RESULT PROOFS

Let $\overline{z}_{\tau t}^*(H_{t-1}, z_t^p)$ be the *type* that works the most in occupation τ and $\underline{z}_{\tau t}^*(H_{t-1}, z_t^p)$ be the type which works the smallest fraction of time in the occupation. Let $\overline{\mu}_\tau = \mu_{it}(\overline{z}_{\tau t}^* \mid H_{t-1}, z_t^P)$ be the beliefs about the type $\overline{z}_{\tau t}$, and $\underline{\mu}_\tau = \mu_{it}(\underline{z}_{\tau t}^* \mid H_t, z_t^p)$ be the beliefs about the type $\underline{z}_{\tau t}^*$.

ASSUMPTION B.1 (Off-equilibrium path)   *if* $\forall z_t^*$, $\sigma_i(a_t \mid H_{t-1}, z_t^*) = 0$, *then* $\mu_{it}(z_t^* \mid H_{t-1}, z_t^p) = \underline{\mu}_\tau$ *if* $h_t < \underline{h}_{\tau t}$, *and* $\mu_{it}(z_t^* \mid H_{t-1}, z_t^P) = \overline{\mu}_\tau$ *if* $h_t > \overline{h}_{t\tau}$

PROOF OF LEMMA 4.1:   We show that it is optimal to accept the highest salary offer. First, given any choice of h_t, the current utility is increasing in salary, and secondly, the given beliefs and any h_t, the continuation value of the worker is non-decreasing if she chooses higher salary. The increase in salary enters the value function through the frisch demand see equation B.5 . Since $\eta_n \lambda_t > 0$ then the current utility is increasing in $S_{i\tau t}(h_t, H_{t-1}, z_t^p)$. Next we show that given choice of h, the continuation value is unchanged by a higher salary. We begin by showing that a higher salary does not affect tomorrow's beliefs. First, given hours choice one occupation is chosen by **Assumption 3.1.** Changing employers within occupation with the same hours worked does not change the the beliefs. Second, we assume salary is not observed by outside employers, hence it is not part of employment history and does not affect beliefs. Therefore, accepting highest salary given hours is optimal.   *Q.E.D.*

PROOF OF PROPOSITION 4.1:   The free-entry condition implies that in equilibrium, the expected value of a vacancy in each occupation at any period, $\pi_{\tau t}$, is zero. Thus $\pi_{\tau t}$ is the continuation value of hiring a new worker in occupation τ in period t. Define $\pi_{t\tau}^e$ by the continuation value of the current employer in occupation τ in period t. That is, $\pi_{\tau t}^e$ 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, $S_{i\tau t}(h_t, H_{t-1}, z_t^p)$, is

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

Therefore,

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

The current employer's profit, substituting the earnings is

(B.2) $\pi_{\tau T}^e = y_{\tau T}(h_T, H_{T-1}, z_T^p) - S_{i\tau T}(h_T, H_T, z_T^P) = \gamma_\tau$

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

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

Thus,

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

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

$$
\begin{aligned}
\pi^e_{\tau T-1} &= y_{\tau T-1}(h_{T-1}, H_{T-2}, z^p_{T-1}) - S_{iT-1\tau}(h_{T-1}, H_{T-1}, z^P_{T-1}) \\
&+ \beta \widetilde{p}_{i\tau T}(z^p_{T-1}, H_{T-1})\pi^e_{\tau T} = \gamma_\tau.
\end{aligned}
$$

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

$$
\begin{aligned}
\pi_{\tau T-s} = y_{\tau T-s}(h_{T-s}, H_{T-s-1}, z^p_{T-s}) &- \gamma_\tau - S_{i\tau T-s}(h_{T-s}, H_{T-s}, z^p_{T-s}) \\
&+ \beta \widetilde{p}_{i\tau T-s+1}(z^P_{T-s}, H_{T-s})\pi^e_{\tau T-s+1} = 0
\end{aligned}
$$

and, therefore,

$$
\text{(B.3)} \qquad S_{i\tau T-s}(h_{T-s}, H_{T-s}, z^p_{T-s}) = y_{\tau T-s}(h_{T-s}, H_{T-s-1}, z^p_{T-s}) - \gamma_\tau(1 - \beta \widetilde{p}_{i\tau T-s+1}(z^P_{T-s}, H_{T-s})).
$$

Therefore, the current employer's profit is

$$
\text{(B.4)} \qquad
\begin{aligned}
\pi^e_{\tau T-s} = y_{\tau T-s}(h_{T-s}, H_{T-s-1}, z^p_{T-s}) &- S_{iT-s\tau}(h_{T-s}, H_{T-s}, z^p_{T-s}) \\
&+ \beta \widetilde{p}_{i\tau T-s+1}(z^P_{T-s}, H_{T-s})\pi^e_{\tau T-s+1} = \gamma_\tau
\end{aligned}
$$

Given the beliefs and worker's strategy to accept the highest offer by Lemma 4.1, and other firm's strategies, equation(4.4) is the competitive salary schedule. $\qquad$ *Q.E.D.*

PROOF OF PROPOSITION 4.2: Using the Bellman principal, the ex-ante value function for an individual who chooses to participate in the labor force in period t and to behave optimally thereafter is $\qquad$ *Q.E.D.*

$$
\text{(B.5)} \qquad V_{i1}(\omega_t)
$$

$$
\begin{aligned}
= \max_{h_t \in (0,1)} \Bigg[&u_1(l_t, H_{t-1}, z_t) + \eta \lambda_t \sum_{\tau=1}^{\Upsilon} I_{\tau t} S_{i\tau t}(h_t, z^p_t, H_t) \\
&+ \beta E_t\{p_{it+1}[V_{i1}(\omega_{t+1}) + \varepsilon_{1t+1}] \mid \omega_t, h_t > 0\} \\
&+ \beta E_t\{(1 - p_{it+1})[V_{i0}(\omega_{t+1}) + \varepsilon_{1t+1}] \mid \omega_t, h_t > 0\} \Bigg].
\end{aligned}
$$

Then $Q(z_t, S_t)$ and $h^*_{it}(z_t, H_t)$ follow directly equations(4.3) -equation(4.6) along with the above equation. $I^0_\tau(z_t, H_t)$ comes follows directly from Assumptions 3.1 and 3.2. Lastly we need to show that given off-equilibrium path beliefs, off-the-equilibrium path by Assumption B.1, on workers who work fewer hours than the minimal (optimal) hours receive the productivity plus the beliefs component attached to the marginal type who works the least hours, and there are no gains from deviation (beliefs will not change), thus, there is no profitable deviation for working less. Same argument applies to working more than the maximum optimal hours, as beliefs are not adjusted to be higher than the beliefs for hours worked more than the highest optimal hours.

PROOF OF LEMMA 4.2: Now we show that the contract which satisfies the zero profit condition is optimal. In order to establish that, we need to show that given other firms' offering the competitive rate, and the worker's strategy and the firm's beliefs, no firm can deviate from the competitive rate and strictly increase its expected profits. First we show that by offering a lower wage rate, the firm cannot increase profit. From Lemma 4.1 workers accept the highest offer, thus a deviation to a lower wage implies the worker rejects the offer and the payoff is zero. Thus, offering a lower salary is not a profitable deviation.

Considering a firm offering a salary, $\widetilde{s}_{i\tau t}$ for h_t such that $\widetilde{s}_{i\tau t} > s^0_{i\tau t}(h_t, H_t, z^p_t)$, where $s^0_{i\tau t}(h_t, H_t, z^p_t)$ is the competitive salary offered by other firms and satisfy the zero profit condition. Worker's state variable are not a function of past salaries therefore, at $t+1$ worker's state variables remain H_t, z^p_t, z_{t+1} and competing firms' offers are unchanged $s^0_{i\tau t+1}(h_{t+1}, H_t, z^p_{t+1})$ as past salary histories are unobserved . Therefore, $Q(z, S^0_t)$, $I^0_\tau(z_t, H_t)$ remain the same. Given the beliefs $\mu_t(z_t \mid H_t, z^p_{nt})$, the probability of participation next period remains unchanged :

$$
\begin{aligned}
\widetilde{p}_{i\tau t+1}(H_t, z^p_t, \widetilde{s}_{\tau t}) &= \widetilde{p}_{i\tau t+1}(H_t, z^p_t) \\
&= \int Q(z^*_{t+1}, H_t, h_{t+1}) I_{\tau t+1}(z^*_{t+1}, H_t, h_{t+1}) \mu_{i,t+1}(z^*_{t+1} \mid H_t, z^P_t)\, \mathrm{d}z^*_{t+1}.
\end{aligned}
$$

Therefore, as established in equationB.3 the continuation expected profit at can be written as

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

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

off-equilibrium beliefs are described in Assumption B.1. Given these beliefs the competitive salary schedule above is optimal.   Q.E.D.

PROOF OF COROLLARY 4.1:   In order to prove this we first show necessity. Suppose there exists an equilibrium in which equation 4.11 does not have a fixed point. Then take any t, τ and i, the probability that a worker remaining in the firm at $t+1$ is either higher or lower than $\widetilde{p}_{i\tau t+1}$. By equation(B.3) and because on the equilibrium path the beliefs are consistent, the zero expected profit condition hold. Since $\widetilde{p}_{i\tau t+1}$ is not equal to the probability of next period participation, the zero profit condition is violated, and hence cannot constitute an equilibrium.

Next we show sufficiency. Suppose equation 4.11 has a fixed point. Then for any t, τ and i, by Lemma 4.1 the competitive salary schedule exists. Given the competitive salary schedule by Proposition 4.2– the worker's strategies of hours and participation and occupation choice and consumption exists and is unique. Hence by Definition 4.1conditions 1,2,3 exist (mutual best responses by construction, beliefs satisfy Bayes' law).   Q.E.D.

PROOF OF PROPOSITION 4.3:   Given the triangular nature of the system of equations in (4.11), it is sufficient to show existence for each equation in its own variable.

Existence of a solution to the worker's consumption and hours problem follows immediately from continuity and strict concavity of the utility function the fact that there is a solution to the worker's problem for any set of contracts offered.

Next, note that any period t,occupation τ and gender i, $\widetilde{p}_{i,\tau,t+1}$ is the solution to

$$
\begin{aligned}
\widetilde{p}_{i\tau t+1} = \int_{z_t^*} \Bigg\{ &\int_{z_{t+1}^*} f(z_{t+1}^* \mid z_t^*) Q(z_{t+1}^*, H_t(\widetilde{p}_{i,\tau,t+1}, ..), h_{t+1}^*(\widetilde{p}_{i\tau t+2}) \\
&I_{t+1}(z_{t+1}^*, H_t(\widetilde{p}_{i\tau t+1}, ..), h_{t+1}^*(\widetilde{p}_{i\tau t+2}), S_{i\tau t+1}(h_{t+1}^*), \underline{h}_{\tau t+1}(H_t, z_t^p), \\
&\overline{h}_{\tau t+1}(H_t, z_t^p) \, dz_{nt+1}^* \Bigg\} \mu_t(z_t^* \mid H_t, z_t^p) \, dz_t^*
\end{aligned}
$$

Here we only make explicit the arguments of interest.

A) Note that $\widetilde{p}_{i\tau t+1} : [0,1]$, 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.

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

Since $I_{\tau t+1}(z_{t+1}^*, H_t(\widetilde{p}_{i\tau nt+1}, ..., \widetilde{p}_{i\tau 2}), h_{t+1}^*, S_{i\tau t+1}(h_{t+1}^*))$ is an indictor function that takes the value 1 when $h_{nt+1}^* \in (\underline{h}_\tau(H_t, z_{nt+1}^p), \overline{h}_\tau(H_t, z_{t+1}^p))$ and 0 otherwise, we can rewrite the integral

$$
\int_{z_{t+1}^*} f(z_{t+1}^* \mid z_t^*) Q_{t+1}(.) I_{t+1}(.) \, dz_{nt+1}^* = \int_{\underline{z}_{\tau t+1}^*}^{\overline{z}_{\tau t+1}^*} f(z_{t+1}^* \mid z_t^*) Q_{t+1}() \, dz_{nt+1}^*
$$

C) Since $h_t(\widetilde{p}_{i\tau t+1})$ is continuous in $\widetilde{p}_{i\tau t+1}$ and $Q(h_t(\widetilde{p}_{i\tau t+1}).)$ is continuous in h_t, we only need to show that the functions $\underline{h}_\tau^{-1}(H_t, z_{nt+1}^p)$ and $\overline{h}_\tau^{-1}(H_t, z_{t+1}^p)$ are continuous in $\widetilde{p}_{i\tau t+1}$.

From the continuity of the production function in each occupation in all factors of production, $\underline{h}_\tau(H_t, z_{t+1}^p)$ and $\overline{h}_\tau(H_t, z_{t+1}^p)$ are continuous in h_t and $h_t(\widetilde{p}_{i\tau t+1})$ is continuous in $\widetilde{p}_{i\tau t+1}$. Hence there inverses are continuous in $\widetilde{p}_{i\tau t+1}$. Therefore, there exists a solution to every period's beliefs separately.   Q.E.D.

APPENDIX C: IDENTIFICATION PROOFS

PROOF OF LEMMA 6.1:   Without loss of generality, assume that

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

The above equation satisfies Assumption(6.1), the explicit functional form simplifies the exposition. Equation () implies that the euler for consumption is

(C.1) $\exp(u_{2c}(c_{nt})) \exp(-u_{2z}(z_{nt})) \exp(-\varepsilon_{2nt}) = \eta_n \lambda_t$

Taking logs of equation(C.1), first differencing the results and rearranging gives us

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

By assumption 6.2 (1) then

(C.3) $E[\Delta u_{2c}(c_{nt})|z_{nt}, z_{nt-1}] = \Delta u_{2z}(z_{nt}) + \log(\lambda_t)$

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

(C.4) $u_{2z}(z_{nt}) = u_{2z}(\underline{z_c}, z_{c'nt}) + \int_{z_c}^{z_{cnt}} \left\{ \frac{\partial E[\Delta u_{2c}(c_{nt})|z_{nt}, z_{nt-1}]}{\partial z_c} \right\} dz_c$

(C.5) $u_{2z}(z_{nt-1}) = u_{2z}(\underline{z_c}, z_{c'nt-1}) + \int_{z_c}^{z_{cnt-1}} \left\{ \frac{\partial E[\Delta u_{2c}(c_{nt})|z_{nt}, z_{nt-1}]}{\partial z_{c-1}} \right\} dz_{c-1}$

Which by Assumption (6.3) 2, and from results in Chesher(2007) is identified. Therefore

(C.6)
$$\Delta \log(\lambda_t) = E[\Delta u_{2c}(c_{nt})|z_{nt}, z_{nt-1}] - \int_{z_c}^{z_{cnt}} \left\{ \frac{\partial E[\Delta u_{2c}(c_{nt})|z_{nt}, z_{nt-1}]}{\partial z_c} \right\} dz_c$$
$$+ \int_{z_c}^{z_{cnt-1}} \left\{ \frac{\partial E[\Delta u_{2c}(c_{nt})|z_{nt}, z_{nt-1}]}{\partial z_{c-1}} \right\} dz_{c-1}$$

and by Assumption(6.2) 1

(C.7) $\log(\lambda_1) = E[\Delta u_{2c}(c_{n1})|z_{n1}] - \int_{z_c}^{z_{cn1}} \left\{ \frac{\partial E[\Delta u_{2c}(c_{nt})|z_{n1}]}{\partial z_c} \right\} dz_c$

Hence is λ_t is identified. Finally by Assumption (6.2) 2 we have

(C.8) $\log(\eta_n) = E_t\{u_{2c}(c_{nt}) - \log(\lambda_t) - u_{2z}(z_{nt})|z_{nt}]$

Using Chesher(2007) result and the that $u_{2c}()$ is assumed known we the results from equation (C.6),(C.7),and(C.8). Q.E.D.

PROOF OF LEMMA 6.2:   This results is show by proving that

(C.9) $\beta \gamma_\tau = \frac{\triangle E_t[d_{nt} I_{n\tau t} S_{n\tau t}|H_{nt}, z_{nt}^p]}{\triangle \widetilde{p}_{\tau t+1}(H_{nt}, z_{nt}^p)},$

(C.10)
$$y_{\tau t}(h_{nt}, H_{nt-1}, z_{nt}^p) = \int_0^{h_{nt}} \left\{ \partial \left[\frac{E_t[d_{nt} I_{n\tau t} S_{in\tau t}|H_{nt}, z_{nt}^p]}{E_t[d_{nt} I_{n\tau t}|H_{nt}, z_{nt}^p, i]} \right] \backslash \partial h \right\} dh$$
$$- \int_0^{h_{nt}} \left\{ \partial \left[\frac{\triangle E_t[d_{nt} I_{n\tau t} S_{n\tau t}|H_{nt}, z_{nt}^p]}{\triangle \widetilde{p}_{\tau t+1}(H_{nt}, z_{nt}^p)} \right] \backslash \partial h \right\} dh$$

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

$$
\begin{aligned}
\gamma_\tau \quad = \quad & \left\{ \int_0^{h_{nt}} \left\{ \partial \left[\frac{E_t[d_{nt}I_{n\tau t}S_{in\tau t}|H_{nt}, z_{nt}^p]}{E_t[d_{nt}I_{n\tau t}|H_{nt}, z_{nt}^p, i]} \right] \backslash \partial h \right\} dh - \frac{E_t[d_{nt}I_{n\tau t}S_{in\tau t}|H_{nt}, z_{nt}^p]}{E_t[d_{nt}I_{n\tau t}|H_{nt}, z_{nt}^p, i]} \right\} \\
& - \left\{ \int_0^{h_{nt}} \left\{ \partial \left[\frac{\triangle E_t[d_{nt}I_{n\tau t}S_{n\tau t}|H_{nt}, z_{nt}^p]}{\triangle \widetilde{p}_{\tau, t+1}(H_{nt}, z_{nt}^p)} \right] \backslash \partial h \right\} dh - \frac{\triangle E_t[d_{nt}I_{n\tau t}S_{n\tau t}|H_{nt}, z_{nt}^p]}{\triangle \widetilde{p}_{\tau, t+1}(H_{nt}, z_{nt}^p)} \right\}
\end{aligned}
$$
(C.11)

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

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

Applying the results from Chesher(2007) then all these parameters are identified because data is informative about $\widetilde{p}_{i\tau, t+1}(H_t, z_t^p)$ by Corollary (4.2).

From Proposition (4.1) the zero profit condition implies that

$$
E_t[d_{nt}I_{n\tau t}(S_{in\tau t} - y_{\tau t}(h_{nt}, H_{nt-1}, z_{nt}^p) + \gamma_\tau - d_{nt+1}I_{n\tau t+1}\beta\gamma_\tau)|H_{nt}, z_{nt}^p, i] = 0
$$
(C.12)

rearranging and noting that

$$
\widetilde{p}_{i\tau, t+1}(H_{nt}, z_{nt}^p) = \frac{E_t[d_{nt}I_{n\tau t}d_{nt+1}I_{n\tau t+1}|H_{nt}, z_{nt}^p, i]}{E_t[d_{nt}I_{n\tau t}|H_{nt}, z_{nt}^p, i]}
$$
(C.13)

We can write the zero profit condition as

$$
\frac{E_t[d_{nt}I_{n\tau t}S_{in\tau t}|H_{nt}, z_{nt}^p]}{E_t[d_{nt}I_{n\tau t}|H_{nt}, z_{nt}^p, i]} = y_{\tau t}(h_{nt}, H_{nt-1}, z_{nt}^p) - \gamma_\tau + \beta\gamma_\tau \widetilde{p}_{i\tau t+1}(H_{nt}, z_{nt}^p)
$$
(C.14)

Taking the difference between equation(C.14) for men and women and rearranging gives

$$
\beta\gamma_\tau = \frac{\triangle E_t[d_{nt}I_{n\tau t}S_{n\tau t}|H_{nt}, z_{nt}^p]}{\triangle \widetilde{p}_{\tau t+1}(H_{nt}, z_{nt}^p)},
$$
(C.15)

which is well defined by Assumption 6.4(1). Substituting equation (C.15) into (C.14) for men and women gives the following system of equation

(C.16)
$$
y_{\tau t}(h_{nt}, H_{nt-1}, z_{nt}^p) - \gamma_\tau = \frac{E_t[d_{nt}I_{n\tau t}S_{in\tau t}|H_{nt}, z_{nt}^p]}{E_t[d_{nt}I_{n\tau t}|H_{nt}, z_{nt}^p, i]} - \frac{\triangle E_t[d_{nt}I_{n\tau t}S_{n\tau t}|H_{nt}, z_{nt}^p] \times \widetilde{p}_{i\tau t+1}(H_{nt}, z_{nt}^p)}{\triangle \widetilde{p}_{\tau t+1}(H_{nt}, z_{nt}^p)}, i = \{m, w\}
$$

Differentiating equation (C.16) with respective to hours and then integrating

$$
\begin{aligned}
y_{\tau t}(h_{nt}, H_{nt-1}, z_{nt}^p) \quad = \quad & y_{\tau t}(0, H_{nt-1}, z_{nt}^p) + \int_0^{h_{nt}} \left\{ \partial \left[\frac{E_t[d_{nt}I_{n\tau t}S_{in\tau t}|H_{nt}, z_{nt}^p]}{E_t[d_{nt}I_{n\tau t}|H_{nt}, z_{nt}^p, i]} \right] \backslash \partial h \right\} dh \\
& - \int_0^{h_{nt}} \left\{ \partial \left[\frac{\triangle E_t[d_{nt}I_{n\tau t}S_{n\tau t}|H_{nt}, z_{nt}^p]}{\triangle \widetilde{p}_{\tau t+1}(H_{nt}, z_{nt}^p)} \right] \backslash \partial h \right\} dh
\end{aligned}
$$
(C.17)

and by Assumption 6.4(2) we have have the results in the lemma. By substituting (C.17) into (C.16) and rearranging, we obtain the final equation in the lemma.   Q.E.D.

PROOF OF LEMMA (6.3):   Define the errors as:

$$
\begin{aligned}
\xi_{i1nt} = & \sum_{s=t}^{\rho(\omega_{nt})} \beta^s [U_{is0}(\omega_{0t}^{(s)}) - U_{is0}(\omega_{1t}^{(s)}) \\
& + p_{0it}^{o(s)} [Q^{-1}(p_{0it}^{o(s)}) + \varphi_1(p_{0it}^{o(s)}) - \varphi_0(p_{0it}^{o(s)})] \\
& - p_{1it}^{o(s)} \{Q^{-1}(p_{1it}^{o(s)}) + \varphi_1(p_{1it}^{o(s)}) - \varphi_0(p_{1it}^{o(s)})\} \\
& - E_t^0 \left\{ \sum_{s=t}^{\rho(\omega_{nt})} \beta^s [U_{is0}(\omega_{0t}^{(s)}) - U_{is0}(\omega_{1t}^{(s)}) \right. \\
& + p_{0it}^{o(s)} [Q^{-1}(p_{0it}^{o(s)}) + \varphi_1(p_{0it}^{o(s)}) - \varphi_0(p_{0it}^{o(s)})] \\
& \left. - p_{1it}^{o(s)} \{Q^{-1}(p_{1it}^{o(s)}) + \varphi_1(p_{1it}^{o(s)}) - \varphi_0(p_{1it}^{o(s)})\} \right\}
\end{aligned}
\tag{C.18}
$$

and

$$
\begin{aligned}
\xi_{i2nt} = & \sum_{s=t}^{\rho(\omega_{nt})} \beta^s \left[\frac{\partial U_{i0s}(\omega_{1t}^{(s)})}{\partial h_t} + \frac{\partial \varphi_0(p_{1it}^{o(s)})}{\partial h_t} \right. \\
& + \frac{\partial p_{1it}^{o(s)}}{\partial h_t} [Q^{-1}(p_{1it}^{o(s)}) + \varphi_1(p_{1it}^{o(s)}) - \varphi_0(p_{1it}^{o(s)})] \\
& + p_{1it}^{(s)} \frac{\partial [Q^{-1}(p_{1it}^{o(s)}) + \varphi_1(p_{1it}^{o(s)}) - \varphi_0(p_{1it}^{o(s)})]}{\partial h_t} \\
& - E_t^o \left\{ \sum_{s=t}^{\rho(\omega_{nt})} \beta^s \left[\frac{\partial U_{is0}(\omega_{1t}^{(s)})}{\partial h_t} + \frac{\partial \varphi_0(p_{1it}^{(s)})}{\partial h_t} \right. \right. \\
& + \frac{\partial p_{1it}^{o(s)}}{\partial h_t} [Q^{-1}(p_{1it}^{o(s)}) + \varphi_1(p_{1it}^{o(s)}) - \varphi_0(p_{1it}^{o(s)})] \\
& \left. \left. + p_{1it}^{o(s)} \frac{\partial [Q^{-1}(p_{1it}^{o(s)}) + \varphi_1(p_{1it}^{o(s)}) - \varphi_0(p_{1it}^{o(s)})]}{\partial h_t} \right\} \right.
\end{aligned}
\tag{C.19}
$$

Given this definition the result follows immediately.   *Q.E.D.*

PROOF OF LEMMA 6.4:   To establish the results we prove that

$$
u_{it0}(z_{nt}) = C_{1it} - E_t^o[Y_{i1nt}|\omega_{nt}] + \frac{1}{2} \int_0^{h_{nt}} \left\{ E_t^o[Y_{i2nt}|\omega_{nt}] + \frac{\partial E_t^o[Y_{i1nt}|\omega_{nt}]}{\partial h} \right\} dh
\tag{C.20}
$$

and

$$
u_{i1t}(z_{nt}, H_{nt-1}, l_{nt}) = C_{2it}(z_{nt}, H_{nt-1}) + \frac{1}{2} \int_0^{h_{nt}} \left\{ E_t^o[Y_{i2nt}|\omega_{nt}] + \frac{\partial E_t^o[Y_{i1nt}|\omega_{nt}]}{\partial h} \right\} dh
\tag{C.21}
$$

where

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

and $C_{2it}(z_{nt}, H_{nt-1}) = u_{i1t}(z_{nt}, H_{nt-1}, 1)$. By applying the results from Chesher (2007) and using the above results we obtain our functional $\mathcal{F}^{-1}(F_{Y|X})$.

Taking expectations of equations (6.13) and (6.14) gives

$$
\begin{aligned}
E_t^o[Y_{i1nt}|\omega_{nt}] = & \; u_{i1}(z_{nt}, H_{nt-1}, 1) - u_{it0}(z_{nt}) - u_{i1}(z_{nt}, H_{nt-1}, l_{nt}^*) \\
& + E_t^o \left\{ \sum_{s=t}^{\rho(\omega_{nt})} \beta^s [u_{i1}(z_{1nt}^{(s)}, H_{1ns-1}^{(s)}, 1) - u_{i1}(z_{0nt}^{(s)}, H_{0ns-1}^{(s)}, 1)] \right\}
\end{aligned}
\tag{C.22}
$$

$$
E_t^o[Y_{i2nt}|\omega_{nt}] = -\frac{\partial u_{i1}(z_{nt}, H_{nt-1}, l_{nt}^*)}{\partial h_t} - E_t^o \left\{ \sum_{s=t}^{\rho(\omega_{nt})} \beta^s \left[\frac{u_{i1}(z_{1nt}^{(s)}, H_{1ns-1}^{(s)}, 1)}{\partial h_t} \right] \right\}
\tag{C.23}
$$

Note that $H_{1ns-1}^{(s)}$ is a function of h_{nt} while $H_{0ns-1}^{(s)}$ is not. Hence take derivative of (C.22) with respect to h_{nt} gives

$$(C.24) \qquad \frac{\partial E_t^o[Y_{i1nt}|\omega_{nt}]}{\partial h_{nt}} = -\frac{\partial u_{i1}(z_{nt}, H_{nt-1}, l_{nt}^*)}{\partial h_t} + E_t^o\left\{\sum_{s=t}^{\rho(\omega_{nt})} \beta^s\left[\frac{u_{i1}(z_{1nt}^{(s)}, H_{1ns-1}^{(s)}, 1)}{\partial h_t}\right]\right\}$$

which implies that

$$(C.25) \qquad E_t^o\left\{\sum_{s=t}^{\rho(\omega_{nt})} \beta^s\left[\frac{u_{i1}(z_{1nt}^{(s)}, H_{1ns-1}^{(s)}, 1)}{h_t}\right]\right\} = \frac{\partial E_t^o[Y_{i1nt}|\omega_{nt}]}{\partial h_{nt}} + \frac{\partial u_{i1}(z_{nt}, H_{nt-1}, l_{nt}^*)}{h_t}$$

Substituting (C.25) into (6.14) gives

$$(C.26) \qquad E_t^o[Y_{i2nt}|\omega_{nt}] = -2\frac{\partial u_{i1}(z_{nt}, H_{nt-1}, l_{nt}^*)}{\partial h_t} - \frac{\partial E_t^o[Y_{i1nt}|\omega_{nt}]}{\partial h_{nt}}$$

Rearranging we get

$$(C.27) \qquad \frac{\partial u_{i1}(z_{nt}, H_{nt-1}, l_{nt}^*)}{\partial h_t} = -\frac{1}{2}\left\{\frac{\partial E_t^o[Y_{i1nt}|\omega_{nt}]}{\partial h_{nt}} + E_t^o[Y_{i2nt}|\omega_{nt}]\right\}$$

Integrating up-to h_{nt} gives

$$(C.28) \qquad u_{i1}(z_{nt}, H_{nt-1}, l_{nt}^*) = u_{i1}(z_{nt}, H_{nt-1}, 1) - \frac{1}{2}\int_0^{h_{nt}}\left\{\frac{\partial E_t^o[Y_{i1nt}|\omega_{nt}]}{\partial h} + E_t^o[Y_{i2nt}|\omega_{nt}]\right\}dh$$

Let

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

then substituting (C.28) into (C.22) and rearranging gives

$$(C.29) \qquad u_{it0}(z_{nt}) = C_{1it} - E_t^o[Y_{i1nt}|\omega_{nt}] + \frac{1}{2}\int_0^{h_{nt}}\left\{E_t^o[Y_{i2nt}|\omega_{nt}] + \frac{\partial E_t^o[Y_{i1nt}|\omega_{nt}]}{\partial h}\right\}dh$$

and obtain the result in the Lemma. $Q.E.D.$

APPENDIX D: ESTIMATION

D.1. *Estimation of Consumption and Earnings Equations*

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

$$(D.1) \qquad E\left[\frac{\partial u_2(c_{nt}, z_{nt}, \varepsilon_{2nt}, \theta_c)}{\partial c_{nt}} - \eta_n\lambda_t \,\middle|\, z_{nt}\right] = 0$$

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

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

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

$$(D.2) \qquad (1-\alpha)^{-1}\triangle\varepsilon_{2nt} = \triangle\ln(c_{nt}) - (1-\alpha)^{-1}\triangle z_{nt}'B_4 + \triangle(1-\alpha)^{-1}\ln(\lambda_t)$$

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

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

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

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

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 (D.2), one can obtain a set of orthogonality conditions,

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

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:

(D.3) $\qquad (1-\alpha)^{-1}\ln(\eta_n) \equiv \sum_{t\in T_1} \left[\ln(c_{nt}) - (1-\alpha)^{-1}z'_{nt}B_4 + (1-\alpha)^{-1}\ln(\lambda_t)\right]/T_1$

The fixed effects estimates of $(1-\alpha)^{-1}\ln(\eta_n)$ are obtained as the simple time averages of the estimated residuals of the consumption equation, which correspond to the sample counterparts of $(1-\alpha)^{-1}\ln(\eta_n)$ defined above. In order to form the sample counterpart of (D.3), we need an estimate of $\left\{(1-\alpha)^{-1}\ln(\lambda_t)\right\}_{t=1}^{T_1}$. From the estimate of Θ_c, however, we can only obtain estimates of $\left\{\triangle(1-\alpha)^{-1}\ln(\lambda_2)\right\}_{t=2}^{T_1}$. 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 $\left\{(1-\alpha)^{-1}\ln(\lambda_t)\right\}_{t=1}^{T_1}$ and $(1-\alpha)^{-1}\ln(\eta_n)$, enabling us to recover α in the third step of our estimation.

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

(D.4) $\qquad E_t\{d_{n\tau t}d_{n\tau t-1}[\triangle S_{nt} - \triangle b_{0\tau t} - b_\tau \triangle HC_{nt} - \triangle z_{nt}^{p\,'}B_{\tau 5} - \beta\gamma_\tau \triangle d_{n\tau t+1}] \mid z_{nt}^p, H_{nt}, h_{nt}^*\} = 0,$

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

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

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

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

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

where $\triangle x'_{\tau nt} = d_{n\tau t}d_{n\tau t-1}(\triangle h_{nt}, \triangle h^2_{nt}, \triangle h_{nt-1}, \ldots, \triangle h_{nt-\rho}, \triangle z^{p'}_{nt}, \triangle d_{n\tau t+1})$. 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 (D.4), one can obtain a set of orthogonality conditions:

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

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

D.2. Estimation of the Final stage

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

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

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

$$
\begin{aligned}
m_{1nt}\left(\Theta u, \Theta^{(N)}_c, \Theta^{(N)}_e, \psi^{(N)}\right) &= \sigma\ln\left[p^{(N)}_{nt}/\left(1-p^{(N)}_{nt}\right)\right] - B_{0t} - z'_{nt}B_1 + z'_{nt}h_{nt}B_2 \\
&\quad + \theta_0\left(1-l^2_{nt}\right) + \sum_{s=1}^{\rho}\theta_s h_{nt}(l_{nt-s}+\beta^s) \\
&\quad - \sigma\sum_{s=1}^{\rho}\beta^s\ln\left(\frac{1-p^{(s)(N)}_{1nt}}{1-p^{(s)(N)}_{0nt}}\right) \\
&\quad - \exp\left((1-\alpha)\phi^{(N)}_{nt}\right)\sum_{\tau=1}^{M}I_{n\tau t}\left[y_{\tau t}\left(h_{nt}, H_{nt-1}, z^p_{nt}, \theta^{(N)}_e\right)\right. \\
&\quad \left. - \gamma^{(N)}_\tau + \beta\gamma^{(N)}_\tau\widetilde{p}^{(N)}_{n\tau t+1}\right]
\end{aligned}
$$

(D.5)

and

$$
\begin{aligned}
m_{2nt}\left(\Theta u, \Theta^{(N)}_c, \Theta^{(N)}_e, \psi^{(N)}\right) &= d_{nt}\left\{\sigma\sum_{s=1}^{\rho}\beta^s\left(1-p^{(s)(N)}_{1nt}\right)^{-1}\nabla_{h_{nt}}p^{(s)(N)}_{1nt}\right. \\
&\quad - z'_{nt}B_2 - 2\theta_0 l_{nt} - \sum_{s=1}^{\rho}\theta_s(l_{nt-s}+\beta^s) \\
&\quad + \exp\left((1-\alpha)\phi^{(N)}_{nt}\right)\sum_{\tau=1}^{M}d_{n\tau t}\left[b^{(N)}_{\tau 1} + 2b^{(N)}_{\tau 2}h_{nt}\right. \\
&\quad \left.\left. + \beta\gamma^{(N)}_\tau\nabla_{h_{nt}}\widetilde{p}^{(N)}_{n\tau t+1}\right]\right\},
\end{aligned}
$$

where $\psi^{(N)} = \left(p^{(N)}_{nt}, p^{(s)(N)}_{0nt}, p^{(s)(N)}_{1nt}, \widetilde{p}^{(N)}_{n\tau t+1}\right)$ 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 (??) and (??). The first is the forecast errors from replacing the expectations of future variables with their realizations. The second is the approximation error that arises from replacing the true values of the conditional choice probabilities, conditional expectation, and time-invariant individual-specific effects with their estimates. Let us define the 2×1 vector $m_{3nt}\left(\Theta_u, \Theta_c^{(N)}, \Theta_e^{(N)}, \psi^{(N)}\right) \equiv \left[m_{1nt}\left(\Theta_u, \Theta_c^{(N)}, \Theta_e^{(N)}, \psi^{(N)}\right), m_{2nt}\left(\Theta_u, \Theta_c^{(N)}, \Theta_e^{(N)}, \psi^{(N)}\right)\right]'$ and let T_3 denote the set of periods for which the hours and participation equations are valid. Define the vector

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

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

(D.6) $\qquad \Theta_u^{(N)} \equiv \arg\min_{\Theta u} \dfrac{\sum_{n=1}^N m_{3n}^{(N)}\left(\Theta_u, \Theta_c^{(N)}, \Theta_e^{(N)}, \psi^{(N)}\right) \left(\Omega_n^{(N)}\right)^{-1} m_{3n}^{(N)}\left(\Theta_u, \Theta_c^{(N)}, \Theta_e^{(N)}, \psi^{(N)}\right)}{N}.$

D.3. *Asymptotic Properties*

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

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

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

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

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

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

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

THEOREM 1   *Under A1–A8 and $\Upsilon(\omega)$ 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 jth 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

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

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

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

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

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

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

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

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.   *Q.E.D.*

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] Aguirregabiria, V. and P. Mira (2005): "Sequential estimation of Dynamic Discrete Games", forthcoming Econometrica.

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

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

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

[7] Altonji J. and C. Pierret " Employer Learning and Statistical Discrimination" *The Quarterly Journal of Economics* Vol. 116, No. 1 (Feb., 2001), pp. 313-350

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

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

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

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

[12] Bajari, P. and H. Hong (2006): "Semiparametric Estimation Game of a Dynamic Incomplete Information", Duke University

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

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

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

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

[17] 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.

[18] Bisin A. and P. Gottardi (2006): "Efficient Competitive Equilibria with Adverse Selection," *Journal of Political Economic, 2006 Vol 114, (3) pp 485-515.*

[19] 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 .*

[20] Blundell and MaCurdy, "Labor Supply: A Review of Alternative Approaches" Handbook, Vol. 3A, Chapter 27

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

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

[23] 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.

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

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

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

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

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

[29] Dey, G. and C. Hill (2007): "Behind the Pay Gap," American Association of University Women Educational Foundation.

[30] 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.

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

[32] 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.*

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

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

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

[36] Gale D. " A Walrasian Theory of Markets with Adverse Selection"   *The Review of Economic Studies*, Vol. 59, No. 2 (Apr., 1992), pp. 229-255

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

[38] Haavelmo, T.M. (1944): "The Probability Approach in Econometrics," Econometrica, 12, Supplement, 118 pp

[39] Heckman, J.J. and T.E. MaCurdy (1980), "A life-cycle model of female labour supply", *Review of Economic Studies,* 47: 47-74.

[40] Hellwig M. "Some Recent Developments in the Theory of Competition in Markets with Adverse Selection" European Economic Review 31 1987 pp 319-25

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

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

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

[44] Jofre-Bonet and Pesendorfer (2003): "Estimation of a Dynamic Auction" *Econometrica* Vol. 71, No. 5 pp. 1443-1489

[45] Jovanovic, B. (1989): " Observable Implications of Models with Multiple Equilibria", Econometrica,

Vol. 57, No. 6. (Nov., 1989), pp. 1431-1437.

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

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

[48] 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.

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

[50] 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*

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

[52] MaCurdy, T.E. (1981), "An empirical model of labour supply in a life-cycle setting", Journal of Political Economy 89: 1059-1085.

[53] Magnac and Thesmar (2002) " Identifying Dynamic Discrete Decision Process., Econometrica, 70, 801-816.

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

[55] 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.

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

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

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

[59] 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.

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

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

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

[63] Pensendofer M and Schmidt-Dengler (2006): "Asymptotic Least Squares Estimators for Dynamic Games", NBER Working Paper w9726

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

[65] Prescott E. and R. Townsend "Pareto Optima and Competitive Equilibria with Adverse Selection and Moral Hazard" Econometrica Vol. 52, No. 1 (Jan., 1984), pp. 21-46

[66] Riley J. "Informational Equilibrium" Econometrica Vol. 47, No. 2 (Mar., 1979), pp. 331-359

[67] Rothschild m. AND J. Stiglitz "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information" *The Quarterly Journal of Economics*, Vol. 90, No. 4 (Nov., 1976), pp. 629-649

[68] Tamer, E. (2003): "Incomplete Simultaneous Discrete Response Model with Multiple Equilibria", *Review of Economic Studies.*, Vol 70, No 1, 2003, pp. 147-167.

[69] Tirole, J. (1996?7): "A Theory of Collective Reputation," *Review of Economic Studies*, 63, 1–22.

[70] 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 (27.7)		4.0 (1.9)	37.8 (10.7)	0.56 (0.82)	0.94 (1.3)	0.85 (0.35)
1969	49.9 (31.1)	7.7 (3.7)	4.0 (1.9)	38.5 (10.9)	0.53 (0.83)	0.93 (1.3)	0.86 (0.35)
1970	50.3 (30.0)	7.7 (3.6)	3.8 (1.8)	38.6 (11.3)	0.49 (0.79)	0.87 (1.3)	0.85 (0.36)
1971	51.0 (31.4)	7.5 (3.5)	3.7 (1.8)	39.0 (11.6)	0.44 (0.76)	0.85 (1.2)	0.83 (0.37)
1972	51.4 (31.5)	7.4 (3.5)	3.7 (1.8)	39.3 (11.8)	0.42 (0.72)	0.80 (1.2)	0.82 (0.38)
1973	53.3 (35.2)	7.4 (3.4)	3.6 (1.8)	39.5 (12.2)	0.39 (0.69)	0.77 (1.1)	0.82 (0.38)
1974	54.3 (35.2)	7.3 (3.4)	3.4 (1.8)	39.8 (12.5)	0.37 (0.68)	0.71 (1.1)	0.82 (0.39)
1975	52.9 (34.6)	6.9 (3.2)	3.3 (1.7)	39.8 (12.6)	0.35 (0.67)	0.66 (1.0)	0.81 (0.39)
1976	53.4 (35.6)	6.8 (3.2)	3.3 (1.7)	39.7 (12.6)	0.35 (0.68)	0.62 (1.0)	0.80 (0.39)
1977	52.1 (35.6)	6.7 (3.3)	3.2 (1.6)	39.7 (12.6)	0.34 (0.68)	0.60 (0.96)	0.79 (0.40)
1978	52.1 (35.6)	6.5 (3.5)	3.2 (1.6)	38.9 (12.7)	0.41 (0.72)	0.51 (0.87)	0.77 (0.42)
1979	55.5 (55.2)	6.7 (3.3)	3.1 (1.5)	39.8 (12.5)	0.34 (0.66)	0.53 (0.88)	0.77 (0.42)
1980	55.1 (39.1)	6.6 (3.3)	3.1 (1.5)	39.9 (12.5)	0.35 (0.69)	0.50 (0.84)	0.78 (0.42)
1981	56.2 (68.0)	6.4 (3.1)	3.1 (1.4)	38.8 (11.9)	0.39 (0.69)	0.50 (0.82)	0.80 (0.39)
1982	54.1 (40.3)	6.3 (3.1)	3.1 (1.4)	38.9 (11.8)	0.38 (0.69)	0.50 (0.82)	0.80 (0.40)
1983	53.1 (39.4)	6.3 (3.1)	3.1 (1.4)	39.0 (11.8)	0.38 (0.67)	0.51 (0.84)	0.80 (0.40)
1984	54.8 (43.2)	6.3 (3.1)	3.1 (1.4)	39.1 (11.7)	0.38 (0.70)	0.51 (0.84)	0.80 (0.40)
1985	57.9 (51.3)	6.5 (3.8)	3.1 (1.4)	39.6 (11.5)	0.37 (0.68)	0.53 (0.85)	0.80 (0.40)
1986	59.2 (48.8)	6.4 (3.2)	3.1 (1.4)	40.2 (11.2)	0.37 (0.69)	0.56 (0.87)	0.81 (0.40)
1987	61.9 (51.4)	6.5 (3.1)	3.1 (1.3)	40.6 (10.8)	0.36 (0.67)	0.57 (0.87)	0.81 (0.40)
1988	64.1 (63.6)	6.6 (3.0)	3.1 (1.3)	41.4 (10.6)	0.34 (0.66)	0.58 (0.87)	0.81 (0.40)
1989	65.9 (69.7)	6.5 (2.9)	3.1 (1.3)	42.1 (10.3)	0.33 (0.63)	0.59 (0.87)	0.81 (0.39)
1990	66.3 (63.7)	6.6 (3.2)	3.1 (1.3)	42.7 (10.1)	0.30 (0.62)	0.61 (0.89)	0.81 (0.39)
1991	61.4 (52.3)	6.4 (3.1)	3.0 (1.4)	43.7 (10.2)	0.29 (0.60)	0.58 (0.87)	0.77 (0.43)
1992	65.6 (62.1)	6.7 (3.8)	3.1 (1.3)	44.0 (9.6)	0.29 (0.61)	0.62 (0.89)	0.82 (0.39)

Standard deviation in parentheses. Household Income and Food Consumption in thousands of year-2000 US$
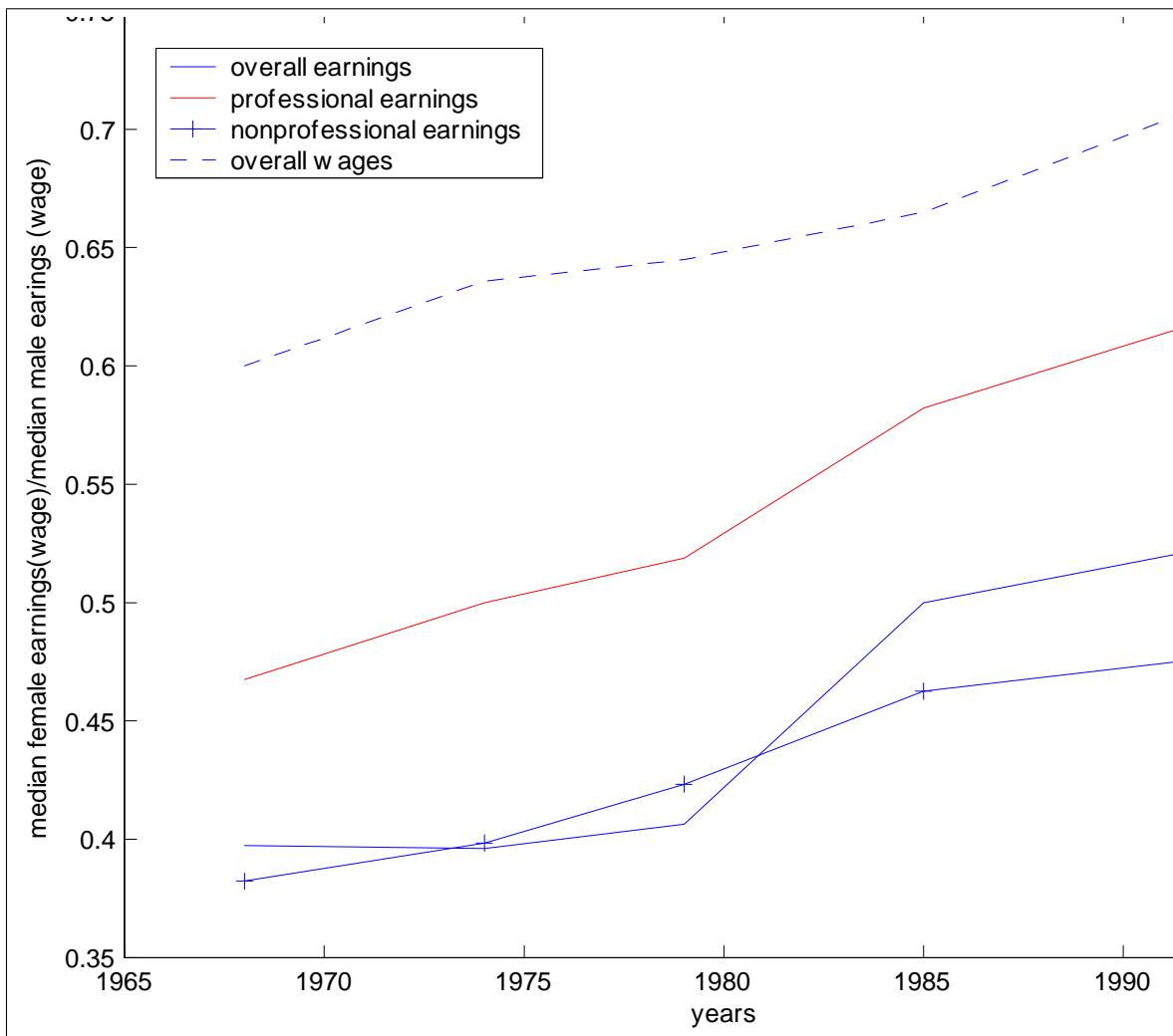
FIGURE 1.— Earnings and Wage Gaps

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
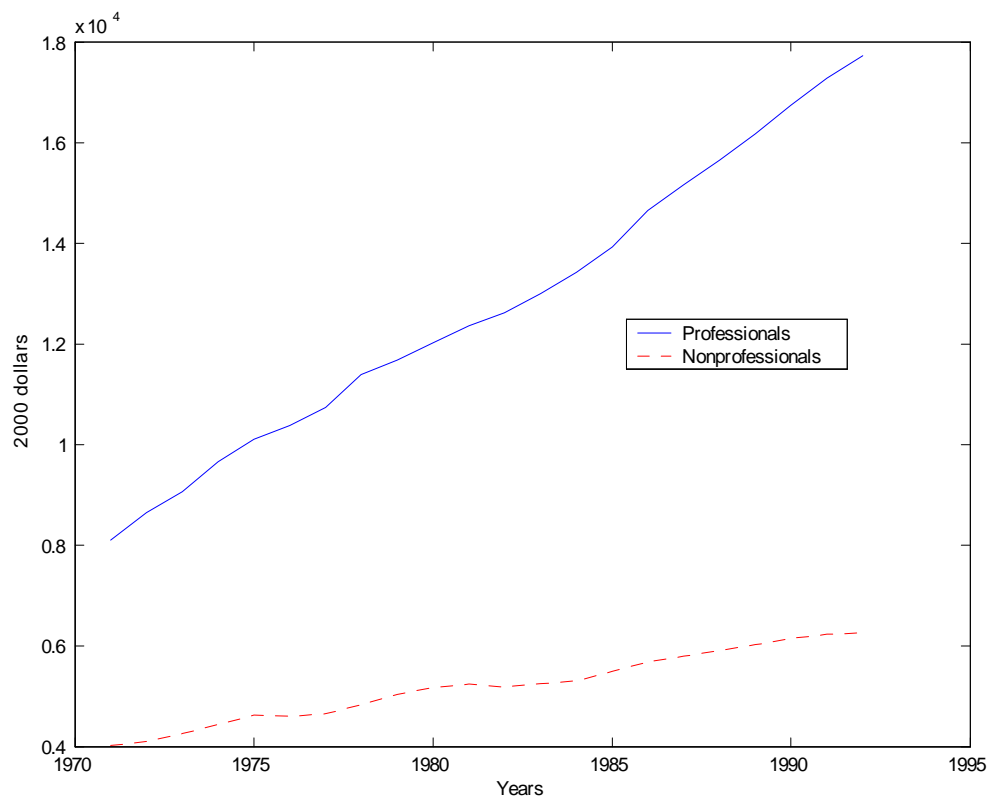
Table 4: Earning Equation

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_{nt}^2	−0.003 (0.008)
$EDUC_{nt}$	0.08 (0.0004)
$EDUC_{nt} \times male\ dummy_{nt}$	−0.03 (0.04)
MS_{nt}	0.205 (0.006)
$SP.EDUC_{nt} \times MS_{nt}$	−0.088 (0.005)
$SP.EDUC_{nt} \times MS_{nt} \times male\ dummy_{nt}$	0.145 (0.003)

Table 6 Utility of Leisure/Home Production

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

Table 7 Nonseparability in Utility of Leisure/Home Production

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

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

Source	Professional	Nonprofessional
Raw	87	76
Predicted	92	81
Explained	95	93
Unexplained	5	7
Decomposition of Explained		Gap
Human Capital[34]	69	74
Beliefs	14	12
Fixed Effect	6	8
Age-education	11	6

Table 9:The Gender Wage Gap
(Median Women Earnings over Median Men Earnings(%))
1974–1978:1984–1988

Source	Professional		Nonprofessional	
	1974:1978	1984:1988	1974:1978	1984:1988
Raw	48	62	38	47
Predicted	52	67	41	50
Hiring Cost	81	96	68	79
symmetric Information	71	82	57	65
Demographic Characteristics	81	89.6	68	75
Participation Cost	81	94.9	68	78.5
Aggregate Production	81	88.5	68	72.5

---

[34]Human Capital includes the effect of current and past hours on the production function.

Table 10: Hours, Participation and Occupation Composition
(1974–1978:1984–1988

Source	1974:1978	1984:1988
	Women	Participation
Raw	62	70
Hiring Cost	51	57
Private Information	56	62
Demographic Characteristics	51	54
Participation Cost	51	51
Aggregate Production	51	53
	Fraction of	Women
Professional		
Raw	34	40
Hiring Cost	28	35
Private Information	30	38
Demographic Characteristics	28	33
Participation Cost	28	35
Aggregate Production	28	30
Nonprofessional		
Raw	48	48
Hiring Cost	42	40
Private Information	45	46
Demographic Characteristics	42	41
Participation Cost	42	42
Aggregate Production	42	40

Table 11:The average Hours worked
1974–1978:1984–1988

Source	Professional				Non-	Professional		
	1974: Women	1978 Men	1984: Women	1988 Men	1974: Women	1978 Men	1984: Women	1988 Men
Raw	1640	2201	1904	2226	1424	1998	1635	2117
Predicted	1702	2100	1960	2123	1490	2030	1700	2070
Hiring Cost	1980	2010	2050	2100	1580	2000	1590	2060
Private Information	1820	2080	1990	2090	1510	1970	1640	1930
Demographics	1980	2100	1980	2150	1580	2000	1610	2080
Participation Cost	1980	2100	2045	2110	1580	2000	1592	2065
Aggregate Production	1980	2100	1820	2156	1580	2000	1575	2065