NBER WORKING PAPER SERIES

HOW LARGE IS THE STOCK COMPONENT OF HUMAN CAPITAL?

Mark Huggett
Greg Kaplan

Working Paper 21238
http://www.nber.org/papers/w21238

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2015

A previous version of this paper circulated under the title "The Money Value of a Man". The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

How Large is the Stock Component of Human Capital?
Mark Huggett and Greg Kaplan
NBER Working Paper No. 21238
June 2015
JEL No. D91,E21,G12,J24

ABSTRACT

This paper examines the value of an individual's human capital and the associated return on human capital using U.S. data on male earnings and financial asset returns. We find that (1) the value of human capital is far below the value implied by discounting earnings at the risk-free rate and (2) the stock component of the value of human capital is smaller than the bond component at all ages. The stock component averages less than 35 percent of the value of human capital at each age.

Mark Huggett
Georgetown University
mh5@georgetown.edu

Greg Kaplan
Department of Economics
Princeton University
Fisher Hall
Princeton, NJ 08544
and NBER
gkaplan@princeton.edu

1 Introduction

A common view is that by far the most valuable asset that most people own is their human capital. We provide a detailed characterization of the value and return to human capital, and their implications for portfolio choice over the lifecycle. We first estimate a statistical model for male earnings and stock returns to describe how earnings move with age, education and a rich structure of aggregate and idiosyncratic shocks. We then embed this statistical model into a decision problem of the type analyzed in the literature on the income-fluctuation problem. The properties of the implied human capital values are calculated by using the stochastic discount factor produced by a solution to this decision problem to value future earnings after taxes and transfers.

We highlight two main findings. First, the value of human capital is far below the value that would be implied by discounting net earnings at the risk-free interest rate. The most important reason for this is the large amount of idiosyncratic earnings risk that we estimate from U.S. data. An agent's stochastic discount factor covaries negatively with this component of earnings risk.

This finding is particularly relevant with respect to the view that various legal impediments, including personal bankruptcy laws, hinder greater skill investment and greater risk sharing in an individual's future earnings. The finding that individual human capital valuations are far below the value implied by discounting earnings at the risk-free rate suggests that individual valuations are well below market valuations. If so, then absent these impediments there is ample scope for alternative financial arrangements to arise to share some of this idiosyncratic risk.[1]

Our second main finding involves decomposing the value of human capital at each age into a stock, bond and orthogonal value component. We find that the stock component is typically below 35 percent of the value of human capital. This holds for two different educational groups (high school or college educated males) under a wide range of attitudes towards risk-aversion. We determine the stock share by projecting the sum of next period's earnings and human capital value onto next period's bond and stock returns. We then value these components using the individual's stochastic discount factor.

This finding is relevant for understanding optimal portfolio allocation. One view is that, in deciding how to allocate financial wealth between stock and bonds, the implicit holdings of stock and risk-free bonds in human wealth are important. This seems particularly relevant if, as we find, the value of human wealth far exceeds the value of financial wealth for most individuals. Thus, a first-order issue for the portfolio allocation literature is to decompose the value of human capital into bond and stock

[1] Krebs, Kuhn and Wright (2013) argue that young, multi-member households are underinsured against mortality risk - one very specific source of human capital risk.

components and determine the features that drive the magnitude of the stock component.

To the best of our knowledge, Benzoni, Collin-Dufresne and Goldstein (2007) is the only paper besides ours that defines the value of human capital as we do and then decomposes this value into components. On the basis of a rough calibration of a joint earnings-stock-returns process, they conclude that the stock component of human capital is 50 percent of the value of human capital at age 20 and remains at 50 percent for the first half of the working lifetime. If correct, then one explanation for why some individuals hold little or no stock in their financial asset portfolios is that they already hold a large implicit position in stock, and risk preferences dictate that stock holdings in overall wealth is less than 50 percent.

We find that the stock component of the value of human capital is positive, but is typically far below the 50 percent value highlighted by Benzoni et al. (2007). A number of model features lead to a positive stock component. For example, social security retirement benefits that are positively linked to the level of average earnings, a left-skewed distribution of idiosyncratic, earnings shocks and a positive conditional correlation between stock returns and the aggregate component of individual earnings all contribute towards a positive stock component. They also have support in US data.

We do not find much support for the claim that cointegration between the aggregate component of earnings and stock returns is key to producing a large stock component, at least when such a relationship is estimated using US data. Cointegration is potentially important as then shock histories with large stock returns will tend to be associated with positive shocks to average earnings. Earnings far in the future will then take on stock-like features. Benzoni et al. (2007) calculate the value of human capital after roughly calibrating such a cointegrated process. In contrast to Benzoni et al. (2007), all our work is based on estimating the relationship between earnings and stock returns.

Our work is most closely related to two literatures. First, a long line of work values human capital by discounting the future earnings stream using a deterministic interest rate or discount factor.[2] Our work differs as discounting is done using an individual's stochastic discount factor, which produces an individual-specific value of human capital. Huggett and Kaplan (2011) is more closely related. They put bounds on individual human capital values using knowledge of the earnings and asset returns process and Euler equation restrictions. Second, there is a vast literature on financial asset allocation decisions over the lifetime. Coco, Gomes and Maenhout (2005), Benzoni et al. (2007), Lynch and Tan (2011), and many others focus on quantitative properties of portfolio decisions. We focus on

[2] See Farr (1853), Dublin and Lotka (1930), Weisbrod (1961), Becker (1975), Graham and Webb (1979), Jorgenson and Fraumeni (1989), Haveman, Bershadker and Schwabish (2003). Some of this work calculates an aggregate value of human capital.

decomposing the value of human capital based on a joint earnings-stock-returns process estimated from micro data.

The remainder of the paper is organized as follows. Section 2 presents the theoretical framework. Section 3 to 5 present our main findings. Section 6 explores the robustness and the key drivers of these findings. Section 7 concludes.

2 Theoretical Framework

This section presents the framework, defines the value of human capital and illustrates the value and return concepts with a simple example.

2.1 Decision Problem

An agent solves Problem P1. Lifetime utility $U(c)$ is determined by a consumption plan $c = (c_1, ..., c_J)$. Consumption at age j is given by a function $c_j : Z^j \rightarrow R^1_+$ that maps shock histories $z^j = (z_1, ..., z_j) \in Z^j$ into consumption. All the variables that we analyze are functions of these shocks.

Problem P1: $\max U(c)$ subject to

(1) $c_j + \sum_{i \in \mathcal{I}} a^i_{j+1} = \sum_{i \in \mathcal{I}} a^i_j R^i_j + e_j$ and $c_j \geq 0, \forall j$

(3) $a^i_{J+1} = 0, \forall i \in \mathcal{I}$

The budget constraint says that period resources are divided between consumption c_j and savings $\sum_{i \in \mathcal{I}} a^i_{j+1}$. Period resources are determined by an exogenous earnings process e_j and by the value of financial assets brought into the period $\sum_{i \in \mathcal{I}} a^i_j R^i_j$. The value of financial assets is determined by the amount a^i_j of savings allocated to each financial asset $i \in \mathcal{I} = \{1, ..., I\}$ and by the gross return $R^i_j > 0$ to each asset i.

2.2 Value and Return Concepts

The value of human capital v_j is defined to equal expected discounted dividends (i.e. net earnings) at a solution $(c^*, a^*) = ((c^*_1, ..., c^*_J), \{(a^{*,i}_1, ..., a^{*,i}_{J+1})\}_{i \in \mathcal{I}})$ to Problem P1. Discounting is done using the agent's stochastic discount factor from the solution to Problem P1. The stochastic discount factor $m_{j,k}$ reflects the agent's marginal valuation of an extra period k consumption good in terms of the period j consumption good. The stochastic discount factor has a conditional probability term $P(z^k|z^j)$

because human capital values are stated using the mathematical expectations operator E.[3]

$$v_j(z^j) \equiv E[\sum_{k=j+1}^{J} m_{j,k}e_k|z^j] \;\; and \;\; m_{j,k}(z^k) \equiv \frac{\partial U(c^*)/\partial c_k(z^k)}{\partial U(c^*)/\partial c_j(z^j)} \frac{1}{P(z^k|z^j)}$$

Given the value concept, we define the gross return R^h_{j+1} to human capital to be next period's value and dividend divided by this period's value: $R^h_{j+1} = \frac{v_{j+1} \, + \, e_{j+1}}{v_j}$. The return to human capital is then well integrated into standard asset pricing theory. Off corners, all returns R_{j+1} satisfy the same type of restriction: $E[m_{j,j+1}R_{j+1}|z^j] = 1$.[4]

2.3 An Interpretation

We now provide an interpretation for v_j. The value v_j is a personalized price to an agent for the non-traded earnings stream that the agent owns. The price v_j is the value of all the shares (total shares are normalized to 1) in the future earnings stream. The price process $\{v_j\}_{j=1}^J$ has the property that if the agent were allowed to change share holdings in this earnings stream at any age at these prices, then the agent would optimally decide not to change share holdings and would make exactly the same consumption and asset choices (c^*, a^*) that were optimal in Problem P1.[5]

2.4 A Simple Example

A simple example illustrates the value and return concepts. An agent's preferences are given by a constant relative risk aversion utility function. Earnings follow an exogenous Markov process. There is a single, risk-free financial asset.[6]

Utility: $U(c) = E[\sum_{j=1}^J \beta^{j-1} u(c_j)|z^1]$, where $u(c_j) = \begin{cases} \frac{c_j^{1-\rho}}{(1-\rho)} & : \quad \rho > 0, \rho \neq 1 \\ \log(c_j) & : \quad \rho = 1 \end{cases}$

Earnings: $e_j = \prod_{k=1}^j z_k$, where $\ln z_k \sim N(\mu, \sigma^2)$ is i.i.d.

[3]The expectations operator *integrates* the relevant age k functions with respect to the distribution $P(z^k|z^j)$. In all of our applied work, the set of partial shock histories is finite and, thus, integration is straightforward summation.

[4]This holds for the return to human capita R^h_{j+1} by construction because $v_j = E[\sum_{k=j+1}^J m_{j,k}e_k|z^j]$ implies $E[m_{j,j+1}(\frac{v_{j+1}+e_{j+1}}{v_j})|z^j] = 1$.

[5]Broadly, the pricing of human capital generalizes the method of pricing a non-traded asset in Lucas (1978). One proposes a second economy where trade in the non-traded asset is allowed and then finds prices that persuade the agent not to do so. The result claimed in the text holds for general utility functions under a concavity asssumption and is not sensitive to the nature of the earnings process or the financial asset returns process. It extends to economies with valued leisure and endogenous earnings. See Huggett and Kaplan (2012, Theorem 1) for a proof. It also holds under a variety of borrowing constraints on financial asset holdings.

[6]The model is a finite-lifetime version of the permanent-shock model analyzed by Constantinides and Duffie (1996).

Decision Problem: $\max U(c)$ subject to

$$(1)\ c_j + a_{j+1} \leq a_j(1+r) + e_j, (2)\ c_j \geq 0, a_{J+1} \geq 0$$

When $1+r = \frac{1}{\beta}\exp(\rho\mu - \frac{\rho^2\sigma^2}{2})$ and initial financial assets are zero, then setting consumption equal to earnings each period is optimal. The stochastic discount factor equals $m_{j,k}(z^k) \equiv \frac{\partial U(c^*)/\partial c_k(z^k)}{\partial U(c^*)/\partial c_j(z^j)} \frac{1}{P(z^k|z^j)} = \frac{\beta^{k-j} u'(e_k(z^k))}{u'(e_j(z^j))}$. This example leads to a closed-form formula where v_j is proportional to earnings e_j and where R_j^h is a time-invariant function of the period shock z_j.

Figure 1 illustrates some quantitative properties. The parameter σ, governing the standard deviation of earnings shocks, varies over the interval $[0, 0.3]$ and $\mu = -\sigma^2/2$. As all agents start with earnings equal to 1, the expected earnings profile over the lifetime is flat and equals 1 in all periods. The lifetime is $J = 46$ periods which can be viewed as covering real-life ages $20 - 65$. The interest rate is fixed at $r = .01$. Thus, the discount factor β is adjusted to be consistent with this interest rate given the remaining parameters: $1 + r = \frac{1}{\beta}\exp(\rho\mu - \frac{\rho^2\sigma^2}{2})$.

Figure 1 shows that the value v_1 of an age 1 agent's human capital falls and that the mean return in any period rises as the shock standard deviation increases. Thus, a high mean return on human capital is the flip side of a low value attached to human capital. These patterns are amplified as the preference parameter ρ increases.

Figure 1 also plots the "naive value" of human capital. The naive value equals earnings discounted at a constant interest rate r that we set equal to the risk-free rate (i.e. $v_1^{naive} = E[\sum_{j=2}^{J} \frac{e_j}{(1+r)^{j-1}}|z^1])$. This follows a traditional empirical procedure that is widely employed in the literature as was mentioned in the introduction. The naive value is exactly the same in each economy in Figure 1 because the risk-free interest rate and the mean earnings profile are unchanged across economies. Our notion of value v_1 differs from v_1^{naive} because the agent's stochastic discount factor is allowed to covary with earnings. Figure 1 shows that negative covariation can be substantial.

Figure 1 plots the total benefit and the marginal benefit of moving from the model consumption plan c to a smooth consumption plan where $c_j^{smooth} = E_1[c_j] = E_1[e_j] = 1$. The benefit function Ω is defined, following Alvarez and Jermann (2004), by the first equation below. The total benefit is $\Omega(1)$ and the marginal benefit is $\Omega'(0)$. The marginal benefit in Figure 1 increases as the standard deviation of the period earnings shock increases.

$$U((1 + \Omega(\alpha))c) = U((1 - \alpha)c + \alpha c^{smooth})$$

$$\Omega'(0) = \frac{\sum_{j=1}^{J}\sum_{z^j}\frac{\partial U(c)}{\partial c_j(z^j)}(c_j^{smooth}(z^j) - c_j(z^j))}{\sum_{j=1}^{J}\sum_{z^j}\frac{\partial U(c)}{\partial c_j(z^j)}c_j(z^j)} = \frac{E[\sum_{j=1}^{J} m_{1,j}c_j^{smooth}|z^1]}{v_1(z^1) + e_1(z^1) + a_1(z^1)(1+r)} - 1$$

The marginal benefit is tightly connected to the value v_1 of human capital. To see this, differentiate the first equation above with respect to α. This implies the leftmost equality in the second equation above. The rightmost equality holds by rearrangement because the individual solves Problem P1.[7] The numerator term in the second equation is pinned down by asset prices so that $E[\sum_{j=1}^{J} m_{1,j}c_j^{smooth}|z^1] = \sum_{j=1}^{J}(\frac{1}{1+r})^{j-1}$, whereas the denominator is determined by the value of human capital plus initial earnings and initial wealth. The only unobservable is the value of human capital. The theory then implies that a high marginal benefit of moving towards perfect consumption smoothing coincides with a low value of human capital. This straightforward point has not, to the best of our knowledge, been noted in the literature on the value of human capital.

3 Empirics: Earnings and Asset Returns

We outline an empirical framework for idiosyncratic earnings shocks, aggregate earnings shocks and stochastic stock returns. Let $e_{i,j,t}$ denote real pretax annual earnings for individual i of age j in year t. We assume that the natural logarithm of earnings consists of an aggregate component $\left(u^1\right)$ and an idiosyncratic component $\left(u^2\right)$ and

$$\log e_{i,j,t} = u_t^1 + u_{i,j,t}^2. \tag{1}$$

In Section 3.1 we describe the structure of the idiosyncratic component of earnings, our estimation procedure and the fit of the estimated model. In Section 3.2 we describe the structure and estimation of the joint process for the aggregate component of earnings and stock returns.

3.1 Idiosyncratic Component of Earnings

The idiosyncratic component of earnings is the sum of four orthogonal components: a common age effect κ_j, an individual-specific fixed effect ξ, a persistent component ζ and a transitory component υ.

$$\begin{aligned} u_{i,j,t}^2 &= \kappa_j + \xi_i + \zeta_{i,j,t} + \upsilon_{i,j,t} \qquad (2)\\ \zeta_{i,j,t} &= \rho\zeta_{i,j-1,t-1} + \eta_{i,j,t} \\ \zeta_{i,0,t} &= 0. \end{aligned}$$

[7] More specifically, convert the period budget constraints in Problem P1 into an age-1 budget constraint, using the fact that the Euler equation holds at a solution to Problem P1. Then the age-1 value of the consumption plan equals the value of human capital, earnings at age 1 and initial wealth.

The common age effect is modeled as a quartic polynomial. The individual fixed effects are assumed to be normally distributed with a constant variance σ_{ξ}^2. The transitory idiosyncratic shocks are assumed to be distributed according to a distribution with zero mean, variance $\sigma_{\upsilon,j}^2$, and third central moment $\mu_{3,\upsilon,j}$. In order to capture life-cycle properties of the variance and skewness of earnings we allow the moments of the transitory component to be age-dependent and model this as a quartic polynomial.

Persistent idiosyncratic shocks are assumed to be distributed according to a distribution with zero mean, variance $\sigma_{\eta,t}^2(X_t)$ and third central moment $\mu_{3,\eta,t}(X_t)$. The variance and skewness have a linear trend, in order to capture low frequency trends over the sample period, and are state dependent via the variable X_t.[8] We model X_t as a two-state process. Specifically, we set $X_t = 1_{\Delta u_t^1 > 0}$ so that X_t is an indicator function taking on the value 1 in booms and 0 in recessions. Thus, aside from a trend term, the variance and skewness of the persistent innovations take on different values in expansions and contractions.

We estimate the idiosyncratic earnings process using data on male annual labor earnings from the Panel Study of Income Dynamics (PSID) from 1967 to 1996.[9] We focus on male heads of households between ages 22 and 60 with real annual earnings of at least $\$1,000$. Our measure of annual gross labor earnings includes pre-tax wages and salaries from all jobs, plus commission, tips, bonuses and overtime, as well as the labor part of income from self-employment. Labor earnings are inflated to 2008 dollars using the CPI All Urban series. We also consider two sub-samples. Individuals with 12 or fewer years of education are included in the High School sub-sample, while those with at least 16 years or a Bachelor's degree are included in the College sub-sample.

Estimation is done in two stages. In the first stage we estimate κ_j by regressing log real annual earnings on a quartic polynomial in age and a full set of year dummies. This is done separately for the three education samples. On the basis of the first-stage results for the PSID, and related results for the Current Population Survey data and NIPA data, we set the contraction years over the time interval 1967-1996 to be 1970, 1974-5, 1979-82, 1989-91 and 1993.

Residuals from this first-stage regression are then used to estimate the remaining parameters of the individual earnings equation: $\left(\rho, \sigma_{\xi}^2, \sigma_{\eta,t}^2(X_t), \mu_{3,\eta,t}(X_t), \sigma_{\nu,j}^2, \mu_{3,\nu,j}\right)$. We compute the auto-covariance function for residual log-earnings up to 10 lags for every age/year combination, as well as the third central moments and third-order auto-covariances. A GMM estimator is then used to estimate the

[8] Allowing for a trend in the shock variances is important for accurately estimating cyclical variation in the variance and skewness. This is because of the well-documented increase in the variance of idiosyncratic earnings shocks over this period. See for example Heathcote, Perri and Violante (2010).

[9] After 1996 the PSID was converted into a biannual survey.

parameters, using the full set of second and third-order autocovariances as moments.[10]

The estimated process delivers a good fit to the variance and third central moment of the earnings distribution as a function of both age and time. The fit of these and other moments for the full sample is displayed in Figure 2. Corresponding results for the College and High School samples are contained in the Appendix.

We highlight three findings from Table 1. First, transitory shocks are left skewed for the full sample and the college sample. Left skewness is needed to match the left skewness of the first-stage residuals as documented in Figure 2. Guvenen, Ozkan and Song (2014) document that male earnings growth rates are left skewed in administrative data. Second, the variance and left skewness of persistent shocks is higher in recessions than in booms. Thus, consistent with the findings in Storesletten, Telmer and Yaron (2004), there is evidence for counter-cyclical risk even when the framework is generalized to account for skewness and a time trend. However, the cyclical variation in risk that we estimate is less dramatic than their findings. Third, the autoregression parameter ρ is higher for the full sample and the college sample compared to the high school sample. Thus, persistent innovations of a given magnitude will be of greater proportional importance for those with a college than a high school education.

The parameter estimates are broadly consistent with those from related specifications (that do not account for skewness), that have been estimated elsewhere in the literature and summarized in Meghir and Pistaferri (2010).[11] We note that our estimate of the variance of the transitory component is approximately 0.1 larger than what has been estimated by others (see for example, Guvenen (2009)). The source of this difference is due entirely to our broader sample selection. Since it is likely that a substantial fraction of this variance is due to measurement error, we make an adjustment when using these estimates as parameters in the structural model.[12]

[10] When computing the auto covariance function, individuals are grouped into 5-year age cells so that when calculating covariances at age j, individuals aged $j \in [j-2, j+2]$ are used. Only cells with at least 30 observations are retained. The moments are weighted by $n_{j,t,l}^{0.5}$ where $n_{j,t,l}$ is the number of observations used to calculate the covariance at lag l in year t for age j. Individuals aged 22 to 60 are used to construct the empirical auto-covariance functions. This means that variances, covariance and third moments from ages 24 to 58 are effectively used in the estimation. Standard errors are calculated by bootstrap with 39 repetitions, thus accounting for estimation error induced by the first-stage estimation.

[11] Our model is estimated using data on log earnings levels. Estimation using data on log earnings growth rates would yield larger estimates of the persistent or permanent shocks. See Heathcote et al. (2010) for a discussion of this issue. We favor the estimation in levels since it allows us to accurately capture the age profile of the cross-sectional variance of earnings.

[12] Using indirect inference on a structural model of consumption and savings behavior, Guvenen and Smith (2011) estimate that the variance of measurement error in male log annual earnings is around 0.02-0.025. Using a validation study of the PSID, French (2004) concludes that the variance of measurement error in the PSID is around 0.01. However, both of their samples are substantially more selected than ours, with a cross-sectional variance about 0.1 lower. Assuming that half of this additional variance is due to measurement error, would suggest that around 0.05-0.06 of the estimated transitory variance is measurement error. Accordingly, we adjust our estimates of the variance $\sigma_{\nu,j}^2$ down by 1/3 at all

3.2 Aggregate Component of Earnings and Stock Returns

The joint process for the dynamics of the aggregate component of earnings and stock returns is modeled as follows. Let $y_t = \begin{pmatrix} u_t^1 & P_t \end{pmatrix}'$, where P_t is an underlying process that generates risky returns. Gross returns on stock R_t^s satisfy $\log R_t^s = \Delta P_t$. We do not assume from the outset that the vector y_t is a stationary process. Rather, we allow for y_t to be a first order integrated process and write it as a Vector Error Correction Model (VECM), a form that is common in the literature on cointegration:[13]

$$\begin{pmatrix} \Delta y_t \\ w_t \end{pmatrix} = \begin{pmatrix} \gamma \\ \beta'\gamma + \rho \end{pmatrix} + \begin{pmatrix} \Gamma & \alpha \\ \beta'\Gamma & 1 + \beta'\alpha \end{pmatrix} \begin{pmatrix} \Delta y_{t-1} \\ w_{t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_t \\ \beta'\varepsilon_t \end{pmatrix} \tag{3}$$

Equation (3) expresses the dynamics of the aggregate states as a coupled system of equations between the growth rates of the aggregate variables y_t and a cointegrating vector w_t defined by $w_t \equiv \beta' y_t + \mu + \rho(t+1)$. To understand this process, first note that when $\alpha = 0$, there is no cointegration and the process collapses to a standard first-order VAR for Δy_t:

$$\Delta y_t = \gamma + \Gamma \Delta y_{t-1} + \varepsilon_t \tag{4}$$

When $\alpha = 0$ the growth rate of log aggregate earnings and stock returns follow a first-order VAR. Mean aggregate earnings growth and mean stock returns are captured by the constant term γ. When $\alpha \neq 0$, the process allows for cointegration. The pattern of cointegration is described by β and a linear time trend ρ in the level of the cointegrating vector w_t. The strength of the cointegration, i.e. how sensitive Δy_t is to w_t, is reflected by α. The error term ε_t is a vector of zero mean IID random variables with covariance matrix Σ.

We estimate (3) and (4) using data on male annual labor earnings from the the Current Population Survey (CPS) from 1967 to 2008. Our sample selection criteria and definition of earnings are the same as those used for the PSID, descibed in section 3.1. We constructe an empirical counterpart to u_t^1 by estimating a median regression (Least Absolute Deviations) of earnings on a full set of age and time dummies. We use median regression rather than OLS since it is more robust to the effects of changes in top coding in the CPS over our sample period. We use the estimates of $\hat{u}_t^1$ from our CPS sample, rather than corresponding estimates from the PSID, as input to the estimation because CPS data has both a longer time dimension and a larger cross-section sample each year compared to PSID data.

ages. We adjust the third central moment at each age so that the coefficient of skewness, $\mu_{3,\nu,j}/\sigma_{\nu,j}^3$ remains unchanged after adjusting the demoninator.

[13] We have used symbols in the specification of the VAR process that are consistent with notation that is common in the literature on cointegration. The parameters α, β, γ and ρ should not be confused with the preference parameters that use the same symbols.

Data on equity and bond returns are annual returns. Equity returns are based on a value-weighted portfolio of all NYSE, AMEX and NASDAQ stocks including dividends. Bond returns are based on 1-month Treasury bill returns.[14] Real returns are calculated by adjusting for realized inflation using the CPI All Urban series.

Table 2 reports results of estimation of (3) and (4). The parameter estimates reveal a moderate degree of persistence in aggregate earnings growth. There is a positive correlation between innovations to earnings growth and innovations to stock returns in all models. This implies that the conditional correlation between earnings growth and stock returns is positive. This is one feature, among others, that will later produce a positive conditional correlation between stock and human capital returns.

The implied steady-state dynamics are reported in Table 3. The estimated model matches the observed correlation structure well. When we input the estimated process into our economic model, we adjust the constants (γ_1, γ_2) estimated in Table 2 so that all models produce in steady state $E[\log R^s] = .041$ and $E[\Delta u^1] = 0$. This facilitates comparisons of human capital value and return properties across models.

In the Appendix we show how the VECM model in (3) can be derived from an underlying VAR(2) for y_t. We present results from standard lag-order selection tests for the order of a more general VAR(p) process and find that for our baseline sample, as well as for college and high-school subsamples, and for alternative measures of aggregate earnings and alternative sample periods, virtually all specifications indicate the presence of two lags, i.e. $p = 2$. This is what leads us to adopt the model in (3). The Appendix also contains tests of the cointegrating rank of (3) based on the methods in Johansen (1995). For all three education samples, tests based on the trace statistic, the maximum eigenvalue or the Schwarz-Bayes information criterion, suggest a cointegrating rank of zero, while the Hannan-Quinn information criterion suggest the presence of one cointegrating vector. We interpret these findings as providing only very weak evidence for cointegration. Hence, we adopt the model without cointegration in (4) as our benchmark specification. However, since these tests may all have relatively little power given the short annual sample period, we present estimates for the model with cointegration and later assess the implications.

[14] All returns come from Kenneth French's data archive. We restrict attention to the period 1967-2008 since this is the period that the CPS earnings data covers.

4 The Benchmark Model

We now use the theoretical framework and the empirical results to quantify the value and return to human capital. The benchmark model has two financial assets. Asset $i = 1$ is riskless and asset $i = 2$ is risky. The agent cannot go short on either financial asset.

Benchmark Model: $\max U(c)$ subject to $c \in \Gamma_1(x, z^1)$

$$\Gamma_1(x, z^1) = \{c = (c_1, ..., c_J) : \exists (a^1, a^2) \ s.t. \ 1 - 2 \ holds \ \forall j\}$$

1. $c_j + \sum_{i \in \mathcal{I}} a^i_{j+1} \leq x$ for $j = 1$ and $c_j + \sum_{i \in \mathcal{I}} a^i_{j+1} \leq \sum_{i \in \mathcal{I}} a^i_j R^i_j + e_j$ for $j > 1$
2. $c_j \geq 0$ and $a^1_{j+1}, a^2_{j+1} \geq 0, \forall j$

The utility function $U(c) = U^1(c_1, ..., c_J)$ is of the type employed by Epstein and Zin (1991). It is defined recursively by applying an aggregator W and a certainty equivalent F. The certainty equivalent encodes attitudes towards risk with α governing risk-aversion. The aggregator encodes attitudes towards intertemporal substitution where ρ is the inverse of the intertemporal elasticity of substitution. We allow for mortality risk via the one-period-ahead survival probability ψ_{j+1}.

$$U^j(c_j, ...c_J) = W(c_j, F(U^{j+1}(c_{j+1}, ..., c_J)), j)$$

$$W(a, b, j) = [(1 - \beta)a^{1-\gamma} + \beta\psi_{j+1}b^{1-\gamma}]^{1/(1-\gamma)} \ and \ F(x) = (E[x^{1-\alpha}])^{1/(1-\alpha)}$$

Table 4 summarizes the parameters in the benchmark model.

Demographics: Agents start economic life at real-life age 22, retire at age 61 and live at most up to age 90. Thus, we set $J = 69$ and $Ret = 40$. Agents face a conditional probability ψ_{j+1} of surviving from period j to period $j + 1$ that is set to estimates for males from the 1989-91 US Decennial Life Tables in NCHS (1992).

Preferences: We set the preference parameters to values estimated from Euler equation restrictions. Vissing-Jorgensen and Attanasio (2003) estimate $1/\gamma = 1.17$ for a prefered specification and conclude that the risk aversion parameter α in the interval $[5, 10]$ can be obtained under realistic assumptions, based on household-level data. Thus, the special case of constant-relative-risk-aversion preferences, where $\gamma = \alpha$, is not the parameter configuration that best fits the Euler equation restrictions. We examine model implications for $1/\gamma = 1.17$ and $\alpha \in \{4, 6, 8, 10, 12\}$. We set the discount factor β so that given all other model parameters the model produces a steady-state wealth to income ratio of 3.5.

Initial Wealth: Initial wealth is set to equal 30 percent of mean earnings at age 22.

Earnings and Asset Returns: Earnings and asset returns in the benchmark model are based on the estimates in Tables 1-3 for the case of no cointegration. The earnings that enter the model are earnings after taxes and transfers. Before the retirement age, model earnings $e_j(z)$ are the process estimated in Tables 1-3 times $(1-\tau)$, where $\tau = .27$. After the retirement age, model earnings $e_j(z)$ equal model social security benefits after taxes.

$$e_j(z) = z_1 g_j(z_2) = \begin{cases} \exp(u^1)\exp(\kappa_j + \xi + \zeta + \upsilon)(1-\tau) & j < Retire \\ \exp(u^1) b(\xi)(1-\tau) & otherwise \end{cases}$$

We group the variables from the statistical model into a state variable $z = (z_1, z_2)$, where $z_1 = \exp(u^1)$ captures the aggregate component of earnings and $z_2 = (\xi, \zeta, \upsilon, \Delta u^1, \log R^s)$ captures the idiosyncratic components (ξ, ζ, υ), the growth in the aggregate component of earnings and the stock return.[15]

The fixed effect ξ is normally distributed with the variance given in Table 1. The transitory shock υ and the persistent shock innovations η follow a Generalized Normal distribution, determined by the first three central moments. The second and third central moments of the persistent shock innovations are state dependent as described in Table 1. The age-dependent second and third moments of the transitory shock distribution are scaled as discussed in footnote 12. See Hosking and Wallis (1997, Appendix A.8) for a discussion of the Generalized Normal distribution.

Social Security: The nature of social security benefits is potentially of great importance for how people value future earnings flows after taxes and transfers. Social security wealth is by some calculations the single most important asset type for many older households.[16] Social security benefits in the model are an annuity payment which is determined by the aggregate earnings level z_1 when the agent reaches the retirement age and by a concave benefit function b. We adopt the benefit function employed by Huggett and Parra (2010) which captures the bend-point structure of old-age benefits in the U.S. social security system. We employ the computationally-useful assumption that the benefit function applies only to an agent's idiosyncratic fixed effect rather than to an average of the agent's

[15] The Appendix discusses how we compute model solutions. The state variable in the model with cointegration is $z = (z_1, z_2)$, where $z_2 = (\xi, \zeta, \upsilon, \Delta u^1, \log R^s, w)$.

[16] Poterba, Venti and Wise (2011) calculate that the capitalized wealth implicit in social security retirement annuities is approximately 33 percent of all wealth for households aged 65-69 and is a much larger percentage of individual wealth for households with low wealth.

past earnings as in the U.S. system. Thus, the model benefit is risky after entering the labor market only because the aggregate component of earnings at the time of retirement is risky. Old-age benefit payments in the U.S. system are indexed to average economy-wide earnings when an individual hits age 60.[17] This is captured within the model by the fact that benefits are proportional to $z_1 \equiv \exp(u^1)$.

Properties of the benchmark model are displayed in Figure 3, which is constructed by simulating many shock histories, calculating allocations along these histories and then taking averages at each age. Figure 3 shows that mean consumption and mean earnings net of taxes and transfers are hump shaped over the lifetime. The college profile is much steeper than the high school earnings profile. One consequence of this is that a larger fraction of young college agents will hold exactly zero financial assets early in life compared to high school agents. This has implications for how strongly college agents discount future earnings early in life.

5 Human Capital Values and Returns: Benchmark Model

We report properties of the benchmark model based on the high school and the college subsamples. Results for the full sample are typically between the results for these education groups.

5.1 Human Capital Values

Figure 4 plots the value of human capital in the benchmark model and a decomposition of this value. The mean value of human capital over the lifetime is hump shaped and is lower for higher values of the risk aversion coefficient.[18] For comparison purposes, we also plot the value of human capital that would be implied by discounting future earnings at the risk-free rate. We label this the naive value.

Our notion of value lies far below the naive value. This occurs because of negative covariation between an agent's stochastic discount factor and earnings and because agents are sometimes on the corner of the risk-free asset choice. Corner solutions occur more frequently early in life for college agents than for high school agents due to differences in the mean earnings profile. When an agent is on the corner, then the agent discounts certain future earnings at more than the risk-free rate.

We now decompose human capital values into a bond, stock and a residual-value component. To do so, we project next period's human capital payout $v_{j+1} + e_{j+1}$ onto the space of conditional asset returns. The decomposition is carried out in the two equations below. The human capital payout

[17] See the Social Security Handbook (2012, Ch. 7).

[18] Figure 3(a)-(b) are constructed by first computing human capital values at each age as a function of the state. We then calculate the sample average of the value at each age, conditional on survival, by simulating many realizations of the state variable over the lifetime. Computational methods are described in the Appendix.

contains a component ($\sum_{i \in \mathcal{I}} \alpha^i_j R^i_{j+1}$) spanned by asset returns and a component (ϵ_{j+1}) orthogonal to asset returns, where α^i_j are the projection coefficients. When the agent is off corners in the holding of asset i then the Euler equation $E[m_{j,j+1} R^i_{j+1} | z^j] = 1$ holds.[19]

$$v_j = E[m_{j,j+1}(v_{j+1} + e_{j+1}) | z^j] = E[m_{j,j+1}(\sum_{i \in \mathcal{I}} \alpha^i_j R^i_{j+1} + \epsilon_{j+1}) | z^j]$$

$$v_j = \sum_{i \in \mathcal{I}} \alpha^i_j E[m_{j,j+1} R^i_{j+1} | z^j] + E[m_{j,j+1} \epsilon_{j+1} | z^j]$$

Figure 4 calculates the value of the bond, stock and orthogonal components as a fraction of the value of human capital at each age and then averages these ratios across the states that occur at each age. The bond component averages more than 80 percent of the value of human capital. This holds for both education groups and for a range of risk-aversion parameters.

Figure 4 shows that the stock component of the value of human capital is positive on average. It averages below 35 percent at all ages for high school agents and below 20 percent for college agents for a wide range of risk-aversion parameters. The stock component is positive, at a given age and state, provided that the sum of next period's earnings and human capital value covaries positively with the return to stock, conditional on this period's state.[20] Our empirical work, as summarized in Table 2, directly relates to the conditional comovement of earnings and stock returns since $cov(\varepsilon_{1,t}, \varepsilon_{2,t}) > 0$ in all the estimated models.

The orthogonal component has a large negative value early in the lifetime. Given that the orthogonal component has a zero mean, this is due to strong negative covariation between the orthogonal component and the stochastic discount factor early in life. This occurs, for example, because consumption and future utility are increasing in the realization of the persistent idiosyncratic earnings shock, other things equal. The persistent shock component is particularly important early in life as the effect of such a shock has many periods over which it impacts future earnings.

While it may seem plausible that the value of human capital is largely bond-like during retirement, it is useful to understand why it is not always 100 percent bond-like. If a retired agent will in all future date-events end up holding positive bonds, then the decomposition will indeed calculate that this agent's human capital in retirement is 100 percent bond-like as social security annuity payments in the model are riskless after retirement. However, if an agent hits the corner of the bond decision in the future under some sequence of risky stock returns, then this is not true. The mean of the agent's

[19] We allow for corners in which case $E[m_{j,j+1} R^i_{j+1} | z^j] \leq 1$.

[20] The Appendix describes our methods for computing the projection coefficients in the value decomposition.

stochastic discount factor will be less than the inverse of the gross risk-free rate in such an event. Thus, the agent discounts future social security transfers at greater than the risk-free rate beyond this date. The value of these transfers at earlier dates takes on a positive stock component provided that a corner solution is induced by low stock return realizations. In summary, while the value of human capital is mostly bond-like in retirement, it is not 100 percent bonds because agents run down financial assets, hit a corner solution on the holdings of the risk-free asset and live off social security transfers.

5.2 Human Capital Returns

Figure 5 plots properties of human capital returns. Mean returns are very large early in the working lifetime. To understand what drives the mean human capital returns, it is useful to return to the main ideas used in the value decomposition. The first equation below decomposes gross returns by decomposing the future payout into a bond, a stock and an orthogonal component. The second equation shows that the conditional mean human capital return always equals the weighted sum of the conditional mean of the bond and stock return.[21]

$$R^h_{j+1} \equiv \frac{v_{j+1} + e_{j+1}}{v_j} = \frac{\alpha^b_j R^b_{j+1} + \alpha^s_j R^s_{j+1} + \epsilon}{v_j}$$

$$E[R^h_{j+1}|z^j] = \frac{\alpha^b_j}{v_j} E[R^b_{j+1}|z^j] + \frac{\alpha^s_j}{v_j} E[R^s_{j+1}|z^j]$$

The weights on the bond and stock return do not always sum to one. When the agent's Euler equation for both stock and bonds hold with equality, then these weights will sum to more than one exactly when the value of the orthogonal component is negative.[22] The value of the orthogonal component of human capital payouts is negative early in the working lifetime. Human capital returns can vastly exceed a convex combination of stock and bond returns when the weights sum to more than one.

The mean return to human capital is near the risk-free rate immediately after retirement but subsequently increases. The high return towards the end of the lifetime might at first seem odd. This should not be surprising, however, as in the penultimate period $v_{J-1} = E[m_{J-1,J} e_J]$ and $1 = E[m_{J-1,J} e_J / v_{J-1}]$. As the payment e_J, conditional on surviving to the last period, is certain, the return is $R^h_J = e_J / v_{J-1} = 1/E[m_{J-1,J}]$. Thus, the return equals the risk-free bond rate when the agent is off the corner (i.e. $R^h_J = 1/E[m_{J-1,J}] = R^b$) but can exceed the risk-free rate when the agent

[21] The orthogonal component drops out as, with a risk-free asset, the mean of the orthogonal component is zero.

[22] In this case, $v_j = E[m_{j+1}(v_{j+1} + e_{j+1})] = \alpha^b_j + \alpha^s_j + E[m_{j+1}\epsilon_{j+1}]$ and $E[m_{j+1}\epsilon_{j+1}] < 0$ imply $\alpha^b_j / v_j + \alpha^s_j / v_j > 1$. Of course, the weights for decomposing returns can and do sum to more than one even when Euler equations do not hold with equality.

is on the corner (i.e. $R_J^h = 1/E[m_{J-1,J}] \geq R^b$). Towards the end of the lifetime an increasing fraction of agents in the model are on this corner and live off their social security annuity.

The positive correlation between human capital returns and stock returns in Figure 5 is based in part on two properties. First, innovations to the aggregate component of earnings growth and to stock returns are positively correlated. This implies that the component of human capital returns related directly to the earnings payout next period covaries positively with stock returns. Second, the old-age transfer benefit formula in the benchmark model is proportional to the aggregate component of earnings at the retirement age. The U.S. social security system has a similar feature as old-age benefits are proportional to a measure of average earnings in the economy when the worker turns age 60, other things equal.

5.3 Portfolio Allocation

Figure 6 describes portfolio allocation for three measures of wealth. Higher values of the risk-aversion coefficient are associated with lower average stock shares of financial wealth. The average stock share increases just before retirement. This is connected to the falling stock share of the value of human capital just before retirement.

The remainder of Figure 6 divides overall wealth (human plus financial wealth) into components. In the model human capital averages more than 50 percent of overall wealth at all ages.[23] Early in the working lifetime and in retirement the value of an agent's human capital makes up on average an overwhelming share of overall wealth. This holds despite the fact that human capital values are far below naive values.

We also divide overall wealth into bond and stock components using the human capital value decomposition. Thus, to account for overall stock holdings, we add together stock directly held in the financial wealth portfolio and the stock position embodied in the value of human capital. For both education groups we find that the bond component of overall wealth exceeds the stock component at all ages. The overall stock share early in life is largely determined by the decomposition analysis presented earlier. This is because financial assets are small in value compared to the value of human capital and negative positions in either financial asset are not allowed.

[23] This is the mean of the shares produced across simulations of a population of individuals, each drawing a sequence of shocks from the stochastic process for aggregate and idiosyncratic shock variables.

6 Discussion: Robustness and Drivers

The previous section documented properties of human capital values in the benchmark model. We find that (1) the value of human capital is far below the value implied by discounting future earnings at the risk-free rate and (2) the stock component of human capital is less than 35 percent of the value of human capital at all ages. This section determines the robustness and drivers of these findings.

6.1 Value of Human Capital

What drives the value of human capital to be substantially below the naive value? To answer this question, we consider a number of perturbations of the benchmark model. For each perturbation, we recalculate human capital values and then plot the results. The benchmark model analyzed in this section is the model in Table 4 that sets risk aversion to $\alpha = 6$ and that sets earnings to the process without cointegration estimated for the full sample.

We first consider perturbations that help agents to better smooth consumption. Intuitively, such perturbations lessen the negative covariation between the stochastic discount factor and earnings. One perturbation starts agents off with an initial wealth of 1 times mean earnings at age 22 rather than the benchmark value of 0.30 times mean earnings. The other perturbations allow agents to hold negative balances in the risk-free asset up to 1.0 times mean earnings or up to the natural borrowing limit. Figure 7 shows that while all perturbations increase the human capital value early in life human capital values remain well below the naive value.

Next we examine the extent to which transitory or persistent idiosyncratic shocks are key drivers. First, we eliminate transitory or persistent shocks. Figure 7 shows that eliminating transitory shocks increases human capital values. However, the quantitative effects are small compared to the massive impact of eliminating persistent idiosyncratic shocks. Eliminating persistent shocks produces more than a tripling of the value of human capital early in life. Lastly, we impose that persistent shocks have no skewness and no cyclical variation in variance or skewness. Figure 7 shows that this increases human capital values and that its effect is more powerful than eliminating transitory shocks. This foreshadows the importance of skewness as a driver of the stock component of human capital values that we find in the next subsection. We note that for each of these three changes we re-estimate the model with the new restrictions imposed.

We examine the effect of altering the preference parameter γ, while keeping relative risk aversion fixed. The value of $1/\gamma$ controls the intertemporal elasticity of substitution (IES). Figure 7 shows that increasing the IES to 2 increases the human capital value slightly, whereas decreasing the IES to 0.5

reduces the value of human capital. Neither change alters the finding that human capital values are substantially below naive values.

6.2 Stock Share of Human Capital Values

What drives the magnitude of the stock share of human capital values? We answer this question by analyzing four perturbations of the benchmark model. We (i) vary the borrowing constraint, (ii) vary the IES, (iii) eliminate the cyclical changes in persistent idiosyncratic shocks and/or eliminate the left-skewness in these shocks and (iv) allow cointegration.

Figure 8 shows the results. Raising the IES to 2 lowers the stock component slightly whereas lowering the IES to 0.5 raises the stock component slightly. Allowing borrowing up to one times mean earnings has almost no impact on the stock component.

Next we eliminate skewness and/or eliminate the cyclical variation in shock distributions. The case of no skewness means that the generalized normal distribution analyzed in the benchmark model is replaced by the normal distribution. In all cases, we re-estimate the idiosyncratic shock process under the new restrictions to best match data moments.

Figure 8 shows that eliminating left-skewness and eliminating cyclicality decreases the stock component of human capital values relative to the benchmark model. This could be stated more positively if one took the benchmark model to be the model with no skewness and no cyclical changes in idiosyncratic shocks. Using that model as a benchmark, implies that allowing left-skewed shocks and allowing the distribution of such shocks to vary cyclically increases the stock component of human capital. Furthermore, Figure 8 shows that the incremental effect of adding left-skewness is substantially larger than adding cyclical variation in a distribution displaying no skewness. Thus, we find that skewness is a quantitatively important factor that increases the stock share of human capital values and, at the same time, acts to lower human capital values.

Figure 8 also addresses whether cointegration between earnings and stock returns is important for the magnitude of the stock share. First, we compare the benchmark model and the model allowing cointegration when both are estimated using CPS 1967-2008 data. We find that allowing cointegration slightly decreases the stock share of human capital values. One might be skeptical that a cointegrated relationship can be precisely estimated over a short time period. For this reason, we repeat the analysis using an aggregate measure of earnings growth to proxy male earnings growth. We use NIPA 1929-2009 data on aggregate wages and salaries and divide this by the labor force to get average earnings. We re-estimate both models from Table 2 using data over the period 1929-2009. Figure 8 shows that the stock share increases somewhat when these estimated models are incorporated into the economic model. However, allowing cointegration does not significantly alter the stock share compared to the case of no cointegration, given the new data.

In summary, Figure 8 shows that two changes to the benchmark model produce a larger stock share. These are to decrease the IES to values below 1 and to re-estimate the model using NIPA data over a longer time period. We also find that skewness is a key driver of the stock component. Abstracting from skewness implies a much smaller stock share of human capital.[24] While one may conjecture that allowing cointegration may substantially increase the stock share of the value of human capital, we do not find support for this when the models are estimated using the same data set.

This last finding may seem surprising in light of the work of Benzoni et al. (2007). The main focus of their work is to examine how cointegration affects portfolio choice as a single parameter κ that controls the strength of adjustment in the cointegrating relationship increases. Stock holding in the financial asset portfolio is zero early in life in their model for a sufficiently large value of κ, when the relative-risk aversion coefficient is set to equal 5. For the parameter configuration that they highlight, the stock component of the value of human capital is 50 percent at age 20 and remains at 50 percent throughout the first half of the working lifetime.

Our model differs from their work in a number of dimensions. For example, we model social security benefits, allow idiosyncratic shocks to be drawn from a skewed distribution and allow for cyclical changes in idiosyncratic shocks. They abstract from all of these features. In addition, the methodology differs. While we estimate the idiosyncratic and aggregate shock structure from micro data, they roughly calibrate a cointegrated aggregate process governing stock returns and the aggregate component of earnings.

To make contact with Benzoni et al. (2007), we simply take the calibrated aggregate process from their work and substitute it into our model. Their calibrated process has the discrete-time approximation of the form indicated below. We take as given their parameter value for $\kappa = 0.15$ and their values governing the variance-covariance structure of the shock terms $(\epsilon_1, \epsilon_2, \epsilon_3)$, but we adjust the constant terms (γ_1, γ_2) so that $E[logR] = .041$ and $E[\Delta u] = 0$.[25]

[24] One might conjecture that increasing the mean growth rate $E[\Delta u_t^1]$ from the sample average of zero might matter. Increasing this mean from zero to 1 or 2 percent substantially increases individual human capital values while leaving the stock share essentially unchanged in the benchmark model.

[25] Appendix A.3 shows how we go from their continuous-time formulation to the discrete-time model and computes steady-state properties. The Benzoni et al. (2007) process produces an unconditional standard deviation of $SD(\Delta u_t) = 0.069$ which is more than twice the values that we calculate in US earnings data as documented in Table A.4 and Table A.5.

$$\begin{pmatrix} \Delta u_{t+1} \\ \log R_{t+1} \\ w_{t+1} \end{pmatrix} = \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & -\kappa \\ 0 & 0 & 0 \\ 0 & 0 & 1-\kappa \end{pmatrix} \begin{pmatrix} \Delta u_t \\ \log R_t \\ w_t \end{pmatrix} + \begin{pmatrix} \epsilon_{1,t+1} \\ \epsilon_{2,t+1} \\ \epsilon_{3,t+1} \end{pmatrix} \tag{5}$$

Figure 9 shows how the stock share of the value of human capital changes across models. Figure 9(a) compares the benchmark model (without cointegration) from this section to the model that results from replacing the aggregate dynamics with the Benzoni et al. (2007) process featuring cointegration, as constructed above. When inserted into our framework, the Benzoni et al. (2007) process produces a stock share that is well below the 50 percent value that they highlight.

Figure 9(b) considers a different comparison. The benchmark model is now modified to exclude skewness and cyclical variation in idiosyncratic risk, but is still estimated to best match data. We then insert the Benzoni et al. (2007) process into this model. Figure 9(b) shows that once again both models produce a stock share that is well below 50 percent over the lifetime. The results in Figure 9 are essentially unchanged if the mean log earnings growth rate is increased to equal 1 percent in all the models rather than the benchmark value of zero.

7 Conclusion

Our analysis highlights two main properties of human capital values based on an analysis of U.S. data on males earnings and financial asset returns: (1) the value of human capital is far below the value implied by discounting future earnings at the risk-free rate and (2) the stock component of the value of human capital averages less than 35 percent at each age over the working lifetime. These properties hold for (i) three different educational groups, (ii) a wide range of parameters characterizing risk aversion, (iii) a range of assumptions on borrowing constraints and (iv) two different statistical models for earnings estimated using male earnings data.

We investigate the main drivers of these two findings. Persistent idiosyncratic shocks and the left skewness of idiosyncratic shocks are two key drivers of low human capital values. In our model framework, an agent's stochastic discount factor falls for larger realizations of idiosyncratic shocks, other things equal. A number of model features lead to a positive stock component of the value of human capital including (i) social security benefits linked to average earnings, (ii) positive conditional correlation between the aggregate component of earnings and stock returns and (iii) left-skewed idiosyncratic shocks. We provide support for all three of these features in US data. We do not find much support for the idea that cointegration between the aggregate component of earnings and stock returns is a key factor driving the size of the stock component of the value of human capital.

References

Alvarez, F. and U. Jermann (2004), Using Asset Prices to Measure the Cost of Business Cycles, *Journal of Political Economy,* 112, 1223-56.

Becker, G. (1975), Human Capital, Second Edition, National Bureau of Economic Research.

Benzoni, L., Collin-Dufresne, P. and R. Goldstein (2007), Portfolio Choice over the Life-Cycle when the Stock and Labor Markets are Cointegrated, *Journal of Finance,* 62, 2123-67.

Coco, J., Gomes, F. and P. Maenhout (2005), Consumption and Portfolio Choice over the Life Cycle, *Review of Financial Studies,* 18, 491-533.

Constantinides, G. and D. Duffie (1996), Asset Pricing with Heterogeneous Consumers, *Journal of Political Economy,* 104, 219-40.

Dublin, L. and A. Lotka (1930), The Money Value of a Man, The Ronald Press Company.

Epstein, L. and S. Zin (1991), Substitution, Risk Aversion and the Temporal Behavior of Consumption and Asset Returns: An Empirical Analysis, *Journal of Political Economy,* 99, 263-86.

Farr, W. (1853), The Income and Property Tax, *Journal of the Statistical Society of London,* 16, 1-44.

French, E. (2004), The Labor Supply Response to (Mismeasured but) Predictable Wage Changes, *Review of Economics and Statistics,* 86(2), 602-613.

Graham, J. and R. Webb (1979), Stocks and Depreciation of Human Capital: New Evidence from a Present-Value Perspective, *Review of Income and Wealth,* 25, 209- 24.

Guvenen, F. (2009), An Empirical Investigation of Labor Income Processes, *Review of Economic Dynamics,* 12(1), 58-79.

Guvenen, F. and A. Smith (2011), Inferring Labor Income Risk from Economic Choices: An Indirect Inference Approach, manuscript.

Guvenen, F., Ozkan, S. and J. Song (2014), The Nature of Countercyclical Income Risk, Journal of Political Economy, 122, 621-60.

Haveman, R., Bershadker, A. and J. Schwabish (2003), Human Capital in the United States from 1975 to 2000: Patterns of Growth and Utilization, Upjohn Institute.

Heathcote, J., Perri, F. and G. Violante (2010), Unequal We Stand: An Empirical Analysis of Economic Inequality in the United States, *Review of Economic Dynamics,* 13(1), 15-51.

Hosking, J. and J. Wallis (1997), Regional Frequency Analysis, Cambridge University Press.

Huggett, M. and G. Kaplan (2011), Human Capital Values and Returns: Bounds Implied By Earnings and Asset Returns Data, *Journal of Economic Theory,* 146, 897-919.

Huggett, M. and G. Kaplan (2012), The Money Value of a Man, NBER Working Paper 18066.

Huggett, M. and J.C. Parra (2010), How Well Does the U.S. Social Insurance System Provide Social Insurance?, *Journal of Political Economy,* 118, 76-112.

Johansen, S. (1995), Likelihood-Based Inference in Cointegrated Vector Autoregressive Models, Oxford University Press.

Jorgenson, D. and B. Fraumeni (1989), The Accumulation of Human and Nonhuman Capital, 1948- 1984, in *The Measurement of Saving, Investment and Wealth,* R. Lipsey and H. Tice, eds. NBER Studies in Income and Wealth.

Krebs, T., Kuhn, M. and M. Wright (2013), Human Capital Risk, Contract Enforcement and the Macroeconomy, manuscript.

Lucas, R. (1978), Asset Prices in an Exchange Economy, *Econometrica,* 46, 1429- 45.

Lynch, A. and S. Tan (2011), Labor Income Dynamics at Business-Cycle Frequencies: Implications for Portfolio Choice, *Journal of Financial Economics,* 101, 333-59.

Meghir, C. and L. Pistaferri (2010), Earnings, Consumption and Lifecycle Choices, NBER Working Paper.

NCHS (1992), *US Decennial Life Tables 1989-1991,* Washington, DC: Hyattsville, Md.

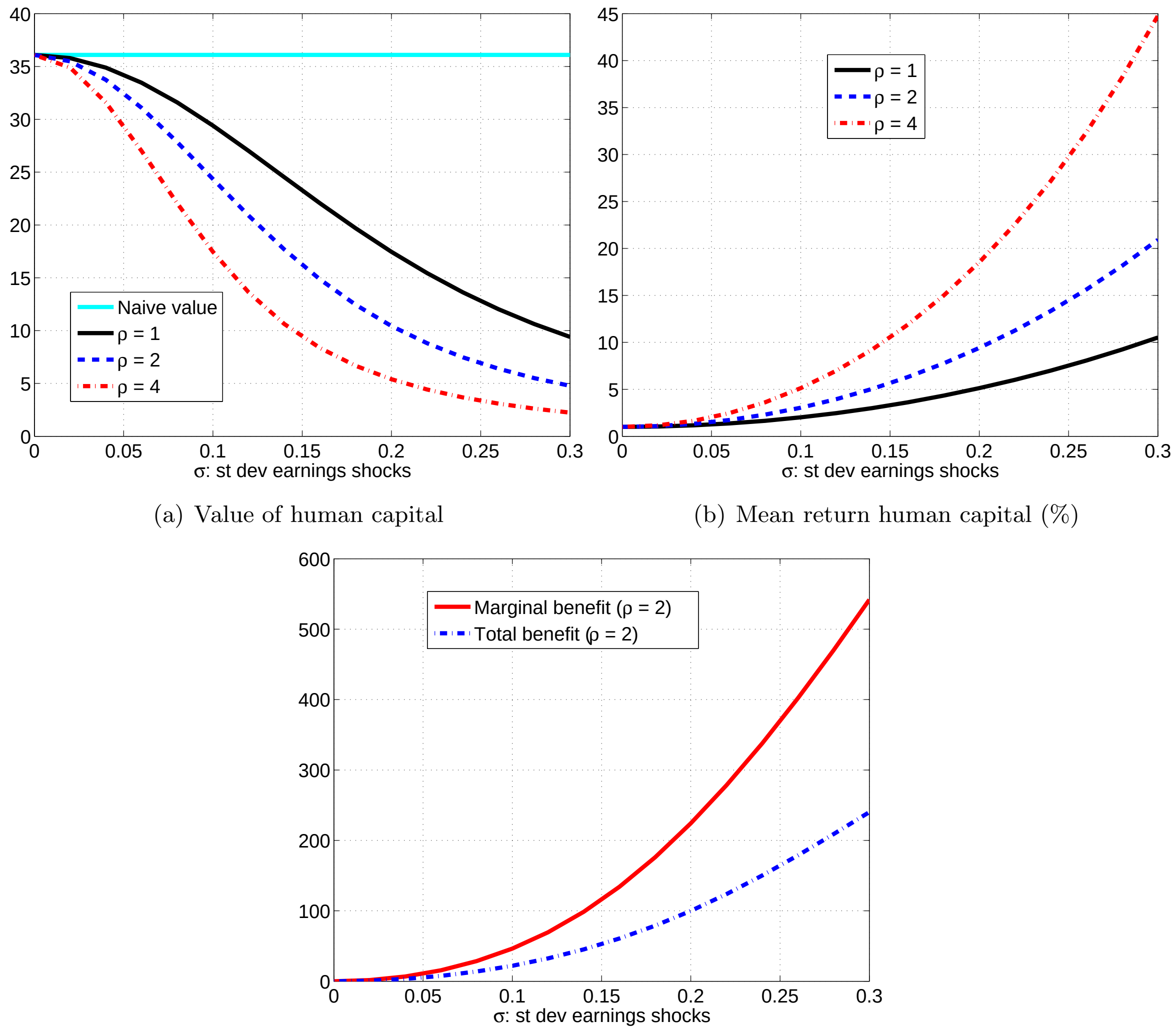
Poterba, J., Venti, S. and D. Wise (2011), The Composition and Drawdown of Wealth in Retirement?, *Journal of Economic Perspectives,* 25, 95-118.

Social Security Handbook (2012), available http://www.ssa.gov.

Storesletten, K., Telmer, C. and A. Yaron (2004), Cyclical Dynamics in Idiosyncratic Labor-Market Risk, *Journal of Political Economy,* 112, 695-717.

Vissing-Jorgensen, A. and O. Attanasio (2003), Stock-Market Participation, Intertemporal Substitution, and Risk-Aversion, *American Economic Review Papers and Proceedings,* 93, 383-91.

Weisbrod, B. (1961), The Valuation of Human Capital, *Journal of Political Economy,* 69, 425-36.

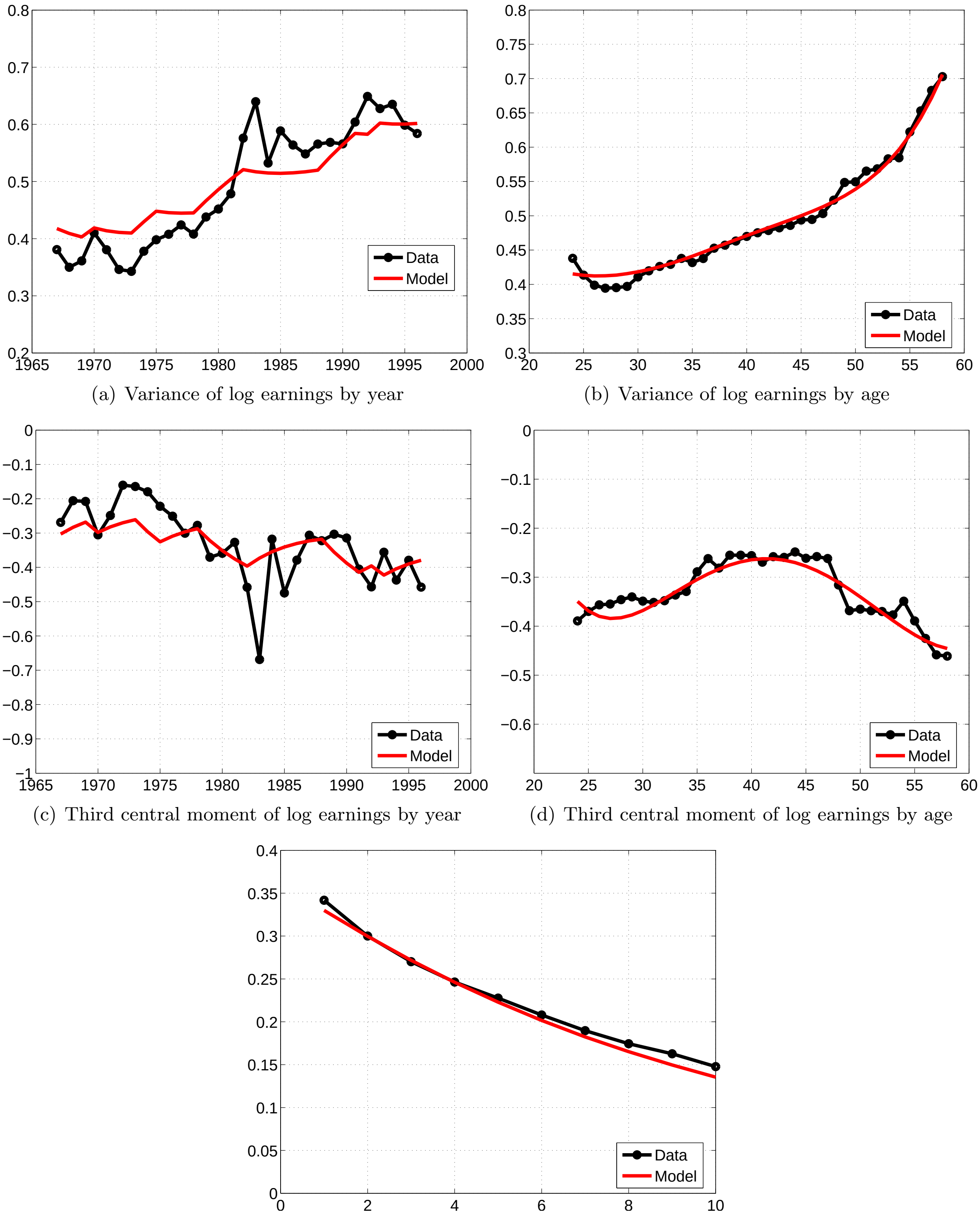


(a) Value of human capital

(b) Mean return human capital (%)

(c) Benefit of moving to a smooth consumption plan (%)

Figure 1: Human capital values and returns: a simple example



(a) Variance of log earnings by year

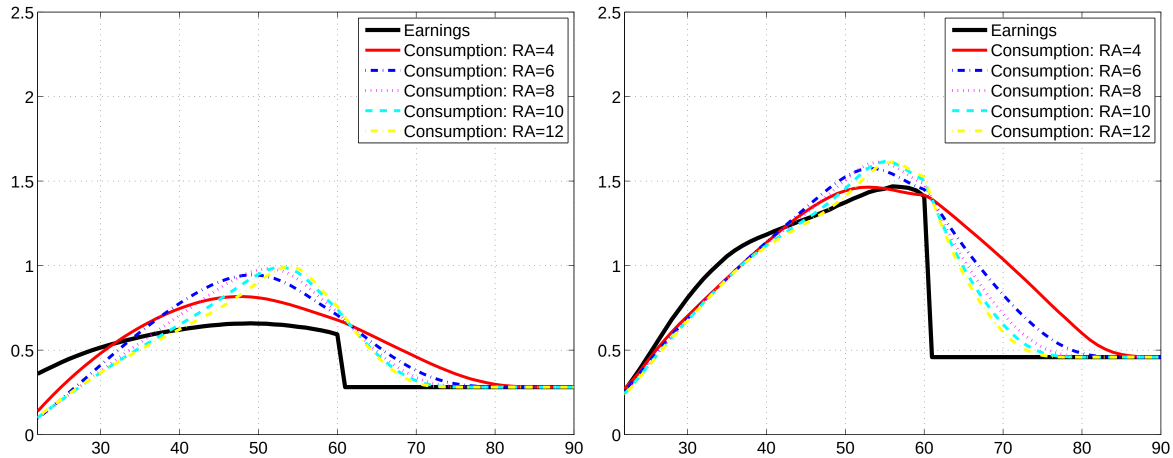
(b) Variance of log earnings by age

(c) Third central moment of log earnings by year

(d) Third central moment of log earnings by age

(e) Average autocovariance function

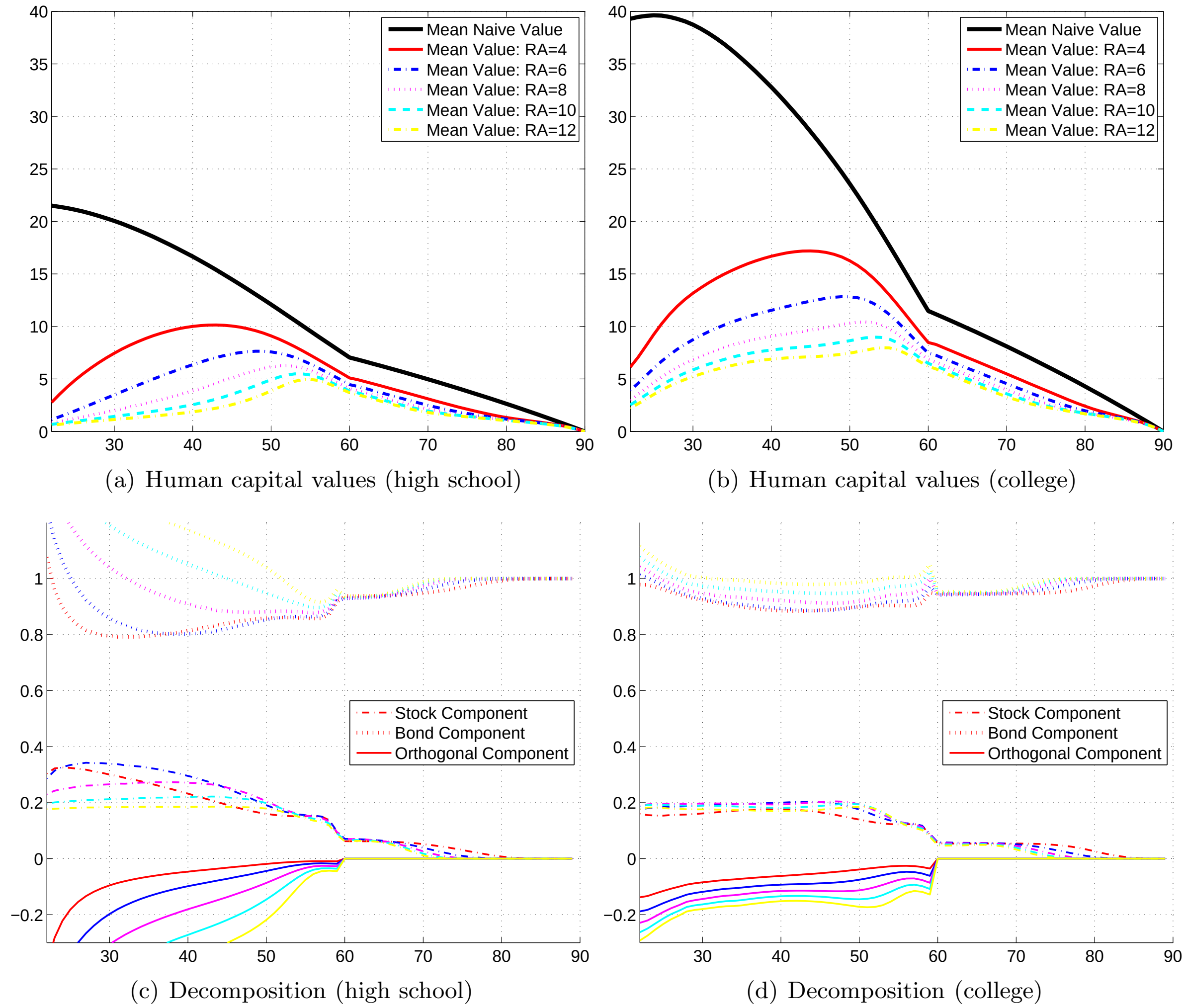
Figure 2: Fit of estimated idiosyncratic earnings model for the full sample



(a) Mean earnings and consumption (high school) (b) Mean earnings and consumption (college)

Figure 3: Life-cycle profiles in the benchmark model

Notes: The vertical scale is in units of 100,000 dollars in year 2008.



(a) Human capital values (high school)

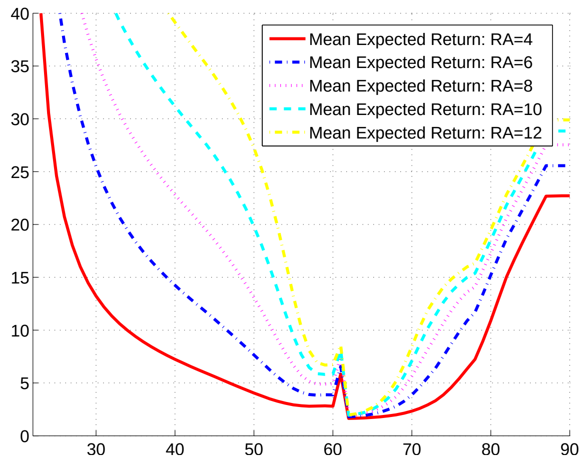
(b) Human capital values (college)

(c) Decomposition (high school)

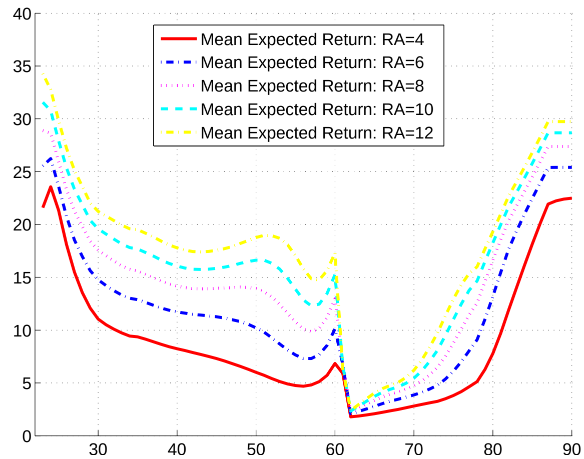
(d) Decomposition (college)

Figure 4: Human capital values and a decomposition

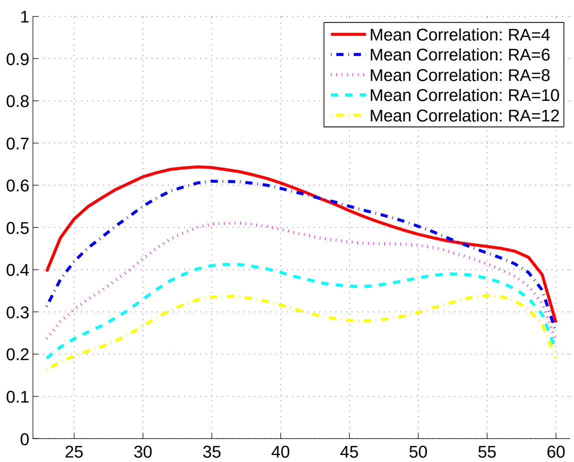
Notes: The vertical scale in Figure 3 (a)-(b) is in units of 100,000 dollars in year 2008.



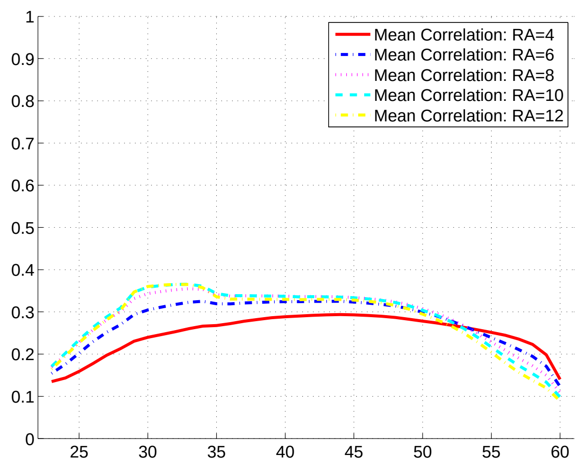
(a) Human capital returns (%) (high school)



(b) Human capital returns (%) (college)

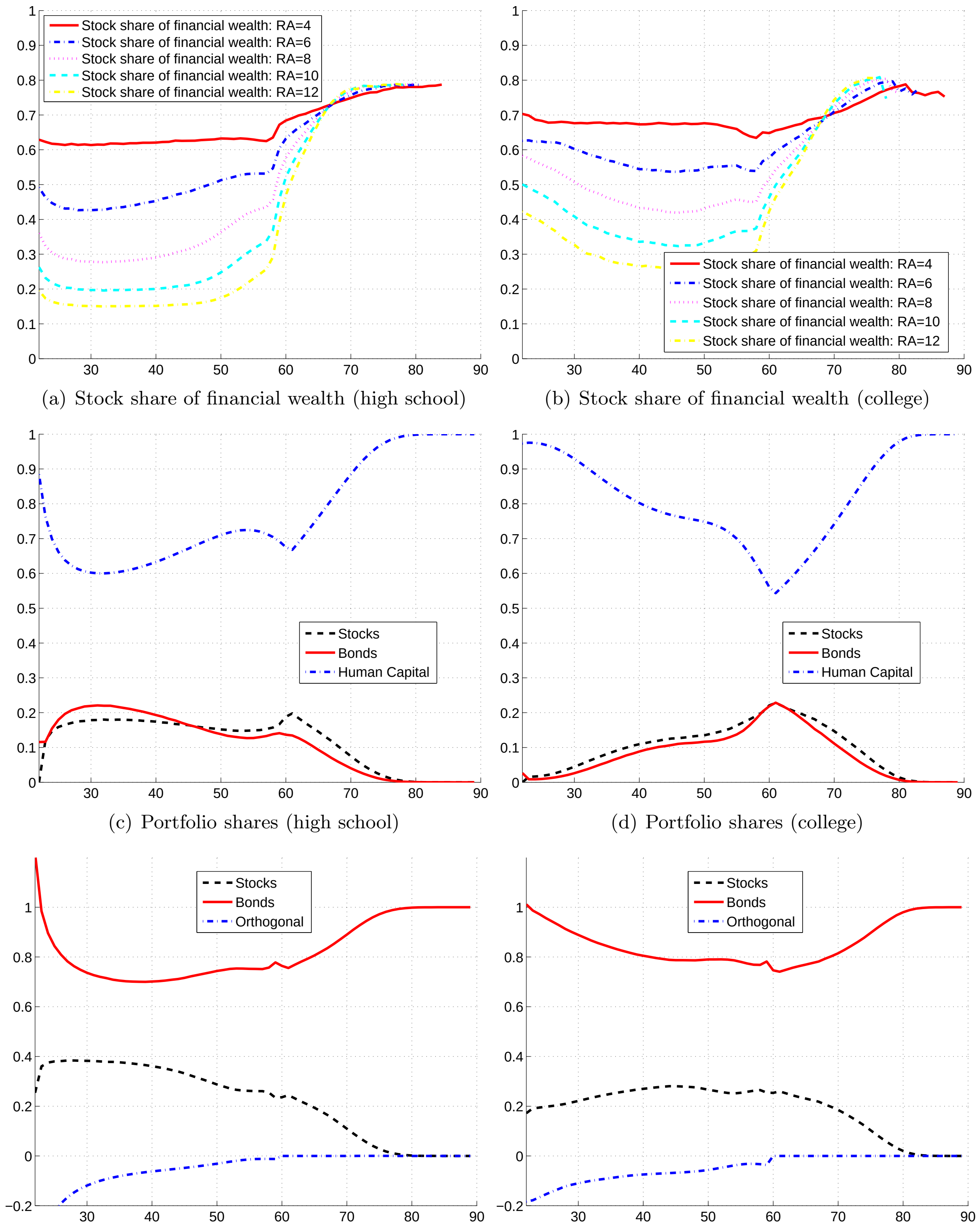


(c) Correlation: HC, stock returns (high school)



(d) Correlation: HC, stock returns (college)

Figure 5: Properties of human capital returns



(a) Stock share of financial wealth (high school)

(b) Stock share of financial wealth (college)

(c) Portfolio shares (high school)

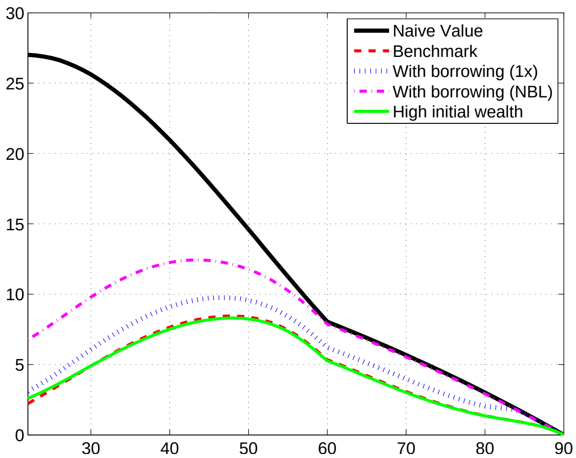
(d) Portfolio shares (college)

(e) Overall portfolio shares (high school)

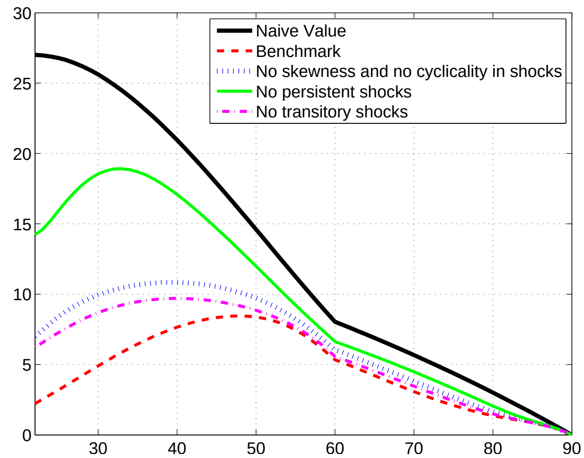
(f) Overall portfolio shares (college)

Figure 6: Portfolio shares in the benchmark model

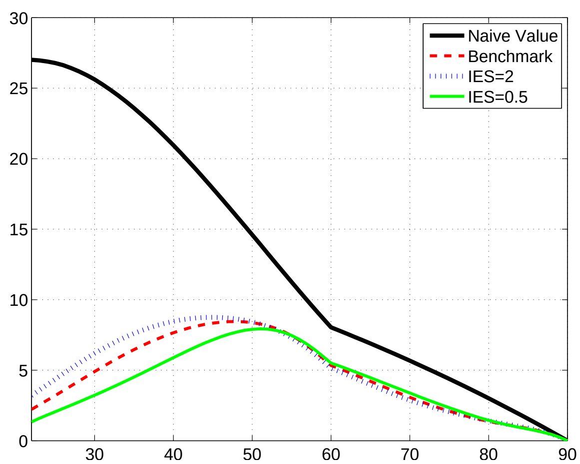
Notes: Financial portfolio shares in panels (a)-(b) are averages over the sub-population with positive asset holdings. Panels (c)-(f) present results setting risk aversion equal to 6.



(a) Human capital values (borrowing, initial wealth)



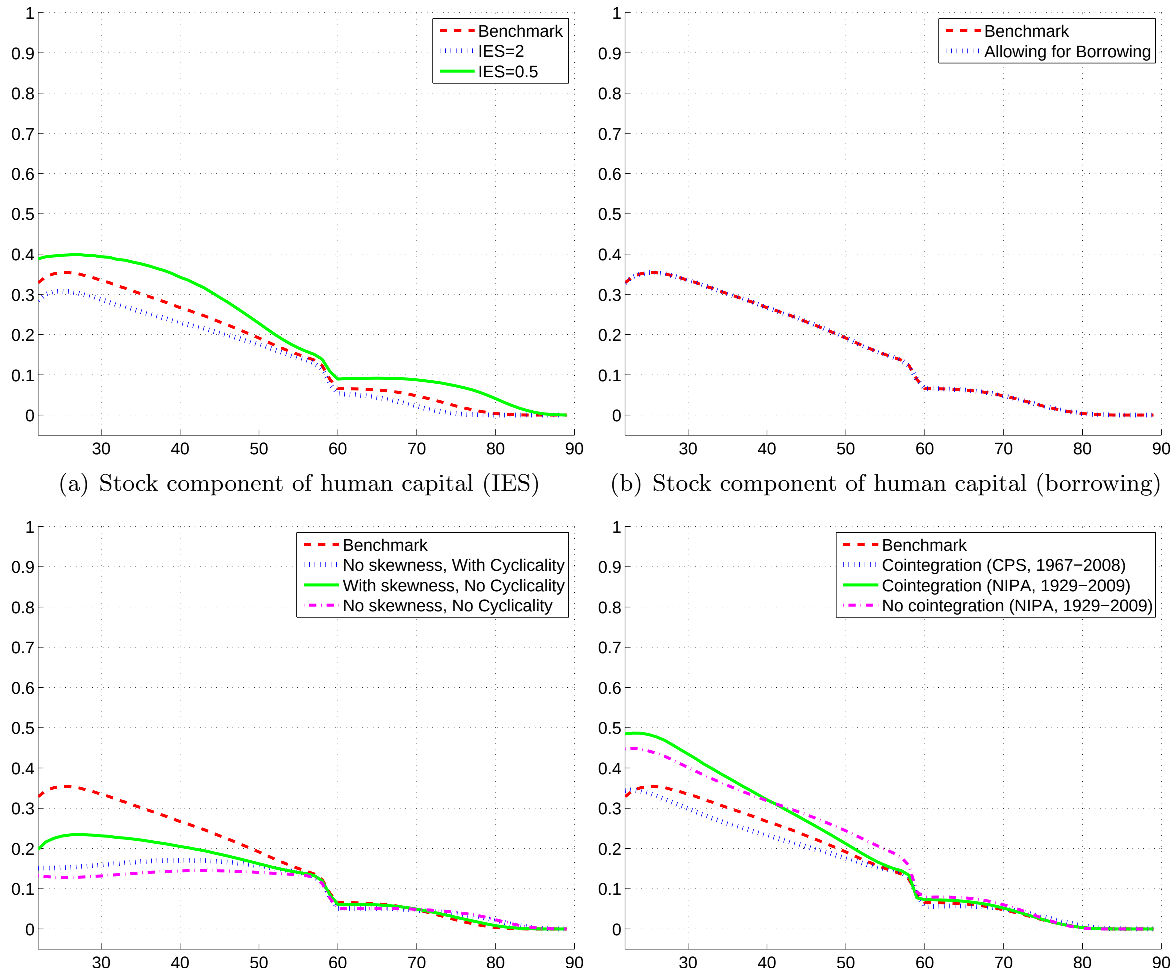
(b) Human capital values (idiosyncratic risk)



(c) Human capital values (IES)

Figure 7: What drives the value of human capital?

Notes: In panel (a) "With borrowing (1x)" refers to the model that allows borrowing up to 1 times average annual earnings, and "With borrowing (NBL)" refers to model that allows borrowing up to the "Natural Borrowing Limits" i.e. limits that impose only that the agent must be able to repay his debt in all states of the world.



(a) Stock component of human capital (IES)

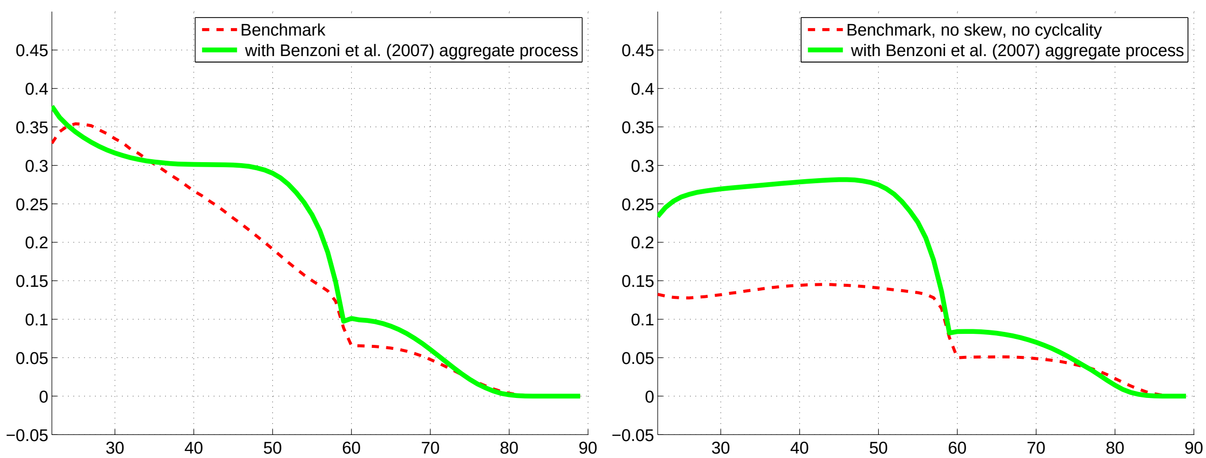
(b) Stock component of human capital (borrowing)

(c) Stock component of human capital (counter-cyclical risk)

(d) Stock component of human capital (cointegration)

Figure 8: What drives the stock component of human capital?

Notes: When comparing different models estimated on the same data set or the same model estimated on different data sets, the constants in all models are reset so that $E[\Delta u_t^i] = 0$ and $E[\log R_t^s] = 0.041$ as previously noted in Table 3.



(a) Stock share of human capital

(b) Stock share of human capital

Figure 9: Analysis of Benzoni et al. (2007)

Table 1: Parameter Estimates for the Idiosyncratic Earnings Process

	Full Sample	College Sample	High School Sample
Fixed Effect			
σ^2_ξ	0.082	0.064	0.137
	(0.013)	(0.015)	(0.032)
Persistent Component			
ρ	0.942	0.951	0.843
	(0.012)	(0.015)	(0.097)
σ^2_η: boom	0.038	0.037	0.047
	(0.005)	(0.006)	(0.023)
σ^2_η: recession	0.058	0.048	0.072
	(0.005)	(0.008)	(0.020)
σ^2_η: linear trend	0.001	0.001	0.002
	(0.000)	(0.000)	(0.002)
$\mu_{3,\eta}$: boom	-0.020	-0.006	-0.149
	(0.014)	(0.015)	(0.181)
$\mu_{3,\eta}$: recession	-0.061	-0.040	-0.190
	(0.013)	(0.019)	(0.170)
$\mu_{3,\eta}$: linear trend	-0.001	-0.001	-0.003
	(0.001)	0.001)	(0.002)
Transitory Component			
σ^2_υ	0.132	0.139	0.128
	(0.005)	(0.006)	(0.023)
$\mu_{3,\upsilon}$	-0.161	-0.162	-0.002
	(0.027)	(0.029)	(0.172)

Notes: Models of the moments of the transitory shock include a fourth-order polynomial in age. The reported moments for transitory shocks are averages over the age range. Standard errors are computed by block bootstrap with 39 repetitions.

Table 2: Parameter Estimates for the Aggregate Stochastic Process

		No Cointegration			With Cointegration		
		Full Sample	College Sample	High School Sample	Full Sample	College Sample	High School Sample
Equation 1: Δu_t^1							
Δu_{t-1}^1	Γ_{11}	0.383	0.260	0.348	0.364	0.12	0.295
		(0.14)	(0.15)	(0.14)	(0.19)	(0.15)	(0.18)
$\log R_{t-1}^s$	Γ_{12}	0.044	0.04	0.057	0.045	0.016	0.058
		(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)
Constant	γ_1	-0.004	-0.003	-0.009	-0.004	-0.005	-0.008
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Equation 2: $\log R_t^s$							
Δu_{t-1}^1	Γ_{21}	-2.149	-2.203	-1.731	0.473	-2.248	0.236
		(1.15)	(1.29)	(0.97)	(1.42)	(1.45)	(1.18)
$\log R_{t-1}^s$	Γ_{22}	0.106	0.153	0.101	0.054	0.145	0.072
		(0.17)	(0.18)	(0.17)	(0.16)	(0.21)	(0.17)
Constant	γ_2	0.032	0.031	0.024	0.00	0.029	0.00
		(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.03)
Var-Cov Matrix							
$var\ (\varepsilon_{1,t}) \times 10^{-4}$		4.42	4.24	6.49	4.42	3.37	6.44
$var\ (\varepsilon_{2,t}) \times 10^{-2}$		3.2	3.24	3.23	2.57	3.24	2.92
$cov\ (\varepsilon_{1,t}, \varepsilon_{2,t}) \times 10^{-3}$		1.23	1.24	1.52	1.28	1.21	2.00
Cointegrating Vector							
$\log R_t^s$	β_2				0.309	-0.211	0.469
					(0.10)	(0.06)	(0.15)
Trend	ρ				-0.019	0.016	-0.026
					(0.01)	(0.00)	(0.01)
Constant	μ				-0.67	0.343	-0.976
Adjustment Parameters							
Δu_t^1	α_1				0.007	-0.196	0.017
					(0.05)	(0.07)	(0.04)
$\log R_t^s$	α_2				-1.04	-0.063	-0.651
					(0.36)	(0.64)	(0.23)

Notes: Standard errors in parentheses.

Table 3: Implied Steady-State Statistics for the Aggregate Stochastic Process

	Full Sample		
	Data	No Cointegration	With Cointegration
$E\left(logR_t^b\right)$	0.012	0.012	0.012
$E\left(\log R_t^s\right)$	0.041	0.045	0.070
$E\left(\Delta u_t^1\right)$	-0.002	-0.004	-0.002
$sd\left(\Delta u_t^1\right)$	0.025	0.025	0.025
$sd\left(\log R_t^s\right)$:	0.187	0.187	0.187
$corr\left(\Delta u_t^1, \log R_t^s\right)$	0.184	0.177	0.156
$corr\left(\Delta u_t^1, \Delta u_{t-1}^1\right)$	0.425	0.441	0.435
$corr\left(\log R_t^s, \log R_{t-1}^s\right)$	0.057	0.055	0.005
$corr\left(\Delta u_t^1 \log R_{t-1}^s\right)$	0.372	0.398	0.394
$corr\left(\log R_t^s, \Delta u_{t-1}^1\right)$	-0.292	-0.270	-0.289
	College Sub-sample		
	Data	No Cointegration	With Cointegration
$E\left(logR_t^b\right)$	0.012	0.012	0.012
$E\left(\log R_t^s\right)$	0.041	0.040	0.045
$E\left(\Delta u_t^1\right)$	0.000	-0.001	-0.001
$sd\left(\Delta u_t^1\right)$	0.023	0.023	0.023
$sd\left(\log R_t^s\right)$:	0.187	0.187	0.186
$corr\left(\Delta u_t^1, \log R_t^s\right)$	0.248	0.251	0.243
$corr\left(\Delta u_t^1, \Delta u_{t-1}^1\right)$	0.346	0.341	0.342
$corr\left(\log R_t^s, \log R_{t-1}^s\right)$	0.057	0.084	0.050
$corr\left(\Delta u_t^1 \log R_{t-1}^s\right)$	0.377	0.387	0.367
$corr\left(\log R_t^s, \Delta u_{t-1}^1\right)$	-0.225	-0.235	-0.229
	High School Sub-sample		
	Data	No Cointegration	With Cointegration
$E\left(logR_t^b\right)$	0.012	0.012	0.012
$E\left(\log R_t^s\right)$	0.041	0.045	0.074
$E\left(\Delta u_t^1\right)$	-0.007	-0.010	-0.008
$sd\left(\Delta u_t^1\right)$	0.030	0.030	0.030
$sd\left(\log R_t^s\right)$:	0.187	0.187	0.186
$corr\left(\Delta u_t^1, \log R_t^s\right)$	0.207	0.194	0.175
$corr\left(\Delta u_t^1, \Delta u_{t-1}^1\right)$	0.386	0.416	0.411
$corr\left(\log R_t^s, \log R_{t-1}^s\right)$	0.057	0.047	0.003
$corr\left(\Delta u_t^1 \log R_{t-1}^s\right)$	0.387	0.420	0.420
$corr\left(\log R_t^s, \Delta u_{t-1}^1\right)$	-0.289	-0.261	-0.276

Notes: Table shows average moments in the data, together with implied steady-state statistics from the corresponding estimated model. Data cover the period 1967-2008. When implementing the estimated processes in the structural model, we adjust the constants (γ_1, γ_2) estimated in Table 2 so that all models have $E[\log R_t^s] = 0.041$ and $E[\Delta u_t^1] = 0$.

Table 4: Parameter Values for the Benchmark Model

Category	Symbol	Parameter Value
Demographics	J, Ret	$(J, Ret) = (69, 40)$
	ψ_{j+1} Survival Probability	U.S. Life Table
Preferences	α Risk Aversion	$\alpha \in \{4, 6, 8, 10, 12\}$
	$1/\gamma$ Intertemporal Substitution	$1/\gamma = 1.17$
	β Discount Factor	see Notes
Returns	R^s, R^b	Table 2 - 3
Earnings	$e_j(z) = \begin{cases} z_1 \exp(\kappa_j + \xi + \zeta + \upsilon)(1-\tau) & \text{if} \quad j < Ret \\ z_1 b(\xi)(1-\tau) & \text{if} \quad j \geq Ret \end{cases}$	$\tau = .27$
	$\zeta' = \rho\zeta + \eta'$ and $\eta' \sim GN(0, \sigma^2_\eta(X), \mu_{3,\eta}(X))$	$b(\cdot)$ see text
	$\xi \sim N(0, \sigma^2_\xi)$ and $\upsilon \sim GN(0, \sigma^2_{\upsilon,j}, \mu_{3,\upsilon,j})$	Table 1-2
Initial Wealth	$\sum_{i\in\mathcal{I}} a^i_1 R^i_1$	$\sum_{i\in\mathcal{I}} a^i_1 R^i_1 = 0.3E[e_1]$

Notes: β is calibrated to generate a steady-state ratio of wealth to income equal to 3.5. All sensitivity analyses are performed by re-calibrating β to generate the same ratio. Survival probabilities are smoothed versions of male values from the 1989-91 US Decennial Life Tables in NCHS (1992). Smoothing is done using a nine point moving average. $E[e_1]$ denotes mean earnings at age 1 in the model.

A Appendix

A.1 Model Fit

The fit of the earnings model for the high school and college samples is provided in Figure A.1 and A.1 respectively.

A.2 Computation

This section describes our methods to compute solutions to the benchmark model and to compute values and returns.

A.2.1 Value Function and Decision Rules

To compute the optimal value function V_j^* and optimal decision rules to the model in section 4, we employ the method of dynamic programming. This involves computing functions V_j solving the Bellman equation (BE). Of course, the idea is that $V_j = V_j^*$. In stating $\hat{\Gamma}_j(x,z)$ in Bellman's equation, we impose the restrictions from the original budget constraint $\Gamma_j(x,z)$. We also use the fact that shocks are Markovian so that the current shock, denoted z, rather than partial histories z^j contain all relevant information. We model the shock $z = (z_1, z_2)$ as stated in section 4.

$$V_j^*(x,z) \equiv \max W(c_j, F(U(c_{j+1}, ..., c_J)), j) \text{ s.t. } c \in \Gamma_j(x,z)$$

$$\text{(BE) } V_j(x,z) = \max W(c_j, F(V_{j+1}(x', z')), j) \ s.t. \ (c, a^1, a^2) \in \hat{\Gamma}_j(x,z)$$

$$\hat{\Gamma}_j(x,z) = \{(c, a^1, a^2) : c + \textstyle\sum_{i \in \mathcal{I}} a^i \leq x, c \geq 0, a^1, a^2 \geq 0\}$$

We compute solutions to Bellman's equation only when the first component of the shock $z = (z_1, z_2)$ takes the value $z_1 = 1$. This is indicated below. To do so requires knowledge of $V_{j+1}(x', z_1', z_2')$ at all values of z_1'. Lemma 1 below shows that $V_j^*(\lambda x, \lambda z_1, z_2) = \lambda V_j^*(x, z_1, z_2), \forall \lambda > 0$ and therefore $V_j^*(x, z_1, z_2) = z_1 V_j^*(\frac{x}{z_1}, 1, z_2)$. In the Algorithm described below, we make use of this key property. In Lemma 1, $\Gamma(x,z)$ is homogeneous provided $c \in \Gamma(x,z) \Rightarrow \lambda c \in \Gamma(\lambda x, \lambda z), \forall \lambda > 0$.

$$V_j(x, 1, z_2) = \max_{(c, a^1, a^2) \in \hat{\Gamma}_j(x, 1, z_2)} W(c_j, F(V_{j+1}(x', z_1', z_2')), j)$$

Lemma 1:

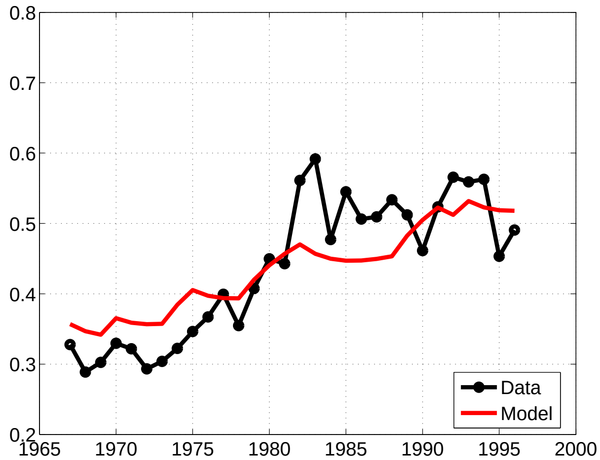
(i) Assume U is homothetic and $\Gamma(x,z)$ is homogeneous. $c^* \in argmax \{U(c) : c \in \Gamma(x,z)\}$ implies $\lambda c^* \in argmax \{U(c) : c \in \Gamma(\lambda x, \lambda z)\}, \forall \lambda > 0$.

(ii) In the benchmark model $V_j^*(\lambda x, \lambda z_1, z_2) = \lambda V_j^*(x, z_1, z_2), \forall \lambda > 0$

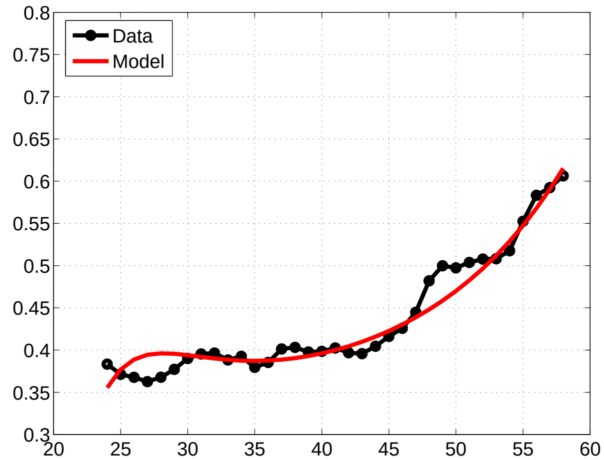
Proof:

(i) obvious

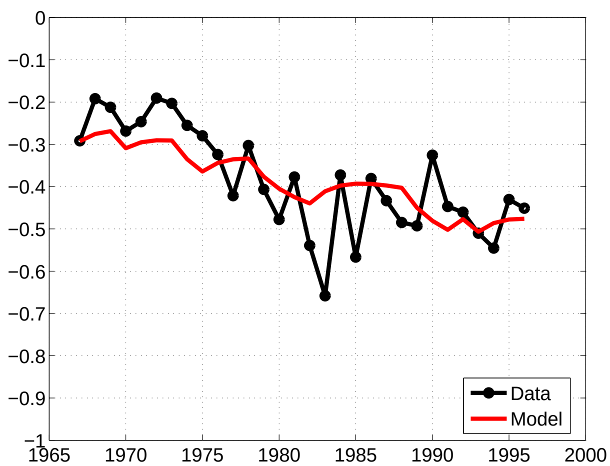
(ii) Follows from Lemma 1(i) after noting two things. First, EZ preferences are homothetic and, in fact, homogeneous of degree 1. Second, $\Gamma_j(x,z)$ is homogeneous in (x, z_1) for any fixed z_2. This is implied because the earnings function from the benchmark model is $e_j = z_1 g_j(z_2)$ and $z_1' = z_1 f_{j+1}(z_2')$,



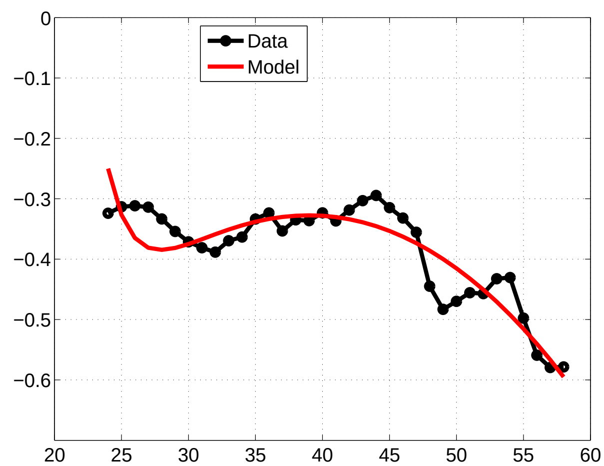
(a) Variance of log earnings by year



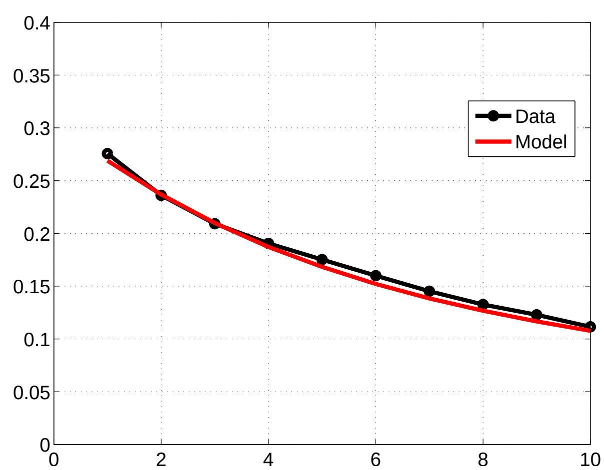
(b) Variance of log earnings by age



(c) Third central moment of log earnings by year

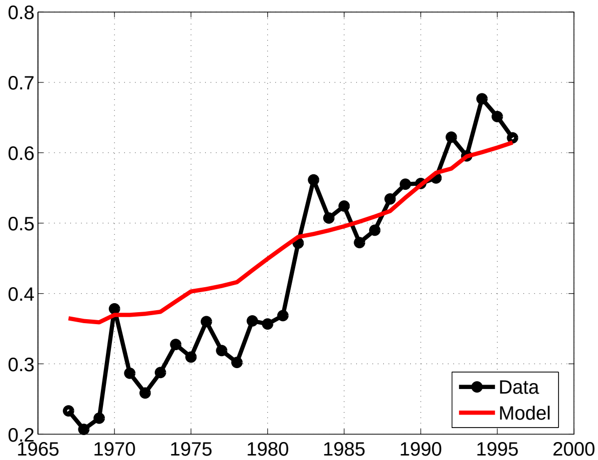


(d) Third central moment of log earnings by age

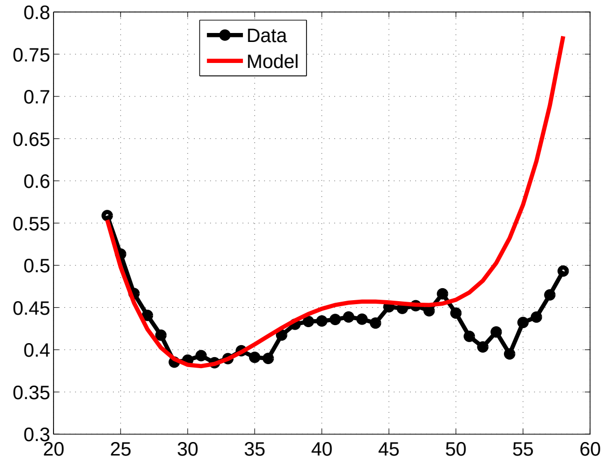


(e) Average autocovariance function

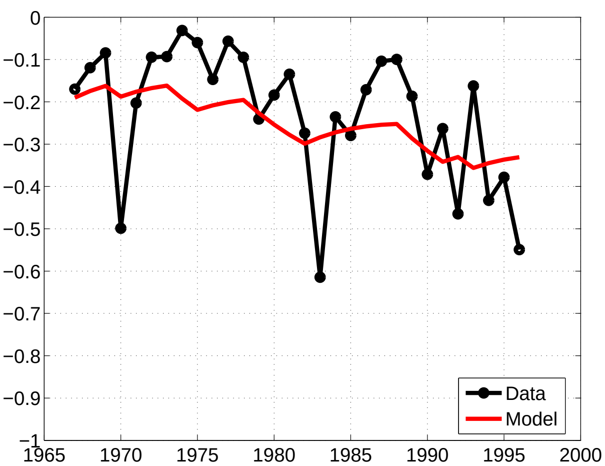
Figure 10: Fit of estimated idiosyncratic earnings model for High School sample



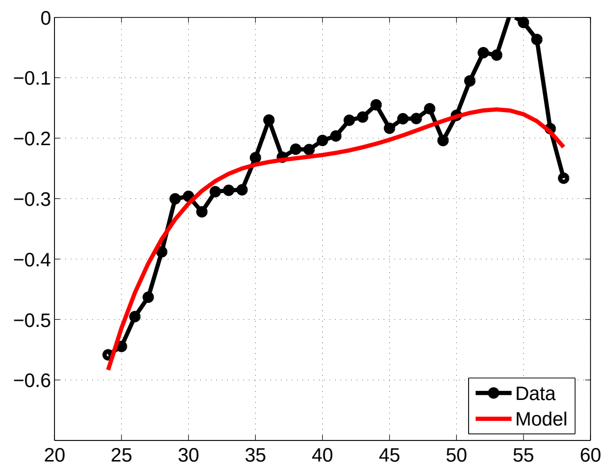
(a) Variance of log earnings by year



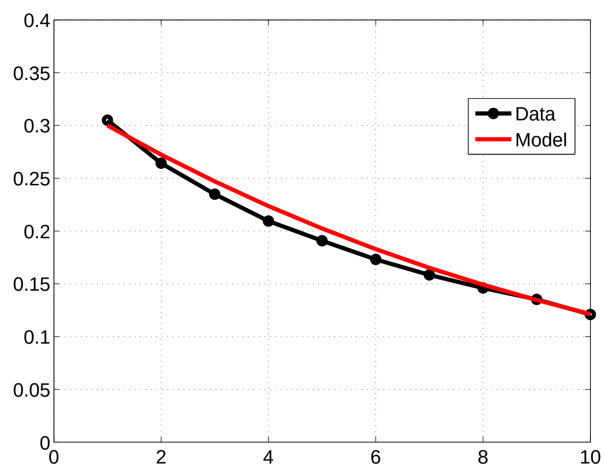
(b) Variance of log earnings by age



(c) Third central moment of log earnings by year



(d) Third central moment of log earnings by age



(e) Average autocovariance function

Figure 11: Fit of estimated idiosyncratic earnings model for College sample

where z_2 is Markov and primes denote next period values. These two properties hold both for the model with and without cointegration. $\diamond$

The Lagrange function corresponding to (BE) is stated below along with first-order conditions.

$$\mathcal{L} = \mathrm{W}(\mathrm{x} - \sum_{\mathrm{i}=1}^{2} \mathrm{a}^{\mathrm{i}}, \mathrm{F}(\mathrm{V}_{\mathrm{j}+1}(\mathrm{x}', \mathrm{z}_1', \mathrm{z}_2')), \mathrm{j}) + \lambda_1[\mathrm{a}^1 - 0] + \lambda_2[\mathrm{a}^2 - 0]$$

$$(1) \quad -W_1 + W_2 dF/da^1 + \lambda_1 = 0$$

$$(2) \quad -W_1 + W_2 dF/da^2 + \lambda_2 = 0$$

$$(3) \quad constraints + \ complementary\ slackness$$

We rewrite equation (1)-(2) below after imposing the functional forms from section 4. The Algorithm is then based on repeatedly solving these Euler equations.

$$(1') \ -1 + \beta\psi_{j+1} E[(\tfrac{c_{j+1}}{c_j})^{-\gamma}(\tfrac{V_{j+1}}{F(V_{j+1})})^{\gamma-\alpha} R^1(z')|x,z] + \lambda_1' = 0$$

$$(2') \ -1 + \beta\psi_{j+1} E[(\tfrac{c_{j+1}}{c_j})^{-\gamma}(\tfrac{V_{j+1}}{F(V_{j+1})})^{\gamma-\alpha} R^2(z')|x,z] + \lambda_2' = 0$$

Algorithm:

1. Set $V_J(x,1,z_2) = W(x,0)$ and $c_J(x,1,z_2) = x$ at grid points (x, z_2).
2. Given $(V_{j+1}(x,1,z_2), c_{j+1}(x,1,z_2))$, compute $(a^1_{j+1}(x,1,z_2), a^2_{j+1}(x,1,z_2))$ at grid points $(x,1,z_2)$ by solving $(1') - (2')$ and (3).
3. Set $c_j(x,1,z_2) = x - \sum_i a^i_{j+1}(x,1,z_2)$ and $V_j(x,1,z_2) = W(c_j(x,1,z_2), F(V_{j+1}), j)$ at grid points.
4. Repeat 2-3 for successive lower ages.

To carry out this Algorithm we mention two points. First, evaluating $(1') - (2')$ involves an interpolation of the first component of the functions (V_{j+1}, c_{j+1}). Second, evaluating $(1') - (2')$ also involves knowledge of (V_{j+1}, c_{j+1}) when the second component of these functions differs from $z_1 = 1$. This is accomplished by using Lemma 1 as indicated below.

$$V_{j+1}(x', z_1', z_2') = z_1' V_{j+1}(\frac{x'}{z_1'}, 1, z_2') \ and \ c_{j+1}(x', z_1', z_2') = z_1' c_{j+1}(\frac{x'}{z_1'}, 1, z_2')$$

$$x' = \sum_i a^i_{j+1}(x,1,z_2) R^i(z') + e_{j+1}(z')$$

A.2.2 Human Capital Values and Returns

We describe how to compute human capital values and returns. Let $(v_j(x,z), R_{j+1}(x,z,z'))$ denote the value and the return to human capital. These functions are recursive versions of the values and returns defined in section 2. Human capital values $v_j(x,z)$ follow the recursion $(**)$, given $v_J(x,z)=0$:

$$(**)\ v_j(x,z) = E[m_{j+1}(x,z,z')(v_{j+1}(x',z') + e_{j+1}(z'))|z]$$

$$R_{j+1}(x,z,z') = \frac{v_{j+1}(x',z') + e_{j+1}(z')}{v_j(x,z)}$$

$$m_{j+1}(x,z,z') = \beta\psi_{j+1}(\frac{c_{j+1}(x',z')}{c_j(x,z)})^{-\gamma}(\frac{V_{j+1}(x',z')}{F(V_{j+1}(x',z'))})^{\gamma-\alpha}$$

$$x' = \sum_i a^i_{j+1}(x,z)R^i(z') + e_{j+1}(z')$$

Although the recursive structure above is a step in the right direction, it is not practical to implement because the aggregate component of earnings z_1 "fans out" over time in the benchmark model. Instead, we compute the functions $(\hat{v}_j, \hat{m}_{j+1})$ defined below and then use Lemma 2 to compute values and returns. $\hat{v}_j$ is defined recursively, given $\hat{v}_J = 0$. To compute $(\hat{v}_j, \hat{m}_{j+1})$, we require as inputs the functions $(c_j, a^1_{j+1}, a^2_{j+1}, V_j)$ from the previous sections computed on the restricted domain. In what follows, we write earnings as $e_j(z) = z_1 g_j(z_2)$ and use the fact that $z'_1 = z_1 f_{j+1}(z'_2)$ which is consistent with the model from section 4. In retirement $z'_1 = z_1 f_{j+1}(z'_2) = z_1$.

$$\hat{v}_j(\hat{x},z_2) = E[\hat{m}_{j+1}(\hat{x},z_2,z'_2)f_{j+1}(z'_2)(\hat{v}_{j+1}(\hat{x}',z'_2) + g_{j+1}(z'_2))|z_2]$$

$$\hat{m}_{j+1}(\hat{x},z_2,z'_2) \equiv \beta\psi_{j+1}(\frac{f_{j+1}(z'_2)c_{j+1}(\hat{x}',1,z'_2)}{c_j(\hat{x},1,z_2)})^{-\gamma}(\frac{f_{j+1}(z'_2)V_{j+1}(\hat{x}',1,z'_2)}{F(f_{j+1}(z'_2)V_{j+1}(\hat{x}',1,z'_2))})^{\gamma-\alpha}$$

$$\hat{x}' \equiv \frac{\sum_i a^i_{j+1}(\hat{x},1,z_2)R^i(z'_2) + f_{j+1}(z'_2)g_{j+1}(z'_2)}{f_{j+1}(z'_2)}$$

Lemma 2 says that the value of human capital is proportional to z_1 other things equal and after correcting for financial asset holdings. It also says that the stochastic discount factor and the return to human capital are independent of the level of z_1, after correcting for financial asset holdings. Lemma 2 and the associated formulas allow the computation of statistics of (v_j, R_j) over the lifetime by means of simulating lifetime draws of z_2 shocks and using $\hat{v}_j$ and the computed decision rules.

Lemma 2: In the benchmark model the following hold when $\hat{x} = x/z_1$:

(i) $m_{j+1}(x,z,z') = \hat{m}_{j+1}(\hat{x},z_2,z'_2)$

(ii) $v_j(x,z) = z_1 v_j(\hat{x},1,z_2) = z_1\hat{v}_j(\hat{x},z_2)$

(iii) $R_{j+1}(x,z,z') = \frac{f_{j+1}(z'_2)(\hat{v}_{j+1}(\hat{x}',z'_2)+g_{j+1}(z'_2))}{\hat{v}_j(\hat{x},z_2)}$.

Proof: (i) The result follows from direct substitution of the consumption and the value function into the definition of $m_{j+1}(x,z,z')$. Here we use Lemma 1 so that $c_j(x,z) = z_1 c_j(\frac{x}{z_1}, 1, z_2)$ and $V_{j+1}(x',z') = z_1' V_{j+1}(\frac{x'}{z_1'}, 1, z_2')$. We also use the fact that $z_1' = z_1 f_{j+1}(z_2')$ and that F is homogeneous of degree 1.

(ii) Lemma 2(ii) holds trivially for $j = J$. We show it holds for j given it holds for $j+1$. The first line below uses the definition and the induction hypothesis. The leftmost equality in the second line follows from the first line, Lemma 2(i) and the induction hypothesis. The rightmost equality follows from the definition of $\hat{v}_j$.

$$v_j(x,z) = E[m_{j+1}(x,z,z')(z_1' v_{j+1}(\hat{x}',1,z_2') + z_1' g_{j+1}(z_2'))|z]$$

$$v_j(x,z) = E[\hat{m}_{j+1}(\hat{x},z_2,z_2')(z_1'\hat{v}_{j+1}(\hat{x}',z_2') + z_1' g_{j+1}(z_2'))|z] = z_1 \hat{v}_j(\hat{x},z_2)$$

(iii) The first line follows from the definition, Lemma 2(ii) and the structure of earnings. The second line follows from the first and $z_1' = z_1 f_{j+1}(z_2')$.

$$R_{j+1}(x,z,z') = \frac{v_{j+1}(x',z') + e_{j+1}(z')}{v_j(x,z)} = \frac{z_1'\hat{v}_{j+1}(\hat{x}',z_2') + z_1' g_{j+1}(z_2')}{z_1 \hat{v}_j(\hat{x},z_2)}$$

$$R_{j+1}(x,z,z') = \frac{f_{j+1}(z_2')(\hat{v}_{j+1}(\hat{x}',z_2') + g_{j+1}(z_2'))}{\hat{v}_j(\hat{x},z_2)}$$

$\diamond$

An algorithm to compute the naive value $v_j^n(z)$ is provided. First, we list some useful points from theory, where $z = (z_1, z_2)$, $e_j(z) = z_1 g_j(z_2)$ and $z_1' = z_1 f_{j+1}(z_2')$. The first two equations are Bellman equations. The third equation is an implication of theory. It follows from the first equation by backwards induction and substituting in for earnings.

$$v_j^n(z) \equiv E[\frac{1}{1+r}(v_{j+1}^n(z') + e_{j+1}(z')|z]$$

$$\hat{v}_j^n(z_2) \equiv E[\frac{f_{j+1}(z_2')}{1+r}(\hat{v}_{j+1}^n(z_2') + g_{j+1}(z_2')|z_2]$$

$$v_j^n(z) = z_1 \hat{v}_j^n(z_2)$$

The algorithm is as follows. Step 1: compute the functions $\hat{v}_j^n(z_2)$ by iterating on Bellman's equation. Step 2: simulate histories of z_2 shocks. Step 3: compute z_1 histories using step 2 and $z_1' = z_1 f_{j+1}(z_2')$. Step 4: compute histories $v_j^n(z)$ using (i) $v_j^n(z) = z_1 \hat{v}_j^n(z_2)$, (ii) $\hat{v}_j^n(z_2)$ from step 1, (iii) shock histories from steps 2-3.

A.2.3 Decomposing Human Capital Values

We decompose the value of human capital into a bond, a stock and a residual component. We then calculate the bond and stock shares of human capital at different ages and states. To do so, apply the Projection Theorem to the payout $y = v_{j+1}(x',z') + e_{j+1}(z')$. By construction, the residual $\epsilon \equiv y - \alpha^b R^b + \alpha^s R^s$ is orthogonal to each asset return.

$$v_j(x,z) = E[m_{j+1}y] = E[m_{j+1}(\alpha^b R^b + \alpha^s R^s + \epsilon)]$$

$$v_j(x,z) = \alpha^b E[m_{j+1}R^b] + \alpha^s E[m_{j+1}R^s] + E[m_{j+1}\epsilon]$$

$$share_j^i(x,z) \equiv \frac{\alpha_j^i(x,z)E[m_{j+1}R^i]}{v_j(x,z)} \; for \; i = s,b$$

Calculate (α^b, α^s) by solving the system below, using the relevant conditional expectation. The system imposes that ϵ is orthogonal to each return.

$$\alpha^b E[R_b^2] + \alpha^s E[R_s R_b] = E[yR_b]$$

$$\alpha^b E[R_b R_s] + \alpha^s E[R_s^2] = E[yR_s]$$

Lemma 3 below is useful in theory and computation. It says that in the decomposition defined above, $share_j^i(x,z)$ is invariant to scaling up or down (x, z_1). Thus, shares can be computed for a single value $z_1 = 1$ to determine the share decomposition for all z_1 values.

Lemma 3: In the benchmark model the following holds for $i = s,b$:

$$share_j^i(\lambda x, \lambda z_1, z_2) = share_j^i(x, z_1, z_2), \forall \lambda > 0$$

Proof: The first line is the definition of the share. The second line uses Lemma 1 and the fact that the solution (α^b, α^s) to the linear system scales linearly in (x, z_1). This latter fact holds as the payout scales linearly in (x, z_1). To show this, write the payoff: $y = v_{j+1}(x', z_1 f_{j+1}(z_2'), z_2') + z_1 f_{j+1}(z_2')g_{j+1}(z_2')$. The payoff scales in (x, z_1) because v_{j+1} scales in its first two components (Lemma 2(ii)) and x' scales in (x, z_1). Lemma 3 then follows if $E[m_{j+1}R^i|\lambda x, \lambda z_1, z_2]$ is constant in λ. This holds because financial asset returns R^i depend only on z_2', z_2 is Markov and m_{j+1} is homogeneous of degree zero in (x, z_1) by Lemma 2(i).

$$share_j^i(\lambda x, \lambda z_1, z_2) = \frac{\alpha_j^i(\lambda x, \lambda z_1, z_2)E[m_{j+1}R^i|\lambda x, \lambda z_1, z_2]}{v_j(\lambda x, \lambda z_1, z_2)}$$

$$share_j^i(\lambda x, \lambda z_1, z_2) = \frac{\lambda \alpha_j^i(x, z_1, z_2)E[m_{j+1}R^i|\lambda x, \lambda z_1, z_2]}{\lambda v_j(x,z)}$$

$$share_j^i(\lambda x, \lambda z_1, z_2) = \frac{\alpha_j^i(x, z_1, z_2)E[m_{j+1}R^i|\lambda x, \lambda z_1, z_2]}{v_j(x,z)}$$

$\diamond$

A.2.4 Discretization of Stochastic Processes

For computation, we discretize the idiosyncratic component of earnings, the aggregate component of earnings and stock returns. We construct our discrete approximations to the estimated stochastic processes as follows.

For the idiosyncratic earnings process, both the persistent and transitory innovations are drawn from a three-parameter Generalized Normal distribution. The (inverse) map from the mean, variance and third central moment to the parameters of the Generalized Normal is unique. Thus, we set model parameters consistent with the values in Table 1 as indicated in Table 4. For a given number of grid points, we choose the upper and lower bounds and the (non-linear) spacing of the grid points for each component to minimize the distance between the average variance and average skewness over the lifecycle of the discretized process, and the corresponding moments from the estimated continuous process, excluding the estimated trend. We fill in transition probabilities by assigning to each grid point the mass implied by the continuous Generalized Normal distribution between the mid-points of adjacent grid points. We construct separate discretizations for booms and recessions. We use 7 grid points for the transitory component, 11 grid points for the persistent component and 3 grid points for the fixed effect. Table A.1 reports moments from the continuous and discretized persistent processes for the baseline model. The moments for the transitory process are matched exactly, for each age, by construction.

For the aggregate earnings process, we assume that innovations are drawn from a joint normal distribution. We use equally spaced grid points in each dimension where the upper and lower bounds are a constant multiple of the unconditional variance in the respective dimensions. We choose this multiple to minimize the distance between the second moments of the underlying estimated continuous process and the discretized process. We use 4 grid points for aggregate earnings growth, 5 grid points for equity returns and in the models that include cointegration we include 3 grid points for the cointegrating vector. Table A.1 reports moments from the continuous and discretized processes for the baseline model.

Table A.1: Model Moments

	Full Sample	College Sample	High School Sample
Idiosyncratic Component			
Variance persistent component: true	0.340	0.339	0.193
Variance persistent component: discretized	0.324	0.329	0.184
Skewness persistent component: true	-1.14	-0.738	-4.73
Skewness persistent component: discretized	-1.20	-0.815	-4.43
Aggregate Component			
Variance aggregate earnings growth: true	0.00632	0.00539	0.00918
Variance aggregate earnings growth: discretized	0.00657	0.00560	0.00953
Variance stock returns: true	0.0349	0.0351	0.0350
Variance stock returns: discretized	0.0333	0.0336	0.0334

We impose a constant lower bound on the possible realizations of the combined process for total earnings, in order to minimize numerical inaccuracies that result from very low probability low earnings realizations.We set this minimum at 5% of average earnings. The minimum does not bind except for the high school sample at young ages where there is large amount of negative skewness in persistent shocks.

A.3 Time Series Model

A.3.1 Full Description of Stochastic Model for Aggregate Variables

We assume the following general VAR model for $y_t = \begin{pmatrix} u_t^1 & P_t \end{pmatrix}$:

$$y_t = v(t) + \sum_{i=1}^{p} A_i y_{t-i} + \varepsilon_t$$

where ε_t is a vector of mean zero IID random variables with covariance matrix Σ. $v(t)$ is a quadratic time trend which is parameterized below. We restrict attention to values of $p \leq 2$ to keep the state space manageable. This model has a general VECM form given by

$$\Delta y_t = v + \delta t + \alpha\beta' y_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta y_{t-i} + \varepsilon_t$$

where the vector β is known as the cointegrating vector. We can split the constant and trend terms into components as follows

$$\begin{aligned} v &= \alpha\mu + \gamma \\ \delta t &= \alpha\rho t + \tau t \end{aligned}$$

where $\gamma'\alpha\mu = 0$ and $\tau'\alpha\rho = 0$. In this case the VECM model can be written as

$$\Delta y_t = \gamma + \tau t + \alpha\left(\beta' y_{t-1} + \mu + \rho t\right) + \sum_{i=1}^{p-1} \Gamma_i \Delta y_{t-i} + \varepsilon_t$$

It is useful to define the one-dimensional object $w_t = \beta' y_t + \mu + \rho(t+1)$ which, in the case of cointegration ($\alpha \neq 0$), is a stationary random variable. By construction, w_t evolves as

$$w_t = w_{t-1} + \beta' \Delta y_t + \rho$$

Since we would like a system where w_t evolves based on variables at $t-1$ or earlier, we can re-write this as

$$\begin{aligned} w_t &= w_{t-1} + \beta'\gamma + \beta'\tau t + \beta'\alpha w_{t-1} + \sum_{i=1}^{p-1} \beta'\Gamma_i \Delta y_{t-i} + \beta'\varepsilon_t \\ &= \left(\beta'\gamma + \rho\right) + \beta'\tau t + \left(1 + \beta'\alpha\right) w_{t-1} + \sum_{i=1}^{p-1} \beta'\Gamma_i \Delta y_{t-i} + \beta'\varepsilon_t \end{aligned}$$

In all of our analyses we assume that $\tau = 0$. The general system can then be written as

$$\begin{pmatrix} \Delta y_t \\ w_t \end{pmatrix} = \begin{pmatrix} \gamma \\ \beta'\gamma + \rho \end{pmatrix} + \begin{pmatrix} \Gamma & \alpha \\ \beta'\Gamma & 1 + \beta'\alpha \end{pmatrix} \begin{pmatrix} \Delta y_{t-1} \\ w_{t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_t \\ \beta'\varepsilon_t \end{pmatrix}$$

Table A.2: Lag-Order Selection Tests for VAR

Lag	P-Value	FPE	AIC	HQIC	SBIC
		Full Sample			
0		.001762	-0.666	-0.635	-0.579
1	0.000	.000020	-5.169	-5.077	-4.911
2	0.004*	.000016*	-5.370*	-5.217*	-4.939*
3	0.160	.000018	-5.333	-5.118	-4.729
		College Sub-sample			
0		.001882	-0.600	-0.569	-0.514
1	0.000	.000015	-5.422	-5.330*	-5.163*
2	0.078*	.000015*	-5.432*	-5.279	-5.001
3	0.374	.000018	-5.248	-4.973	-4.744
		High School Sub-sample			
0		.003469	0.012	0.042	0.098
1	0.000	.000028	-4.801	-4.709	-4.543
2	0.014*	.000025*	-4.919*	-4.766*	-4.488*
3	0.132	.000028	-4.837	-4.561	-4.257

Notes: Lag order selected by each criteria is denoted by *. P-values are from likelihood ratio tests of the null that true lag length is $p-1$ or less.

We adopt the Johansen (1995) normalization for β, which implies for the two variable case that we can write

$$\Delta w_t = \Delta u_t^1 + \beta_2 \log R_t^s + \rho$$

The model with $p = 2$ becomes

$$\begin{pmatrix} \Delta u_t^1 \\ \log R_t^s \\ w_t \end{pmatrix} = \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_1 + \beta_2\gamma_2 + \rho \end{pmatrix} + \begin{pmatrix} \Gamma_{11} & \Gamma_{12} & \alpha_1 \\ \Gamma_{21} & \Gamma_{22} & \alpha_2 \\ \Gamma_{11} + \beta_2\Gamma_{21} & \Gamma_{12} + \beta_2\Gamma_{22} & 1 + \alpha_1 + \beta_2\alpha_2 \end{pmatrix} \begin{pmatrix} \Delta u_{t-1}^1 \\ \log R_{t-1}^s \\ w_{t-1} \end{pmatrix} \quad (6)$$

$$+ \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{1t} + \beta_2\varepsilon_{2t} \end{pmatrix} \quad (7)$$

A.3.2 Estimation of Aggregate Stochastic Process

We test for the lag order of the underlying VAR. Table A2 reports results from likelihood ratio tests and a number of commonly used statistical information criteria. All criteria suggest a lag length of $p = 2$. The college and high-school sub-samples, and alternative measures of aggregate earnings and alternative sample periods all indicate the presence of two lags.

We also test for the presence of cointegration. Table A3 reports results from tests of the cointegrating rank based on the methods in Johansen (1995). Our results suggest only very weak evidence for cointegration. Alternative variable definitions, specifications and time periods lead to similar results.

Table A4 presents the average moments in the data together with the implied steady-state statistics from the model for three different data samples: the full sample from the CPS 1967-2008, NIPA 1967-2008 and NIPA 1929-2009. The NIPA measure of earnings growth is the change in the log of total wages and salaries per member of the labor force.

A.3.3 Benzoni et. al. (2007)

We describe the construction of the system of equations underlying the results in Figure 9 from section 6. The three equations below are equation 2, 8 and 14 from Benzoni et al. (2007), where y_t is log

Table A.3: Cointegration Rank Selection Tests

Maximum Rank	Trace Statistic	5% Critical Value	Eigenvalue	SBIC	HQIC
Full Sample					
0	9.95*	15.41	0.220*	-5.05*	-5.21
1	0.00	3.76	0.000	-5.02	-5.26*
College Sub-sample					
0	9.49*	15.41	0.200*	-5.08*	-5.247
1	0.55	3.76	0.014	-5.03	-5.274*
High School Sub-sample					
0	9.91*	15.41	0.218*	-4.66*	-4.824
1	0.07	3.76	0.002	-4.63	-4.875*

Notes: Rank of cointegration by each criteria is denoted by *. Trace statistic criteria is obtained by selecting the lowest rank that cannot be rejected.

Table A.4: Implied Steady-State Statistics: Alternative Data Sources and Sample Periods

	No Cointegration					
	CPS: 1967-2008		NIPA: 1967-2008		NIPA: 1929-2009	
	Data	Model	Data	Model	Data	Model
$E\left(\log R_t^s\right)$	0.041	0.045	0.041	0.044	0.068	0.069
$E\left(\Delta u_t^1\right)$	-0.002	-0.004	0.004	0.001	0.012	0.012
$sd\left(\Delta u_t^1\right)$	0.025	0.025	0.024	0.025	0.029	0.029
$sd\left(\log R_t^s\right)$:	0.187	0.187	0.187	0.187	0.178	0.178
$corr\left(\Delta u_t^1, \log R_t^s\right)$	0.184	0.177	0.234	0.216	0.070	0.071
$corr\left(\Delta u_t^1, \Delta u_{t-1}^1\right)$	0.425	0.441	0.429	0.460	0.398	0.384
$corr\left(\log R_t^s, \log R_{t-1}^s\right)$	0.057	0.055	0.058	0.057	-0.054	-0.052
$corr\left(\Delta u_t^1 \log R_{t-1}^s\right)$	0.372	0.398	0.640	0.685	0.680	0.675
$corr\left(\log R_t^s, \Delta u_{t-1}^1\right)$	-0.292	-0.270	-0.189	-0.194	-0.096	-0.096
	With Cointegration					
	CPS: 1967-2008		NIPA: 1967-2008		NIPA: 1929-2009	
	Data	Model	Data	Model	Data	Model
$E\left(\log R_t^s\right)$	0.041	0.070	0.041	0.070	0.068	0.045
$E\left(\Delta u_t^1\right)$	-0.002	-0.002	0.004	0.005	0.012	0.004
$sd\left(\Delta u_t^1\right)$	0.025	0.025	0.024	0.024	0.029	0.026
$sd\left(\log R_t^s\right)$:	0.187	0.187	0.187	0.182	0.178	0.172
$corr\left(\Delta u_t^1, \log R_t^s\right)$	0.184	0.155	0.234	0.178	0.070	-0.020
$corr\left(\Delta u_t^1, \Delta u_{t-1}^1\right)$	0.425	0.435	0.429	0.435	0.398	0.260
$corr\left(\log R_t^s, \log R_{t-1}^s\right)$	0.057	0.005	0.058	-0.007	-0.054	-0.096
$corr\left(\Delta u_t^1 \log R_{t-1}^s\right)$	0.372	0.394	0.640	0.664	0.680	0.660
$corr\left(\log R_t^s, \Delta u_{t-1}^1\right)$	-0.292	-0.283	-0.189	-0.213	-0.096	-0.176

Notes: Table shows average moments in the data, together with implied steady-state statistics from the corresponding estimated model. NIPA data is total wage and salaries per member of the labor force. When implementing the estimated processes in the structural model, we adjust the constants (γ_1, γ_2) so that all models have $E[\log R_t^s] = 0.041$ and $E[\Delta u_t^1] = 0$.

Table A.5: Implied Steady-State Statistics: Two Models

	Benchmark Model	Benzoni Process
$sd\left(\Delta u_t^1\right)$	0.025	0.069
$sd\left(\log R_t^s\right)$:	0.187	0.160
$corr\left(\Delta u_t^1, \log R_t^s\right)$	0.177	0.000
$corr\left(\Delta u_t^1, \Delta u_{t-1}^1\right)$	0.441	0.327
$corr\left(\log R_t^s, \log R_{t-1}^s\right)$	0.055	0.000
$corr\left(\Delta u_t^1 \log R_{t-1}^s\right)$	0.398	0.347
$corr\left(\log R_t^s, \Delta u_{t-1}^1\right)$	-0.270	0.000

Notes: Table shows implied steady-state statistics. When implementing the estimated processes in the structural model, we adjust the constant terms so that all models have $E[\log R_t^s] = 0.041$ and $E[\Delta u_t^1] = 0$.

dividends, R_t is the gross stock return and u_t is the log of the common component of earnings. The parameter κ is the key adjustment parameter controlling the strength of cointegration.

$$dy_t = (g - \sigma^2/2)dt + \sigma dz_3$$

$$R_t - 1 = \mu dt + \sigma dz_3$$

$$d(u_t - y_t - \bar{uy}) = -\kappa(u_t - y_t - \bar{uy})dt + \nu_1 dz_1 - \nu_3 dz_3$$

The three equations below are a discrete-time approximation of this continuous-time process, where $(z_{1,t}, z_{3,t})$ are independent standard normal random variables. We rewrite this system of equations as system (8) below, using $\Delta u_t \equiv u_t - u_{t-1}$ and $w_t \equiv (u_t - y_t - \bar{uy})$.

$$y_{t+1} - y_t = g - \sigma^2/2 + \sigma z_{3,t+1}$$

$$\log R_{t+1} = \mu + \sigma dz_{3,t+1}$$

$$(u_{t+1} - y_{t+1} - \bar{uy}) - (u_t - y_t - \bar{uy}) = -\kappa(u_t - y_t - \bar{uy}) + \nu_1 dz_{1,t+1} - \nu_3 dz_{3,t+1}$$

$$\begin{pmatrix} \Delta u_{t+1} \\ \log R_{t+1} \\ w_{t+1} \end{pmatrix} = \begin{pmatrix} g - \sigma^2/2 \\ \mu \\ 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & -\kappa \\ 0 & 0 & 0 \\ 0 & 0 & 1-\kappa \end{pmatrix} \begin{pmatrix} \Delta u_t \\ \log R_t \\ w_t \end{pmatrix} + \begin{pmatrix} \nu_1 z_{1,t+1} + (\sigma - \nu_3) z_{3,t+1} \\ \sigma z_{3,t+1} \\ \nu_1 z_{1,t+1} - \nu_3 z_{3,t+1} \end{pmatrix} \tag{8}$$

System (8) produces the steady-state statistics listed in Table A5. This occurs when we set $(\kappa, g, \nu_1, \nu_3, \sigma, \mu) = (.15, .018, .05, .16, .16, .07)$, which are the benchmark parameter values used by Benzoni et al. (2007), and when we alter the constant terms in (8) to produce the same steady-state mean values $(E[\log R], E[\Delta u])$ as in both of our benchmark models. This leaves the variance-covariance properties of the model unchanged.