Human Capital and Earnings Distribution Dynamics

Mark Huggett, Gustavo Ventura and Amir Yaron[*][†]

September 2004

Abstract

Earnings heterogeneity plays a crucial role in modern macroeconomics. We document that mean earnings and measures of earnings dispersion and skewness all increase in US data over most of the working life-cycle for a typical cohort as the cohort ages. We show that (i) a human capital model can replicate these properties from the right distribution of initial human capital and learning ability, (ii) differences in learning ability are essential to produce an increase in earnings dispersion over the life cycle and (iii) differences in learning ability account for the bulk of the variation in the present value of earnings across agents. These findings emphasize the need to further understand the role and origins of initial conditions.

JEL Classification: D3, J24, J31

Keywords: Earnings Distribution, Human Capital, Heterogeneity.

1 Introduction

Recent work in macroeconomics has explored the quantitative implications of dynamic models for the distribution of consumption, income and wealth. This work takes earnings or wages as an *exogenous* random process and then proceeds to characterize the distributional implications of optimal consumption-savings and labor-leisure behavior.[1] These models would appear to be attractive for assessing the distributional effects of changes in government policy since they are able to produce many of the quantitative features of the actual distribution of consumption, income and wealth.[2]

A critical issue for this research agenda is to integrate deeper foundations for the determinants of earnings and wages into these models by allowing earnings to be *endogenous*. We list two reasons for why this is important. First, we note that when earnings are exogenous there is no channel for policy to affect consumption and welfare through earnings. This channel is arguably of first order importance. In fact, a dominant theme in the earnings distribution literature is that earnings profiles are determined by the optimal investment of time and resources into the accumulation of skills. As a result, these investment decisions will not be invariant to changes in government policies. Second, a key issue for the purposes of assessing many government policies is the degree to which the variation in the present value of earnings is due to differences established early in life versus shocks received over the life cycle. If the former is responsible for the bulk of the variation in earnings, then policies directed towards these initial differences are of first-order importance.

This paper takes a first step towards developing deeper foundations by examining, at a quantitative level, the earnings distribution dynamics of a well-known and widely-used human capital model. More specifically, we document properties of how the US earnings distribution evolves for a typical cohort of individuals as the cohort ages. We then assess the ability of the model to replicate these properties. This assessment serves to highlight the potential role and importance of differences in initial conditions for understanding the dynamics of the earnings distribution.

The specific properties of the US earnings distribution that we focus on relate to how average earnings, and measures of earnings dispersion and skewness change

---

[1]See, for example, Cagetti (2002), Carroll (1997), Castañeda et. al. (2002), Deaton (1992), De Nardi (2002), Domeij and Klein (2001), Gourinchas and Parker (2002), Heathcote, Storesletten and Violante (2003), Hubbard et. al. (1994), Huggett (1996), Krueger and Perri (2002), Krueger and Fernandez-Villaverde (2001), Quadrini (2000) and Storesletten, Telmer and Yaron (2004).

[2]These models have been widely applied. Focusing solely on the issue of social security reform, the literature includes Deaton et. al. (2002), De Nardi et. al. (1999), Fuster (1999), Huggett and Ventura (1999), Imrohoroglu et. al. (1995), Storesletten, Telmer and Yaron (1999) among others.

for a typical cohort as the cohort ages. To characterize these age effects, we use earnings data for US males and employ a methodology, described later in the paper, for separating age, time and cohort effects in a consistent way for a variety of earnings statistics. Our findings, summarized in Figure 1, are that average earnings, earnings dispersion and earnings skewness increase with age over most of the working life-cycle.

[Insert Figure 1 a-c Here]

We assess the ability of the Ben-Porath (1967) human capital model to replicate the patterns in Figure 1. This framework is the natural candidate for our study. The Ben-Porath model is well-known and widely-used, and has been the basis for both theoretical and empirical analyses of human capital (e.g., its prominence in the literature is reflected in recent surveys, such as Mincer (1997) and Neal and Rosen (1999)).[3] In our version of this model, each agent is endowed with some immutable learning ability and some initial human capital. Each period an agent divides available time between market work and human capital production. Human capital production is increasing in learning ability, current human capital and time allocated to human capital production. An agent maximizes the present value of earnings, where earnings in any period is the product of a rental rate, human capital and time allocated to market work.

Our assessment focuses on the dynamics of the cohort earnings distribution produced by the model from different initial joint distributions of human capital and learning ability across agents. Our findings are striking. We establish that the earnings distribution dynamics documented in Figure 1 can be replicated quite well by the model from the right initial distribution. In addition, the model produces the key properties of the cross-sectional earnings distribution. These conclusions are not sensitive to the precise value of the elasticity parameter in the human capital produc-

---

[3]Earnings distribution facts have long been interpreted as being qualitatively consistent or inconsistent with specific human capital models. This is standard in the earnings and wage regression literature (e.g. Card (1999)), in the many excellent reviews of human capital theory (e.g. Weiss (1986), Mincer (1997) and Neal and Rosen (2000)) and in work that simulates properties of human capital models (e.g. von Weizsacker (1993)). In contrast, Heckman (1975, 1976), Haley (1976), Rosen (1976) and a number of related papers provide a quantitative assessment. However, distributional implications were not addressed because model parameters were estimated so that the age-earnings profile produced by one agent in the model best matches the earnings data. Our work is closest to the work by Heckman, Lochner and Taber (1998) and Andolfatto, Gomme and Ferrall (2001) who use human capital models with agent heterogeneity to analyze a number of distributional issues. The former focuses on time variation in the skill premium, whereas the latter focuses on earnings, income and wealth profiles.

tion function, nor are they sensitive to the age at which human capital accumulation process articulated by the model begins.

The initial distributions which replicate the patterns in Figure 1 rely crucially on differences in learning ability across agents. Age-earnings profiles for agents with high learning ability are steeper than the profiles for agents with low learning ability. This is the key mechanism for how the model produces increases in earnings dispersion and skewness for a cohort as the cohort ages. Earnings profiles are steeper for high ability agents since early in life they allocate a relatively larger fraction of their time to human capital production and thus have low earnings, while their time allocation decisions and high learning ability imply that later in the life-cycle they have higher levels of human capital and, hence, earnings. This mechanism is consistent with regularities long discussed in the human capital literature such as the fact that time allocated to skill acquisition is concentrated at young ages, that age-earnings profiles are steeper for people who choose high amounts of schooling and that the present value of earnings increases in a measure of learning ability.[4]

It is important to mention that it is *not* the case that the model can always match a set of life-cycle earnings distribution facts, provided that one can choose an infinite number of parameters characterizing the initial distribution. Proposition 1 in section 3 shows that when all agents are born with the same learning ability, but different initial human capital, the model *always* generates a counterfactual pattern of decreasing earnings dispersion no matter how one chooses the distribution of human capital across agents. Intuitively, one can always exactly match any distribution of earnings at the end of the working life-cycle provided one can choose the distribution of initial human capital freely. However, the ability to match the facts documented in Figure 1 requires that one exactly matches the earnings distribution in the end of the working life cycle *as well as* in all previous periods. Thus, having an infinite number of parameters to choose in the form of an unrestricted initial distribution does not guarantee that one can match the patterns in Figure 1.

We close the paper by contrasting the implications of the model with some evidence on persistence in individual earnings. The model implies that over time both individual earnings levels and earnings growth rates are strongly positively correlated. Evidence from US data shows that earnings levels are positively correlated but that earnings growth rates one year apart are negatively correlated. This and related evidence suggests that there is potentially an important role for idiosyncratic shocks that lead to mean reversion in earnings. These shocks are by construction absent from the

---

[4]Mincer (1997) summarizes evidence on the first point and Lillard (1977) provides evidence on the last two points.

benchmark model. A critical issue for future work is to determine the importance of both initial conditions and shocks over the life-cycle in models in which the earnings distribution is endogenous.[5] We believe that this issue can be usefully pursued by investigating both the distributional dynamics of earnings and consumption over the life cycle.

The paper is organized as follows. Section 2 describes the data and our empirical methodology. Section 3 presents the model. Section 4 discusses parameter values. Section 5 presents the central findings of the paper. Section 6 concludes.

2 Data and Empirical Methodology

2.1 Data

The findings presented in the introduction are based on earnings data from the PSID 1969-1992 family files. We utilize earnings of males who are the head of the household. We consider two samples. We define a *broad* sample to include all males who are currently working, temporarily laid off, looking for work but are currently unemployed, students, but does not include retirees. The *narrow* sample equals the broad sample less those unemployed or temporarily laid off. We note that the theoretical model we analyze is not a model of unemployment or lay offs. This would suggest that the narrow sample is more relevant. However, since the results are not sensitive to the choice of sample we present the results for the broad sample.

We consider males between the ages of 20 and 58. This is motivated by several considerations. First, the PSID has many observations in the middle but relatively fewer at the beginning or end of the working life cycle. By focusing on ages 20-58, we have at least 100 observations in each age-year bin with which to calculate age and year-specific earnings statistics. Second, near the traditional retirement age there is a substantial fall in labor force participation that occurs for reasons that are abstracted from in the model we analyze. This suggests the use of a terminal age that is earlier than the traditional retirement age. We also restrict the sample to those with strictly positive earnings. This is not essential to our methodology but it does allow us to take logs as a convenient data transformation. This restriction almost never binds.[6] Finally, we exclude the Survey of Economic Opportunities (SEO) sample which is

---

[5]Keane and Wolpin (1997) address this issue in the context of a model with an occupational choice decision. Storesletten et. al. (2004) do so in a model of exogenous earnings.

[6]Most of those who report being laid off, unemployed or students turn out to have some earnings during the year.

a subsample of the PSID that over samples the poor. Given all the above sample selection criteria, the average and standard deviation of the number of observations per panel-year are 2137 and 131 respectively.

2.2 Construction of Age Profiles

We focus the analysis on cohort-specific earnings distributions. Let $e_{j,t}^p$ be the real earnings at percentile p of the earnings distribution of agents who are age j at time t. These agents are from cohort $s = t - j$ (i.e., agents who were born in year $t - j$).[7] We assume that the percentiles of the earnings distribution $e_{j,t}^p$ are determined by cohort effects α_s^p, age effects β_j^p and shocks $\epsilon_{j,t}^p$. The relationship between these variables is given below both in levels and in logs, where the latter is denoted by a tilde.

$$e_{j,t}^p = \alpha_s^p \beta_j^p \epsilon_{j,t}^p$$

$$\tilde{e}_{j,t}^p = \tilde{\alpha}_s^p + \tilde{\beta}_j^p + \tilde{\epsilon}_{j,t}^p$$

This formulation is consistent with the theoretical model that we present in the next section. In particular, in a steady state of the model with a constant growth rate of the rental rate of human capital, $e_{j,t}^p$ is produced by a cohort effect α_s^p that is proportional to the rental rate in cohort year s, a time-invariant age effect β_j^p and no shocks (i.e. $\epsilon_{j,t}^p \equiv 1$ and $\tilde{\epsilon}_{j,t}^p \equiv 0$). Expressed somewhat differently, in steady state the cross-sectional, age-earnings distribution just shifts up proportionally each period.

We use ordinary least squares to estimate the coefficients $\tilde{\alpha}_s^p$ and $\tilde{\beta}_j^p$ for various percentiles p of the earnings distribution.[8] In Figure 2 we graph the age effects of different percentiles of the *levels* of the earnings distribution by plotting β_j^p. The age effects β_j^p are scaled so that each graph passes through the geometric average value at age $j = 40$ of $e_{j,t}^p$ across all cohorts and so that mean earnings equal 100 at the end of the working life cycle.[9] The percentiles considered in Figure 2 range from a

---

[7]Real values are calculated using the CPI. To calculate $e_{j,t}^p$ we use a 5 year bin centered at age j. For example, to calculate earnings percentiles of agents age $j = 30$ in year $t = 1980$ we use data on agents age $28 - 32$ in 1980. We also use a 5 year bin centered at ages 20 and 58. To do this we use data on agents age 18-22 and 56-60.

[8]Each regression has $J \times T$ dependent variables regressed on $J + T$ cohort dummies and J age dummies. T and J denote the number of time periods in the panel and the number of distinct age groups, which in our case equal $J = 58 - 20$ and $T = 1992 - 1969$.

[9]More specifically, we plot $\beta_j^p e_{40}^p / \beta_{40}^p$, where e_{40}^p is the geometric average real earnings at age 40 and percentile p in the data. We then scale all profiles by a common factor to normalized mean earnings to 100.

low of $p = .025$ (earnings such that 2.5 percent of the agents are below this value) to a high of $p = .99$ (earnings such that 99 percent of the agents are below this value). We calculate 23 different percentiles $p = .025, .05, .10, ..., .90, .925, .95, .975, .99$, but for visual clarity display only a subset of these in Figure 2.

[Insert Figure 2 Here]

The findings in Figure 1a-c in the introduction are all calculated directly from the results graphed in Figure 2. Figure 1a shows that average earnings increase with age over most of the working life cycle. Early in the life cycle this follows because earnings at all percentiles in Figure 2 shift up with age. Later in the life cycle this follows from the strong increase with age at the highest percentiles of the earnings distribution despite the fact that earnings at the median and lower percentiles are already decreasing with age. The increase in earnings dispersion in Figure 1b, using the Gini coefficient as a measure of earnings dispersion, follows from the general fanning out of the distribution which is a striking feature of Figure 2. The increase in the skewness measure with age in Figure 1c is implied by the strong fanning out at the top of the distribution observed in Figure 2.

2.3 Alternative Views of Age Effects

A more general specification of the regression equation used in the last subsection would allow the percentiles of the earnings distribution to be determined by time effects γ_t^p in addition to age β_j^p and cohort α_s^p effects as in the equation below. Once again, a logarithm of a variable is denoted by a tilde. Time effects can be viewed as effects that are common to all individuals alive at a point in time. An example would be a temporary rise in the rental rate of human capital that increases the earnings of all individuals in the period.

$$e_{j,t}^p = \alpha_s^p \beta_j^p \gamma_t^p \epsilon_{j,t}^p$$

$$\tilde{e}_{j,t}^p = \tilde{\alpha}_s^p + \tilde{\beta}_j^p + \tilde{\gamma}_t^p + \tilde{\epsilon}_{j,t}^p$$

The linear relationship between time t, age j, and birth cohort $s = t - j$ limits the applicability of the regression specification above. Specifically, without further restrictions the regressors in this system are co-linear and these effects cannot be estimated. This identification problem is well known in the econometrics literature.[10]

---

[10]See, for example, Weiss and Lillard (1978), Hanoch and Honig (1985) and Deaton and Paxson (1994) among others.

In effect any trend in the data can be arbitrarily reinterpreted as a year (time) trend or alternatively as trends in ages and cohorts.

Given this problem, our approach is to determine how sensitive the age effects in Figure 1 and 2 are to alternative restrictions on the coefficients $(\tilde{\alpha}_s^p, \tilde{\beta}_j^p, \tilde{\gamma}_t^p)$. One view, which we label the cohort dummies view, comes from constructing Figure 2 by setting time effects to zero (i.e. $\tilde{\gamma}_t^p = 0$) as was done in the last subsection. A second view, which we label the time dummies view, comes from constructing Figure 2 by setting cohort effects to zero (i.e. $\tilde{\alpha}_s^p = 0$).[11] A third view, which is intermediate to both previous views, comes from constructing Figure 2 after allowing age, cohort and time effects but with the restriction that time effects are mean zero and are orthogonal to a time trend.[12] This restriction implies that time trends are attributed to cohort and age effects rather than time effects. We label this last view the restricted time dummies view.

[Insert Figure 3 (a-c) Here]

Figure 3 highlights the age effects on average earnings, earnings dispersion and earnings skewness using these three views. The results are that all three views lead to the same qualitative results. Quantitatively, the cohort dummies view is almost indistinguishable from the restricted time dummies view. The time dummies view produces a flatter profile of earnings dispersion as compared to the cohort dummies or restricted time dummies view. In the remainder of the paper we focus on the results from the cohort dummies view highlighted in Figure 1.

2.4   Related Empirical Work

Our empirical work is related to previous work both at a substantive and a methodological level. At a substantive level, labor economists have examined patterns in mean earnings and measures of earnings dispersion and skewness at least since the work of Mincer (1958, 1974), where the focus was on cross-section data. A common finding from cross-section data is that mean earnings is hump-shaped with age and that measures of earnings dispersion tend to increase with age. A number of studies (e.g. Creedy and Hart (1979), Shorrocks (1980), Deaton and Paxson (1994), Storresletten et. al. (2002)) have examined the pattern of earnings dispersion in cohort

---

[11]Each regression has $J \times T$ dependent variables regressed on T time dummies and J age dummies. This regression has J less regressors than the regression incorporating cohort effects.

[12]Formally, this normalization requires that $\frac{1}{T}\sum_{t=1}^{T} \tilde{\gamma}_t = 0$ and $\frac{1}{T}\sum_{t=1}^{T} \tilde{\gamma}_t t = 0$. Appendix A provides more details on how we carry out this estimation.

or repeated cross-section data and have found that dispersion tends to increase with age.[13] Schultz (1975), Smith and Welch (1979) and Dooley and Gottschalk (1984) present evidence that dispersion profiles are U-shaped in that a measure of dispersion decreases early in the life cycle and then later increases with age. We find a slight U-shape in the dispersion profile when dispersion is measured by the Gini coefficient.

At a methodological level, our work and a number of the studies cited above go beyond the early work based on a single cross-section. In particular, these studies separate age effects from cohort and/or time effects using panel data or repeated cross-sections. For example, Deaton and Paxson (1994) focus on how the variance of log earnings and the variance of log consumption in household-level data evolves over the life cycle. Their main results are based on regressing the variance of log earnings of a cohort on age and cohort dummies. They use the estimated age coefficients to highlight the effect of aging. The methodology that we employ is broadly similar. However, since we are interested in several earnings statistics there is the issue that if we were to employ this procedure on each separate statistic of interest then age and cohort effects would be extracted in a different way for each statistic. Our proposed solution is to employ the same procedure directly on the percentiles of the age and cohort specific earnings distributions. This procedure produces the age effects graphed in Figure 2. Using Figure 2, one can calculate the resulting age effects for *any* statistic of interest, knowing that cohort and/or time effects have been extracted in a consistent way.

3 Human Capital Theory

An agent maximizes the present value of earnings over the working lifetime by dividing available time between market work and human capital production.[14] This present value is given in the decision problem below, where r is a real interest rate and earnings in a period equal the product of the rental rate of human capital w_j, the agent's human capital h_j and the time spent in market work $(1 - l_j)$. The stock of human capital increases when human capital production offsets the depreciation of current human capital. Human capital production $f(h_j, l_j, a)$ depends on an agent's

---

[13]Creedy and Hart (1979) and Shorrocks (1980) use individual-level data, whereas Deaton and Paxson (1994) and Storresletten et. al. (2002) use household-level data.

[14]We note that utility maximization implies present value earnings maximization in the absence of a labor-leisure decision and liquidity constraints. Hence, nothing is lost for the study of human capital accumulation and the implied earnings dynamics if one abstracts from consumption and asset choice over the life-cycle.

learning ability a, human capital h_j and the fraction of available time l_j put into human capital production. Learning ability is fixed at birth and thus does not change over time.

$$\max \sum_{j=1}^{J} w_j h_j (1 - l_j)/(1 + r)^{j-1}$$

subject to $l_j \in [0, 1]$, $h_{j+1} = h_j(1 - \delta) + f(h_j, l_j, a)$.

We formulate this decision problem in the language of dynamic programming. The value function $V_j(h; a)$ gives the maximum present value of earnings at age j from state h when learning ability is a. The value function is set to zero after the last period of life (i.e. $V_{J+1}(h; a) = 0$). Solutions to this problem are given by optimal decision rules $h_j(h; a)$ and $l_j(h; a)$ which describe the optimal choice of human capital carried to the next period and the fraction of time spent in human capital production as functions of age j, human capital h and learning ability a.

$$V_j(h; a) = \max_{l, h'} w_j h(1 - l) + (1 + r)^{-1} V_{j+1}(h'; a)$$

subject to $l \in [0, 1]$, $h' = h(1 - \delta) + f(h, l, a)$.

We focus on a specific version of the model described above that was first analyzed by Ben-Porath (1967). In this model, the human capital production function is given by $f(h, l, a) = a(hl)^\alpha$. Proposition 1 below presents key results for this model.

Proposition 1: Assume $f(h, l, a) = a(hl)^\alpha, \alpha \in (0, 1)$, the depreciation rate $\delta \in [0, 1)$, the rental rate equals $w_j = (1 + g)^{j-1}$ and the gross interest rate $(1 + r)$ is strictly positive. Then

(i) $V_j(h; a)$ is continuous and increasing in h and a, is concave in h and $h_j(h; a)$ is single-valued.

(ii) If in addition $aA_j(a)^\alpha + (1 - \delta)A_j(a) \geq A_{j+1}(a)$, then the optimal decision rules are as follows:

$$h_j(h; a) = \begin{cases} aA_j(a)^\alpha + (1 - \delta)h & \text{for } h \geq A_j(a) \\ ah^\alpha + (1 - \delta)h & \text{for } h \leq A_j(a) \end{cases}$$

10

$$l_j(h;a) = \begin{cases} A_j(a)/h & \text{for } h \geq A_j(a) \\ 1 & \text{for } h \leq A_j(a) \end{cases}$$

$$A_j(a) \equiv \left(\frac{a\alpha(1+g)}{1+r}\right)^{\frac{1}{1-\alpha}} \left(\sum_{k=0}^{J-j}\left[\frac{(1+g)(1-\delta)}{(1+r)}\right]^k\right)^{\frac{1}{1-\alpha}}$$

(iii) Let the initial distribution of human capital and ability be such that all agents have the same ability $a > 0$ but different human capital levels and that all agents earnings are strictly positive. Also let $aA_j(a)^\alpha + (1-\delta)A_j(a) \geq A_{j+1}(a)$. Then the Lorenz curve for both human capital and earnings produced by the model becomes more equal for a cohort as the cohort ages.

Proof: See the Appendix.

We now comment on the implications of Proposition 1. First, the fact that $V_j(h;a)$ is concave in human capital means that each period the decision problem is a concave programming problem. Thus, standard techniques can be used to compute solutions regardless of any further restrictions on the parameters of the model. Our methods for computing solutions, which are described in the Appendix, employ these techniques.

Second, if the parameters of the model are restricted then a simple, closed-form solution exists. The solution has the property that an agent spends all time in human capital accumulation provided that current human capital is below an age and ability dependent cutoff $A_j(a)$. The restrictions in Proposition 1(ii) amount to the assumption that once an agent with ability a stops full-time schooling (i.e current human capital is above the cutoff level $A_j(a)$) then the agent never returns to full-time schooling (i.e. future human capital remains above future cutoff levels $A_{j+1}(a)$). The parameter values used in this paper turn out to satisfy these restrictions at all ability levels for the initial distributions of learning ability and human capital that best match the facts documented in Figure 1.

Third, the fact that the decision rule for human capital $h_j(h;a)$ in Proposition 1(ii) is increasing in both current human capital and learning ability has a number of implications. For example, at the end of the working life cycle the agents who are high earners are precisely those who started off with high initial human capital and/or ability. This is true since at the end of the life cycle earnings are proportional to human capital. Similar reasoning implies the greater the dispersion in earnings at the end of the working life cycle the greater is the required dispersion in human capital or learning ability at the beginning of the life cycle. This is key for this paper as it focuses on characterizing the nature of initial agent heterogeneity that is critical for replicating observed earnings distribution dynamics.

11

Fourth, Proposition 1 (iii) highlights a key property of the model. Specifically, if all agents within an age group have the same learning ability, then as these agents age both human capital dispersion and earnings dispersion must *decrease* for any dispersion measure consistent with the Lorenz order. More precisely, the Lorenz curves for human capital and earnings can be ordered in the sense that the Lorenz curve for age j lies strictly below the corresponding Lorenz curve for age $j + 1$ and so on. This follows from the fact that agents with the lower human capital have higher human capital growth rates and the fact that agents with lower earnings have higher earnings growth rates. This result implies that differences in learning ability are absolutely fundamental for this model to be able to produce even the qualitative pattern of growing earnings dispersion documented in Figure 1.

Finally, we comment on one form of heterogeneity that we abstract from. Individuals may conceivably face different rental rates for the same human capital services. This could be motivated by racial or gender discrimination. We note that adding exogenous differences in rental rates would not by itself produce either the increase in earnings dispersion or skewness with age that we document in US data. More specifically, if rental rates differ proportionally over the life cycle across agents, holding initial human capital and learning ability equal, then earnings dispersion and skewness would be counterfactually constant. This follows from Proposition 1, as such differences do not alter human capital decisions even though they have proportional effects on earnings.

4 Parameter Values

The findings of this paper are based on the parameter values indicated in Table 1. The time period in the model is a year. An agent's working lifetime is taken to be either 39 or 49 model periods, which corresponds to a real life age of 20 to 58 and 10 to 58 respectively. These two values allow us to explore different views about when the human capital accumulation mechanism highlighted by the model begins. The real interest rate is set to 4 percent. The rental rate of human capital equals $w_j = (1 + g)^{j-1}$ and the growth rate is set to $g = .0014$. This growth rate equals the average growth rate in average real earnings per person over the period 1968-92 in our PSID sample.[15] Within the model the growth rate of the rental rate equals the growth rate of average earnings, when rental growth and population growth are constant and when the initial distribution of human capital and ability is time invariant. Given

---

[15]The growth rate of average wages (e.g. total labor earnings divided by total work hours) over 1968-92 in our PSID sample equals .0017.

the growth in the rental rate, we set the depreciation rate to $\delta = 0.0114$ so that the model produces the rate of decrease of average real earnings at the end of the working life cycle documented in Figure 1.[16] The model implies that at the end of the life cycle negligible time is allocated to producing new human capital and, thus, the gross earnings growth rate approximately equals $(1 + g)(1 - \delta)$. When we choose the depreciation rate on this basis the value lies in the middle of the estimates in the literature surveyed by Browning, Hansen, and Heckman (1999).

Estimates of the elasticity parameter α of the human capital production function are surveyed by Browning et. al. (1999). These estimates range from 0.5 to almost 1.0. We note that this literature estimates α so that the earnings profile produced by one agent in the model best fits the earnings data. Thus, the maintained assumption is that everyone is identical at birth so that the initial distribution of learning ability and human capital across agents is a point mass.[17] We note that this initial distribution is unrestricted by the theory and therefore treat it as a free parameter in our work. Thus, we remain agnostic about the value of α and assess the model for values between 0.5 and 1.0.

Table 1: Parameter Values

Model Periods	Interest Rate	Rental Growth	Depreciation Rate	Production Function
$J = 39, 49$	$r = .04$	$g = .0014$	$\delta = .0114$	$\alpha \in [0.5 - 1.0)$

5 Findings

5.1 Earnings Distribution Dynamics

Earnings distribution dynamics implied by the model are determined in two steps. First, we compute the optimal decision rule for human capital for the parameters described in Table 1. Second, we choose the initial distribution of the state variable to best replicate the properties of US data documented in Figure 1. The Appendix describes how these steps are carried out.

---

[16]We use a rate of growth in earnings at the end of the life cycle equal to -0.01. The growth rate in mean earnings at the end of the life-cycle from Fig. 1 is -0.0107 and -0.0078 for age groups 55-58 and 50-58 respectively.

[17]Heckman et. al. (1998) allow for agent heterogeneity. They estimate model parameters so that earnings of one agent in the model best match earnings data for individuals sorted by a measure of ability and by whether or not they went to college.

We consider both parametric and non-parametric approaches for choosing the initial distribution. In the parametric approach this distribution is restricted to be jointly, log-normally distributed. This class of distributions is characterized by 5 parameters. In the non-parametric approach, we allow the initial distribution to be any histogram on a rectangular grid in the space of human capital and learning ability. In practice, this grid is defined by 20 points in both the human capital and ability dimensions and thus, there are a total of 400 bins used to define the possible histograms. In both approaches we search over the vector of parameters that characterize these distributions so as to minimize the distance between the model and data statistics for mean earnings, dispersion and skewness.[18]

The results are presented in Figure 4 and 5 for the parametric and non-parametric case under the assumption that human capital accumulation starts at a real life age of 10 and 20, respectively. Note that the model implications are very similar for these two different starting ages. For a better visual presentation, we graph in all cases results for only the central value of $\alpha = 0.7$. We emphasize that similar quantitative patterns emerge for all values of α between .5 and .9. These figures demonstrate that the model is able to replicate the qualitative properties of the US earnings distribution dynamics presented in Figure 1 both when the initial distribution is chosen parametrically and non-parametrically. Moreover, the results for the non-parametric case are quite striking: the model replicates to a surprising degree the quantitative features of US earnings distribution dynamics.[19]

[Insert Figure 4 (a-c) Here]

[Insert Figure 5 (a-c) Here]

As a measure of the goodness of fit, we present in Table 2 the average (percentage)

---

[18]More precisely, we find the parameter vector γ characterizing the initial distribution that solves the minimization problem below, where m_j, d_j, s_j are the statistics of means, dispersion and inverse skewness constructed from the PSID data, and $m_j(\gamma), d_j(\gamma), s_j(\gamma)$ are the corresponding model statistics.

$$\min_{\gamma} \sum_{j=1}^{J} ([\log(m_j/m_j(\gamma))]^2 + [\log(d_j/d_j(\gamma))]^2 + [\log(s_j/s_j(\gamma))]^2)$$

This form of the objective ensures that the numerical solution to the problem is not affected by the units of measurement of the statistics in question.

[19]As we explained in the introduction, this ability of the model to replicate the facts does *not* rely upon the possibility of choosing an infinite number of parameters characterizing the initial distribution.

14

deviation, in absolute terms, between the model implied statistics and the data.[20] By this measure, on average the model implied statistics differ from the data by 2.5% to 3.8% in the non-parametric case for different values of the elasticity parameter of the production function. In the parametric case, the fit is naturally not as good; in this case the model differs from the data by 5% to 7.5%. Graphically, the parametric case produces too much earnings skewness in each age group. Nonetheless, a parsimonious representation of the initial distribution can go a long way towards reproducing the dynamics of the US age-earnings distribution.

Table 2: Mean Absolute Deviation (%)

Case	$\alpha = 0.5$	$\alpha = 0.6$	$\alpha = 0.7$	$\alpha = 0.8$	$\alpha = 0.9$
Panel A: Accumulation starts at Age 10					
Non-Parametric	3.5	3.2	2.6	2.5	2.8
Parametric	7.5	6.4	5.9	5.2	6.2
Panel B: Accumulation starts at Age 20					
Non-Parametric	3.1	3.5	2.8	3.9	3.8
Parametric	6.8	7.0	5.2	5.0	6.4

To close this section, we note that the benchmark human capital model is also successful in an alternative dimension. Specifically, features of the cross-section earnings distribution implied by the model are roughly in line with the corresponding features in cross-section data. We construct the cross-section earnings distribution implied by the data using the cohort-specific earnings percentiles in Figure 2 together with the assumption that the population growth rate is 1%. The resulting cross-sectional earnings distribution has a Gini coefficient of 0.33, a skewness measure of 1.16 and a fraction of earnings in the upper 20%, 10%, 5% and 1% of 40.2%, 25.1%, 15.5% and 4.7% respectively. The model for $\alpha = 0.7$ in the non-parametric case implies a cross-sectional earnings distribution with a Gini coefficient of 0.327, a skewness measure of 1.18, with corresponding fractions of earnings in the upper tail of 40.9%, 27.0%, 17.5% and 6.1%.

5.2 Importance of Ability and Human Capital Differences

The previous section demonstrated that the US earnings distribution dynamics documented in Figure 1 can be fairly well matched by the model from the right initial

---

[20]The goodness of fit measure is $[\sum_{j=1}^{J} |\log(m_j/m_j(\gamma))| + |\log(d_j/d_j(\gamma))| + |\log(s_j/s_j(\gamma))|]/(3J)$.

distribution of human capital and learning ability. Which features of this initial distribution are critical? We know from Proposition 1 in section 3 that differences in learning ability have to exist across agents if the model is to produce any increase in earnings dispersion for a cohort as the cohort ages. Thus, learning ability differences are essential. But could the model produce the patterns in Figure 1 with only differences in learning ability and no human capital differences early in life?

To answer this question, we place a grid on values of learning ability, and search for the distribution of learning ability and the common, fixed value of initial human capital that best reproduces the facts presented in Figure 1.[21] Our findings are presented in Figure 6 where the model begins to operate when agents are at a real life age of 10. Starting the model later than this age produces even more strongly counterfactual implications.

We find that the model generates a much more pronounced U-shaped pattern for earnings dispersion than is present in the data. To understand why this occurs recall that all agents start life with the same level of human capital. Optimal accumulation then dictates that early in the life cycle agents with high learning ability devote most of their time or all available time to accumulating human capital. Thus, early in the life cycle the earnings of high ability agents are lower than those of their low ability counterparts. This follows from Proposition 1(ii) in section 3. The bottom of the U-shape occurs where earnings of high ability agents overtake those of lower ability agents. This occurs at about age 24 for the distribution which best matches the data. After this age, earnings dispersion increases as high ability agents have more steeply sloped age-earnings profiles than low ability agents. Thus, we conclude that while differences in learning ability are essential, differences in human capital early in the life cycle are also important. The next section goes on to show that a positive correlation between learning ability and initial human capital is a feature of the initial distributions which best match the data. This positive correlation lifts up the age-earnings profiles of high ability agents relative to low ability agents and, thus, reduces the strong U-shape in the dispersion profile displayed in Figure 6.

[Insert Figure 6 a-c Here]

Another way of assessing the importance of ability versus human capital differences is to ask to what extent is the dispersion in the present value of earnings accounted for by differences in learning ability alone. To answer this question, we undertake a simple variance decomposition exercise. We calculate the variance of the present

---

[21]We put a grid of 20 values of learning ability (as in the general case) to search for the best distribution.

16

value of earnings (as of age 20), and report the percentage of this variance that can be attributed to learning ability differences.[22] Table 3 below shows that, once again, differences in learning ability are key. They account for most of the variance in the present value of earnings. In all cases, ability differences account for more than 60% of the total variance. The residual variance is due to human capital differences at fixed ability levels.

Table 3: Percentage Variance in PV of Earnings Due to Learning Ability Differences

Statistic	$\alpha = 0.5$	$\alpha = 0.6$	$\alpha = 0.7$	$\alpha = 0.8$	$\alpha = 0.9$
Panel A: Accumulation starts at Age 10					
Non-Parametric	78.2	73.8	70.5	68.4	73.1
Parametric	77.8	72.9	68.0	77.1	79.2
Panel B: Accumulation starts at Age 20					
Non-Parametric	62.7	63.1	63.5	64.4	65.1
Parametric	73.0	65.4	72.8	70.8	78.3

5.3 Properties of Initial Distributions

Tables 4 and 5 characterize properties of the initial distributions that produce the earnings distribution implications highlighted in Figures 4 and 5. Several regularities are apparent. First, the properties of means, dispersion, skewness and correlation in Table 4 for the non-parametric case are similar to those in Table 5 for the parametric case. Thus, the economic content of what the model and the data in Figure 1 impose on the initial distribution appears not to be too sensitive to whether or not one restricts this initial distribution in a parsimonious way.

Second, initial human capital and learning ability are positively correlated when the human capital accumulation process articulated by the model starts at age 10 but are much more highly correlated when the process starts at age 20. This finding is implied by the dynamics of the model. In particular, distributions which at age 10 have low correlation induce more highly correlated distributions in each successive period as agents age. This occurs, according to Proposition 1, since in each period

---

[22]More formally, we proceed as follows. Let $PV(a, h_1)$ denote the present value of earnings from initial condition (a, h_1). The results in Table 3 report the ratio $[\sigma^2(E(PV(a, h_1)|a))/\sigma^2(PV(a, h_1))] \times 100$. The numerator is the variance across learning ability levels of the mean present value of earnings, conditional on learning ability. The denominator is total variance in the present value of earnings.

high ability agents produce more human capital than low ability agents, holding initial human capital at the beginning of life equal. Thus, when the initial distribution is chosen to best match the data the model implies that the correlation between human capital and learning ability increases as agents age.

Third, when the human capital accumulation process starts at age 10, the model implies that for a cohort average human capital at the beginning of the life cycle is less than at the end of the life cycle. Thus, there is net human capital accumulation for a cohort over the life cycle. To see this point recall that mean earnings at age 58 is normalized to equal 100 and that the rental rate of human capital is set to equal $w_j = 1.0014^{j-1}$. The implication is that mean human capital must be slightly less than 100 at age 58 to match the earnings data at that age. Since mean human capital early in life is less than this level the conclusion follows.

Table 4: Ability and Human Capital at Birth (Non-Parametric Case)

Statistic	$\alpha = 0.5$	$\alpha = 0.6$	$\alpha = 0.7$	$\alpha = 0.8$	$\alpha = 0.9$
Panel A: Accumulation starts at Age 10					
Mean (a)	0.466	0.319	0.209	0.139	0.087
Coef. of Variation (a)	0.601	0.463	0.358	0.243	0.212
Skewness (a)	1.303	1.190	1.183	1.168	1.103
Mean (h_1)	69.6	71.4	74.9	76.0	83.5
Coef. of Variation (h_1)	0.456	0.453	0.422	0.397	0.261
Skewness (h_1)	1.152	1.146	1.151	1.155	1.142
Correlation (a, h_1)	0.10	0.205	0.305	0.397	0.418
Panel B: Accumulation starts at Age 20					
Mean (a)	0.453	0.320	0.210	0.134	0.089
Coef. of Variation (a)	0.669	0.504	0.365	0.324	0.168
Skewness (a)	1.251	1.188	1.147	1.131	1.111
Mean (h_1)	86.8	88.1	93.4	94.5	99.6
Coef. of Variation (h_1)	0.475	0.486	0.510	0.457	0.501
Mean (h_1)	86.8	88.1	93.4	94.5	99.6
Skewness (h_1)	1.148	1.163	1.167	1.135	1.124
Correlation (a, h_1)	0.621	0.689	0.781	0.792	0.741

Table 5: Ability and Human Capital at Birth (Parametric Case)

Statistic	$\alpha = 0.5$	$\alpha = 0.6$	$\alpha = 0.7$	$\alpha = 0.8$	$\alpha = 0.9$
Panel A: Accumulation starts at Age 10					
Mean (a)	0.499	0.322	0.207	0.139	0.089
Coef. of Variation (a)	0.514	0.436	0.353	0.235	0.198
Skewness (a)	1.125	1.092	1.061	1.027	1.010
Mean (h_1)	64.0	69.2	74.7	75.1	78.6
Coef. of Variation (h_1)	0.454	0.453	0.434	0.403	0.184
Skewness (h_1)	1.100	1.100	1.090	1.077	1.071
Correlation (a, h_1)	0.070	0.145	0.171	0.333	0.351
Panel B: Accumulation starts at Age 20					
Mean (a)	0.467	0.321	0.209	0.136	0.088
Coef. of Variation (a)	0.613	0.474	0.347	0.257	0.158
Skewness (a)	1.191	1.109	1.058	1.033	1.012
Mean (h_1)	86.7	89.5	92.3	96.6	100.1
Coef. of Variation (h_1)	0.427	0.439	0.481	0.468	0.459
Skewness (h_1)	1.088	1.092	1.109	1.105	1.100
Correlation (a, h_1)	0.600	0.621	0.781	0.792	0.796

Fourth, mean learning ability declines as the curvature parameter α increases, while the opposite is true for mean initial human capital. To gain intuition, note that for given learning ability and initial human capital a higher value of α lowers earnings early in life and raises earnings later in life – in effect rotating individual age-earnings profiles counter-clockwise. This follows, see Proposition 1, since as α increases time spent working early in life decreases whereas end of life human capital increases. Raising mean initial human capital and lowering mean learning ability serves to rotate the age-earnings profiles clockwise to counteract the effect of increasing α.

5.4 Persistence in Individual Earnings

So far we have looked at how the earnings distribution changes as agents age. However, it is possible that different theoretical models may all be able to replicate the patterns of means, dispersion and skewness in US cohort data, but differ in their implications for earnings persistence. The latter is a topic that has spawned considerable attention in the labor, consumption, and income distribution literatures and

for which the benchmark model has strong implications. In addition, it is of independent interest to investigate the performance of the benchmark model in terms of a number of facts that we did not force it to match.

We now characterize the extent to which measures of persistence in the model are consistent or inconsistent with the corresponding measures from US data. We consider two measures of persistence in cohort data: (1) the correlation of individual earnings levels across periods and (2) the correlation of individual earnings growth rates across periods. Table 6 shows the results for various age groups within the model, when the initial distribution of human capital and ability is selected using the non-parametric methodology. For ease of exposition, we report results only for the case when accumulation starts at age 10 and $\alpha = 0.7$. The findings are that both earnings levels and growth rates are very highly correlated across model periods.

Table 6: Persistence in Individual Earnings $\alpha = 0.7$

Statistic	Age $(j) = 45$	Age $(j) = 40$
Panel A: Correlation - Levels		
Correlation(E_j, E_{j-1})	0.9999	0.9997
Correlation(E_j, E_{j-5})	0.9966	0.9854
Correlation(E_j, E_{j-10})	0.9679	0.8671
Panel B: Correlation - Growth Rates		
Correlation(z_j, z_{j-1})	0.9995	0.9994
Correlation(z_j, z_{j-5})	0.9960	0.9652
Correlation(z_j, z_{j-10})	0.9750	0.5229

E_j and $z_j = \log(E_j/E_{j-1})$ denote earnings and earnings growth rates, respectively.

We now compare the results in Table 6 with estimates from US data. The correlation of earnings levels has been examined in US data by Parsons (1978) and Hyslop (2001) among others. They find that earnings among US males are positively correlated for all horizons considered and that the correlation typically falls as the horizon increases. Hyslop finds that the average correlation is 0.83 for a one year horizon and 0.59 for a six year horizon. Parsons finds that correlations are typically higher for older age groups. These results are qualitatively consistent with those from the human capital model.

A different picture emerges for the correlation of growth rates. Abowd and Card (1989) estimate the correlation in earnings growth rates for US males. They find that the average correlation of earnings growth rates one year apart is *negative* and equal to about -0.34, and close to zero when the growth rates are more than one year apart. Baker (1997) reports similar findings. Storesletten et. al. (2004) report high but stationary persistence in log-earnings which imply slightly lower negative auto-correlations of growth rates one year apart. Processes with similar dynamics have also been estimated by McCurdy (1982) and Hubbard, Skinner and Zeldes (1994). The results estimated from the data are thus clearly inconsistent with those implied by the model. These results are suggestive of a key ingredient present in stochastic models of the earnings distribution, namely, shocks that cause earnings to be mean reverting. In the conclusion we outline some candidates for shocks that can be incorporated into human capital theory.

6 Conclusion

We assess the degree to which a widely-used, human-capital model is able to replicate the age dynamics of the US earnings distribution documented in Figure 1. We find that the model can account quite well for these age-earnings dynamics. In addition, we find that the model produces a cross sectional earnings distribution closely resembling that implied by the age-earnings dynamics documented in Figure 2. Our findings indicate that differences in learning ability across agents are key. In particular, in the model high ability agents have more steeply sloped age-earnings profiles than low ability agents. These differences in earnings profiles in turn produce the increases in earnings dispersion and skewness with age that are documented in Figures 1 and 2. These findings are robust to the age at which the human capital accumulation mechanism described by the model begins and to different values of the elasticity parameter of the human capital production function. We also find that, despite its relative success in replicating these facts, the model is inconsistent with evidence related to the persistence of individual earnings.

We mention two areas in which future work seems promising. The first has to do with the fact that the distribution of agents by initial human capital and ability is unrestricted by the model. Models of the family can provide restrictions on this initial distribution. For this class of models, an assessment of the ability to replicate the facts of age-earnings dynamics and intergenerational earnings correlations is a natural next step.

The second area for future work deals with the fact that the model examined

here abstracts from many seemingly important features. Three such features are the absence of a leisure decision, an occupational choice decision and shocks that make human capital risky. We comment on this last feature. First, allowing for risky human capital would be one way of integrating deeper foundations for earnings risk into the standard consumption-savings problem considered by the literature on the life-cycle, permanent-income hypothesis. This literature has examined in detail the determinants of consumption and financial asset holdings over the life cycle, but no comparable effort has been put into investigating the accumulation of human capital. Second, while there seems to be agreement that human capital is risky there is relatively little work that analyzes different sources of risk and then determines their quantitative importance.[23] It is clear from this paper that a richer set of facts is needed to identify both initial conditions and shocks in a model with risky human capital, given that human capital theory can explain the patterns in Figure 1 without shocks. Two interesting questions for a theory with risky human capital are (i) can such a model account for both the distributional dynamics of earnings and consumption over the life cycle? and (ii) what fraction of the dispersion in lifetime earnings is accounted for by initial conditions versus shocks? We plan to explore these questions in future work.

---

[23]Within a human capital model, shocks can no longer be modeled as exogenous shocks to earnings. Instead, they must be modeled at a deeper level as shocks to the depreciation of human capital, to learning ability, to the employment match, to rental rates and so on. Each one of these alternatives poses different modeling as well as empirical challenges.

References

Andolfatto, D, Gomme, P. and C. Ferrall (2000), Human Capital Theory and the Life-Cycle Pattern of Learning and Earning, Income and Wealth, mimeo.

Ben-Porath, Y. (1967), The Production of Human Capital and the Life Cycle of Earnings, *Journal of Political Economy,* 75, 352-65.

Browning, M., Hansen, L. and J. Heckman (1999), Micro Data and General Equilibrium Models, in *Handbook of Macroeconomics,* ed. J.B. Taylor and M. Woodford, (Elsevier Science B.V, Amsterdam).

Cagetti. M. (2002), Wealth Accumulation over the Life-cycle and Precautionary Savings, mimeo, University of Virginia.

Carroll, C. (1997), Buffer Stock Saving and the Life-Cycle Permanent-Income Hypothesis, *Quarterly Journal of Economics*, 92, 1-56.

Card, D. (1999), The Causal Effect of Education on Earnings, in *Handbook of Labor Economics,* Volume 3, ed. O. Ashenfelter and D. Card.

Castañeda, A., Diaz-Jimenez, J. and V. Rios-Rull (2002), Accounting for Earnings and Wealth Inequality, forthcoming *Journal of Political Economy.*

Creedy, J. and P. Hart (1979), Age and the Distribution of Earnings, *Economic Journal,* 89, 280-93.

Deaton, A. (1992), *Understanding Consumption*, Clarendon Lectures in Economics, Oxford University Press.

Deaton, A. and C. Paxson (1994), Intertemporal Choice and Inequality, *Journal of Political Economy,* 102, 437-67.

Deaton, A., Gourinchas, P. and C. Paxson (2002), Social Security and Inequality over the Life Cycle, in *Distributional Aspects of Social Security and Social Security Reform*, Martin Feldstein and Jeffrey Liebman, eds, The University of Chicago Press.

De Nardi, C. (2002), Wealth Inequality and Intergenerational Links, University of Minnesota, mimeo.

De Nardi, C., Imrohoroglu, S. and T.Sargent (1999), Projected U.S. Demographics and Social Security, *Review of Economic Dynamics*, 2, 575-615.

Domeij, D. and P. Klein. (2002) Public Pensions: To What Extent Do They Account for Swedish Wealth Inequality?, *Review of Economic Dynamics*, 5,503-34.

Dooley and Gottschalk (1984), Earnings Inequality Among Males in the United States: Trends and the Effect of Labor Force Growth, *Journal of Political Economy,* 92, 59- 89.

Fuster, L. (1999), Is Altruism Important for Understanding the Long Run Effects of Social Security?, *Review of Economic Dynamics*, 2, 616-37.

Gourinchas, P. and J. Parker (2002), Consumption Over the Life Cycle, *Econometrica*, 70, 47-89.

Haley, W. (1976), Estimation of the Earnings Profile from Optimal Human Capital Accumulation, *Econometrica,* 44, 1223-38.

Hanoch, G, and M. Honig (1985), "True" Age Profiles of Earnings: Adjusting for Censoring and for Period and Cohort Effects, *Review of Economics and Statistics,* 67, 383-94.

Heathcote, J., Storesletten, K. and G. Violante (2003), The Cross-Sectional Implications of Rising Wage Inequality in the United States, working paper.

Heckman, J. (1975), Estimates of a Human Capital Production Function Embedded in a Life Cycle Model of Labor Supply, in *Household Production and Consumption,* ed. N. Terleckyj, Columbia University Press, New York.

Heckman, J. (1976), A Life-Cycle Model of Earning, Learning and Consumption, *Journal of Political Economy,* 84, S11-S44.

Heckman, J., Lochner, L. and C. Taber (1998), Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents, *Review of Economic Dynamics,* 1, 1-58.

Hubbard G. R., J. Skinner, and S. Zeldes (1994), The Importance of Precautionary Motives in Explaining Individual and Aggregate Savings, *Carnegie Rochester Conference Series on Public Policy,* 40, 59-126.

Huggett, M. (1996), Wealth Distribution in Life-Cycle Economies, *Journal of Monetary Economics*, 38, 469-94.

Huggett, M. and G. Ventura (1999), On the Distributional Effects of Social Security Reform, *Review of Economic Dynamics*, 2, 498-531.

Hyslop, D. (2001), Rising US Earnings Inequality and Family Labor Supply: The Covariance Structure of Intrafamily Earnings, *American Economic Review,* 91, 755- 77.

Imrohoroglu, A., Imrohoroglu, S. and D. Joines (1995), A Life Cycle Analysis of Social Security, *Economic Theory*, 6, 83-114.

Keane, M. and Wolpin, K., The Career Decisions of Young Men, *Journal of Political Economy*, 105(3), 473-522.

Krueger, D. and Fernandez-Villaverde, J. (2001), Consumption and Saving Over the Life-Cycle: How Important are Consumer Durables?, Stanford University, mimeo.

Krueger, D. and F. Perri (2002) Does Income Inequality Lead to Consumption Inequality? Evidence and Theory, Stanford University, mimeo.

Lillard, L. (1977) Inequality: Earnings vs. Human Wealth, *American Economic Review,* 67, 42-53.

MaCurdy T. (1982) The Use of Time Series Processes to Model the Error Structure of Earnings in Longitudinal Data Analysis *Journal of Econometrics,* 18, 83-118.

Mincer, J. (1958), Investment in Human Capital and Personal Income Distribution, *Journal of Political Economy,* 66, 281- 302.

Mincer, J. (1974), *Schooling, Experience and Earnings,* Columbia University Press, New York.

Mincer, J. (1997), The Production of Human Capital and the Life Cycle of Earnings: Variations on a Theme, *Journal of Labor-Economics,* 15 (1), Part 2.

Neal, D. and S. Rosen (2000), Theories of the Distribution of Earnings, in *Handbook of Income Distribution,* ed. A. Atkinson and F. Bourguignon, North Holland Publishers.

Parsons, D. (1978), The Autocorrelation of Earnings, Human Wealth Inequality, and Income Contingent Loans, *Quarterly Journal of Economics* , 92, 551-69.

Press, W. et. al. (1992), *Numerical recipes in FORTRAN,* Second Edition, Cambridge University Press.

Quadrini, V (2000), Entrepreneurship, Savings and Social Mobility, *Review of Economic Dynamics*, 3, January.

Rosen, S. (1976), A Theory of Life Earnings, *Journal of Political Economy,* 84, S45-S67.

Schultz, T. (1975), Long-Term Change in Personal Income Distribution: Theoretical Approaches, Evidence, and Explanations, in *The "Inequality" Controversy: Schooling and Distributive Justice,* ed. D. Levine and M. Bane, (New York, Basic Books).

Shorrocks, A. (1980), Income Stability in the United States, in *Statics and Dynamics of Income,* ed. Klevmarken and Lybeck.

Smith and Welch (1979), Inequality: Race Differences in the Distribution of Earnings, *International Economic Review,* 20, 515-26.

Stokey, N. and R. Lucas with E. Prescott (1989), Recursive Methods in Economic Dynamics, (Harvard University Press, Cambridge).

Storesletten, K., Telmer, C, and A. Yaron (1999), The Risk Sharing Implications of Alternative Social Security Arrangements, *Carnegie-Rochester Conference Series on Public Policy*, 50, 312-59.

Storresletten, K., Telmer, C. and A. Yaron (2004), Consumption and Risk Sharing Over the Life Cycle, *Journal of Monetary Economics*, Volume 51, Issue 3, 609-633.

Weiss, Y. and L. Lillard (1978), Experience, Vintage, and Time Effects in the Growth of Earnings: American Scientists, 1960-1970, *Journal of Political Economy,* 86, 427-47.

Weiss, Y. (1986), The Determination of Life-Cycle Earnings, in *Handbook of Labor Economics,* Volume I , ed. O. Ashenfelter and R. Layard, Elsevier Science Publishers.

Weizsacker, R. von (1993), *A Theory of Earnings Distribution,* Cambridge University Press.

A Appendix

A.1 Proposition 1

To prove Proposition 1 it is useful to reformulate the dynamic programming problem by expressing earnings as a function of future human capital and ability. The resulting earnings function is denoted $G(h, h', a; j)$.

$$V_j(h; a) = \max_{h'} G(h, h', a; j) + (1+r)^{-1} V_{j+1}(h'; a)$$

$$h' \in \Gamma(h, a) \equiv [h(1-\delta) + f(h, 0, a), h(1-\delta) + f(h, 1, a)]$$

Proof of Proposition 1:

(i) The continuity of the value function follows by repeated application of the Theorem of the Maximum starting in the last period of life. To apply the Theorem of the Maximum, we make use of the continuity of $G(h, h', a; j)$ and the fact that the constraint set is a continuous and compact-valued correspondence. These are easily verified. To show that the value function increases in h and a, note that this holds in the last period since $V_J(h; a) = w_J h$. Backward induction establishes the result for earlier periods using the fact that $G(h, h', a; j)$ increases in h and a.

The concavity of the value function in human capital follows from backwards induction by applying repeatedly the argument used in Stokey and Lucas (1989, Thm. 4.8). To apply this argument, we make use of three properties. First, the graph of the constraint set $\{(h, h') : h \in R_+^1, h' \in \Gamma(h, a)\}$ is a convex set for any given ability level a. This follows from the fact that the human capital production function is concave in current human capital. Second, $G(h, h', a; j)$ is jointly concave in (h, h'). This can be easily verified. Third, the terminal value function $V_{J+1}(h; a) \equiv 0$ is concave in human capital.

The decision rule $h_j(h; a)$ is single-valued since the objective function is strictly concave and the constraint set, for given $(h; a)$, is convex. The objective function is strictly concave because the value function is concave and because $G(h, h', a; j)$ is strictly concave in h'.

(ii) Define $V_j(h; a)$ recursively, given $V_J(h; a) = (1+g)^{J-1} h$, as follows:

$V_j(h; a) = [(1+g)^{j-1} \sum_{k=0}^{J-j} [\frac{(1+g)(1-\delta)}{(1+r)}]^k] h + C_j(a)$ for $h \geq A_j(a)$.

$V_j(h; a) = \frac{1}{(1+r)} V_{j+1}(h(1-\delta) + ah^\alpha; a)$ for $h \leq A_j(a)$

28

$$C_j(a) \equiv (1+r)^{-1}(C_{j+1}(a) + D_{j+1}aA_j(a)^\alpha) - (1+g)^{j-1}A_j(a), \text{ where } C_J(a) = 0$$

$$D_j \equiv [(1+g)^{j-1} \sum_{k=0}^{J-j}[\tfrac{(1+g)(1-\delta)}{(1+r)}]^k]$$

Now verify that the functions $(V_j(h;a), h_j(h;a))$ satisfy Bellman's equation. Verification amounts to checking that $h_j(h;a)$ satisfies Bellman's equation without the max operation and that it achieves the maximum in the right-hand-side of Bellman's equation. Since the first part is routine, the proof focuses on the second part. A sufficient condition for an interior solution is given in the first equation below. The second equation follows from the first after substituting the relevant functions evaluated at $h' = h_j(h;a)$. Here we make use of the assumption on the cutoff values $A_j(a)$ in Prop 1(ii) since we substitute for $V'_{j+1}(h';a)$ assuming interior solutions obtain in future periods. Rearrangement of the second equation implies that $A_j(a)$ is defined as in Prop 1(ii).

$$-G_2(h, h', a; j) = (1+r)^{-1}V'_{j+1}(h';a)$$

$$(1+g)^{j-1}(1/(a\alpha))A_j(a)^{1-\alpha} = \frac{(1+g)^j}{(1+r)} \sum_{k=0}^{J-j-1}[\frac{(1+g)(1-\delta)}{(1+r)}]^k$$

It remains to consider the possibility of a corner solution. The first equation below gives a sufficient condition for a corner solution. The second equation follows from the first after substitution. Since V_{j+1} is concave in human capital, it is clear that V'_{j+1} is bounded below by the derivative above the cutoff human capital level $A_{j+1}(a)$. Thus, from the interior solution case, the second equation holds whenever $h \leq A_j(a)$.

$$-G_2(h, h', a; j) \leq (1+r)^{-1}V'_{j+1}(h';a)$$

$$(1+g)^{j-1}(1/(a\alpha))h^{1-\alpha} \leq \frac{1}{(1+r)}V'_{j+1}(h_j(h;a);a)$$

(iii) Focus first on the Lorenz curve for human capital. From Proposition 1(ii) an individual's growth rate of human capital decreases as current human capital increases. Thus, the growth rate of aggregate human capital for agents above the pth percentile of human capital is no greater than the growth rate of those below the pth percentile. The height of the period j Lorenz curve at percentile p must be weakly lower than that of the period $j + 1$ Lorenz curve. As this holds at all percentiles p, the claim follows.

Focus now on the Lorenz curve for earnings. Earnings equal $e_j = w_j(h_j - A_j(a)) = w_j(h_{j-1}(1-\delta) + aA_{j-1}(a)^\alpha - A_j(a))$. Differentiate the expression below with respect to human capital. Note that the growth rate falls as human capital increases and, thus, as earnings increase. Repeat the argument used for the human capital Lorenz curve to get that the height of the period j earnings Lorenz curve at percentile p must be weakly lower than that of the period $j+1$ Lorenz curve. As this holds at all percentiles p, the claim follows.

$$e_j/e_{j-1} = (w_j/w_{j-1})[(h_{j-1}(1-\delta) + aA_{j-1}(a)^\alpha - A_j(a))/(h_{j-1} - A_{j-1}(a))]$$

■

A.2 Computation

We sketch the computation algorithm for the non-parametric case.

Step 1: Calculate the optimal decision rule $h_j(h;a)$.

Step 2: Put a grid on learning ability and initial human capital (h,a) and calculate life-cycle profiles of human capital, hours and earnings from these grid points.

Step 3: Find the initial distribution.

To calculate the optimal decision rule in step 1, for any value of learning ability a, we put a non-uniform grid on human capital of 300 points on $[0, h^*]$, where the choice of h^* may be revised depending on the results of step 3. We calculate the optimal decision rule for human capital at gridpoints starting from period $j = J - 1$ by solving the dynamic programming problem starting from period $J - 1$, given $V_J(h;a) = w_J h$. Since the value function is concave in human capital each period, the dynamic programming problem is a concave programming problem. Golden section search (see Press et al (1992), ch. 10) is used to calculate $h_j(h;a)$ at gridoints. To carry this out, we calculate the value function off gridpoints using linear interpolation. Backward recursion on Bellman's equation produces $h_j(h;a)$ for $j = 1, ..., J - 1$.

In step 2 we put a grid of 20 points on $[0, a^*]$ and 20 points on $[0, h^*]$. This implies a total of 400 points (h, a). Using the decision rule from step 1, we simulate life-cycle profiles of labor earnings from any initial pair (h, a). Since decision rules are computed at gridpoints of human capital holdings, but its values are not restricted to lie on these gridpoints, we use linear interpolation to calculate values off gridpoints.

In step 3 we use the Simplex algorithm, as described by Press et al (1992, ch. 10), to find the 400 values of the histogram over $[0, h^*] \times [0, a^*]$ that minimizes the distance between model and data statistics. For any trial of the vector describing the initial distribution, we calculate the mean, dispersion and skewness statistics at each age using the calculated life-cycle profiles and the guessed initial distribution. The calculation of decision rules and the posterior life-cycle simulation are independent of the initial distribution. This reduces computation time as life-cycle profiles are calculated only once and stored to be used later in the calculation of the relevant statistics in all the trials required by the simplex method. If the histogram that best matches the data puts strictly positive weight on (h, a) pairs where $a = a^*$ or $h = h^*$, then the upper bounds are increased and steps 1-3 are repeated.

A.3 Data

A.3.1 Restricted Time Effects

Below we provide details for implementing the restricted time effects discussed in section 2.3. Let $X = [\alpha_s, \beta_j, \gamma_t]$ be the matrix of cohort, age and time dummies with the number of rows equal to the number of $\{j, t\}$ pairs available for earnings $e_{j,t}$, where for simplicity we omit the dependence on percentile p. Define the unrestricted regression $e = Xb + \epsilon$. Note, here e is the vector of all possible $e_{j,t}$, and b is the vector of unrestricted dummy coefficients. The problem is that due to the linear relationship between time, age, and cohort, the matrix $(X'X)^{-1}$ is singular, so this unrestricted version can not be implemented.

Let ca be the number of age and cohort dummies available, and T be the number of time dummies available. Define the matrix R and W as indicated below. The matrix R has two rows corresponding to our two restrictions of setting the mean time dummies to zero and setting time dummies be orthogonal to trend. Let b^* be the corresponding vector of b that obeys this normalization. That is $Rb^* = \mathbf{0_{2 \times 1}}$. It turns out that in spite of the fact that $(X'X)^{-1}$ is singular the matrix W below is non-singular.

$$R = \begin{bmatrix} \mathbf{0}_{1 \times ca} & \mathbf{1}_{1 \times T} \\ \mathbf{0}_{1 \times ca} & 1, 2, ...T \end{bmatrix} \tag{1}$$

$$W = \begin{bmatrix} X'X & -R' \\ R & 0_{2 \times 2} \end{bmatrix}$$

Define the vectors V and s as indicated below, where λ is the Lagrange multiplier on the restricted least square residual (or moment conditions). Then the moment

31

conditions corresponding to this minimization is $V = Wd$. Since W is invertible we have $d = W^{-1}V$ and therefore have the solution to the desired restricted set of estimates b^*.
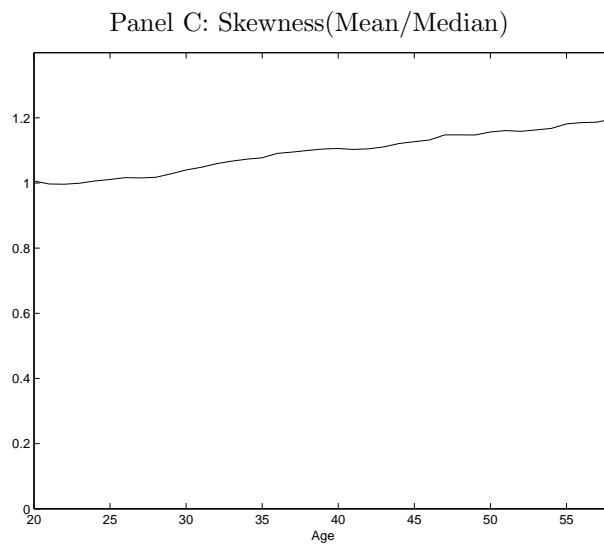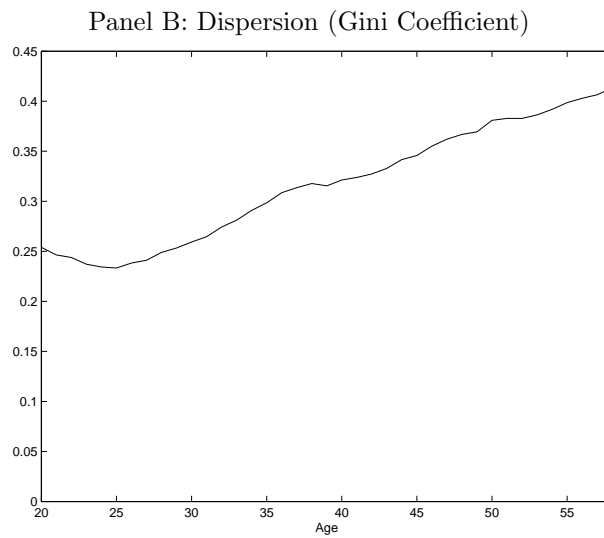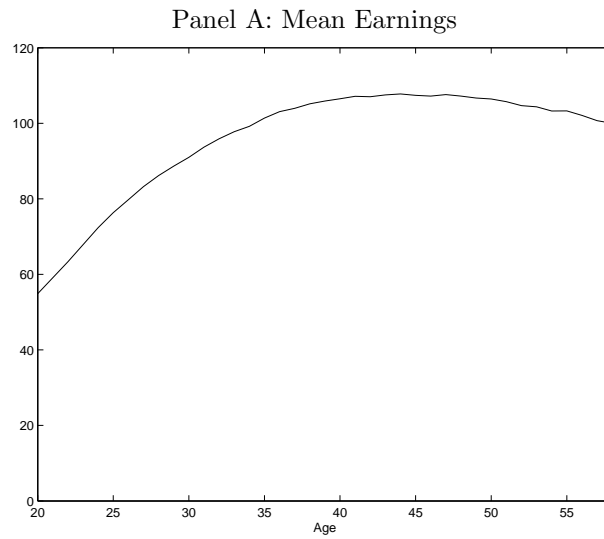
$$V = \begin{bmatrix} X'e \\ \mathbf{0_{2\times 1}} \end{bmatrix}$$
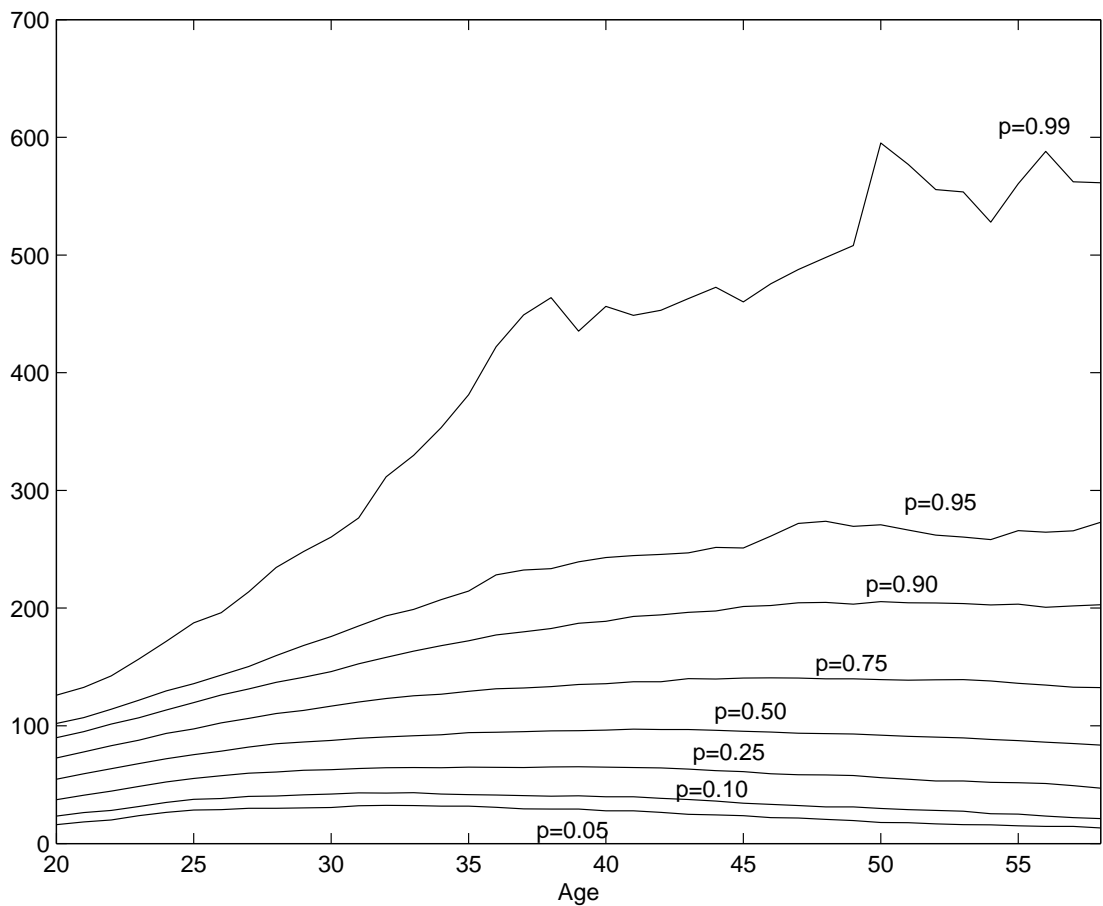
$$d = \begin{bmatrix} b^* \\ \lambda \end{bmatrix}$$

Figure 1 : Earnings Distribution Dynamics – PSID Data

This figure plots mean, dispersion, and skewness in earnings by age using PSID data. The age-profiles are based on the percentile estimation procedure described in section 2.2.

Panel A: Mean Earnings

Panel B: Dispersion (Gini Coefficient)

Panel C: Skewness(Mean/Median)

Figure 2 : Earnings Percentiles (0.05 to 0.99) – PSID Data

This figure plots the age percentiles of earnings using the methodology described in section 2.2. The line corresponding to the $p-th$ percentile shows the level of earnings such that p-percent of individuals earn below this level at each age. Earnings levels are normalized so that mean earnings at age 58 are 100. Although Figure 1 is based on 23 percentiles (see text), Figure 2 displays only the following 8 percentiles (0.05,0.10,0.25,0.50,0.75,0.90,0.95,0.99).
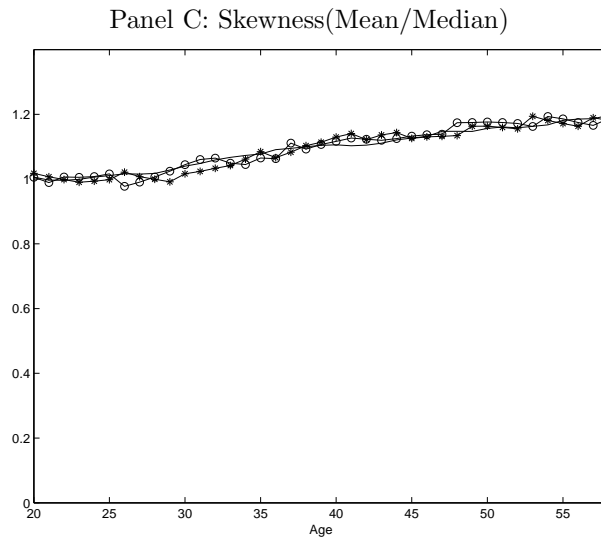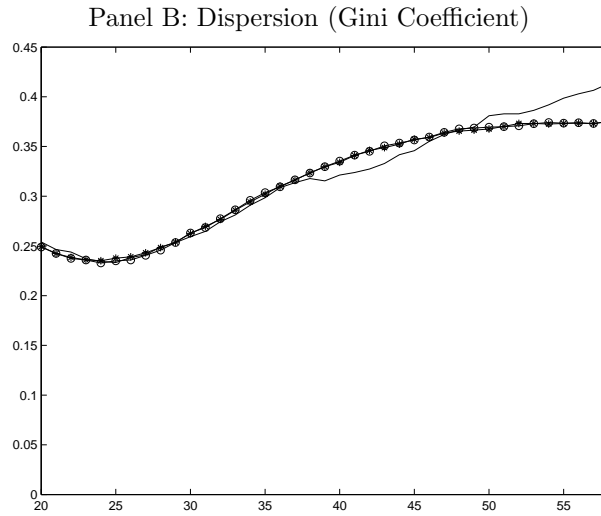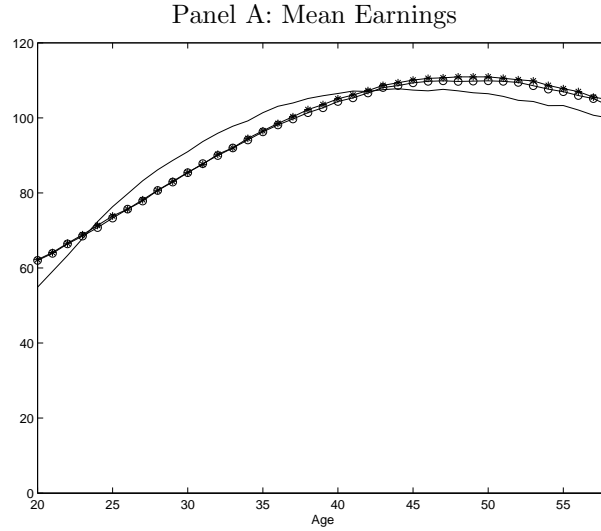
Figure 3 : Earnings Distribution Dynamics: Alternative Age Effects

This figure plots mean, dispersion, and skewness in earnings by age for PSID data using the estimated methods described in sections 2.2 and 2.3. Each figure displays three alternative ways to capture age effects. The line denoted by (−) corresponds to cohort dummies, the line denoted (∘) corresponds to time dummies, and the line with (⋆) corresponds to restricted time dummies. Note that the cohort dummies and restricted dummies are almost indistinguishable.
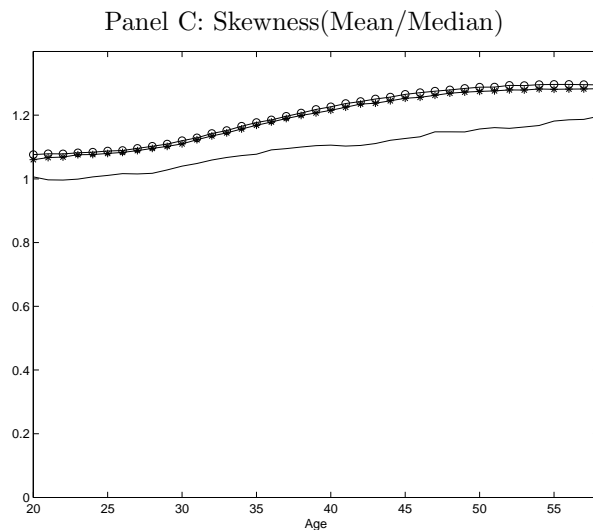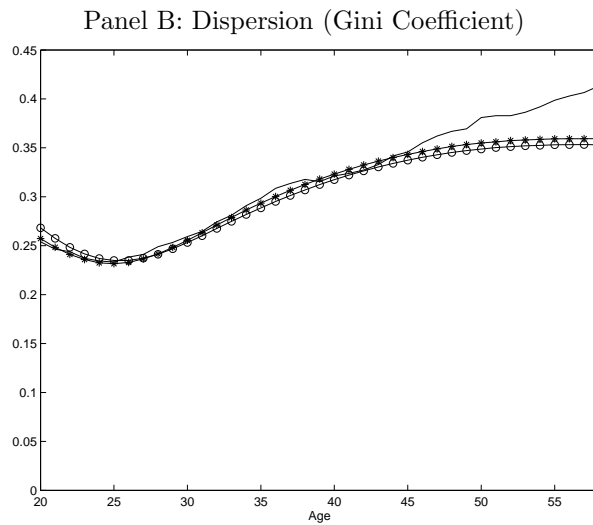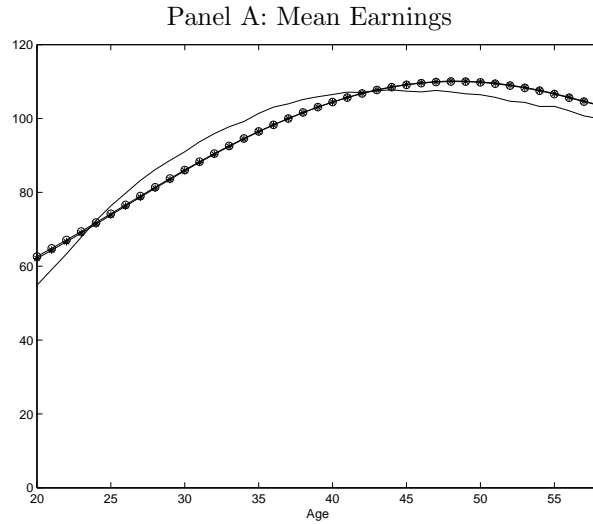
Panel A: Mean Earnings

Panel B: Dispersion (Gini Coefficient)

Panel C: Skewness(Mean/Median)
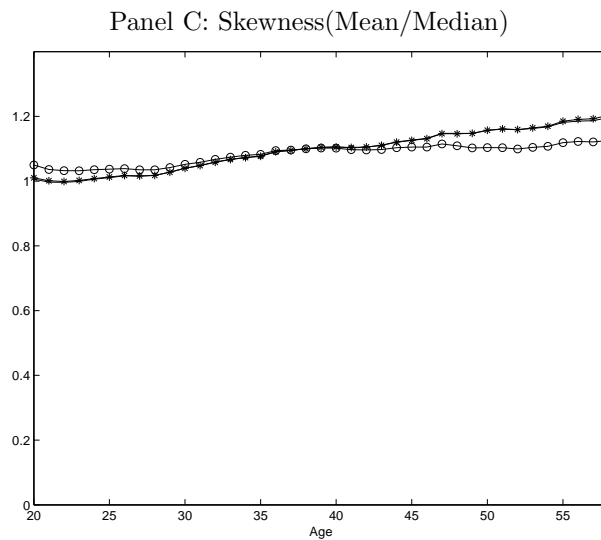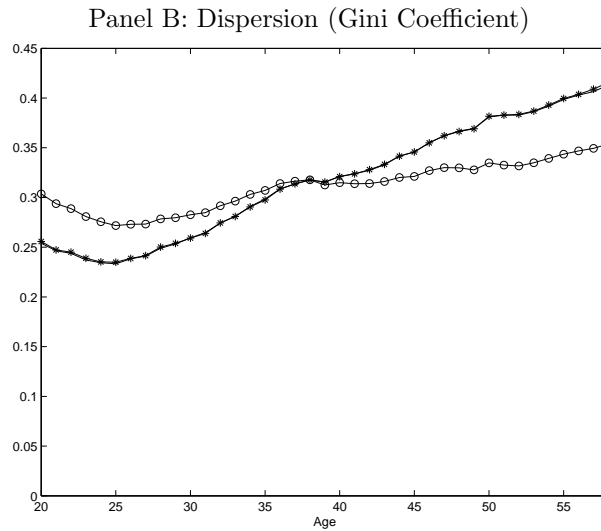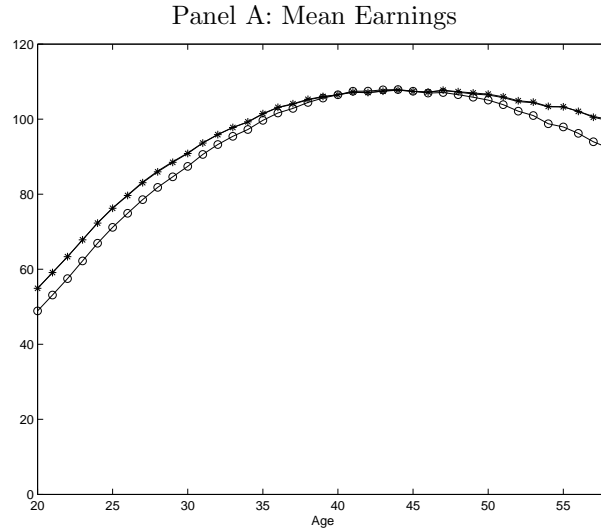
Figure 4 : Earnings Distribution Dynamics: Non-Parametric Case

The figures below plot the model implied mean, dispersion, and skewness in earnings by age. All panels are based on the non-parametric case for the distribution of initial human capital, h_1, and learning ability, a, when the curvature parameter, α, is 0.7. The symbol $(-)$ denotes the data, the symbol $(\star)$ denotes the model when accumulation starts at age 10, $(\circ)$ denotes the model when accumulation starts at age 20.

Panel A: Mean Earnings

Panel B: Dispersion (Gini Coefficient)

Panel C: Skewness(Mean/Median)

Figure 5 : Earnings Distribution Dynamics: Parametric Case

The figures below plot the model implied mean, dispersion, and skewness in earnings by age. All panels are based on the parametric case (bivariate log-normal) for the distribution of initial human capital, h_1, and learning ability a, when the curvature parameter, α is 0.7. The line $(-)$ denotes the data, the symbol $(\star)$ denotes the model when accumulation starts at age 10, $(\circ)$ denotes the model when accumulation starts at age 20.

Panel A: Mean Earnings

Panel B: Dispersion (Gini Coefficient)

Panel C: Skewness(Mean/Median)

37

Figure 6 : Earnings Distribution Dynamics: Fixed Initial Human Capital

The figures below plot the model implied mean, dispersion, and skewness in earnings by age. All panels are based on the non-parametric case for the distribution of learning ability a, and a common initial human capital level, h_1, when the curvature parameter, α, is 0.7. The symbol $(-)$ denotes the data, the symbol $(\circ)$ denotes the model when accumulation starts at age 10.

Panel A: Mean Earnings

Panel B: Dispersion (Gini Coefficient)

Panel C: Skewness(Mean/Median)

38