NBER WORKING PAPER SERIES

THE MONEY VALUE OF A MAN

Mark Huggett
Greg Kaplan

Working Paper 18066
http://www.nber.org/papers/w18066

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2012

The Money Value of a Man
Mark Huggett and Greg Kaplan
NBER Working Paper No. 18066
May 2012
JEL No. D91,E21,G12,J24

ABSTRACT

This paper posits a notion of the value of an individual's human capital and the associated return on human capital. These concepts are examined using U.S. data on male earnings and financial asset returns. We decompose the value of human capital into a bond, a stock and a residual value component. We find that (1) the bond component of human capital is larger than the stock component at all ages, (2) the value of human capital is far below the value implied by discounting earnings at the risk-free rate, (3) mean human capital returns exceed stock returns early in life and decline with age and (4) human capital returns and stock returns have a small positive correlation over the working lifetime.

Mark Huggett
Georgetown University
mh5@georgetown.edu

Greg Kaplan
Department of Economics
University of Pennsylvania
160 McNeil Building
3718 Locust Walk
Philadelphia, PA 19104
and NBER
gkaplan@sas.upenn.edu

Our basic objective is to figure out the value of any stream of uncertain cash flows.

- from Asset Pricing by John Cochrane (2001, p. 6)

1 Introduction

A common view is that by far the most valuable asset that most people own is their human capital. Our goal is to figure out the value of an individual's human capital and the associated return on human capital. The value and return concepts that we analyze are uncontroversial. The value of human capital equals the expected discounted value of future dividends, where discounting is done using the individual's stochastic discount factor. The return to human capital is the future value and dividend divided by the current value.

To the best of our knowledge, there is little work which adopts these concepts and then undertakes a detailed empirical analysis of the value and return to an individual's human capital. This seems remarkable given that the basic objective of asset pricing is to figure out the value of any stream of uncertain cash flows. Looking forward, we believe that these value and return concepts will become central in connecting literatures that many economists currently view as being disconnected.

We highlight four literatures where an empirical understanding of human capital values and returns at the level of the individual is relevant. First, consider the literature on portfolio allocation over the lifetime. This literature tries to understand portfolio choices and to give practical advice on portfolio allocation. If human capital is by far the most valuable asset that most individuals own, then a good starting point for giving portfolio advice would be to understand the relative magnitude of the stock and bond positions implicit in the value of an individual's human capital. Thus, it is key to decompose human capital values into useful components.

Second, there is a literature on the international diversification puzzle. Baxter and Jermann (1997) argue that (i) human capital is more valuable than physical capital, (ii) the return to domestic physical and human capital are very highly correlated and (iii) a very large negative position in domestic assets is a very good hedge on the return to domestic human capital. They conclude that a diversified international portfolio involves a negative position in domestic marketable assets for all the developed countries that they analyze. This reasoning, in their estimation, deepens the puzzling lack of international financial diversification. Although this line of argument is based on the value and return to a claim to economy-wide earnings, an individual directly holds only a claim to his or her own future labor earnings. Thus, a first step towards reevaluating this line of argument is to analyze the degree to which individual human capital returns are like domestic stock returns.

Third, Alvarez and Jermann (2004) argue that theory implies that the marginal benefit of moving towards a smoother consumption plan (e.g. by implementing social welfare policies that effectively complete markets) is an upper bound to the total benefit. We note that marginal benefit calculations are closely tied to the value of an individual's human capital. Specifically, the marginal benefit for an individual is given by the ratio of the value of the smooth consumption plan to the value of human capital plus initial financial assets. By this logic, low values of human capital are associated with high values to perfect consumption smoothing. Thus, the value of an individual's human capital connects with long-standing debates about the magnitude of the potential gains to business-cycle smoothing or to perfect consumption insurance.

Fourth, there is a literature that estimates parameters of Epstein-Zin preferences governing risk aversion and intertemporal substitution based on Euler equations. Epstein and Zin (1991) argue that the future utility terms that enter the Euler equation can be replaced by a function of the return on the overall wealth portfolio. In practice, they proxy this return with stock returns but acknowledge that this return should also reflect the return to human capital. Vissing-Jorgensen and Attanasio (2003) apply these ideas to estimate these preference parameters using household-level data. They assume that the unobserved expected return on human capital is a weighted average of stock and bond returns, where the weights are age and state invariant. In our view, an analysis of human capital values and returns would provide a useful perspective on such an empirical strategy.

This paper has two main contributions. First, we provide a theoretical justification for defining the value of human capital as future dividends discounted using an individual's stochastic discount factor. Specifically, we describe a large class of decision problems for which allowing an agent the option to buy or sell shares in his or her future dividend stream at this defined value results in the agent optimally deciding to continue to hold all of these human capital shares. This result clarifies the theoretical foundations for the value and return concepts that we employ and the scope for applications.[1]

Second, we provide a detailed characterization of the value and return to human capital using male earnings data and financial asset return data. To do this we use a two-step procedure. We 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. This statistical model is then embedded into a decision problem of the type analyzed in the literature on the income-fluctuation

[1] The idea that one can price a non-traded asset using marginal rates of substitution that result from best decisions, absent trade in this asset, is not new. Lucas (1978), Bodie, Merton and Samuelson (1992) and Svensson and Werner (1993) price non-traded assets in this way. We revisit this issue as we have not seen this idea pursued at the depth that is needed to clarify to empirical researchers both the scope for applications as well as the limits of this notion of the value of human capital.

problem. We characterize the properties of the implied human capital values and returns by using the stochastic discount factor produced by a solution to this decision problem to discount future earnings.

We find that the value of human capital is far below the value that would be implied by discounting future earnings at the risk-free interest rate. One 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 risk. We decompose the value at each age into three components: a bond, a stock and an orthogonal component. This is done by projecting future human capital pay outs (i.e. the sum of next period's earnings and human capital value) onto next period's bond and stock returns. The bond and stock components, stated as a ratio to the value of human capital, are both positive on average but the bond component greatly exceeds the stock component at each age. This holds when the earnings data is for males with a high school or with a college education. The stock component as a ratio to the value of human capital is on average larger for the high school than for the college education group.

The value of the orthogonal component is strongly negative early in life but tends to zero as retirement approaches. The negative value early in life is mostly due to the presence of a highly persistent idiosyncratic component of earnings variation which is orthogonal to stock and bond returns. This component covaries negatively with the stochastic discount factor as a positive earnings shock leads to an increase in consumption.

We find that the mean return to human capital falls with age over the working lifetime. Moreover, the mean return greatly exceeds the return to stock early in the lifetime. Both of these results are related to our findings on the value of the orthogonal component. We show that the mean human capital return always equals a weighted sum of the mean stock and bond returns. The weights are determined by the projection coefficients from the value decomposition. These weights sum to more than one exactly when the value of the orthogonal component is negative. We also find that human capital returns and stock returns have only a small positive correlation over the working lifetime. This correlation is higher for high school than for college-educated males.

The remainder of the paper is organized as follows. Section 2 outlines the literature most closely related to our work. Section 3 presents the theoretical framework. Section 4 presents our analysis of human capital values and returns. Section 5 summarizes our main findings and discusses the sensitivity of our benchmark results on the decomposition of human capital values into bond and stock components.

2 Related Literature

A long line of research calculates the "money value of a man" by discounting an individual's future dividends d_k, usually measured as earnings, at a deterministic interest rate r as follows:[2] $v_j = E_j[\sum_{k>j} \frac{d_k}{(1+r)^{k-j}}]$. Our definition of the value of human capital differs because we allow for covariation between an individual's stochastic discount factor and dividends, whereas the literature referenced above does not. The definition of value $v_j = E_j[\sum_{k>j} \frac{d_k}{(1+r)^{k-j}}]$ is problematic for analyzing human capital returns. Specifically, the mean return to human capital implied by this value always equals the deterministic interest rate used to discount dividends: $E_j[R_{j+1}] \equiv E_j[\frac{v_{j+1}+d_{j+1}}{v_j}] = 1 + r$.

The finance literature has analyzed the return to human capital. We mention two lines of work. First, Campbell (1996), Baxter and Jermann (1997) and many others characterize the return to human capital using aggregate data. Our work differs as we value a claim to an individual's earnings rather than a claim to aggregate, economy-wide earnings. Second, Huggett and Kaplan (2011) characterize human capital values and returns using individual-level earnings data. They put bounds on values and returns based on (i) non-parametric restrictions of stochastic discount factors, (ii) knowledge of the process governing individual earnings and asset returns and (iii) the assumption that Euler equations hold. The bounds turn out to be quite loose. They conclude that a structural approach, like that carried out in the present paper, is critical to gain a sharper understanding of individual-level human capital values and returns.

There is a vast literature on the Mincerian return, which is measured as the coefficient on additional years of schooling in a Mincerian earnings regression. The Mincerian return has sometimes been viewed as a rough measure of a marginal return to schooling.[3] Instead of analyzing marginal returns on specific marginal decisions, we analyze the return on a claim to all the dividends received by an individual.

Our work also connects to the literature on portfolio allocation decisions over the lifetime (e.g. Campbell, Cocco, Gomes and Maenhout (2001), Benzoni, Collin-Dufresne and Goldstein (2007), Lynch and Tan (2011) among many others). Whereas this literature analyzes portfolio choice, our focus is on characterizing properties of the value and return to human capital when individual earnings are es-

[2] This procedure goes at least as far back as the work of Farr (1853, Table VII). It is also used by Dublin and Lotka (1930), Weisbrod (1961), Becker (1975), Lillard (1977), Graham and Webb (1979), Jorgenson and Fraumeni (1989) and Haveman, Bershadker and Schwabish (2003) among others. One objective of this line of research is to determine the aggregate value of human capital and compare it to the aggregate value of physical capital.

[3] Heckman, Lochner and Todd (2006) review this literature. Palacios-Huerta (2003) constructs an empirical approximation to the marginal returns to an additional year of schooling using group earnings data for groups differing in educational attainment.

timated so as to display the rich sources of variation due to age, education as well as aggregate and idiosyncratic shocks. Section 5 discusses how our results relate to the portfolio allocation results of the latter two papers.

3 Theoretical Framework

This section defines the value of human capital, gives a justification for this value and provides a simple example economy to illustrate the value and return concepts. The justification offered is not sensitive to the precise mechanism by which labor market earnings are determined. Thus, empirical researchers can pursue the same notion of the value and return to human capital while investigating vastly different theories of earnings.

3.1 Decision Problem

An agent solves Problem P1. The objective is to maximize lifetime utility. Lifetime utility $U(c,n)$ is determined by lifetime consumption and leisure plans (c,n), where $c = (c_1, ..., c_J)$ and $n = (n_1, ..., n_J)$. Consumption and leisure at age j are given by functions $c_j : Z^j \to R^1_+$ and $n_j : Z^j \to [0,1]$ that map shock histories $z^j = (z_1, ..., z_j) \in Z^j$ into the corresponding values of these variables. All the variables that we analyze are functions of these shocks.

Problem P1: $\max U(c,n)$ 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 $c_j \geq 0$

(2) $e_j = G_j(y^j, n^j, z^j)$ and $0 \leq n_j \leq 1$

(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 earnings 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 $i \in \mathcal{I} = \{1, ..., I\}$ in the previous period and by the gross return $R^i_j > 0$ to each asset i.

Problem P1 puts loose restrictions on the way in which earnings are determined. Earnings e_j are a kitchen-sink function G_j of age j, shocks z^j, leisure decisions $n^j = (n_1, ..., n_j)$ and other decisions $y^j = (y_1, ..., y_j)$. This formulation captures models where earnings (i) are exogenous, (ii) equal the product of work time and an exogenous wage and (iii) are determined by many different human capital theories. For example, within human capital theory a standard formulation (see Ben-Porath (1967) or Heckman (1976)) is that earnings $e_j = w_j h_j l_j$ equal the product of an exogenous rental rate w_j, human capital or skill $h_j = H_j(y^{j-1}, n^{j-1})$ and work time $l_j = L(y_j, n_j)$. Problem P1 captures this standard formulation (i.e. $G_j(y^j, n^j, z^j) = w_j H_j(y^{j-1}, n^{j-1}) L(y_j, n_j)$) among other possibilities.[4]

3.2 Value and Return Concepts

We now define and justify our notion of the value of human capital. To do so, we confront the agent who faces Problem P1 with a different market structure given by Problem P2. In this market structure there is a firm that is mandated to follow the human capital decisions that shape earnings in a solution to Problem P1 and to sell units of leisure to the agent at a price p_j. This firm produces a dividend d_j in period j equal to the value of earnings and leisure produced in the period. The agent has the option of buying or selling shares in this firm at a price of v_j and can buy leisure time at price p_j. The prices (v_j, p_j) are personalized prices for the agent and not prices at which other agents can trade. Owning shares in this firm (total shares are normalized to 1) at the end of a period gives the owner a proportionate claim to future dividends.

Problem P2: $\max U(c, n)$ subject to

(1) $c_j + \sum_{i \in \mathcal{I}} a^i_{j+1} + s_{j+1} v_j + p_j n_j = \sum_{i \in \mathcal{I}} a^i_j R^i_j + s_j(v_j + d_j)$ and $c_j \geq 0$

(2) $0 \leq n_j \leq 1$

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

To complete the description of Problem P2, we specify human capital values, dividends and leisure prices (v_j, d_j, p_j). Let $(c^*, n^*, e^*, y^*, a^*)$ denote the solution to Problem P1, where unsubscripted variables are lifetime plans (e.g. $c^* = (c^*_1, ..., c^*_J)$). The value of human capital v_j is then defined to equal expected discounted dividends. Discounting is done using the agent's stochastic discount factor $m_{j,k}$ 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 $P(z^k | z^j)$ as, following the literature, we find it convenient to express human capital values in terms of the mathematical expectations operator.

[4] In human capital theory, earnings are determined by decisions beyond a leisure decision. For example, the "other decisions" variable $y_j = (y_{1j}, y_{2j}, y_{3j})$ can capture time allocated to skill production y_{1j}, time devoted to work y_{2j} and market goods input into skill production y_{3j}. If there are market goods inputs, then the concept of earnings in G_j is earnings net of the value of market goods input. One can view the fact that leisure time enters G_j simply as a way to determine which of the remaining uses of one's time are feasible.

The dividends to human capital d_j are the sum of earnings and the value of leisure. The leisure price p_j is the agent's intratemporal marginal rate of substitution in a solution to Problem P1.

$$v_j(z^j) \equiv E[\sum_{k=j+1}^{J} m_{j,k} d_k | z^j] \quad and \quad m_{j,k}(z^k) \equiv \frac{dU(c^*, n^*)/dc_k(z^k)}{dU(c^*, n^*)/dc_j(z^j)} \frac{1}{P(z^k|z^j)}$$

$$d_j(z^j) \equiv e_j^*(z^j) + p_j(z^j) n_j^*(z^j) \quad and \quad p_j(z^j) \equiv \frac{dU(c^*, n^*)/dn_j(z^j)}{dU(c^*, n^*)/dc_j(z^j)}$$

Theorem 1 asserts that when the agent is given the ability to trade away the value implicit in his future earnings and leisure streams, then the agent optimally decides not to do so. Thus, the agent starts every period owning all of these human capital shares and optimally decides to end the period owning all of these shares. The valuation of human capital persuades the agent to do this and to optimally make exactly the consumption, leisure and asset allocation decisions that were optimal in Problem P1.

Assumption A1:

(i) The utility function U is increasing, concave and differentiable in (c, n) and is strictly increasing in its first component.

(ii) For all j, the set Z^j is finite and the probability $P(z^j)$ is strictly positive $\forall z^j \in Z^j$. The gross return $R_j^i(z^j)$ is strictly positive $\forall i \in \mathcal{I}, \forall j, \forall z^j \in Z^j$.

Theorem 1: Assume A1. If $(c^*, n^*, e^*, y^*, a^*)$ solves Problem P1 and c^* is strictly positive, then (c^*, n^*, a^*, s^*) solves Problem P2, where $s_j^* \equiv 1, \forall j$, when the agent takes the value of human capital v_j, dividends d_j and the price of leisure p_j as exogenous.

Proof: See the Appendix.

One important assumption in Theorem 1 is that preferences are concave over consumption-leisure plans. This is key as then Problem P2 is a concave programming problem. This means that necessary conditions to Problem P1 are straightforward sufficient conditions for the conjectured solution to be optimal in Problem P2. This holds even though Problem P1 is not in general a concave program.

Comments:

1. Given the value concept, we define the gross return to human capital in the same way it is done for any marketed asset: $R^h_{j+1} = \frac{v_{j+1} + d_{j+1}}{v_j}$. The return to human capital is then well integrated into standard asset pricing theory in that all returns satisfy the same type of restriction: $E[m_{j,j+1}R_{j+1}|z^j] = 1$. This holds for all financial assets in Problem P1 and P2 by standard Euler equation arguments and for the return to human capital by construction.[5]

2. The value of human capital can be decomposed into a component capturing only future earnings and a component capturing only the value of future leisure. If one sets $v^e_j(z^j) = E[\sum_{k=j+1}^{J} m_{j,k}e^*_k|z^j]$ and $v^n_j(z^j) = E[\sum_{k=j+1}^{J} m_{j,k}p_k n^*_k|z^j]$, then a modified version of Theorem 1 holds. This result offers a justification for valuing the earnings and leisure components separately in empirical work.

3. One might ask whether Theorem 1 can be extended to apply to situations in which financial markets feature additional restrictions that are not present in Problem P1. It is clear that a version of Theorem 1 will hold as long as the resulting version of Problem P2 is a concave program. For example, adding to Problem P1 a short sales constraint or a borrowing constraint restriction requiring that total savings are always above some specified level each period fits this condition. Thus, to adopt the value and return notions advocated in this paper and the justification provided for them does not require that one posit a rich financial structure or that the financial structure is devoid of financial frictions that are often analyzed in applied work.

3.3 Relevance of the Notion of Human Capital Value

The introduction mentioned several literatures for which a quantitative understanding of human capital values and returns at the level of the individual is relevant. We now articulate in concrete terms how these literatures connect to calculations of human capital values and returns.

Portfolio Allocation

The introduction stated that if human capital is by far the most valuable asset that most people own, then a good starting point for giving portfolio advice would be to have an empirical understanding of the relative magnitude of the bond and stock positions implicit in the value of an individual's human capital. We now indicate how to decompose human capital values v_j into useful components. The first equation below holds by the definition of the human capital value and by projecting future human capital payouts $(v_{j+1}+d_{j+1})$ onto the space spanned by future conditional financial asset returns. This

[5] $v_j = E[\sum_{k=j+1}^{J} m_{j,k}d_k|z^j]$ implies $E[m_{j,j+1}(\frac{v_{j+1}+d_{j+1}}{v_j})|z^j] = 1$.

divides the payouts into an asset market component and a component orthogonal to asset returns. Under the assumption that the individual solves Problem P1, the Euler equation $E[m_{j,j+1}R^i_{j+1}|z^j] = 1$ holds for each financial asset that the individual is allowed to hold. Thus, human capital values are decomposed into an asset market component $\sum_{i\in\mathcal{I}} \alpha^i_j$ and the value of the orthogonal component $E[m_{j,j+1}\epsilon_{j+1}|z^j]$. Section 4 of the paper carries out this type of decomposition, where the financial assets are stocks and bonds and the possibility of a corner solution is allowed.

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

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

It is useful to understand how the value decomposition changes as earnings are perturbed. Theorem 2 provides a transparent result on a class of pertubations where the value $v_j(z^j)$ is unchanged in all histories across the two problems yet the decomposition changes. Specifically, the decomposition changes only at age j in history $\bar{z}^j$. In this history the decomposition implies that in the modified problem the agent effectively holds an extra γ^i units of financial asset $i \in \mathcal{I}$ implicit in future earnings. The agent ends up reducing the holdings of financial asset i by γ^i. In section 5 of the paper, we analyze a number of perturbations of earnings in which the portfolio composition of financial assets also moves in the opposite direction from the movements implied by the value decomposition.

Theorem 2: Assume A1. Consider Problem P1 with exogenous earnings e and no leisure. Modify earnings so that $\hat{e}$ *equals* e *except at age* $j+1$ *following history* $\bar{z}^j$*, where* $\hat{e}_{j+1}(\bar{z}^j, z_{j+1}) = e_{j+1}(\bar{z}^j, z_{j+1}) + \sum_{i\in\mathcal{I}} \gamma^i R^i_{j+1}(\bar{z}^j, z_{j+1}) \geq 0, \forall z_{j+1}$ *for any* $(\gamma^1, ..., \gamma^I)$ *satisfying* $\sum_{i\in\mathcal{I}} \gamma^i = 0$.

(i) If (c, a) *solves P1, then* $(c, \hat{a})$ *solves the modified version of P1, where* $\hat{a}$ *equals* a *except in history* $\bar{z}^j$ *and* $\hat{a}^i_{j+1}(\bar{z}^j) = a^i_{j+1}(\bar{z}^j) - \gamma^i, \forall i \in \mathcal{I}$.

(ii) $v_j(z^j)$ *is the same in both problems.*

Proof:

(i) The set of consumption allocations satisfying the constraints in Problem P1 is the same as the set of consumption allocations satisfying the constraints in the modified version of P1. Thus, c solves both problems. If a finances c in P1, then $\hat{a}$ clearly finances c in the modified version of P1.

(ii) By Thm. 2(i), the stochastic discount factors $m_{j,k}$ are the same in both problems. The result then follows from the definition of $v_j(z^j)$. $\diamond$

Welfare Gains Literature

The introduction claimed that the marginal benefit of moving towards a smoother consumption plan is closely tied to the value of human capital. To see this, define the benefit function $\Omega(\alpha)$ using the first equation below. This follows Alvarez and Jermann (2004). The leftmost equality in the second equation follows from differentiating the first equation, whereas the rightmost equality holds because the individual solves Problem P1.[6]

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

$$\Omega'(0) = \frac{\sum_{j=1}^{J}\sum_{z^j}\frac{dU(c,n)}{dc_j(z^j)}(c_j^{smooth}(z^j) - c_j(z^j))}{\sum_{j=1}^{J}\sum_{z^j}\frac{dU(c,n)}{dc_j(z^j)}c_j(z^j)} = \frac{E[\sum_{j=1}^{J} m_{1,j}c_j^{smooth}|z^1]}{v_1^e(z^1) + e_1(z^1) + \sum_{i\in\mathcal{I}} a_1^i(z^1)R^i(z^1)} - 1$$

The result is that the marginal benefit is determined by the ratio of the value of the smooth consumption plan to the value of human capital plus initial financial wealth. Moreover, the marginal benefit function is an upper bound to the total benefit (i.e. $\Omega'(0) \geq \Omega(1)$) under straightforward conditions.[7] The theory is attractive because in some applications financial asset returns pin down the value of the smooth consumption plan. Thus, the numerator of the marginal benefit function can be determined without making parametric assumptions on an individual's preferences. The key remaining part of such a calculation is the value of an individual's human capital based on earnings, which is denoted v_1^e - see Comment 2 to Theorem 1.

Estimating Preference Parameters

There is a literature that estimates risk aversion and intertemporal substitution parameters of Epstein-Zin utility functions. These utility functions in life-cycle applications are generated by recursively applying the aggregator function W and the certainty equivalent F stated below. Epstein and Zin (1991) estimated preference parameters using the representative-agent abstraction. One problem that they faced is that the Euler equations would appear to be intractable as the stochastic discount factor $m_{j,j+1}$ involves not only consumption but also future utility. They claimed that the return R_{j+1}^w on the agent's overall wealth portfolio could be used as a theoretical proxy for the terms involving future utility. This allowed them to restate the Euler equation as indicated below.

[6] More specifically, this follows from converting 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.

[7] Alvarez and Jermann (2004, Proposition 3) establish that Ω is concave when U is increasing, concave and homothetic in the consumption plan. When Ω is concave, then clearly $\Omega'(0) \geq \Omega(1)$.

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

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

$$E[m_{j,j+1}R^i_{j+1}|z^j] = 1 \ and \ m_{j,j+1} = \beta(\frac{c_{j+1}}{c_j})^{-\rho}(\frac{U^{j+1}}{F(U^{j+1})})^{\rho-\alpha}$$

$$E[\beta^{\frac{1-\alpha}{1-\rho}}(\frac{c_{j+1}}{c_j})^{\frac{\rho(\alpha-1)}{1-\rho}}(R^w_{j+1})^{\frac{\rho-\alpha}{1-\rho}}R^i_{j+1}|z^j] = 1$$

Vissing-Jorgensen and Attanasio (2003) face the same basic issues when they attempt to estimate the preference parameters (ρ, α) using household-level consumption data. The return R^w_{j+1} that is theoretically relevant is the value-weighted average of the returns on all financial assets and the return to human capital: $R^w_{j+1} = \sum_{i=1}^{I+1} \tilde{\alpha}^i_j R^i_{j+1}$.[8] In practice, Vissing-Jorgensen and Attanasio (2003) carry out estimation by assuming that the expected human capital return is an age and state invariant weighted average of stock and bond returns. Our work offers some perspective as we show analytically how the expected human capital return is related to a weighted average of expected stock and bond returns and indicate how these weights vary with age and financial wealth.

3.4 A Simple Example

We analyze a simple example to illustrate the value and return concepts. The example is a parametric decision problem which is a special case of Problem P1. An agent's preferences are given by a constant relative risk aversion utility function, earnings are exogenous and there is a single, risk-free asset. As leisure does not enter the utility function, the value of human capital is determined solely by earnings.[9]

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.

Risk-free return: $R^f = (1+r) > 0$

[8] We allow for I financial asset returns and let asset $I+1$ denote human capital. The relevant weights are $\tilde{\alpha}^i_j = a^i_{j+1}/(\sum_{i=1}^{I} a^i_{j+1} + v_j)$ for all financial assets and $\tilde{\alpha}^{I+1}_j = v_j/(\sum_{i=1}^{I} a^i_{j+1} + v_j)$ for human capital, where a^i_{j+1} are the asset choices in a solution to Problem P1 and v_j is the value of human capital.

[9] 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

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

Theorem 3 establishes that there is a solution to the decision problem in which consumption equals earnings. At this solution the value of human capital at any age is proportional to earnings. The return to human capital is lognormally distributed. The mean return to human capital always exceeds the risk-free return for any positive magnitude of the shock variance. Furthermore, the distribution of one-period returns to human capital is invariant to age and the history of past shocks.[10]

Theorem 3: If $1+r = \frac{1}{\beta}\exp(\rho\mu - \frac{\rho^2\sigma^2}{2})$ *and initial assets are zero, then* $(c_j, a_{j+1}) = (e_j, 0), \forall j$ *solves the decision problem. Furthermore, at this solution:*

(i) the value of human capital is $v_j = f_j e_j$*, where* $f_j = \sum_{k=j+1}^{J} \beta^{k-j} \exp((k-j)[(-\rho+1)\mu + (-\rho+1)^2\frac{\sigma^2}{2}])$.

(ii) the return to human capital satisfies (a)-(c):

(a) $R^h_{j+1} = (\frac{1+f_{j+1}}{f_j})z_{j+1}$

(b) $E[R^h_{j+1}|z^j] = \frac{1}{\beta}\exp(\mu\rho + \frac{\sigma^2}{2}(1-(1-\rho)^2))$

(c) $E[R^h_{j+1} - R^f|z^j] = \frac{\exp(\mu\rho - \frac{\sigma^2\rho^2}{2})}{\beta}[exp(\sigma^2\rho) - 1]$

Proof: See the Appendix.

It is natural to view this decision problem within the framework of a pure exchange economy with idiosyncratic earnings risk. This viewpoint is useful because the risk-free interest rate is then endogenously determined. In an equilibrium analysis of this economy, there is always an equilibrium where the gross risk-free rate $R^f = 1+r$ is precisely that given in the "if" part of Theorem 3: $1+r = \frac{1}{\beta}\exp(\rho\mu - \frac{\rho^2\sigma^2}{2})$. This claim holds simply because no agent at any age or in any state chooses to borrow or to lend. Thus, at the interest rate stated in Theorem 3 the credit and goods markets clear.

We now illustrate some quantitative properties of the simple example. Figure 1 documents these properties as the parameter σ governing the standard deviation of earnings shocks and the parameter ρ governing risk aversion and intertemporal substitution vary. In Figure 1, the parameter σ varies over

[10] Theorem 3 uses the following standard properties of the lognormal distribution (e.g. Green (1997, p. 71)): $\ln x \sim N(\mu, \sigma^2)$ implies $E[x] = exp(\mu + \frac{\sigma^2}{2})$ and $var(x) = exp(2\mu + \sigma^2)(exp(\sigma^2) - 1)$.

the interval $[0, 0.3]$ and $\mu = -\sigma^2/2$. Thus, 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 for all the economies in Figure 1 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(a)-(b) shows that the value of an agent's human capital falls and that the mean return on an agent's human capital rises as the shock standard deviation increases. Thus, in the economies analyzed a high mean return on human capital is the flip side of a low value attached to human capital. Figure 1 also shows that these patterns are amplified as the parameter ρ increases.

Figure 1(a) also plots a notion of value that we label the "naive value". The naive value is calculated by discounting earnings at a constant interest rate r set equal to the risk-free interest rate in the model (i.e. $value = E[\sum_{j=2}^{J} \frac{e_j}{(1+r)^{j-1}}|z^1])$. This follows a traditional procedure employed in the empirical literature as discussed in section 2. The naive value of a young agent's human capital is exactly the same in each economy in Figure 1 simply because the risk-free interest rate and the mean earnings profile are unchanged across economies. Our notion of the value of human capital differs from the naive value because the agent's stochastic discount factor covaries negatively with earnings. Figure 1 shows that this negative covariance is substantial as the standard deviation of the shock increases.

Figure 1(c) plots the total benefit and the marginal benefit of moving from the model consumption plan to a perfectly smooth consumption plan where $c_j^{smooth} = E_1[c_j] = E_1[e_j] = 1$. Section 3.3 provided theory under which the total benefit $\Omega(1)$ and the marginal benefit $\Omega'(0)$ satisfy the restriction below. In the simple example the numerator is known without parametric assumptions on preferences. It is simply $\sum_{j=1}^{J}(\frac{1}{1+r})^{j-1}$ for any value of the standard deviation of earnings shocks. Figure 1(c) indicates that the marginal benefit increases as the standard deviation of the period earnings shocks increases. Thus, in the simple example a high marginal benefit of moving towards perfect consumption smoothing coincides with a low value of human capital.

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

4 Analysis of Human Capital Values and Returns

4.1 The Benchmark Model

We now use the theoretical framework to quantify the value and return to human capital. Our benchmark model has two financial assets $\mathcal{I} = \{1, 2\}$ (one riskless a^1 and one risky a^2) and has tighter asset holding restrictions than considered in Problem P1 from section 3. The asset holding restriction states that going short on the risky asset is not allowed and that there is a maximum leverage ratio of $p > 0$ on the risky asset.

Benchmark Model: $\max U(c)$ subject to $c \in \Gamma_1(x, z_1)$

$$\Gamma_1(x, z_1) = \{c = (c_1, ..., c_J) : \exists (a^1, a^2) \; s.t. \; 1-3 \; holds \; \forall j\}$$

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. $e_j = G_j(z_j)$ and $c_j \geq 0$
3. $a^2_{j+1} \geq 0$ and $a^2_{j+1} \leq p \sum_{i \in \mathcal{I}} a^i_{j+1}$

The utility function $U(c) = U^1(c_1, ..., c_J)$ is of the type employed by Epstein and Zin (1991). This utility function is defined recursively by repeatedly 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}.

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

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

Our choice of the benchmark model reflects several considerations. First, the choice of two assets is in part motivated by the computational burden of solving the model. However, household portfolios have often been characterized in terms of the holdings of risky versus low risk assets. In addition, much of the discussion of portfolio choice in the existing literature is framed in terms of the holdings of a risky and a low risk asset. Second, the benchmark model is analyzed in partial equilibrium. This is natural as the main goal of the paper is to figure out the value and return to human capital in light of the properties of earnings and asset returns data.[11] Third, earnings are exogenous. This is helpful at this early stage of analysis as the statistical properties of earnings and asset returns in the model will then closely mimic data properties. Future work on human capital values and returns will likely consider deeper models of earnings (e.g. Huggett, Ventura and Yaron (2011)) and account for leisure.

[11] A general equilibrium approach could address the underlying sources of the movements in earnings and asset returns.

4.2 Empirics: Earnings and Asset Returns

We now describe the structure of earnings and asset returns in the benchmark model. We start by outlining an empirical framework for the dynamic relationship between the idiosyncratic and aggregate components of earnings and the return on the risky asset. The framework incorporates a number of features that have been hypothesized to be important in the existing literature, including counter-cyclical idiosyncratic risk, return predictability, and cointegration. We are not aware of any previous studies that have used micro data to estimate an earnings process with this set of features.

4.2.1 Empirical Framework

Let $e_{i,j,t}$ denote real 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 $\left(u^2\right)$:

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

The idiosyncratic component is the sum of four orthogonal components: (i) a common age effect κ, (ii) an individual-specific fixed effect α, (iii) an idiosyncratic persistent component ζ and (iv) an idiosyncratic transitory component υ. The common age effect, κ, is modeled as a quartic polynomial.

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

The individual fixed effects are assumed to be normally distributed with a constant variance. The two shocks are assumed to be normally distributed with time-dependent variances that are functions of a set of time-varying aggregate variables, X_t:

$$\alpha_i \sim N\left(0, \sigma_\alpha^2\right), \ \eta_{i,j,t} \sim N\left(0, \sigma_{\eta,t}^2\left(X_t\right)\right), \ \upsilon_{i,j,t} \sim N\left(0, \sigma_{\upsilon,j,t}^2\left(X_t\right)\right) \tag{3}$$

This structure implies that aggregate conditions affect both the mean and variance of earnings. In our empirical implementation we set $X_t = \Delta u_t^1 \equiv u_t^1 - u_{t-1}^1$. In order to capture life-cycle properties of the variance of earnings we allow the variance of the transitory component to be age-dependent. This dependence is modeled as a quartic polynomial.

The joint dynamics of equity returns and the aggregate component of earnings are 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 assume a vector autoregression (VAR) model for y_t:

$$y_t = v(t) + \sum_{i=1}^{p} A_i y_{t-i} + \varepsilon_t \tag{4}$$

where ε_t is a vector of zero mean IID random variables with covariance matrix Σ. $v(t)$ is a quadratic time trend. We do not impose that this process is stationary. Rather, we assume that y_t is a first order integrated, $I(1)$, process. One reason for assuming that $y_t \sim I(1)$ is that it allows us to connect with the literature on cointegration. In the Appendix, we show that (4) implies the following stationary VAR process for Δy_t and a cointegrating vector w_t defined by $w_t \equiv \beta' y_t + \mu + \rho(t+1)$:

$$\begin{pmatrix} \Delta y_t \\ w_t \end{pmatrix} = \begin{pmatrix} \gamma \\ \beta'\gamma + \rho \end{pmatrix} + \sum_{i=1}^{p-1} \begin{pmatrix} \Gamma_i \\ \beta'\Gamma_i \end{pmatrix} \Delta y_{t-i} + \begin{pmatrix} \alpha \\ 1 + \beta'\alpha \end{pmatrix} w_{t-1} + \begin{pmatrix} \varepsilon_t \\ \beta'\varepsilon_t \end{pmatrix} \tag{5}$$

When $p = 2$, this process takes the simple form

$$\begin{pmatrix} \Delta y_t \\ w_t \end{pmatrix} = \begin{pmatrix} \gamma \\ \beta'\gamma + \rho \end{pmatrix} + \begin{pmatrix} \Gamma & \alpha \\ \beta'\Gamma & 1 + \beta'\alpha \end{pmatrix} \begin{pmatrix} \Delta y_{t-1} \\ w_{t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_t \\ \beta'\varepsilon_t \end{pmatrix} \tag{6}$$

When there is no cointegration (i.e. $\alpha = 0$) the process collapses to a standard VAR for Δy_t:

$$\Delta y_t = \gamma + \sum_{i=1}^{p-1} \Gamma_i \Delta y_{t-i} + \varepsilon_t \tag{7}$$

4.2.2 Data Sources

For estimating individual earnings dynamics we use data on male annual labor earnings from the Panel Study of Income Dynamics (PSID) from 1967 to 1996. We restrict attention to 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. Full details can be found in the Appendix.

We also consider two sub-samples based on education. We divide the sample into High School and College sub-samples, based on their maximum observed completed years of education. 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. We hence make no distinction between high-school dropouts and high-school graduates; and our College group does not include college dropouts.

The model for idiosyncratic earnings risk is estimated in two stages. In the first stage we use OLS to estimate the age profile $\hat{\kappa}_j$ and the aggregate component $\hat{u}_t^1$. Residuals from the first stage are then used to obtain GMM estimates of the remaining parameters in (2) and (3), where the moments included are the elements of the auto-covariance function for each age/year combination. Full details of the estimation procedure can be found in the Appendix.

Although the PSID is an ideal data set for studying the auto-correlation structure of individual earnings, its relatively small sample size and the fact that after 1996 it was converted into a biannual survey means that it is less suited to studying dynamics in the aggregate component of earnings. Our approach is to retain the PSID as our data source for the idiosyncratic component of earnings but to also analyze an alternative measure of the aggregate component of labor earnings u_t^1 estimated using Current Population Survey (CPS) data from 1967 to 2008. We estimate this component from CPS data in the same way as we do in the PSID: we run a first-stage regression of individual log earnings on a polynomial in age and on time dummy variables. In order to minimize the effect on our estimates of changes in top-coding in the CPS over time, we use a median regression (Least Absolute Deviations) to extract these time and age effects. The Appendix documents that our measure of the aggregate component of earnings produces contraction years that are closely related to those based on three alternative measures from the National Income and Product Accounts.

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, whereas bond returns are based on Treasury bill returns.[12] Real returns are calculated by adjusting for realized inflation using the same CPI All Urban series that was applied to the earnings data.

4.2.3 Estimation Results: Idiosyncratic Process

Our benchmark model for idiosyncratic earnings dynamics allows for time variation in the variance of the persistent component. This variation takes the form of a linear trend to capture low frequency trends over the sample period, as well as a two-state process that captures cyclical variation.[13] Modeling cyclical variation as a two-state process has the advantages of being comparable to estimation results in the literature (Storesletten et al. (2004)), and of being easily implemented in our computational model. The two states are chosen to reflect periods of expansions and contractions. To determine expansions and contractions, we compared the periods of positive and negative earnings growth in our aggregate variables. Thus X_t is an indicator variable that is defined by whether or not

[12] All returns come from Kenneth French's data archive. Returns data are available from 1927-2010. We restrict attention to the period 1967-2008 since this is the period that the CPS earnings data covers.

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

$\Delta u_t^1 > 0$.[14]

Parameter estimates are shown in Table 1. These results are broadly consistent with estimates from similar specifications that have been estimated elsewhere in the literature summarized in Meghir and Pistaferri (2010).[15] 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, in particular the fact that we do not impose restrictions on hours worked. 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 below.[16]

We highlight three findings from Table 1. First, we find evidence of counter-cyclical risk. The variance of persistent shocks increases by about 0.02 in contractions relative to expansions. This is broadly consistent with the findings in Storesletten et al. (2004) who find an increase of around 0.03 when they jointly estimate these variances as well as the autoregression parameter ρ. Second, we find some differences across education groups. The High School subsample has less persistent shocks but larger fixed effects and stronger counter-cyclical risk compared to the College subsample. Third, when ρ is set to one, as is sometimes imposed in the literature, this also substantially reduces the average shock variance.

4.2.4 Estimation Results: Aggregate Process

In the Appendix we present results from standard lag-order selection tests for the order p of the underlying VAR in (4). We find that for our baseline sample, as well as for college and high-school sub-samples, and for alternative measures of aggregate earnings and alternative sample periods, virtually all specifications indicate the presence of two lags in (4). We thus focus our attention on a model with one lag in the VAR in first differences, as in (6). The Appendix also contains tests of the cointegrating rank of (5) based on the methods in Johansen (1995). For all three samples, tests based on the trace statistic, the maximum eigenvalue or the Schwarz-Bayes information criterion, suggest a cointegrating

[14] The Appendix lists contraction years implied by CPS data and other data sources.

[15] 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.

[16] Using indirect inference on a structural model of consumption and savings behavior, Guvenen and Smith (2010) 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 estimate transitory variance is measurement error. We adjust our estimates down by one third at all ages accordingly.

rank of zero, while the Hannan-Quin information criterion suggest the presence of one cointegrating vector. We interpret these findings as providing only very weak evidence for cointegration and hence we adopt the model without cointegration as our benchmark specification. However, since these tests may all have relatively little power given the short annual sample period, we also present estimates for the model with cointegration.

Table 2 reports results of estimation of (6) and (7). The parameter estimates for the benchmark model, with and without cointegration, reveal a moderate degree of persistence in aggregate earnings growth. There is a positive correlation between innovations to earnings growth and innovations to returns. This implies that the conditional correlation between earnings growth and stock returns is positive. The implied steady-state dynamics are reported in Table 3. The estimated model matches the observed correlation structure well. When we later input the estimated process into our economic model, we adjust the parameters (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.

4.3 Parameter Values in the Benchmark Model

Table 4 summarizes the parameters in the benchmark model. Agents live for a maximum of $J = 69$ model periods and retire in period $Ret = 40$. This corresponds to starting economic life at age 22, retiring at age 61 and living at most up to age 90.

We set the preference parameters governing intertemporal substitution and risk aversion to values estimated from Euler equation restrictions based on household-level data. Vissing-Jorgensen and Attanasio (2003) conclude that the substitution elasticity $(1/\rho)$ is likely to be above 1 and estimate $1/\rho = 1.17$ for a prefered specification. They conclude that the risk aversion parameter (α) in the interval $[5, 10]$ can be obtained under realistic assumptions. Thus, the special case of constant-relative-risk-aversion preferences, where $\rho = \alpha$, is not the parameter configuration that best fits the Euler equation restrictions. We examine model implications for $1/\rho = 1.17$ and $\alpha \in \{4, 6, 8, 10\}$. Agents face a conditional probability ψ_{j+1} of surviving from period j to period $j + 1$. Survival probabilities are set to estimates for males from the 1989-91 US Decennial Life Tables in NCHS (1992). 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.

Earnings and asset returns in the benchmark model are based on the estimates from the previous section for the case of no cointegration. Later in the paper we examine the model implications for the case of cointegration. 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 $z_2 = (\alpha, \zeta, \upsilon, \Delta u^1, \log R^s)$ captures the idiosyncratic components of earnings, the growth in the aggregate component of earnings and the stock return. For the case of cointegration, we add an extra component to this state variable: $z_2 = (\alpha, \zeta, \upsilon, \Delta u^1, \log R^s, w)$. The law of motion for z_1 is $z_1' = z_1 f_{j+1}(z_2')$, where $f_j(z_2) = \exp(\Delta u^1)$ for $j \leq Ret$ and $f_j(z_2) = 1$ for $j > Ret$. We will see shortly that the law of motion after retirement helps to capture social security payments. The Appendix describes our methods to compute solutions to the decision problem and the implied human capital values and returns.

Before the retirement age, earnings in the model is $e_j(z) = z_1 g_j(z_2)(1 - \tau)$. The first component of the state $z_1 = \exp(u^1)$ captures the aggregate component of earnings. The second component $g_j(z_2) = \exp(\kappa_j + \alpha + \zeta + \upsilon)$ captures variation due to idiosyncratic shocks and age. The benchmark parameter values are those estimated in Table 1 and Table 2 for the case of no cointegration.[17] Earnings are taxed at a proportional rate τ.

Model earnings are viewed as earnings after taxes and transfers. The literature has argued that the tax-transfer system can substantially impact the degree to which pretax-earnings shocks can be smoothed. This has implications for the degree to which idiosyncratic earnings risk acts to lower human capital values. Additionally, the social security system has potentially major implications for how the bond and stock components of the value of human capital change as agents approach retirement. This in turn has implications for how the financial asset portfolio changes with age. Work that imputes the magnitude of social security wealth from data concludes that social security wealth is a major component of total wealth. For example, 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 larger percentage of wealth for households with low wealth.

Earnings in retirement are given by a social security transfer. Transfers in the model are an annuity payment which is determined by the aggregate earnings level z_1 when the agent retires 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 α rather than to an average of the agent's past earnings as in the U.S. system.

Old-age benefit payments in the U.S. system are proportional to average economy-wide earnings when

[17] We lower the estimate of the transitory shock variance to account for measurement error as described in section 4.2.3.

an individual hits age 60. This is captured within the model by the fact that transfers are proportional to z_1. The model implies that after entering the labor market, an agent's social security transfers are risky only because the aggregate component of earnings at the time of retirement is risky. This feature of the model social security system together with the estimated correlation structure of stock returns and the aggregate component of earnings will produce a positive correlation between human capital returns and stock returns just prior to retirement.

The benchmark model has two assets: stock and bonds. The stock return follows the process estimated in Table 2. In all experiments, we adjust the constant γ estimated in Table 2 so that the mean stock return ($E[\log R_t^s]$) is 4.1% and the mean growth rate of aggregate earnings ($E[\Delta u_t^1]$) is zero, while retaining the estimated properties for all second moments of the stochastic process. The bond return is risk-free. The bond return in the model is set to 1.2%, which is the average real return over the period 1967-2008 based on the French data set and the CPI-U price index. We set the leverage ratio on stock holding to $p = 1$. This implies that negative positions in stock or bonds are not allowed.

Life-cycle properties of the benchmark model are displayed in Figure 2. The figure is constructed by simulating many shock histories, calculating allocations along these histories and then taking averages. Initial wealth is set at 30 percent of mean earnings at age 22 in all the figures in this section. Figure 2(a)-2(b) show that the profile for mean consumption and for mean earnings net of taxes and transfers are hump shaped over the lifetime. The earnings profile for college males is much steeper over the working lifetime than the corresponding profile for males with only a high school education. 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. Figure 2(c)-2(d) show that the average fraction of financial wealth held in stock tends to fall as the risk aversion coefficient α increases.

4.4 Human Capital Values and Returns

We report the value and return properties of the benchmark model. In this section we highlight only the results based on the high school and the college earnings data as the results for the full sample are typically between the results implied by the two education groups.

4.4.1 Human Capital Values

Figure 3 plots the value of human capital in the benchmark model and a decomposition of this value. Figure 3 shows that the mean value of human capital over the lifetime is hump shaped and that the mean value 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 in Figure 3. The naive value is quite far away from our notion of value. The substantial differences in Figure 3 between the naive value and our notion of value is due largely to negative covariation between the agent's stochastic discount factor and earnings.[19] We will very shortly highlight sources of this negative covariation.

It is useful to decompose human capital values. To do so, we project next period's human capital pay out $v_{j+1} + e_{j+1}$ onto the space of asset returns as described in section 3.3. Thus, next period's pay out is uniquely decomposed into the component in the space spanned by the conditional bond and stock returns and into an orthogonal component. This then implies that the value of human capital can be decomposed into three components: a bond, a stock and a residual value component. This decomposition is valid regardless of whether or not the agent's Euler equation holds with equality.

Figure 3 shows the results of carrying out this decomposition. This figure plots the value of the bond, stock and orthogonal components as a fraction of the value of human capital at each age when averaged across the states that occur at each age. We find that the bond component is on average more than 80 percent of the value of human capital at each age over the lifetime. This holds for all values of the risk aversion coefficient that we examine and for both education groups. We also find that the stock share of the value of human capital is slightly larger early in life for the high school than for the college agents. This leads high school agents to hold a lower average share of stock in the financial asset portfolio early in life compared to college agents.

While it may seem plausible that the value of human capital is largely bond-like during retirement, it is useful to understand why the value of human capital in retirement is not always 100 percent bonds. 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 bonds as social security transfer payments in retirement are certain. However, if the agent hits the corner of the bond decision in the future under some sequence of risky stock returns, then this is not true as the mean of the agent's stochastic discount factor will be less than the inverse of the gross risk-free rate in such an event. This then implies that the value of future social security transfers to the agent is low at this date. Thus, the value of these transfers at earlier dates takes on a positive stock component when

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

[19] Another reason why $v_j < v_j^{naive}$ is that agents are sometimes on the corner of the risk-free asset choice (i.e. $E[m_{j,j+1}] < \frac{1}{1+r}$). Quantitatively, the vast majority of the gap between the naive value and the value in Figure 3 is due to negative covariation.

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 may run down financial assets, hit a corner solution on the holdings of the risk-free asset and live off social security transfers.

Figure 3 shows that the stock component of the value of human capital in the benchmark model is positive on average and accounts for less than 20 percent of the value of human capital over the lifetime. 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. Specifically, Table 2 shows that the innovations to the growth rate of the aggregate component of male earnings and stock returns are always positively correlated. The conditional correlation of log earnings growth and log stock returns is slightly above 0.3 in all the models estimated in Table 2.

The stock component of the value of human capital falls sharply across ages just before retirement in Figure 3. This is associated with a rise in the average share of stock held in the financial asset portfolio across the same ages as was documented in Figure 2. The logic behind this is that agents adjust their financial asset portfolio so that overall stock holdings (i.e. stock in financial assets plus the stock position in human wealth) as a fraction of overall wealth is roughly constant on average across ages. We document this pattern in Figure 5 later on in this section.

Figure 3 shows that the value of the orthogonal component is strongly negative early in the lifetime. Given that the orthogonal component has a zero mean, this large negative value is due to strong negative covariation between the orthogonal component and the stochastic discount factor. This occurs, for example, when the agent's consumption and future utility is 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. The value of the orthogonal component of human capital payouts will play an important role in our analysis of whether human capital returns look very much like the return to stock, like the return to bonds or like neither of these assets.

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

4.4.2 Human Capital Returns

Figure 4 plots properties of human capital returns. Expected human capital returns are very large early in the working lifetime and decline with age over most of the working lifetime. The large return to human capital early in the lifetime may at first seem surprising in light of the results of the human capital value decomposition. The decomposition showed a heavy average weight on bonds over the lifetime. This fact may lead some to conjecture that early in life the mean return should be near the bond return or between the mean return to stock and bonds. Neither conjecture is correct.

To understand what drives the mean human capital returns profile over the lifetime, 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]

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

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

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] Figure 3 documented that the value of the orthogonal component of human capital payouts is strongly negative early in the working lifetime. This occurs because persistent idiosyncratic shocks are especially important early in life as they signal higher earnings many periods into the future. When the decomposition weights on expected returns sum to more than one then human capital returns can vastly exceed a convex combination of stock and bond returns.

At this point, it is useful to return to an issue raised in the introduction and in section 3.3. We mentioned that Vissing-Jorgensen and Attanasio (2003) estimate preference parameters of Epstein-Zin utility functions from micro data. Their methods involve proxying the unobserved expected return to a household's human capital by an age and state invariant weighted average of bond and

[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.

stock returns. Our analysis shines light on how this approximation may be problematic. While expected human capital returns do in theory equal a weighted average of financial asset returns in the benchmark model, the theoretical weights vary by age and state. Figure 4 highlights the point that the mean weights fall with age. We also note that the weights in the above decomposition early in life are sensitive to the level of initial financial wealth: lower financial wealth tends to increase mean human capital returns, other things equal, and thus increases the sum of the weights. Thus, our results suggest that proxying mean human capital returns with an age and state-invariant average of stock and bond returns may be more problematic when the data includes relatively young households or households with low financial wealth.

The mean return to human capital during retirement is near the risk-free rate except at the very end of the lifetime. The high return at the very end of the lifetime might at first seem odd since agents in the benchmark model receive a real annuity after retirement. This should not be surprising, however, as in the penultimate period $v_{J-1} = E[me_J]$ and $1 = E[me_J/v_{J-1}]$. As the return conditional on surviving to the last period is certain, the return is $R^h_J = e_J/v_{J-1} = 1/E[m]$. Thus, the return on human capital equals the risk-free bond rate when the agent is off the corner on the risk-free asset choice (i.e. $R^h_J = 1/E[m] = R^b$) but can exceed the risk-free rate when the agent is on the corner (i.e. $R^h_J = 1/E[m] \geq R^b$). Towards the end of the lifetime an increasing fraction of agents in the model for both education groups are on corners. They run their financial assets down to zero and live off their social security annuity. This accounts for the high mean returns towards the end of the lifetime.

The pattern of mean returns over the lifetime in Figure 4 can be contrasted with the patterns for the simple example economy analyzed in section 3.3. The simple example featured permanent idiosyncratic risk but no aggregate sources of risk. The distribution of the return to human capital in the simple example economy was the same at each age over the lifetime. Thus, the substantial fall in mean returns in Figure 4 is due to differences in model features other than the presence of persistent idiosyncratic risk. Three key differences are that (i) the model has a retirement period, (ii) agents receive a social security annuity during retirement and (iii) agents hit the corner of their asset holding decision. The first difference implies that as agents approach retirement, persistent shocks become in effect just like transitory shocks. The first two differences imply that an agent's stochastic discount factor is less variable around retirement compared to early in the lifetime. The third difference means that the mean return to human capital early in life can be very large when an agent holds very little or no financial assets. A higher fraction of college agents hold no financial assets early in life in the model economy compared to high school agents. This is driven by differences in the mean earnings profiles.

The correlation between human capital returns and stock returns over the lifetime in the benchmark model is summarized in Figure 4(c)-4(d). The correlation is positive but low early in life. The correlation is slightly larger for High School males than for College males. However, for neither case do we find that individual-level human capital returns are very much like U.S. stock returns in terms of the correlation over the working lifetime. Some studies of human capital returns that are based on valuing a claim to aggregate, economy-wide earnings (see Baxter and Jermann (1997)) conclude that human capital returns are more highly correlated with stock returns than our findings indicate.

The positive correlation that we do find for individual-level returns before retirement is based on several data properties. First, we find in Table 3 that the aggregate component of earnings growth for males in CPS data is positively correlated with stock returns. This fact helps to shape the coefficients in our VAR model including the slightly larger positive conditional correlation of earnings growth and 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. In the U.S. social security system a similar feature holds as old-age benefits are proportional to a measure of average earnings in the economy when the worker turns age 60, other things equal. This model feature implies that the value of human capital for an agent will be approximately proportional to aggregate earnings at retirement.[23] These two properties are behind the high mean return to human capital and the positive correlation of human capital and stock returns in the last period of the working lifetime that is documented in Figure 4 for both education groups.

4.4.3 Portfolio Allocation

Figure 5 provides two related views of how overall wealth is divided into useful parts over the lifetime. An individual's overall wealth is defined as the value of the individual's human capital plus the value of financial assets. Figure 5(a)-5(b) divide overall wealth into three types of assets: human wealth and the wealth directly held in stock and bonds, for the economy with risk aversion set equal to $\alpha = 6$. Early in life, from age 22 to well past age 30, the value of human capital is more than 90 percent of overall wealth for both education groups. 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. Even at the retirement age, the human capital share of an individual's wealth exceeds on average the share either in bonds or in stock.

Figure 5 also provides a view of the composition of the overall wealth portfolio in which the value of

[23]Lemma 2 in Appendix A.2 establishes that the value of human capital is precisely proportional to the aggregate earnings component, after an adjustment for financial wealth, because preferences are homothetic.

human capital is decomposed into bond and stock components as well as the value of the orthogonal component. This allows us to characterize the average overall wealth share held in stock and bonds. 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.

The mean overall bond share greatly exceeds the overall stock share throughout the working lifetime. This holds despite the fact, see Figure 2, that agents hold on average more stock than bonds in the financial asset portfolio. The overall stock share early in life in Figure 5 is largely determined by the decomposition analysis presented in Figure 3. This is because financial assets are small in value compared to the value of human capital and agents are not allowed to hold negative positions in either financial asset.

The overall stock share averages over 20 percent and displays little variation with age in Figure 5 for much of the working lifetime. In the standard two-period portfolio problem with constant-relative-risk-aversion (CRRA) preferences, the optimal stock share equals $\frac{1}{\alpha}\frac{E[R^x]}{Var(R^x)} \doteq \frac{1}{6}\frac{.048}{.04} = .2$, where R^x is the excess return of the risky asset over the risk-free asset and α is the coefficient of relative risk aversion. Applying this formula to the mean and the variance of excess returns within our model results in a 20 percent stock share when $\alpha = 6$.[24] While background risk, correlated returns, multiple-period horizons and corner solutions are present in the benchmark model and invalidate this formula (see Gollier (2001)), it is interesting to see that this formula is not wildly at odds with average overall portfolio shares in the benchmark model. It is valuable to keep this point in mind in the next section where we analyze a number of earnings perturbations. Perturbations that substantially increase the stock share of the value of human capital above 20 percent coincide with dramatically lower stock shares in the financial asset portfolio.

5 Discussion

The previous section documented properties of human capital values and returns in the benchmark model. Using the estimated process, based on male earnings and stock return data, we find that (1) the stock component of the value of human capital averages below 20 percent each period over the working lifetime, (2) the value of human capital is far below the value implied by discounting future earnings at the risk-free rate, (3) the mean human capital return is substantially larger than the mean stock return early in life and declines with age and (4) human capital returns have a small positive correlation with stock returns. These findings hold for a range of risk-aversion parameters. They also

[24] The mean and variance in the formula were calculated based on a lognormal distribution, using the mean and variances of the log returns from US data in Table 3.

hold based on estimates from three separate male earnings data samples - a high school, a college and a full sample. We now examine the sensitivity of the first of these findings in two directions.

5.1 Counter-Cyclical Idiosyncratic Risk

There is evidence that idiosyncratic risk is larger in contractions than expansions. Specifically, Storesletten et. al. (2004) estimate an earnings process that allows the innovation variance of the persistent, idiosyncratic earnings shock to take one value in expansions and another in contractions. They find that the variance is higher in contractions than expansions. They hypothesize that this data feature is important for portfolio choice, the cost of business cycles and a number of other issues. Our estimates of the autocorrelation parameter ρ and the standard deviation in aggregate good times $\sigma(H)$ and bad times $\sigma(L)$, are close to their estimates. We estimate that $(\rho, \sigma(H), \sigma(L)) = (.957, .160, .212)$ for the full sample, whereas Storesletten et. al. (2004, Table 2, row C) estimate $(\rho, \sigma(H), \sigma(L)) = (.952, .125, .211)$.[25]

How important is counter-cyclical risk for the stock share of the value of human capital? This can be addressed from at least two perspectives. One perspective asks how our findings change if one restricts the empirical framework to keep the innovation variance constant in aggregate good and bad times and then re-estimates the remaining parameters of the earnings process using the same data set. Another perspective asks how our findings change if parameters controlling counter-cyclical risk are altered, other things equal. This latter perspective is more focused on understanding the model, given that parameters are estimated with error, rather than in interpreting data.

We consider two experiments. Experiment A1 shuts down counter-cyclical risk by imposing that the innovation variance is the same in good and bad times and re-estimates all model parameters related to earnings. Experiment A2 alters the ratio of counter-cyclical risk from $\sigma(L)/\sigma(H) = .212/.160$ to $\sigma(L)/\sigma(H) = .257/.076$, while fixing all other parameters of the earnings process. This corresponds to the same average variance of the persistent shock but a difference of the variance from bad to good times of .06 (assuming equal weight on good and bad times). The benchmark model is the estimated model based on earnings data from the full sample as this model allows the most direct comparison with the empirical results from Storesletten et. al. (2004). In both experiments the discount factor β is re-calibrated to generate a steady-state wealth to income ration of 3.5.

Figure 6 displays the results. Experiment A1 shows that eliminating counter-cyclical risk implies that the average stock share of the value of human capital falls and that the average stock share in

[25]Our estimates are based on male earnings data from the PSID, whereas their estimates are based on household earnings data from the PSID adjusted for household size.

the financial asset portfolio increases compared to the benchmark model with counter-cyclical risk. Experiment A2 shows that increasing the magnitude of counter-cyclical risk beyond that estimated in the benchmark model increases the stock share of the value of human capital and decreases the average stock share in the financial asset portfolio. The experiments show that the stock share of the value of human capital increases as the degree of counter-cyclical risk increases.

These results broadly agree with the work of Lynch and Tan (2011). They analyze how the financial asset portfolio changes as the magnitude of counter-cyclical risk is varied, other things equal. Within their model the change in the stock share of the financial asset portfolio, after turning off counter-cyclical risk, is in the same direction but is larger than what occurs within our model. One possible reason for the larger magnitude in their work is that they impose that $\rho = 1$ while using the magnitudes of ratios of counter-cyclical risk estimated by Storesletten et al (2004) and using the average variance of persistent shocks equal to $\sigma^2 = .0225$ from a separate study.[26] We have also analyzed the model using parameter values closer to those employed in Lynch and Tan (i.e. $\rho = 1, \sigma(L) = .184, \sigma(H) = .105$). The results for these parameter values are in line with those for experiment A2 in Figure 6.

5.2 Cointegration

The benchmark model does not allow for cointegration between the sum of log stock returns and the log of the aggregate component of male earnings. This possibility may be important for the value of human capital because it allows for the possibility that histories with a sequence of large positive stock returns are associated with a sequence of large positive shocks to the common component of earnings. Mechanically, this holds under cointegration since a linear combination of these two random variables is then a stationary process. Such a relationship may strengthen the negative covariation between stochastic discount factors and future earnings when agents are young and have many future earnings periods, leading the value of human capital to move more in sync with current stock returns.

We analyze this possibility as well as three experiments. The new benchmark model uses the estimated process with cointegration and counter-cyclical risk from Table 2 based on the full sample of all males from the CPS over the period 1967-2008. Experiment B1 alters the new benchmark model by using the estimated process based on data over the period 1929-2009. The measure of the growth rate in the common component of earnings impacting all individuals is now the change in the log of aggregate wages and salaries paid to all workers from NIPA per member of the labor force.[27]

[26] Our estimation results from Table 1 show that if one imposes $\rho = 1$ then the average variance of the persistent shock is $\sigma^2 = .014$ and $\sigma^2(L) - \sigma^2(H) = .011$ for the full sample.

[27] Table A.4 in the Appendix documents that this NIPA measure of earnings growth has similar properties to the measure that we estimate from CPS data.

Experiment B2 alters the new benchmark model so that the process governing the common component of earnings growth and stock returns is an approximation to the process used by Benzoni et. al. (2007). The Appendix constructs this approximation and presents summary statistics implied by the process. The main focus of Benzoni et. al. (2007) is to examine how cointegration affects portfolio choice as a single parameter (the parameter κ in their model) that controls the strength of adjustment in the cointegrating relationship increases. They find that stock holding is zero early in life for their preferred value of κ when relative risk aversion is sufficiently large. They view that this occurs because a cointegrating relationship makes the value of human capital look more like stock as the adjustment parameter κ increases.

Figure 7 presents the results. The new benchmark model has a lower stock share of the value of human capital compared to the previous benchmark model. Thus, simply allowing for cointegration, while estimating the process using the same data, does not increase the stock share of the value of human capital. Of course, such a relationship may be difficult to estimate precisely with a short times series. For this reason experiment B1 is valuable because it examines evidence from a longer time period using NIPA data to proxy the common component of earnings growth. The stock share of the value of human capital in experiment B1 increases relative to the new benchmark model but still averages slightly under 20 percent in each period over the working lifetime. In experiment B2 the stock share of the value of human capital rises substantially above the levels in either benchmark model to more than 30 percent for some risk-aversion parameters. In addition, the average stock share in the financial asset portfolio is fairly sensitive to the risk-aversion parameter: the share is near zero early in life for risk-aversion parameters above 6. This result is roughly consistent with the main finding in Benzoni et. al. (2007).

What is driving the substantial difference between experiment B2 on the one hand and the results for either benchmark model on the other hand? This is not due to differences in the mean log return to stock or the mean log earnings growth rate as we adjust the constants in each process analyzed in this paper to produce the same means to facilitate comparisons across models. One difference is that the unconditional standard deviation of log earnings growth is $SD(\Delta u^1) = .069$ in experiment B2 but is $SD(\Delta u^1) = .025$ in both benchmark models.[28] Thus, aggregate earnings risk differs substantially across models.

Are the differences between the results of experiment B2 and the benchmark models primarily due to the nature of cointegration or due to differences in the variance of aggregate earnings growth, absent

[28] Table A.4 in the Appendix shows that $SD(\Delta u^1) = .025$ using CPS data over 1967-2008. The NIPA measure is $SD(\Delta u^1) = .024$ over 1967-2008 and $SD(\Delta u^1) = .029$ over 1929-2009.

any role played by cointegration? Experiment B3 addresses this issue. We construct a process with no cointegration but which generates an unconditional contemporaneous covariance structure and first-order auto-covariance structure that is identical to the process in experiment B2. The Appendix describes this process. The results for experiment B3 in Figure 7 show that an important part, but not all, of the difference between experiment B2 and either benchmark model is driven by a greater magnitude of short-run aggregate labor market risk, not by the presence of cointegration. The labor market risk in experiments B2 and B3 is larger than what we find in U.S. data.

References

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

Baxter, M. and U. Jermann (1997), The International Diversification Puzzle Is Worse Than You Think, *American Economic Review,* 87, 170-80.

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

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

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.

Bodie, Z., Merton, R. and W. Samuelson (1992), Labor Supply Flexibility and Portfolio Choice in a Life Cycle Model, *Journal of Economic Dynamics and Control,* 16, 427- 49.

Campbell, J. (1996), Understanding Risk and Return, *Journal of Political Economy,* 104, 298-345.

Campbell, J., Cocco, J., Gomes, F. and P. Maenhout (2001), Investing Retirement Wealth: A Life-Cycle Model, in Campbell and Feldstein, eds: *Risk Aspects of Social Security Reform* (University of Chicago Press).

Cochrane, J. (2001), Asset Pricing, Princeton University Press.

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.

Gollier, C. (2001), The Economics of Risk and Time, The MIT Press.

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.

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.

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

Heckman, J., Lochner, L. and P. Todd (2006), Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond, Handbook of the Economics of Education, Volume 1, eds. Hanushek and Welch.

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

Huggett, M., Ventura, G. and A. Yaron (2011), Sources of Lifetime Inequality, *American Economic Review,* 101, 2923- 54.

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.

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

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.

Palacios-Huerta, I. (2003), An Empirical Analysis of the Risk Properties of Human Capital Returns, *American Economic Review,* 93, 948- 64.

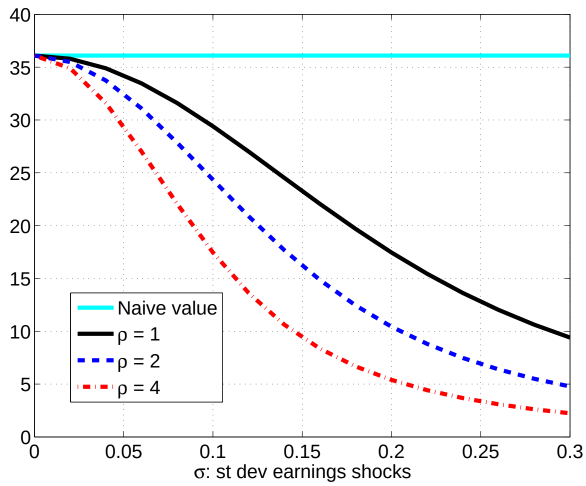
Poterba, J., Venti, S. and D. Wise (2011), The Composition and Drawdown of Wealth in Retirement?, *Journal of Economic Perspectives,* 25, 95-118.

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

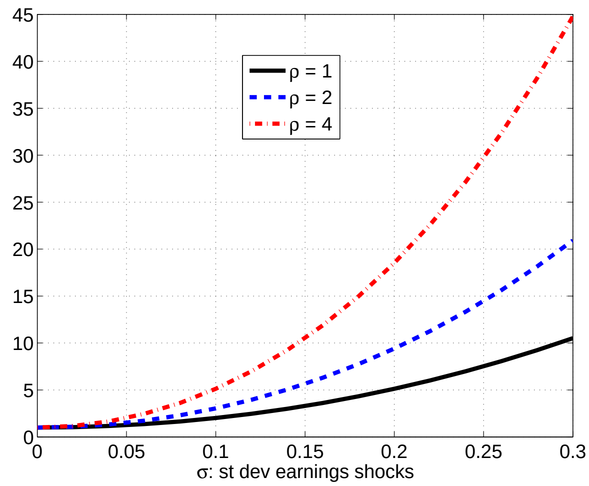
Svensson, L and I. Werner (1993), Nontraded Assets and Incomplete Markets: Pricing and Portfolio Choice, *European Economic Review,* 37, 1149- 68.

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

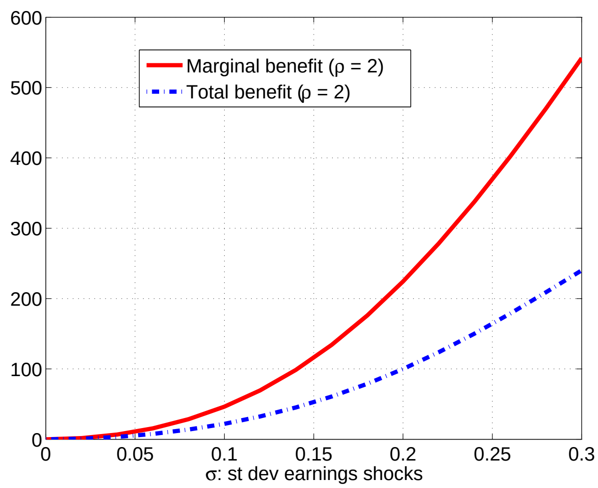
Weisbrod, B. (1961), The Valuation of Human Capital, *Journal of Political Economy,* 69, 425-36.



(a) Value of human capital



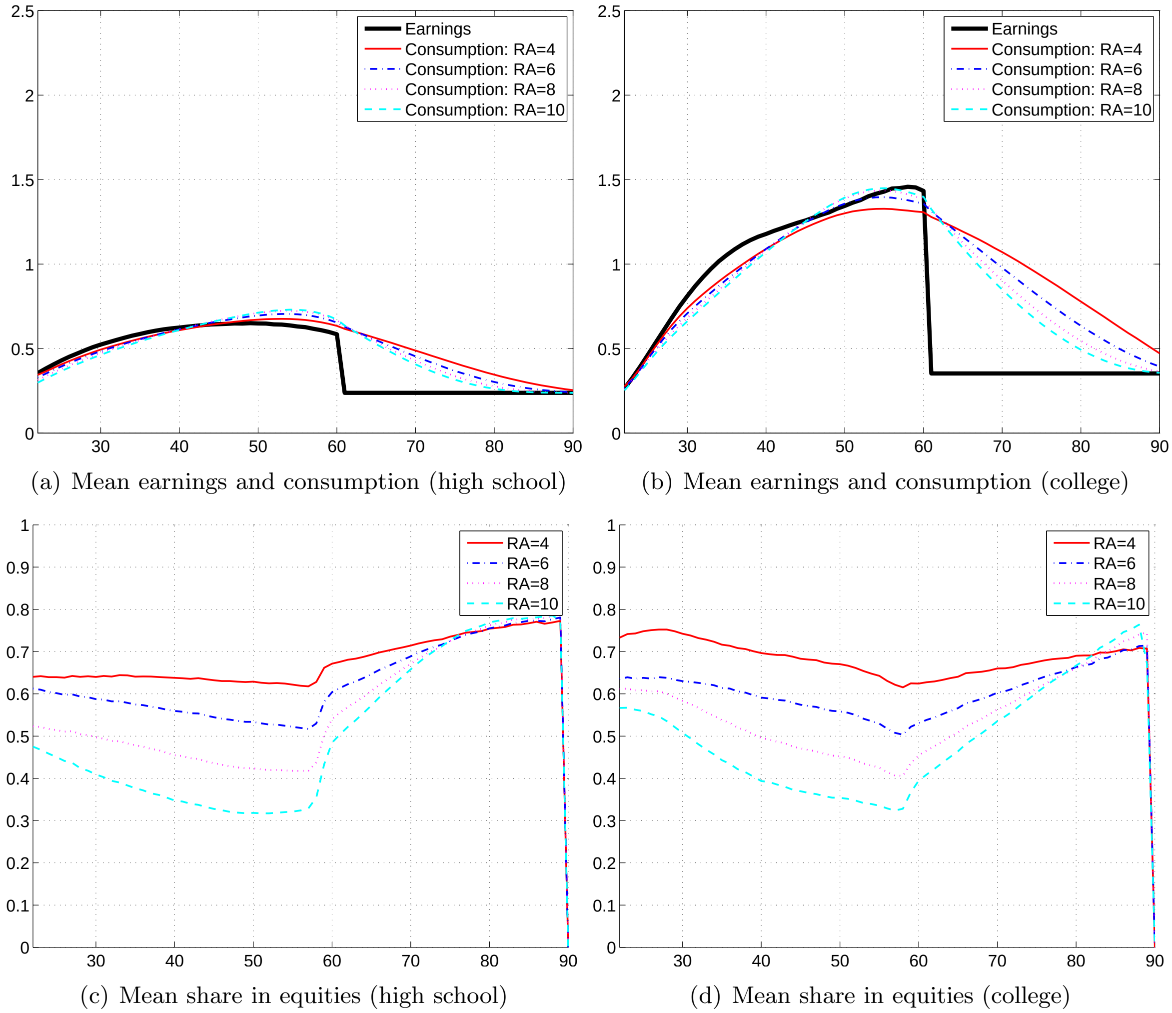
(b) Mean return human capital (%)



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

Figure 1: Human capital values and returns: simple example

Notes:



(a) Mean earnings and consumption (high school)

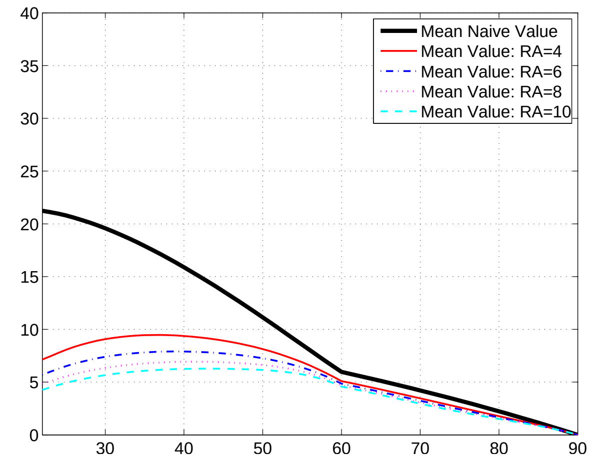
(b) Mean earnings and consumption (college)

(c) Mean share in equities (high school)

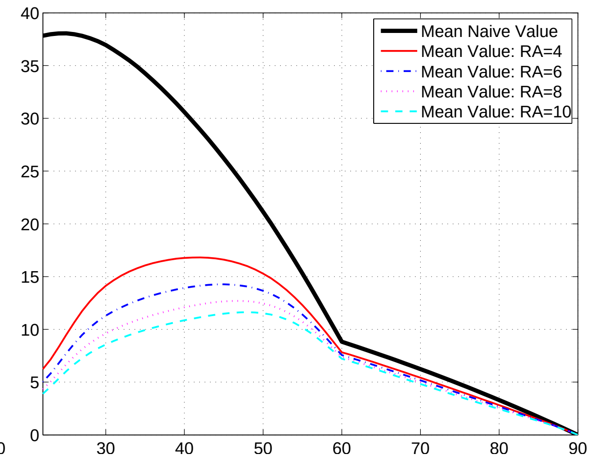
(d) Mean share in equities (college)

Figure 2: Life-cycle profiles in the benchmark model

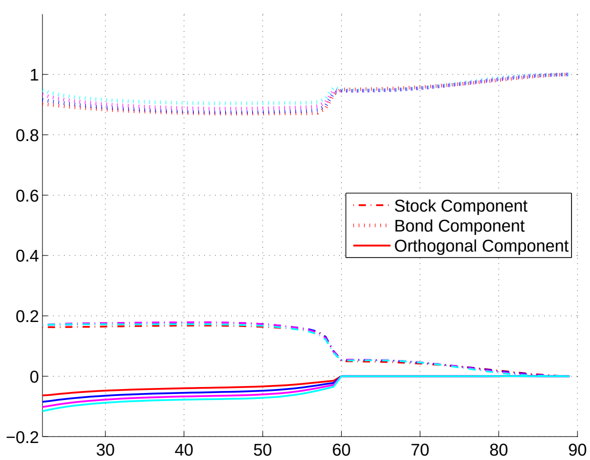
Notes: Portfolio shares are averages over the sub-population with positive asset holdings.



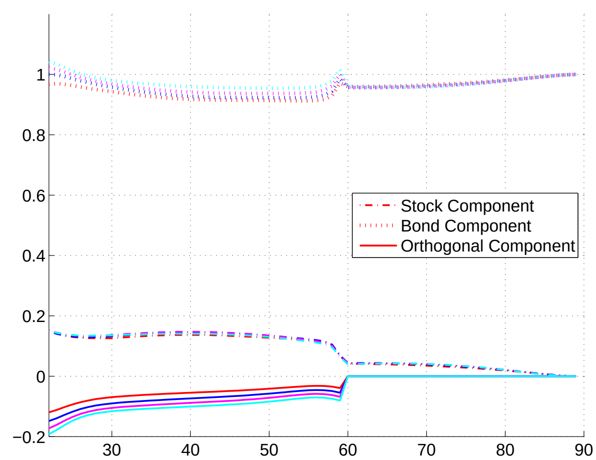
(a) Human capital values (high school)



(b) Human capital values (college)



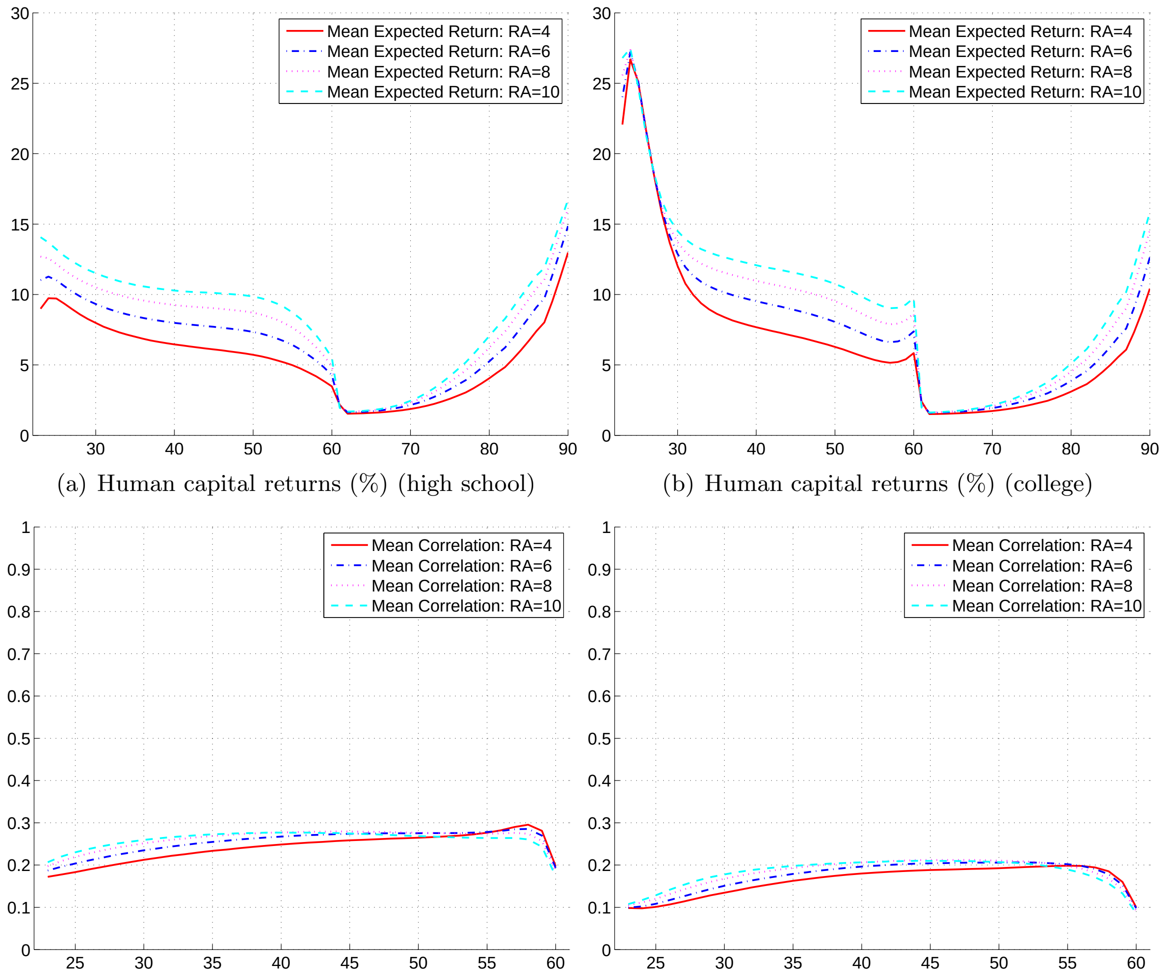
(c) Decomposition (high school)



(d) Decomposition (college)

Figure 3: Human capital values and decomposition

Notes:



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

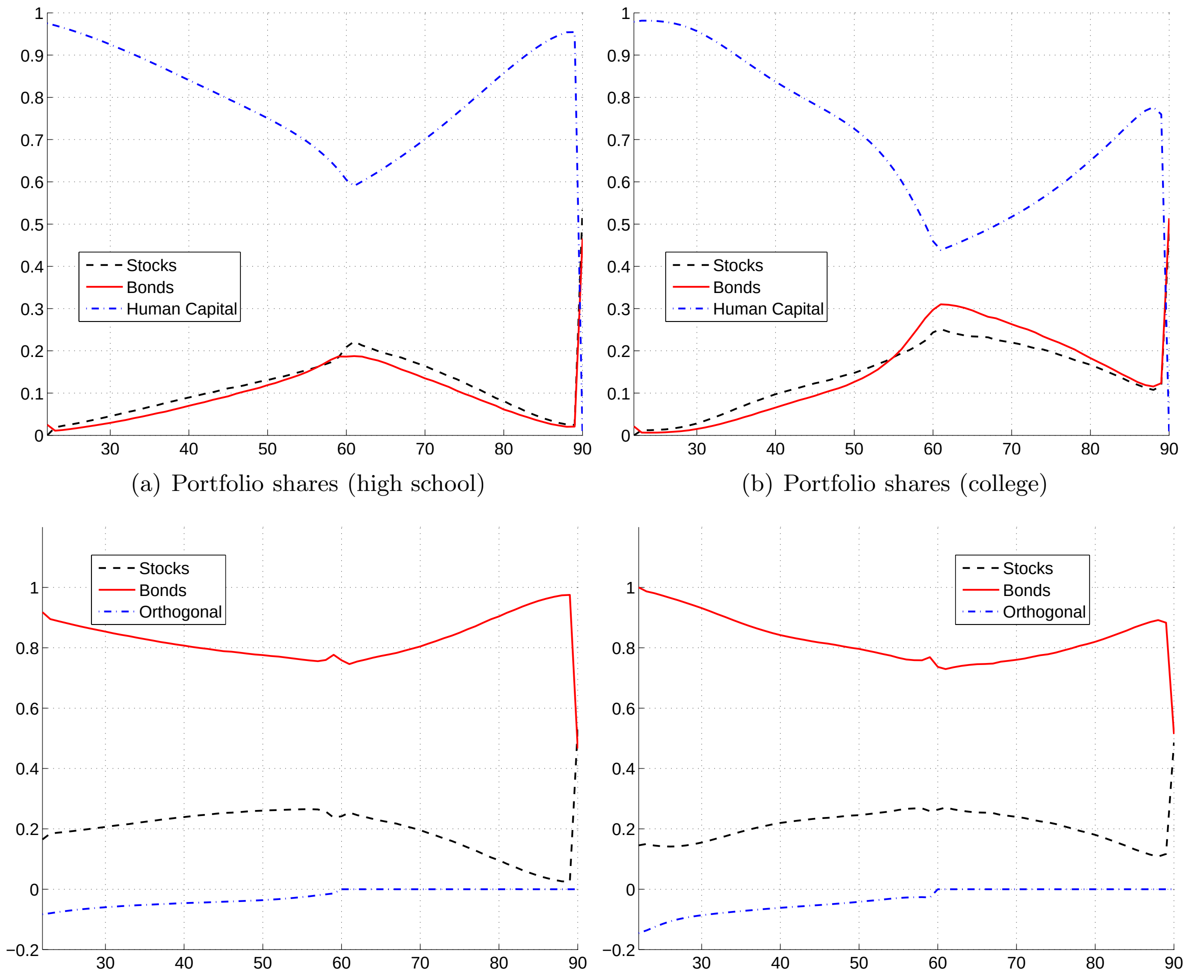
(b) Human capital returns (%) (college)

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

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

Figure 4: Properties of human capital returns

Notes:

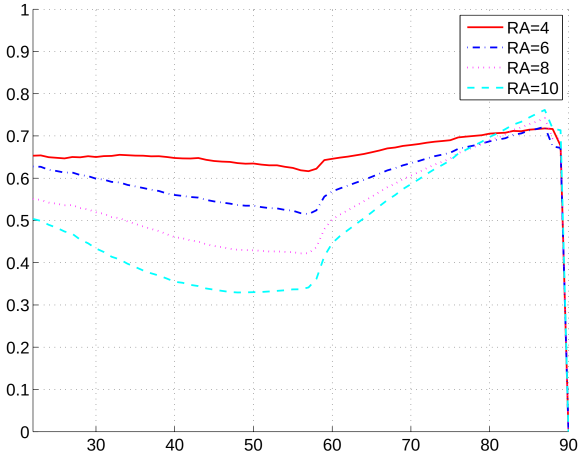


(a) Portfolio shares (high school) (b) Portfolio shares (college)

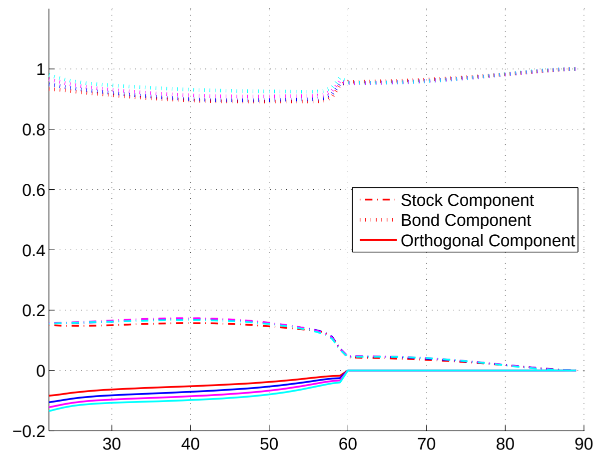
(c) Overall portfolio shares (high school) (d) Overall portfolio shares (college)

Figure 5: Portfolio shares in the benchmark model: risk aversion = 6

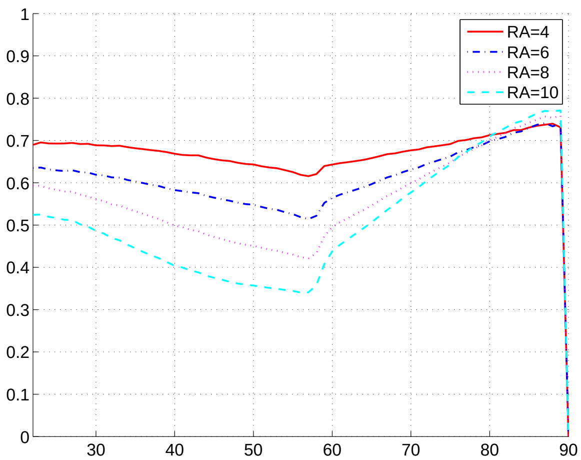
Notes:



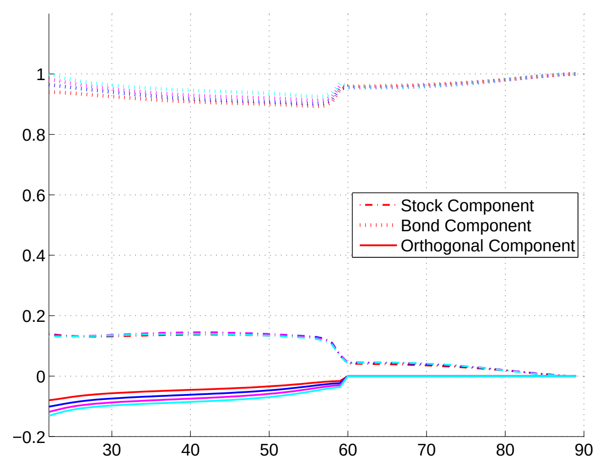
(a) Portfolio shares (benchmark)



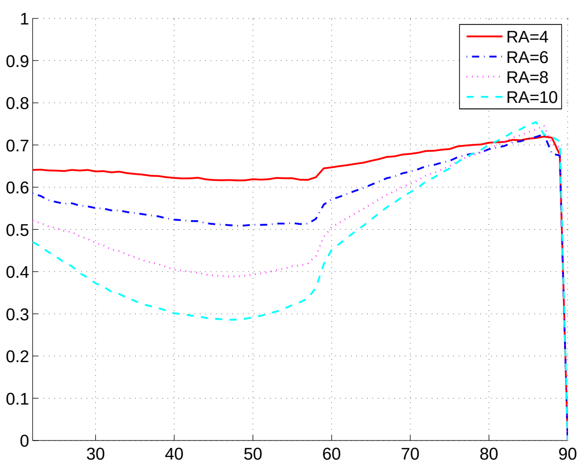
(b) HC decomposition (benchmark)



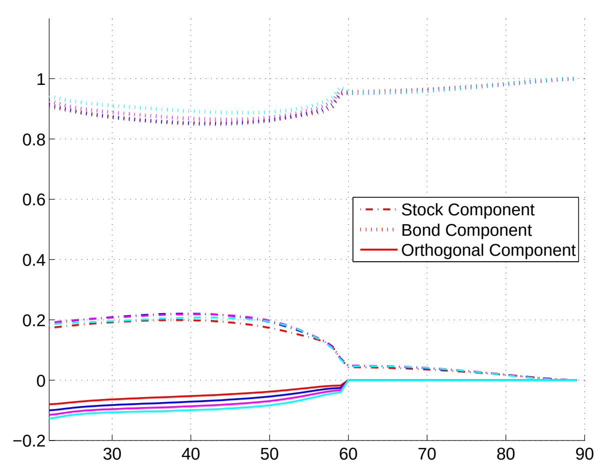
(c) Portfolio shares (experiment A1)



(d) HC decomposition (experiment A1)



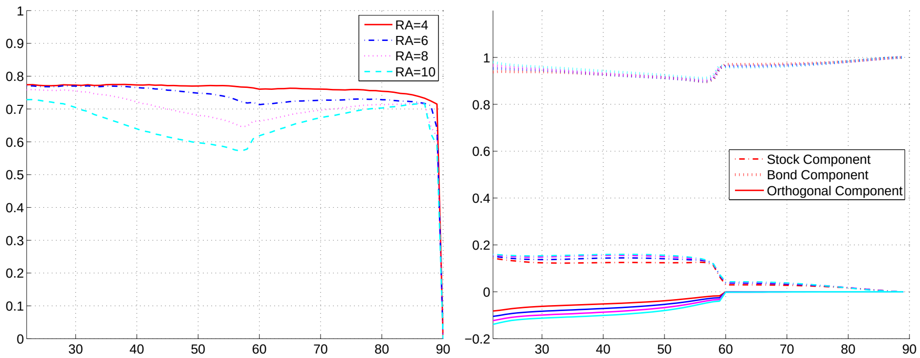
(e) Portfolio shares (experiment A2)



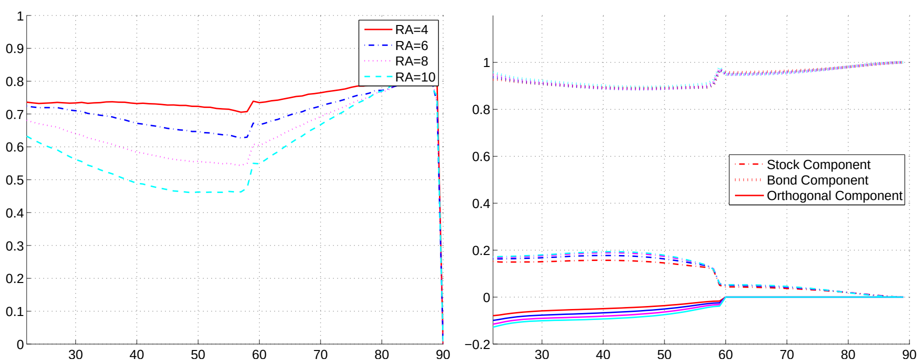
(f) HC decomposition (experiment A2)

Figure 6: Effect of counter-cyclical risk on portfolio shares and human capital decomposition

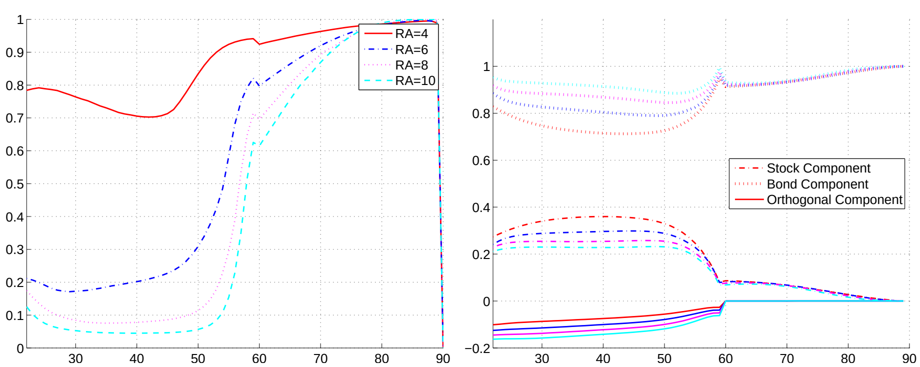
Notes: Portfolio shares are averages over the sub-population with positive asset holdings.



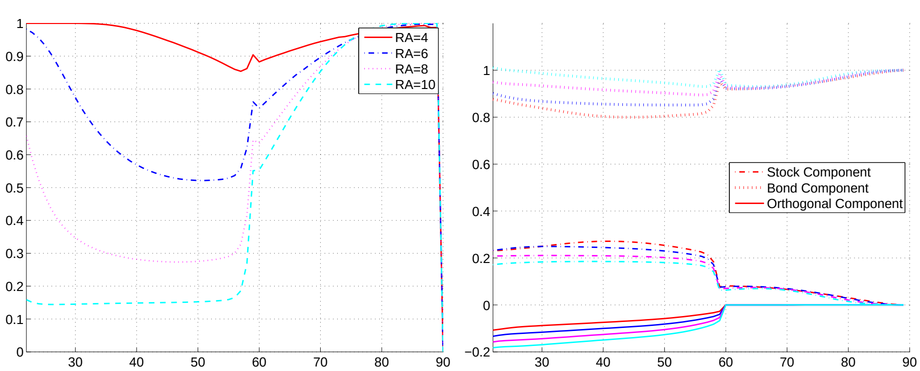
(a) Portfolio shares (benchmark cointegration) (b) HC decomposition (benchmark cointegration)



(c) Portfolio shares (experiment B1) (d) HC decomposition (experiment B1)



(e) Portfolio shares (experiment B2) (f) HC decomposition (experiment B2)



(g) Portfolio shares (experiment B3) (h) HC decomposition (experiment B3)

Figure 7: Effect of cointegration on portfolio shares and human capital decomposition

Notes: Portfolio shares are averages over the sub-population with positive asset holdings.

Table 1: Parameter Estimates for the Idiosyncratic Earnings Process

Parameters	Full Sample	College Sub-sample	High School Sub-sample	Full Sample $\rho = 1$
ρ	0.957	0.959	0.902	1
	(0.009)	(0.014)	(0.061)	
σ_α^2	0.110	0.092	0.121	0.150
	(0.009)	(0.023)	(0.019)	(0.010)
av. σ_η^2	0.033	0.033	0.039	0.014
	(0.004)	(0.005)	(0.015)	(0.001)
$\sigma_\eta^2(L) - \sigma_\eta^2(H)$	0.020	0.012	0.025	0.011
	(0.005)	(0.007)	(0.008)	(0.003)
av. σ_ε^2	0.150	0.151	0.147	0.178
	(0.008)	(0.015)	(0.016)	(0.005)
Linear time trend in σ_η^2	0.0011	0.0014	0.0013	0.0006
	(0.0002)	(0.0004)	(0.0009)	(0.0001)

Notes: All models include a fourth-order polynomial in age in the variance of the transitory shock σ_ε^2. Reported variance are averages over age range. Standard errors computed by block bootstrap with 39 repetitions.

Table 2: Parameter Estimates for the Aggregate Stochastic Process

		No Cointegration			With Cointegration		
		Full Sample	College Sub-sample	High School Sub-sample	Full Sample	College Sub-sample	High School Sub-sample
Equation 1: Δu_t^1							
Δu_{t-1}^1	Γ_{11}	0.383	0.260	0.348	0.364	0.120	0.295
		(0.14)	(0.15)	(0.14)	(0.19)	(0.15)	(0.18)
$\log R_{t-1}^s$	Γ_{12}	0.044	0.040	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	0.000	0.029	0.000
		(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
$var\ (\varepsilon_{2,t}) \times 10^{-2}$		3.20	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				-0.670	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

	Data	No Cointegration	With Cointegration
	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
Lifetime	J, Ret	$(J, Ret) = (69, 40)$
Preferences	α Risk Aversion	$\alpha \in \{4, 6, 8, 10\}$
	$1/\rho$ Intertemporal Substitution	$1/\rho = 1.17$
	ψ_{j+1} Survival Probability	U.S. Life Table
	β Discount Factor	see Notes
Earnings	$e_j(z) = \begin{cases} z_1 g_j(z_2)(1-\tau) & \text{if} \quad j < Ret \\ z_1 b(\alpha)(1-\tau) & \text{if} \quad j \geq Ret \end{cases}$	Table 1 - 2
		$\tau = .27$
		$b(\cdot)$ see text
Returns	R^s, R^b	Table 2 - 3
Leverage	p	$p = 1$

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.

A Appendix

A.1 Proof of Theorems

Theorem 1: Assume A1. If $(c^*, n^*, e^*, y^*, a^*)$ solves Problem P1 and c^* is strictly positive, then (c^*, n^*, a^*, s^*) solves Problem P2, where $s_j^* \equiv 1, \forall j$, when the agent takes the value of human capital v_j, dividends d_j and the price of leisure p_j as exogenous.

Proof:

We develop necessary conditions to P1 in two steps. First, consider a restricted version of Problem P1 where earnings are exogenously given by e^* and leisure n is restricted to be no more than n^*. Since $(c^*, n^*, e^*, y^*, a^*)$ solves P1, it is clear that (c^*, n^*, a^*) also solves the restricted version of P1. Second, the Lagrangian and necessary conditions $P1-1$ to $P1-4$ for an interior solution to the restricted version of P1 are stated below and apply for each age j and shock history z^j. The Lagrangian omits the non-negativity constraint for consumption as by hypothesis c^* is strictly positive. Age subscripts are dropped (e.g. $c_j(z^j)$ is denoted $c(z^j)$) to simplify notation.

$$L = U(c,n) + \sum_j \sum_{z^j} \lambda(z^j)[\sum_{i \in \mathcal{I}} a^i(z^{j-1})R^i(z^j) + e^*(z^j) - c(z^j) - \sum_{i \in \mathcal{I}} a^i(z^j)]$$
$$+ \sum_j \sum_{z^j} \gamma(z^j)[n(z^j) - 0] + \sum_j \sum_{z^j} \rho(z^j)[n^*(z^j) - n(z^j)]$$

P1- 1: $dU(c,n)/dc(z^j) - \lambda(z^j) = 0$

P1- 2: $dU(c,n)/dn(z^j) + \gamma(z^j) - \rho(z^j) = 0$

P1- 3: $-\lambda(z^j) + \sum_{z_{j+1}} \lambda(z^j, z_{j+1})R^i(z^j, z_{j+1}) = 0, \forall i \in \mathcal{I}$

P1- 4: all constraints and complementary slackness conditions

Kuhn-Tucker conditions P2-1 to P2-5 are sufficient for a maximum as Problem P2 is a concave program. The Lagrangian function for Problem P2 is stated below.

$$L = U(c,n) + \sum_j \sum_{z^j} \lambda(z^j)[s(z^{j-1})(v(z^j) + d(z^j)) + \sum_{i \in \mathcal{I}} a^i(z^{j-1})R^i(z^j) - c(z^j) - \sum_{i \in \mathcal{I}} a^i(z^j)$$
$$-s(z^j)v(z^j) - p(z^j)n(z^j)] + \sum_j \sum_{z^j} \gamma(z^j)[n(z^j) - 0] + \sum_j \sum_{z^j} \delta(z^j)[1 - n(z^j)]$$

P2- 1: $dU(c,n)/dc(z^j) - \lambda(z^j) = 0$

P2- 2: $dU(c,n)/dn(z^j) - \lambda(z^j)p(z^j) + \gamma(z^j) - \delta(z^j) = 0$

P2- 3: $-\lambda(z^j) + \sum_{z_{j+1}} \lambda(z^j, z_{j+1})R^i(z^j, z_{j+1}) = 0, \forall i \in \mathcal{I}$

P2- 4: $-v(z^j) + \sum_{z_{j+1}} \frac{\lambda(z^j, z_{j+1})}{\lambda(z^j)}(v(z^j, z_{j+1}) + d(z^j, z_{j+1})) = 0$

P2- 5: all constraints and complementary slackness conditions

It remains to show that P1-1 to P1-4 evaluated at (c^*, n^*, a^*) imply that P2-1 to P2-5 hold at (c^*, n^*, a^*, s^*). Observe that (1) P1-1 implies P2-1, (2) P2-1, the definition of $p(z^j)$ and setting $\gamma(z^j) = \delta(z^j) = 0$ implies P2-2, (3) P1-3 implies P2-3, (4) substituting P2-1 into P2-4, it is straightforward to verify, from the definition of v, that P2-4 holds. It remains to show that (c^*, n^*, a^*, s^*) satisfies all constraints and complementary slackness conditions. The budget constraint in P2 holds by the budget constraint in the restricted version of P1 holding at (c^*, n^*, a^*), by $s_j^* \equiv 1$ and by the definition of dividends d. The leisure restrictions clearly hold at n^* and the complementary slackness conditions on leisure hold by setting $\gamma(z^j) = \delta(z^j) = 0$. $\diamond$

Theorem 3: If $1+r = \frac{1}{\beta}exp(\rho\mu - \frac{\rho^2\sigma^2}{2})$ and initial assets are zero, then $(c_j, a_{j+1}) = (e_j, 0), \forall j$ solves the decision problem. Furthermore, at this solution:

(i) the value of human capital is $v_j = f_j e_j$, where $f_j = \sum_{k=j+1}^{J} \beta^{k-j} exp((k-j)[(-\rho+1)\mu + (-\rho+1)^2 \frac{\sigma^2}{2}])$.

(ii) the return to human capital satisfies (a)-(c):

(a) $R_{j+1}^h = (\frac{1+f_{j+1}}{f_j})z_{j+1}$

(b) $E[R_{j+1}^h | z^j] = \frac{1}{\beta}exp(\mu\rho + \frac{\sigma^2}{2}(1 - (1-\rho)^2))$

(c) $E[R_{j+1}^h - R^f | z^j] = \frac{exp(\mu\rho - \frac{\sigma^2\rho^2}{2})}{\beta}[exp(\sigma^2\rho) - 1]$

Proof:

The conjectured solution $(a_1, ..., a_{J+1}) = (0, ..., 0)$ gives utility $E[\sum_{j=1}^{J} F_j(a_j, a_{j+1}, e_j)|z^1]$, where $F_j(a_j, a_{j+1}, e_j) \equiv \beta^{j-1}u(a_j(1+r) + e_j - a_{j+1})$. We show that the utility difference D between the conjectured plan and any alternative feasible plan $(a'_1, ..., a'_{J+1})$ is positive:

$$D \equiv E[\sum_{j=1}^{J}(F_j(a_j, a_{j+1}, e_j) - F_j(a'_j, a'_{j+1}, e_j))|z^1]$$

$$D \geq -E[\sum_{j=1}^{J}(F_{1j}(a_j, a_{j+1}, e_j)(a'_j - a_j) + F_{2j}(a_j, a_{j+1}, e_j)(a'_{j+1} - a_{j+1}))|z^1]$$

$$D \geq -E[F_{2J}(a_J, a_{J+1}, e_J)(a'_{J+1} - a_{J+1})|z^1] = -E[F_{2J}(a_J, a_{J+1}, e_J)a'_{J+1}|z^1] \geq 0$$

The first equation defines the utility difference. The second equation follows because F_j is concave and differentiable in the first two arguments. These properties of F_j are implied by u being concave and differentiable. The leftmost inequality in the third equation follows from two facts. First, the Euler equation $E[F_{2j}(a_j, a_{j+1}, e_j) + F_{1j+1}(a_{j+1}, a_{j+2}, e_{j+1})|z^j] = 0$ holds at the conjectured solution. This allows one to cancel out all but two terms inside the expectation in the second equation. Second, the fact that initial assets are equal in all feasible asset plans (i.e. $a_1 = a'_1$) implies that the only remaining non-zero term inside the expectation is that stated in the third equation. The equality in the third equation holds as $a_{J+1} = 0$. The remaining inequality holds as $F_{2J} < 0$ and as $a'_{J+1} \geq 0$ in any feasible plan.

(i) The result follows from the definition of the value of human capital:

$$v_j(z^j) = E[\sum_{k=j+1}^{J} m_{j,k}e_k|z^j] = E[\sum_{k=j+1}^{J} \beta^{k-j}\frac{e_k^{-\rho}}{e_j^{-\rho}}e_k|z^j] = E[\sum_{k=j+1}^{J} \beta^{k-j}\prod_{l=j+1}^{k} z_l^{-\rho+1}]e_j = f_j e_j$$

$$f_j \equiv E[\sum_{k=j+1}^{J} \beta^{k-j}\prod_{l=j+1}^{k} z_l^{-\rho+1}] = E[\sum_{k=j+1}^{J} \beta^{k-j}\prod_{l=j+1}^{k} exp((-\rho+1)\ln(z_l))]$$

$$f_j = \sum_{k=j+1}^{J} \beta^{k-j} exp((k-j)[(-\rho+1)\mu + (-\rho+1)^2\frac{\sigma^2}{2}])$$

(ii) Result (a) follows from the definition. Result (b) follows from (a) and the fact that $f_j = E[\beta z_{j+1}^{-\rho+1}(1 + f_{j+1})]$. Result (c) follows from (b) and the restriction on the risk-free rate.

$$(a)\ R^h_{j+1} \equiv \frac{e_{j+1} + v_{j+1}}{v_j} = \frac{e_{j+1} + f_{j+1}e_{j+1}}{f_j e_j} = (\frac{1 + f_{j+1}}{f_j})z_{j+1}$$

$$(b)\ E[R^h_{j+1}|z^j] = (\frac{1 + f_{j+1}}{f_j})exp(\mu + \frac{\sigma^2}{2}) = (\frac{1}{\beta exp((1-\rho)\mu + \frac{(1-\rho)^2\sigma^2}{2})})exp(\mu + \frac{\sigma^2}{2})$$

$$(b)\ E[R^h_{j+1}|z^j] = \frac{1}{\beta}exp(\mu\rho + \frac{\sigma^2}{2}(1 - (1-\rho)^2))$$

◇

A.2 Computation

This section describes our methods to compute solutions to the model in section 4 and to compute values and returns.

A.2.1 Value Function and Decision Rules

We want to compute the optimal value function V_j^* and optimal decision rules. We employ the method of dynamic programming. This involves computing functions V_j solving the Bellman equation (BE). Of course, the idea is that these two value functions coincide $V_j = V_j^*$. In stating $\hat{\Gamma}_j(x,z)$ in Bellman's equation, we impose all the restrictions from the original budget constraint $\Gamma_j(x,z)$ as well as the solvency restrictions.[29]

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

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

$$\hat{\Gamma}_j(x,z) = \{(c,a^1,a^2) : (i)-(ii)\ hold\}$$

$$(i)\ c + \textstyle\sum_{i\in\mathcal{I}} a^i \leq x, c \geq 0, a^2 \geq 0, a^2 \leq p \sum_{i\in\mathcal{I}} a^i$$

$$(ii)\ x' = \textstyle\sum_{i\in\mathcal{I}} a^i R^i(z') + G_{j+1}(z') \geq \underline{x}_{j+1}(z'), \forall z' \in Z$$

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 $e_j = G_j(z) = z_1 H_j(z_2)$ and $z_1' = z_1 f_{j+1}(z_2')$, 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.

$$\mathcal{L} = \mathrm{W(c_j, F(V_{j+1}(x', z_1', z_2')), j)} + \lambda_1[\mathrm{a}^2 - 0] + \lambda_2[\mathrm{p} \sum_{\mathrm{i}\in\mathcal{I}} \mathrm{a^i} - \mathrm{a}^2]$$

$$(1)\quad -W_1 + W_2 dF/da^1 + \lambda_2 p = 0$$

$$(2)\quad -W_1 + W_2 dF/da^2 + \lambda_1 + \lambda_2(p-1) = 0$$

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

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.

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

$$(2')\ -1 + \beta\psi_{j+1} E[(\tfrac{c_{j+1}}{c_j})^{-\rho}(\tfrac{V_{j+1}}{F(V_{j+1})})^{\rho-\alpha} R^2(z')|x,z] + \lambda_1' + \lambda_2'(p-1) = 0$$

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 $(a_{j+1}^1(x,1,z_2), a_{j+1}^2(x,1,z_2))$ at grid points $(x,1,z_2)$ by solving $(1')$–$(2')$ and (3).

[29] Solvency requires that asset holding restrictions are consistent with positive consumption at all dates and shock histories. This is expressed in $\hat{\Gamma}_j(x,z)$ in constraint (ii) which says that asset holdings always keep future financial wealth above minimum levels $\underline{x}_{j+1}(z')$ consistent with solvency. These minimum levels are defined via $\underline{x}_J(z) \equiv 0$ and the recursion $\underline{x}_j(z) = inf\{x : \hat{\Gamma}_j(x,z) \neq \emptyset\}$. The recursion implies that $\underline{x}_j(z) = 0, \forall j, \forall z \in Z$, given that earnings are non-negative (i.e. $G_{j+1}(z') \geq 0$).

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.

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

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

A.2.2 Human Capital Values and Returns

We describe how to compute human capital values and returns in the benchmark model. 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 3. Human capital values $v_j(x,z)$ follow the simple recursion $(**)$, given $v_J(x,z) = 0$:

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

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

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

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

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 $e_j = z_1 H_j(z_2)$ and use the fact that $z_1' = z_1 f_{j+1}(z_2')$ which is consistent with the formulation in Table 4.

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

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

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

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

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 $\hat{v}_j$.

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

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

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

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

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

$\diamond$

An algorithm to compute the naive value $v_j^n(z)$ is provided. First, we list some useful points from theory, where $z = (z_1, z_2)$, $e_j(z) = z_1 H_j(z_2)$ 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.

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

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

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

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.

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

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

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: $y = v_{j+1}(x', z_1 f_{j+1}(z_2'), z_2') + z_1 f_{j+1}(z_2') H_{j+1}(z_2')$. 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).

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

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

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

◇

A.3 Data Appendix

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. $v\left(t\right)$ 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

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

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

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

The model with $p = 2$ becomes

$$\begin{pmatrix} \Delta u_t^1 \\ \log R_t^s \\ w_t \end{pmatrix} = \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_1 + \beta_2\gamma_2 + \rho \end{pmatrix} + \begin{pmatrix} \Gamma_{11} & \Gamma_{12} & \alpha_1 \\ \Gamma_{21} & \Gamma_{22} & \alpha_2 \\ \Gamma_{11} + \beta_2\Gamma_{21} & \Gamma_{12} + \beta_2\Gamma_{22} & 1 + \alpha_1 + \beta_2\alpha_2 \end{pmatrix} \begin{pmatrix} \Delta u_{t-1}^1 \\ \log R_{t-1}^s \\ w_{t-1} \end{pmatrix} \tag{8}$$

$$+ \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{1t} + \beta_2\varepsilon_{2t} \end{pmatrix} \tag{9}$$

No cointegration implies the restriction that $\alpha_1 = \alpha_2 = 0$ in which case (8) becomes a two variable VAR. When the underlying VAR in levels includes only one lag, so that $p = 1$, the restriction is that $\Gamma_{ij} = 0$. To implement this in our model we treat the aggregate component of the state vector as $s_t = \begin{pmatrix} \Delta u_t^1 & \log R_t^s & w_t \end{pmatrix}'$ and we construct a discrete Markov chain that approximates (8).

A.3.2 Data and Sample Selection

We use the core PSID sample from waves 1968 to 1997, which refers to earnings in years 1967 to 1996. After 1997 the PSID became a bi-annual survey, hence we exclude the more recent waves. We restrict attention to male heads of household between the ages of 22 and 60 with annual labor income of at least \$1000 in 2008 dollars. Our measure of annual labor income 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. Our final sample contains $63,541$ observations on $5,622$ individuals. The median number of annual observations per individual is 9. We construct three education samples: one comprising all males (Full Sample), one comprising males with 12 or fewer years of education (High School Sub-sample) and one comprising males with at least 16 years of education (College Sub-sample).

Our CPS data comes from the IPUMS database of March Outgoing Rotation Groups. We use data on earnings from 1967 to 2008 and impose the same age and minimum earnings selection criteria as for the PSID. The aggregate components of labor earnings for each subsample are measured as the coefficients on year dummies in a regression that is analogous to the one described below.

A.3.3 Estimation of Idiosyncratic Earnings Model

Estimation is done in two stages. In the first stage we estimate u_t^1 and 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. Residuals from the first-stage regression are then used to estimate the remaining parameters of the individual earnings equation, $\left(\rho, \sigma_\alpha^2, \sigma_{\eta,t}^2(X_t), \sigma_{\nu,j,t}^2(X_t)\right)$. The auto-covariance function for residual log-earnings is calculated for up to 10 lags for every age/year combination. For this purpose, 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. A GMM estimator is then used to estimate the parameters, where the moments included are the elements of the auto-covariance function. 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 and covariance from ages 24 to 58 are effectively used in the estimation. Standard errors are calculated by bootstrap with 250 repetitions, thus accounting for estimation error induced by the first-stage estimation.

When estimating models that allow for cyclical variation in σ_η^2 and σ_ν^2, we base our choice of growth/contraction years on the sign of $\Delta\hat{u}_t^1$. Table A1 shows the contraction years between 1967 and 1996 as implied by the estimates from the CPS sample, the PSID sample, and aggregate measures from the National Income and Product Accounts. Based on these numbers, we set the contraction years to be 1970, 74-75, 79-82, 89-91 and 93.

Table A.1: Contraction Years by Data Source 1967-1996

Data Source	Contraction Years
CPS: Labor earnings	70,74-75,79-82,87,89-91,93
PSID: Labor earnings	70,74-75,79-83,89-91,92-93
NIPA: GDP per capita	70,74-75,79-82,90-91,93,95
NIPA: GNP per capita	70,74-75,79-82,90-93,95
NIPA: Wage and salaries per capita	70-71,74-75,79-82,89-91,93

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.

A.3.4 Estimation of Aggregate Stochastic Process

We test for the lag order (p) of the underlying VAR in (4). 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 in (4). We thus focus our attention on a model with one lag in the VAR in first differences, as in (6).

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 1667-2008 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.3.5 Experiment B2 and B3

Section 5 of the paper analyzes the importance of cointegration. This Appendix describes the construction of the system of equations underlying experiment B2 and B3 described in section 5. The three equations below are equation 2, 8 and 14 from Benzoni et al (2007), where y_t is log 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 that Benzoni et al (2007) highlight in their analysis.

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

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

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

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 for the Aggregate Stochastic Process with Alternate 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 for Experiments B2 and B3

	Benchmark Model	Experiment B2	Experiment B3
$sd\left(\Delta u_t^1\right)$	0.025	0.069	0.069
$sd\,(\log R_t^s)$:	0.187	0.160	0.160
$corr\left(\Delta u_t^1, \log R_t^s\right)$	0.177	0.000	0.000
$corr\left(\Delta u_t^1, \Delta u_{t-1}^1\right)$	0.441	0.327	0.330
$corr\,(\log R_t^s, \log R_{t-1}^s)$	0.055	0.000	0.000
$corr\left(\Delta u_t^1 \log R_{t-1}^s\right)$	0.398	0.347	0.350
$corr\left(\log R_t^s, \Delta u_{t-1}^1\right)$	-0.270	0.000	0.000

Notes: Table shows implied steady-state statistics from Benchmark model and experiments B2 and B3. 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 $E[\Delta u_t^1] = 0$.

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 (10) below, using $\Delta u_t \equiv u_t - u_{t-1}$ and $w_t \equiv (u_t - y_t - \bar{uy})$.

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

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

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

$$\begin{pmatrix} \Delta u_{t+1} \\ \log R_{t+1} \\ w_{t+1} \end{pmatrix} = \begin{pmatrix} g - \sigma^2/2 \\ \mu \\ 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & -\kappa \\ 0 & 0 & 0 \\ 0 & 0 & 1-\kappa \end{pmatrix} \begin{pmatrix} \Delta u_t \\ \log R_t \\ w_t \end{pmatrix} + \begin{pmatrix} \nu_1 z_{1,t+1} + (\sigma - \nu_3) z_{3,t+1} \\ \sigma z_{3,t+1} \\ \nu_1 z_{1,t+1} - \nu_3 z_{3,t+1} \end{pmatrix} \tag{10}$$

System (10) produces the steady-state statistics listed in Table A5 for experiment B2. 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 (10) 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.

The goal of experiment B3 is to eliminate cointegration from system (10), by setting $\kappa = 0$, while at the same time preserving a number of properties. The resulting process B3 has the same (i) steady-state means, (ii) steady-state contemporaneous variance and covariances and (iii) steady-state first-lagged autocorrelations. The system takes the form of (11) below.

$$\begin{pmatrix} \Delta u_{t+1} \\ \log R_{t+1} \end{pmatrix} = \begin{pmatrix} g - \sigma^2/2 \\ \mu \end{pmatrix} + \begin{pmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{pmatrix} \begin{pmatrix} \Delta u_t \\ \log R_t \end{pmatrix} + \begin{pmatrix} \nu_1 z_{1,t+1} + (\sigma - \nu_3) z_{3,t+1} \\ \sigma z_{3,t+1} \end{pmatrix} \tag{11}$$

When we set $(\Gamma_{11}, \Gamma_{12}, \Gamma_{21}, \Gamma_{22}, \nu_1, \nu_3, \sigma) = (.33, .151, 0, 0, 0.061, 0.16, 0.16)$, system (11) generates the same steady-state contemporaneous variance and covariances and first-lagged autocorrelations as system (10).