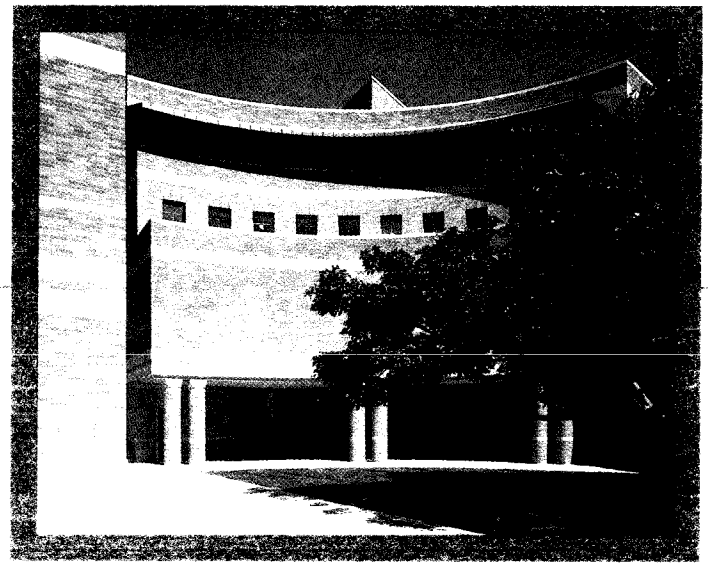
THE ANDERSON SCHOOL AT UCLA



#32-00

Demographics and Expected Returns

November 2000

Jiang Luo
Anderson Graduate School of Management
University of California, Los Angeles

Finance Working Paper
Sponsored by:

Comments Welcome

Demographics and Expected Returns*

Jiang Luo†

University of California, Los Angeles

November 6, 2000

Abstract

This paper studies the relation between demographics and the equity risk premium in a dynamic overlapping generations (OLG) equilibrium model. Investors have both labor and investment income. The labor income and the dividend processes are correlated. Investors trade stocks for consumption purposes and to hedge against the risk of labor income. The per capita stock supply is normalized to unity, and the demographic structure is time varying. In equilibrium, the equity risk premium is linear in the real per capita stock price, the dividend yield and the dividend payout ratio, but the coefficients of the linear relation are time varying because of demographic change. Proxying the coefficients by linear functions of the change in the share of population in the age range 40-64, we derive a non-linear predictive regression for the equity risk premium, which is not only significant in the empirical tests using post-1947 data but also improves significantly on previous predictive relations.

*I gratefully acknowledge the guidance and advice of Michael Brennan. Helpful comments and suggestions were also received from Antonio Bernardo, Amit Goyal, John Riley and Avanidhar Subrahmanyam. All remaining errors are mine.

†Corresponding Address: Department of Finance, The Anderson Graduate School of Management; University of California, Los Angeles; 110 Westwood Plaza; Los Angeles; CA 90095-1481. Phone: (310) 825-8155. Fax: (310) 206-5455. E-mail: jiang.luo@anderson.ucla.edu.

1 Introduction

The late 1990s is a particularly challenging period for models that predict the equity risk premium. The stock market has had a strong upwards movement, while the dividend yield and the dividend payout ratio have been relatively low. These two variables, which are significant in predicting the equity risk premium in the period 1947-1994 (Fama and French (1988) and Lamont (1998)), lose their forecasting power when the late 1990s data are included. During this period of extraordinary stock market performance, in which previous predictive relations have failed, the baby boomers entered their middle age. Is this just a coincidence? If not, it is important to understand how stock price movements are related to variations in the demographic structure, and whether demographic change has caused the simple predictive relations between the dividend yield and/or the dividend payout ratio and the equity risk premium to break down.

This paper develops a model which predicts that the break down is caused by demographic change. The intuition for the relation between demographics and the equity risk premium can be easily derived from life cycle portfolio theories, which were pioneered by Modigliani and Brumberg (1954), Friedman (1957) and Samuelson (1969). They assume that the objective of an investor's consumption-investment decision is to smooth consumption over time to maximize lifetime utility, which induces a life cycle pattern in investment behavior. Suppose that the supply of risk assets is constant or inelastic. If an investor's demand for risky assets is affected by his age, one might expect that when the population is dominated by the investors of the age at which the demand for risky assets is highest, the stock price will rise to ensure market clearing and the equity risk premium will fall.[1] Moreover, when the population is dominated by the investors of the age at which the demand for risky assets is the most sensitive to certain macroeconomic factors, the stock price will become sensitive to these factors as well.

In this paper, we develop a dynamic overlapping generations (OLG) equilibrium model,

[1]This effect is similar to the common wisdom of "too much money chasing too few deals", which has been found in the venture capital investment by Gompers and Lerner (2000). They show that inflows of capital into venture funds increase the valuation of these funds' new investments, which is consistent with competition for a limited number of attractive investments being responsible for rising prices.

which allows us to study the relation between demographics and the equity risk premium. We assume that the per capita stock supply is constant,[2] and that the demographic structure is deterministic and exogenous but time varying. We are primarily interested in the information for stock returns contained in the demand side of the economy, which is represented by the demographic structure in our case.

In the model, agents' income derives from two sources: labor and capital investments. Labor income is received in the form of wages. Capital is assumed to pay dividends. The labor income process and the dividend process are correlated. Investors' demand for stocks is affected by their desire to hedge against the risk of labor income. In equilibrium, the excess return on stocks follows an AR(1) process (with deterministic time varying coefficients), and the expected excess return is linear in the labor income state variables. We show that investors with exponential utility over consumption are more risk tolerant and, everything else equal, hold more risky assets on average when young than when old.[3] The sensitivity of investors' demand for risky assets to the labor income state variables is also age specific. When the demographic structure is time varying, the per capita demand for stocks would also be time varying if the excess stock return were independent of the demographic structure. Therefore, with a constant per capital supply of stocks, the excess stock return must be related to the demographic structure to ensure that the per capita demand for stocks equals the per capita supply.

A non-linear relation between the equity risk premium and demographic variables, the real per capita stock price, the dividend yield and the dividend payout ratio is derived from the equilibrium price function. The equity risk premium is linear in the real per capita stock price, the dividend yield and the dividend payout ratio; however, the coefficients of this linear relation depend on the demographic structure, which is time varying. These coefficients are functions of moments of the current and all the future demographic structures. Calibration

[2]It is for simplicity to assume that the per capita stock supply is constant. Thus, the supply of risk assets is proportional to the population. To allow for a stochastic stock supply will introduce another state variable, which affects the stock price and the expected excess return. But our results about the predictive relation for the equity risk premium still hold.

[3]Samuelson (1991) shows that this life cycle pattern in investment behavior holds when investors have exponential utility over consumption and the excess return on stocks is mean-reverting with constant coefficients, which is equivalent to an AR(1) process. Here, we show that this investment behavior holds even when the coefficients of the AR(1) process are time varying.

results show that the historical demographic change could induce significant variations in these coefficients.

We test the significance of the empirical relation between the equity risk premium and the real per capita stock price, the dividend yield, the dividend payout ratio and the demographic variables. Since the demographic variables in the predictive relation are not observable, we proxy them by linear functions of the change in the share of population in the age range 40-64. Calibration results suggest that households in this age range have the highest labor income, but that their labor income is the least sensitive to the macroeconomic risk, as measured by the unemployment rate. An increase in the share of population in this age range reduces the macroeconomic risk in the whole economy, reducing the importance of stocks as a hedging vehicle. However, since households in this age range also have the highest level of idiosyncratic labor income risk, and their risk aversion is relatively high, the increase in this share of population also makes the population more risk averse, which tends to make the role of stocks as a hedging vehicle more important. Calibration results suggest that the second effect dominates. This justifies the previous research for the equity risk premium, for example, Poterba (2000), in using this share as an explanatory variable in predictive regressions, though he focuses on the predictive power of the level of this share for real stock returns, instead of the equity risk premium, and finds no significant results. Calibration results also suggest that the demographic variables in the predictive relation are correlated with the *change* in this share of population.

Writing the demographic variables in the predictive relation as linear functions of the change in the share of population in the age range 40-64, we express the equity risk premium as a linear function of the change in this share of population, the real per capita stock price, the dividend yield, the dividend payout ratio and the products of the change in this share of population with the real per capita stock price, the dividend yield and the dividend payout ratio. Empirical evidence supports the derived predictive regression for the equity risk premium, whether or not the late 1990s data are included. The inclusion of demographic variables significantly improves on the previous predictive regressions, in which the equity risk premium is regressed on the dividend yield and/or the dividend payout ratio. In particular, replication of the regression by Lamont (1998) for the period 1947-1994, in

which the quarterly equity risk premium is regressed on the dividend yield and the dividend payout ratio, yields R^2 of 11.7%, while adding the change in this share of population to the regression raises R^2 to 14.7%. For the period 1947-1999, the previous predictive regressions are no longer significant, while our predictive regression is significant with R^2 of 13.6%.

Our model salvages the predictive power of the dividend yield and the dividend payout ratio, but suggests a changing structure for the predictive relation, which is caused by demographic change. When demographic change is smooth, as in the period 1947-1994, one can ignore the structural change when studying the predictive power of the dividend yield and the dividend payout ratio, even though this usually yields lower R^2. But in late 1990s, the structural change is large, and this must be taken into account in the regression if meaningful results are to be obtained. Not surprisingly, simple regressions of the equity risk premium on the dividend yield and/or the dividend payout ratio with constant coefficients for the period 1947-1999 are not significant.

Prior research on the predictability of the equity risk premium either ignores demographic considerations or focuses solely on the effects of demographic change on stock returns. These two strands of research on the equity risk premium are totally distinct in spite of the fact that they are both attempting to explain the same variable. They are based on different theoretical frameworks, and their explanatory variables seldom overlap.

The first strand of research investigates the information in stock prices, dividends and/or earnings for stock returns. It has identified several explanatory variables that are statistically important in predicting stock returns, for example, the dividend yield (Fama and French (1988)), the earnings yield (Shiller (1984)) and, most recently, the dividend payout ratio (Lamont (1998)) and etc..[4] The relation between the equity risk premium and explanatory variables is usually derived from the present value formula, which expresses the current stock price as the discounted value of future dividends. However, Bossaerts and Hillion (1999), while confirming the presence of in-sample predictability across several different national markets, find that even the best prediction models have no out-of-sample forecasting power. In fact, if data for the late 1990s are included, even the in-sample tests fail to be significant.

[4]See Cochrane (1997) for survey.

Several explanations have been offered for this failure. Most focus on the statistical biases of empirical tests.[5]

Several authors, including Viceira (1997), Goyal and Welch (1999) and Pesaran and Timmermann (1995), have pointed to the possibility of a changing structure in the predictive regression for the equity risk premium. Viceira (1997) tests if there is a structural break in the relation between the dividend yield and the stock return, but fails to detect one. Goyal and Welch (1999) provide a "learning market hypothesis", under which investors' attempts to take advantage via market-timing strategies of the dividend yield's forecasting ability drive its forecasting power to zero. Pesaran and Timmermann (1995) find that the predictive power of various economic factors over stock returns changes through time and tends to vary with the volatility of returns. In addition, the timing of the episodes where many of the regressors get included in the forecasting model seems to be linked to macroeconomic events. They suggest using forecasting procedures that allow for possible regime changes in analyzing stock return predictability.

The second strand of research on the equity risk premium investigates the relation between demographics and stock returns.[6] The theoretical research in this strand, such as Yoo (1994a), Brooks (1999) and Abel (1999), focuses on whether a shifting age structure can significantly affect equilibrium asset returns and asset prices. These authors present simulation or analytic results, which suggest that demographic change can affect equilibrium returns.

The empirical research in this strand is typically based on life cycle portfolio selection theories,[7] and proceeds to test either the direct relation between (changes in) shares of population between certain ages and stock returns, like Erb, Harvey and Viskanta (1999), Poterba (2000), Macunovich (1997) and Yoo (1994b), or the relation between changes in moments of the population age structure and stock returns, like Bakshi and Chen (1994). Bakshi and Chen (1994) claim that the extraordinary stock market performance in the

[5]For example, Goetzman and Jorion (1993) and Nelson and Kim (1993) on small sample bias; Hodrick (1992) on bias of test statistics in long-horizon forecasting; Goetzman and Jorion (1995) on survivorship bias; Stambaugh (1999) on bias arising from near-nonstationarity in regressors; and etc..

[6]See Poterba (2000) for survey.

[7]For theoretical work, see Modigliani (1986), Canner, Mankiw and Weil (1997), Samuelson (1989, 1991), Bodie, Merton and Samuelson (1992), Kimball (1993), Brennan (1998) and Xia (2000); for empirical evidence, see Mankiw and Weil (1989), Bossons (1973), Poterba and Samwick (1997) and Heaton and Lucas (1999).

1990s is caused by the fact that the baby boomers entered their middle age and thereafter the population average age increases at the same time. Some researchers, including Siegel (1998) and Schieber and Shoven (1997), have predicted that stock prices will drop when the baby boomers retire later in this century.

The demographic part of our analysis is most closely related to Bergantino (1998), who studies the U.S. stock market. He first develops estimates of age specific asset demands, and then uses these demands along with the changing demographic structure to construct time varying estimates of the demand for financial assets. The findings suggest clear relations between the level of age specific asset demand and the level of stock prices, and between the difference in "demographic demand" and difference in asset prices over multi-year horizons. Our model is different from that of Bergantino in that we consider the effects of demographics on stock returns in a general equilibrium framework, in which investors anticipate the effects of their future demands on stock prices and returns when choosing their portfolios. Bergantino simply takes investors' life cycle pattern of investment behavior as exogenous.

Although it is commonly agreed that demographics affect stock returns,[8] the predictive power of demographic variables is not high, especially in forecasting short-term returns. One reason for this is that the second strand of research ignores the information in dividends and/or earnings for stock returns. Ideally, with sufficiently precise high-frequency demand data, the demographic data in this case, it should be possible to provide a good estimate of the equity risk premium, which might be consistent with that conditional on the dividend yield and/or the dividend payout ratio. But the time series of demographic data are rather limited. For example, for U.S. demographic data, we only have annual observations in 5 or 10 year intervals in people's age. In addition, demographic change is quite smoother than the stock return. In general, it is difficult to predict a noisy time series by a slow moving time series. Not surprisingly, regressions of short-term, for example, quarterly, returns on demographic variables have low R^2s. This greatly limits the application of demographics in practice. Finally, a question for any partial equilibrium study is: if investors adjust their demands when they realize that demographic factors affect stock prices and therefore stock

[8]See Bakshi and Chen (1994), Macunovich (1997), Erb, Harvey and Viskanta (1997), Bergantino (1998), Poterba (2000) and etc..

returns, will the life cycle investment pattern and the relation between demographic change and stock returns still hold? The partial equilibrium models claim that, given the return process, investors' investment behavior will have a life cycle pattern. Then, they conclude that demographic change should be correlated with stock prices and returns. In another words, they first assume that stock returns are exogenously given, then conclude that stock returns should be endogenously determined. Although the intuition is plausible, the logic is defective.

Our paper unifies these two strands of research on the predictability of the equity risk premium using an OLG general equilibrium model. We consider both the information in stock prices, dividends and earnings, and the information in demographics for stock returns, and provides a clear description of the determination of the equity risk premium. The non-linear predictive relation between the equity risk premium and the change in the share of population in the age range 40-64, the real per capita stock price, the dividend yield and the dividend payout ratio not only confirms the previous findings of the predictive power of the dividend yield and the dividend payout ratio, but also leads to significant improvements on previous empirical research.

The remainder of the paper is organized as follows: Section 2 presents the model; Section 3 studies the investors' optimization problem and obtains the equilibrium; Section 4 further simplifies the economy and calibrates it; Section 5 derives the non-linear predictive regression for the equity risk premium and conducts the empirical test; Section 6 concludes.

2 The Model

Consider a simple economy with a single good that can be either consumed or invested. The economy is defined as follows.

2.1 Time parameters and the Demographic Structure

Let τ, $\tau = 0, 1, ..., T$, denote the age of an investor who is assumed to live for $T + 1$ periods. For simplicity, T is assumed to be a constant.[9] The number of investors alive at time t is $G(t)$. The fraction of investors of age τ at time t is $g(\tau; t)$. $G(t)$ and $g(\tau; t)$ are assumed to be exogenous and deterministic, and are common knowledge. Thus, there is no uncertainty about the demographic structure among investors. Investors can fully anticipate the effects of demographic change on the economy when making consumption-investment decisions.

2.2 Preferences

All investors are assumed to have identical constant absolute risk aversion (CARA) preferences defined on their consumption. At time t, an investor of age τ maximizes expected utility of the following form

$$E_t\left[-\sum_{s=t}^{\bar{\tau}} \beta^{\tau+s-t} \exp(-\alpha c_s)\right] \tag{2.1}$$

where E_t is the expectations operator conditional on his information at time t; c_s is his consumption in period s; $\bar{\tau} = t + T - \tau$ is the date of which he will leave the market.[10]

2.3 The labor income process

Investors derive their income from two sources: labor and investments. Labor income is defined to include all the income other than that obtained from financial assets.[11] At time t, an investor of age τ has labor income

[9] We ignore uncertainty about life-span, which has been studied by Hubbard, Skinner and Zeldes (1994). To incorporate uncertain life-span in the model may introduce second order effects, but is unlikely to change our main results.

[10] We ignore bequests in the model. Including bequests has only a negligible effect on investors' utility functions. Also, since we have a fairly broad definition of labor income, the non-financial bequests are already captured in the labor income process.

[11] We ignore endowments in the model. If agents receive endowment only when they are born, the model still holds. Also, since we have a fairly broad definition of labor income, non-financial endowments are already captured in the labor income process.

$$Y_{\tau,t} = h_\tau + n_\tau t + \frac{1}{2} Z_t^T \omega_\tau Z_t + \epsilon_{\tau,t} \tag{2.2}$$

Labor income has a time and age dependent deterministic component, $h_\tau + n_\tau t$, which captures the secular trend in labor income and the deterministic element of the life cycle. In addition, it has two stochastic components, $\frac{1}{2} Z_t^T \omega_\tau Z_t$ and $\epsilon_{\tau,t}$. The first stochastic component, $\frac{1}{2} Z_t^T \omega_\tau Z_t$, where Z_t is an $N \times 1$ column vector of normal variables and ω_τ is a symmetric $N \times N$ matrix dependent on τ, captures the common stochastic component in labor income. The labor income state variable vector, Z_t, can be interpreted as a vector of macroeconomic variables that affect all investors' labor income, such as changes in the business cycle and the production technology. Its effects on the labor income of investors of different age, which are measured by ω_τ, may be different. The quadratic form, $\frac{1}{2} Z_t^T \omega_\tau Z_t$, can be calibrated to capture a rich class of characteristics of labor income, for example, the fat tail, skewness and even the time variation in the risk of labor income.[12] The second stochastic component, $\epsilon_{\tau,t}$, is an idiosyncratic temporary shock, i.e. $\epsilon_{\tau,t} \perp \epsilon_{\varrho,s}$ for $\tau \neq \varrho$ or $t \neq s$, and has a normal distribution $\epsilon_{\tau,t} \sim N(0, \sigma_\tau^2)$. The size of the shocks, $\epsilon_{\tau,t}$, as measured by the variance σ_τ^2, may depend on age.

Z_t is assumed to follow an AR(1) process

$$Z_t = a_Z Z_{t-1} + \epsilon_{Z,t} \tag{2.3}$$

where a_Z is an $N \times N$ constant matrix and all its eigenvalues are less than 1; $\epsilon_{Z,t}$ is an $N \times 1$ column vector, and has an independent and identical normal distribution. The process of Z_t can be calibrated to capture the short-term persistence in the aggregate labor income process, which is documented by Campbell (1996) and Pischke (1995).

[12] The non-trivial quadratic form of $\frac{1}{2} Z_t^T \omega_\tau Z_t$ is essential to our model. Otherwise, we will have a degenerated equilibrium, in which the stock price does not depend on the labor income state variable vector, Z, and the risk premium depends only on the real per capita stock price.

2.4 Financial assets

There are two publicly traded assets in the economy, a riskless asset and a risky asset (stock). The riskless asset is assumed to be in infinitely elastic supply at a positive constant rate of return r. Its gross rate of return is thus $R = 1 + r$. We normalize the per capita supply of stocks to unity.[13] Thus, the number of shares outstanding at time t is $G(t)$. Each share of stocks pays a dividend D_t at time t, and D_t is governed by the process

$$D_t = \bar{D} + F_t + \epsilon_{D,t} \tag{2.4}$$

where $\bar{D}$ is a constant. $\epsilon_{D,t}$ is a temporary shock to D_t. It has an independent and identical normal distribution. The dividend state variable, F_t, also follows an AR(1) process

$$F_t = a_F F_{t-1} + \epsilon_{F,t} \tag{2.5}$$

where a_F is a constant with $0 \leq a_F < 1$. $\epsilon_{F,t}$ is a temporary shock to F_t. It has an independent and identical normal distribution. Thus, the dividend is the sum of a constant and two random components, one transitory and one permanent. The permanent random component can be calibrated to capture the short-term persistence in dividends.

The "fundamental" of stocks, F_t, can be empirically interpreted as a value proportional to the demeaned earnings. Denote the earnings at time t by B_t and its long run level by $\bar{B}$. F_t can be expressed as $F_t = a_B(B_t - \bar{B})$, where a_B is a constant. For simplicity, we use F_t in the derivation of the equilibrium. In the later calibration and empirical work, we follow the tradition of using earnings to predict stock returns, and use B_t in place of F_t.

[13]This normalization is for simplicity. To allow for a stochastic stock supply will introduce another state variable, which affects the stock price and the excess return. But the predictive relation for the risk premium, which is derived later, still holds.

2.5 Distributional assumptions

The stochastic dimension of the economy is determined by the rank of the covariance matrix of $\{\epsilon_{\tau,t}\}_{\tau=0}^{T}$, $\epsilon_{D,t}$, $\epsilon_{F,t}$ and $\epsilon_{Z,t}$. For simplicity, we assume that $\epsilon_{\tau,t}$ is cross-sectionally uncorrelated with all other random variables, $\{\epsilon_{s,t}\}_{s\neq\tau}$, $\epsilon_{D,t}$, $\epsilon_{F,t}$ and $\epsilon_{Z,t}$. Define $\xi_t \equiv (\epsilon_{D,t}, \epsilon_{F,t}, \epsilon_{Z,t})^T$. It is shown later that the stochastic dimension of the asset return structure is determined by the rank of the covariance matrix of ξ_t. We assume that ξ_t has a multivariate independent and identical normal distribution $\xi_t \sim N(0_{(N+2)\times 1}, \Sigma)$, where Σ is nonsingular and, in general, is not diagonal. Thus, dividends and labor income can either be positively or negatively correlated. Since labor income is given exogenously, stocks provide a vehicle for investors to hedge against the risk of labor income.

2.6 Informational assumptions

The structure of the economy is common knowledge. At time t, investors observe D_t, F_t and Z_t. The Markovian structure of the economy implies that past information for dividends and labor income is redundant for investors' consumption-investment decision making. Therefore, investors have identical sufficient information sets $\Im_t \equiv \{D_t, F_t, Z_t\}$.

The assumptions that are essential to the tractability of the model are the following. First, the riskless interest rate is constant. Secondly, dividends follow a Gaussian process, and labor income is a quadratic form of Gaussian processes. Thirdly, utility over consumption is negative exponential. The assumptions of a constant riskless interest rate, a Gaussian dividend process and a quadratic Gaussian labor income process, and negative exponential utility are restrictive. Under this framework, investors' demands for stocks do not depend on their wealth. Thus, the aggregate demand is independent of the wealth distribution. This framework drastically simplifies our derivation and makes the over-lapping generations equilibrium solvable.

3 Equilibrium

In this section, we solve for the equilibrium of the economy defined in Section 2. The method is similar to that of Campbell and Kyle (1993). We first conjecture an equilibrium price function. Based on the conjectured price function, we solve the investors' optimization problem. Market clearing is then imposed to verify the conjectured price function.

As a prelude to the conjecture-verification procedure, consider first the optimization problem, at time t, of an investor of age τ. Let W be his wealth, c his consumption, and X the number of shares of stocks he holds. His optimization problem is then

$$J(W_t, \Im_t, \tau; t) \equiv \max_{c_t, X_t} \mathrm{E}_t \left[-\sum_{s=t}^{\bar{\tau}} \beta^{\tau+s-t} \exp(-\alpha c_s) \right] \tag{3.1}$$

subject to

$$W_{t+1} = (W_t - c_t)R + X_t Q_{t+1} + Y_{\tau+1,t+1} \tag{3.2}$$

where $Q_{t+1} \equiv P_{t+1} + D_{t+1} - RP_t$ is the excess return on one share of stocks. Q_{t+1} is different from the rate of return, which is defined as the excess return on one dollar invested in stocks. One has to divide the share return by the share price to get the rate of return.

3.1 The Constant Demographics Case: $g(\tau; t) \equiv g(\tau)$

In what follows, we assume that the demographic structure is constant over time and solve for the stationary equilibrium. The equilibrium is stationary in that, even though the equilibrium price function and the equilibrium value function depend on the state variables and the time trend, the coefficients of the functions are independent of time. These coefficients depend only on investors' age. Intuitively, when the demographic structure is constant over time, the economy is characterized by the Markovian processes of the labor income state variable vector, Z_t, and the dividend state variable, F_t. The functional form of the equilibrium should be the same for all periods in an infinite horizon setting.

While the constant demographic structure does not induce any time variation in the economy, it is still worth studying. First, if one wants to compare the effects of two demographic structures on two segmented economies, such as the cross-national study by Erb, Harvey and Viskanta (1999), a comparative analysis of the stationary case is informative. Secondly, intuition about investors' life cycle pattern in investment behavior under the stationary demographic structure applies to the time varying case as well. Since the equilibrium is stationary, the intuition is clearer and easier to derive. Thirdly, we use the stationary case to derive the stationary equilibrium, which we use as the boundary condition of the backward induction in the later numerical example in Section 4.

The equilibrium price function is stated in proposition 3.1.

Proposition 3.1 *When the demographic structure is constant over time, i.e. $g(\tau;t) \equiv g(\tau)$, the economy defined in Section 2 has a steady-state equilibrium in which the equilibrium price function is*

$$P_t = p_0 + p_F F_t + p_Z Z_t \tag{3.3}$$

where p_0 is a constant; $p_F = \dfrac{a_F}{R - a_F}$; p_Z is a $1 \times N$ row vector of constants.

Proof: See Appendix A.

The first term of the price function, p_0, can be divided into three parts: the discounted value of the constant level of dividends, $\bar{D}$, which is $\dfrac{\bar{D}}{r}$; the discount in the price to compensate for the risk of dividends; and the discount in the price to compensate for the risk of labor income. The level of the discount in the price to compensate for the risk of either dividends or labor income can be obtained by considering the economy in which labor income is deterministic, i.e. $\{\omega_\tau\}_{\tau=0}^{T} = 0_{N\times N}$ and $\{\sigma_\tau\}_{\tau=0}^{T} = 0$. It is easy to show, from the proof of proposition 3.1, that the economy with deterministic labor income has an equilibrium price function $P_t = p_0^* + p_F^* F_t$, where $p_F^* = \dfrac{a_F}{R - a_F}$ and p_0^* is a constant. Thus, the discount in the price to compensate for the risk of dividends is $p_0^* - \dfrac{\bar{D}}{r}$. The discount to compensate for the risk of labor income is $p_0 - p_0^*$.

The second term, $p_F F_t$, is the discounted value of the stochastic part of dividends. As new information for dividends, F, arrives, the stock price adjusts to fully reflect it. The third item, $p_Z Z_t$, reflects the role of stocks in hedging the risk of labor income. As shown later, the expected excess share return on stocks is correlated with labor income. Thus, when new information for labor income, Z, arrives, investors rebalance their portfolio to hedge against the risk of labor income. The stock price adjusts to accommodate the rebalance motivated trade. Therefore, the equilibrium stock price depends not only on the information for the future dividend, F, but also on the information for the background risk, Z.

Define $\Pi_t \equiv (1, Z_t)^T$. Given the equilibrium price function, the excess share return on stocks is

$$Q_{t+1} \equiv P_{t+1} + D_{t+1} - RP_t = \Theta\Pi_t + \Phi\xi_{t+1} \tag{3.4}$$

where $\Theta = (\bar{D} + (1-R)p_0, p_Z a_Z - Rp_Z)$ and $\Phi = (1, 1 + p_F, p_Z)$. ξ_{t+1} which was defined in Section 2 as $\xi_{t+1} \equiv (\epsilon_{D,t+1}, \epsilon_{F,t+1}, \epsilon_{Z,t+1})^T$ represents the uncertainty of asset returns. Note that the expected excess share return, $\Theta\Pi_t$, is not affected by the "fundamental" of stocks, F. When new information for the fundamental, F, arrives, the stock price adjusts to fully reflect it. The expected excess share return stays the same. However, when new information for labor income, Z, arrives, the stock price the expected excess share return adjust to reflect the changed asset demands.

Given the excess share return on stocks and the labor income process, we can derive the investors' optimal consumption-investment policy.

Theorem 3.2 *Problem (3.1)-(3.2) has the following solution*

$$J(W_t, \Im_t, \tau; t) = -\beta^\tau \exp(-\gamma_\tau W_t - \mu_\tau t - \frac{1}{2}\Pi_t^T \nu_\tau \Pi_t) \tag{3.5}$$

where γ_τ and μ_τ are functions of τ; ν_τ is a $(N+1) \times (N+1)$ symmetric matrix dependent on τ. The optimal demand for stocks, X_t, and consumption, c_t, are

$$X_t = \frac{1}{\gamma_{\tau+1}}\Gamma_{\tau+1}E_t(Q_{t+1}) - \frac{1}{\gamma_{\tau+1}}\kappa_{\tau+1}\Pi_t \tag{3.6}$$

$$\begin{aligned} c_t &= \bar{c}_\tau + \frac{R\gamma_{\tau+1}}{\alpha + R\gamma_{\tau+1}}W_t + \frac{\gamma_{\tau+1}n_{\tau+1} + \mu_{\tau+1}}{\alpha + R\gamma_{\tau+1}}t \\ &\quad + \frac{1}{2(\alpha + R\gamma_{\tau+1})}\Pi_t^T m_{\tau+1}\Pi_t \end{aligned} \tag{3.7}$$

where $\Gamma_{\tau+1}$ and $\bar{c}_\tau$ are functions of τ; $\kappa_{\tau+1}$ is a $1 \times (N+1)$ row vector dependent on τ; $m_{\tau+1}$ is a $(N+1) \times (N+1)$ matrix dependent on τ.

Proof: See Appendix A.

The value function retains the negative exponential form of the utility function. It depends not only on wealth, but also on the time trend, t, and a quadratic form of Π. The value function depends on t because of the secular growth in labor income. Since labor income grows with time, everything else equal, investors nowadays have higher utility than the investors of the same age in old days. Note that Π is obtained by augmenting Z with 1. Investors trade stocks both for consumption purposes and to hedge against the risk of labor income. The incentive to trade is much clearer if one considers investors' demand for stocks, X. X consists of two components. The first is a myopic portfolio which reflects the trade off between expected return and risk. The second is a hedge portfolio. Since the expected excess share return on stocks is correlated with labor income, stocks provide a vehicle to hedge against the risk of labor income.

The coefficients of the value functions and consumption-investment policies are age specific. This introduces the life cycle pattern in investment behavior. It is difficult to tell how the demand of an investor of certain age responds to the labor income state variable vector, Z. However, the risk aversion of the value function, γ_τ, provides an informative way to study investors' average holdings of stocks. γ_τ is the denominator in the demand function, X. When γ_τ is lower, investors will be more risk tolerant and, everything else equal, hold more stocks on average.

[Insert Figure 1 Here]

From equation (A.9) in Appendix A, we have $\gamma_\tau = \dfrac{\alpha R \gamma_{\tau+1}}{\alpha + R\gamma_{\tau+1}}$. Figure 1 plots γ_τ as a function of age for $\alpha = 0.05$, $r = 1.2\%$ and $T = 54$. These parameter values are the same as those used in the numerical example in Section 4, in which we consider the population in the age range 20-74. The economic age of an investor of calendar age 20 is $\tau = 0$. To be consistent to the later calibrations, we let the x-axis be of calendar age. We have tried other parameter values. The pattern of γ_τ is robust.

γ_τ is monotonic increasing and convex in τ. After a certain age, it becomes steeper. Intuitively, since the excess share return follows a mean-reverting process, young investors with longer investment horizons could hold more risk assets than old investors when the return is low, and wait until the return reverts. In the sense of risk aversion, investors will be more risk taking when young than when old. This result is in accordance with Samuelson (1991). The convexity of γ_τ comes from the fact that a same calendar change shortens an old investor's investment horizon in percentage much more than it does to a young investor's investment horizon.

In sum, young investors are more risk taking. Everything else equal, they tend to hold more stocks than old investors. When an investor gets older, he gradually sells his holdings of stocks, but after a certain age, he accelerates his selling.

3.2 The Time Varying Demographics Case: $g(\tau; t)$

When the demographic structure is time varying, the equilibrium price function, the value function and the investors' optimal consumption-investment policy have the same functional forms as when the demographic structure is constant. But because of demographic change, the coefficients of the equilibrium depend not only on investors' age, but also on time. Demographic change and the equilibrium are related in a highly non-linear way.

The equilibrium price function is stated in proposition 3.3.

Proposition 3.3 *The economy defined in Section 2 has an equilibrium in which the equilibrium price function is*

$$P_t = p_{0,t} + p_F F_t + p_{Z,t} Z_t \tag{3.8}$$

where $p_{0,t}$ *is a function of* t*;* $p_F = \dfrac{a_F}{R - a_F}$*;* $p_{Z,t}$ *is a* $1 \times N$ *row vector dependent on* t*.*

Proof: See Appendix A.

Most of the analysis of the components of the price function in the economy with a constant demographic structure still holds. But now $p_{0,t}$ and $p_{Z,t}$ depend on time t because of the changing demographic structure. The discount in the price to compensate for the risk of dividends and labor income, and the sensitivity of the price to the risk of labor income are time varying.

For example, consider the simplest case of an economy with a time varying demographic structure, in which the demographic structure is $g(\tau; 0)$ at time 0 and $g(\tau; 1)$ at time 1 and later. From the prior analysis, we know that the equilibrium is stationary from time 1 on. At time 0, investors fully anticipate the next period price function. But since $g(\tau; 0) \not\equiv g(\tau; 1)$, the equilibrium price function at time 1 will not hold for time 0. The coefficients of the price function at time 0 are specific to the demographic change from $g(\tau; 0)$ to $g(\tau; 1)$. In general, the stock price is linear in F and Z, but the coefficients of this linear relation depend on the current and all the future demographic structures.

Given the equilibrium price function, the excess share return on stocks is

$$Q_{t+1} \equiv P_{t+1} + D_{t+1} - RP_t = \Theta_{t+1}\Pi_t + \Phi_{t+1}\xi_{t+1} \tag{3.9}$$

where $\Theta_{t+1} = (p_{0,t+1} + \bar{D} - Rp_{0,t}, p_{Z,t+1}a_Z - Rp_{Z,t})$ and $\Phi_{t+1} = (1, 1 + p_F, p_{Z,t+1})$. Similar to the price function, the excess share return keeps the same functional form as in the economy with a constant demographic structure, but the coefficients are time varying. Since it is assumed that the demographic structure, both the current and the future, is common knowledge, investors fully anticipate the effects of future demographic change on future stock prices and excess share returns when making consumption-investment decisions. Therefore, both Θ and Φ are determined by the current and all the future demographic structures.

Given the excess share return on stocks and labor income process, we can derive the investors' optimal consumption-investment policy.

Theorem 3.4 *Problem (3.1)-(3.2) has the following solution*

$$J(W_t, \Im_t, \tau; t) = -\beta^\tau \exp(-\gamma_\tau W_t - \mu_\tau t - \frac{1}{2}\Pi_t^T \nu_{\tau,t}\Pi_t) \tag{3.10}$$

where γ_τ and μ_τ are functions of τ; $\nu_{\tau,t}$ is a $(N+1)\times(N+1)$ symmetric matrix dependent on τ and t. The optimal demand for stocks, X_t, and consumption, c_t, are

$$\begin{aligned} X_t &= \frac{1}{\gamma_{\tau+1}}\Gamma_{\tau+1,t+1}E_t(Q_{t+1}) - \frac{1}{\gamma_{\tau+1}}\kappa_{\tau+1,t+1}\Pi_t & (3.11)\\ c_t &= \bar{c}_{\tau,t} + \frac{R\gamma_{\tau+1}}{\alpha + R\gamma_{\tau+1}}W_t + \frac{\gamma_{\tau+1}n_{\tau+1} + \mu_{\tau+1}}{\alpha + R\gamma_{\tau+1}}t \\ &\quad + \frac{1}{2(\alpha + R\gamma_{\tau+1})}\Pi_t^T m_{\tau+1,t+1}\Pi_t & (3.12) \end{aligned}$$

where $\Gamma_{\tau+1,t+1}$ and $\bar{c}_{\tau,t}$ are functions of τ and t; $\kappa_{\tau+1,t+1}$ is a $1\times(N+1)$ row vector dependent on τ and t; $m_{\tau+1,t+1}$ is a $(N+1)\times(N+1)$ matrix dependent on τ and t.

Proof: See Appendix A.

Both the value function and the investors' optimal consumption-investment policy are time varying because of demographic change and the secular growth in labor income. In particular, different from the constant demographics case, their coefficients are also time varying because of demographic change. Investors in our model fully anticipate the effect of demographic change on stock returns. This rational expectation of our model is in contrast to other work on the relation between demographics and stock returns, which relies on partial equilibrium models of life cycle portfolio selection.[14]

Demographics can be viewed as an additional set of state variables that affect the determination of stock prices, the investors' optimal investment-consumption policy, and etc..

[14]See Bergantino (1998).

3.3 Derivation of the Predictive Relation for the Equity Risk Premium

The equity risk premium is obtained by dividing the expected excess share return by the stock price. From equation (3.9), the equity risk premium, $R_{m,t+1} - R$, can be written as

$$R_{m,t+1} - R = \frac{\Theta_{t+1}\Pi_t}{P_t} = \theta_{0,t+1}\frac{1}{P_t} + \theta_{Z,t+1}\frac{Z_t}{P_t} \tag{3.13}$$

where $\theta_{0,t+1} = \Theta_{t+1,11}$ is the first element of Θ_{t+1}; $\theta_{Z,t+1} = (\Theta_{t+1,12}, \Theta_{t+1,12}, ..., \Theta_{t+1,1N})$ is the vector obtained by taking the 1st element of Θ_{t+1} away.

The econometrician can not directly observe the labor income state variable vector, Z. It must be inferred from the available information, such as the stock price, P, the dividends, D, and the dividend state variable, F. The dividend state variable, F, is directly observable because we have assumed that it is a value proportional to the demeaned earnings, i.e. $F_t = a_B(B_t - \bar{B})$, where the earnings, B, are observable. From the normality assumption of the model, the labor income state variable vector, Z, can be expressed as

$$Z_t = \eta_{0,t} + \eta_{1,t}P_t + \eta_{2,t}D_t + \eta_{3,t}F_t + \upsilon_t \tag{3.14}$$

where $\eta_{0,t}$, $\eta_{1,t}$, $\eta_{2,t}$ and $\eta_{3,t}$ are $N \times 1$ matrices dependent on t. υ_t is the unexpected component of Z conditional on P, D and F. Its variance also depends on t. The coefficients are time dependent because the coefficients, $p_{0,t}$ and $p_{Z,t}$, of the price function, $P_t = p_{0,t} + p_F F_t + p_{Z,t}Z_t$, are time varying because of demographic change, and thereafter the correlation matrix of Z, P, D and F is time varying. We follow the tradition of using earnings to predict stock returns. Substitute the expression for F into equation (3.14),

$$Z_t = \eta^*_{0,t} + \eta_{1,t}P_t + \eta_{2,t}D_t + \eta^*_{3,t}B_t + \upsilon_t \tag{3.15}$$

where $\eta^*_{0,t} = \eta_{0,t} - a_B\bar{B}\eta_{3,t}$ and $\eta^*_{3,t} = a_B\eta_{3,t}$. Substituting the expression for Z from equation (3.15) into equation(3.13), we have the predictive relation for the equity risk premium

$$R_{m,t+1} - R = \lambda_{0,t} + \lambda_{1,t}\frac{1}{P_t} + \lambda_{2,t}\frac{D_t}{P_t} + \lambda_{3,t}\frac{B_t}{P_t} + \varepsilon_{t+1} \tag{3.16}$$

where $\lambda_{0,t} = \theta_{Z,t+1}\eta_{1,t}$, $\lambda_{1,t} = \theta_{0,t+1} + \theta_{Z,t+1}\eta^*_{0,t}$, $\lambda_{2,t} = \theta_{Z,t+1}\eta_{2,t}$, $\lambda_{3,t} = \theta_{Z,t+1}\eta^*_{3,t}$ and $\varepsilon_t = \frac{\theta_{Z,t+1}\upsilon_t}{P_t}$. Equation (3.16) is the one of the main results of this paper. In equilibrium, the equity risk premium is linear in the real per capita stock price, the dividend yield and the earnings yield, but the coefficients of this linear relation are time varying because of demographic change. These coefficients are highly non-linear functions of all the moments of the current and the future demographic structures. Without further assumptions, we can not determine their signs and magnitude.

In what follows, we try different strategies to study this predictive relation. In Section 4, we calibrate the model by further simplifying the economy, and obtain the dynamics of the coefficients of the predictive relation. We are primarily interested in whether demographic change can induce significant variations in these coefficients. In Section 5, we infer a demographic variable that summarizes the demographic information for stock returns from the calibration results. We then linearize the predictive relation and conduct empirical tests. We compare the performance of the predictive relation to those of previous empirical research.

4 Calibration

In this section, we further simplify the economy and calibrate the model. We want to show that demographic change can induce significant variations in the coefficients of the linear predictive relation so that simple linear predictive regressions that ignore the structural change may produce misleading results. We prespecify that, among the innovations to dividends, $\epsilon_{D,t}$, to the dividend state variable, $\epsilon_{F,t}$, and to the labor income state variable vector, $\epsilon_{Z,t}$, only $\epsilon_{F,t}$ and $\epsilon_{Z,t}$ are correlated. Therefore, the information in dividends, D, for the labor income state variable vector, Z, is contained in that in the dividend state variable, F. Note that F is a deterministic linear function of earnings, B. The conditional distribution of Z on P, D and B is equivalent to that on P and B. Under this specification, $\lambda_{2,t}$ equals zero. The equity risk premium becomes linear only in the inverse of real per

capita stock price and the earnings yield with time varying coefficients. We would focus on the dynamics of the coefficients, $\lambda_{0,t}$, $\lambda_{1,t}$ and $\lambda_{3,t}$ of this simplified relation implied by the historical demographic change.

Because of data limitations, all the calibrations are with respect to annual data. The derivation of the equilibrium uses the backward induction procedure. We first solve the equilibrium at a future date, the year 2050 in this case, by assuming that the demographic structure is constant after this date, and use this as the boundary condition for the time varying equilibrium which we solve recursively.

[Insert Figures 2-3 Here]

The demographic data were collected from Citibase. We only consider the population in the age range 20-74, with the average population from 1982 to 1984 normalized to unity. Therefore, the economic age of an investor of calendar age 20 is $\tau = 0$. The density function of the time varying demographic structure of 1947-2050, $g(\tau; t)$, is plotted in figure 2. Additional information on the variables is given in Appendix C. As shown in figure 2, the demographic structure has undergone drastic changes in the post-1947 period. From the 1960s to the 1980s, there was a boom in the share of the young generation, which are usually called "baby boomers". Baby boomers entered their middle age in the 1990s, and will retire in the late 2010s. As they age, the demographic structure changes accordingly. Figure 3 plots two moments of the demographic structure, the share of population in the age range 40-64, the population average age, and their changes. The share of population in the age range 40-64 is the lowest around 1987, then sharply increases and attains its highest level in 2007. It gradually decreases until 2028. The population age profile has the similar pattern as this share of population.

[Insert Figure 4 Here]

We use the family questionnaire of the Panel Study of Income Dynamics (PSID) to estimate the labor income process, which is specified by equations (2.2) and (2.3). The PSID provides a panel of annual observations of individual and family income and other

variables from 1967 to 1992. In using the PSID, we take age of family head as a measure of τ and the average of family labor income per family member, adjusted to the Standard and Poor's (S&P) Composite Index, as a measure of $Y_{\tau,t}$.[15] We deflated $Y_{\tau,t}$ by CPI to obtain the real terms.

Prespecify the dimension of the labor income state variable, N, to be 1, i.e. Z_t becomes a scalar. We proxy the labor income state variable, Z, by the de-meaned unemployment rate scaled up by 100. For a given τ, we regressed the time series of $Y_{\tau,t}$ on a constant, time t and $\frac{1}{2}Z_t^2$ to obtain the estimates of h_τ, n_τ, ω_τ and σ_τ, which are plotted in figure 4. Due to data limitations, these estimates are volatile for even a small change in age. This contradicts the economic intuition that the coefficients of the labor income process should be smooth in age. Therefore, we use cubic polynomial to approximate these estimates. We will use the cubic polynomial approximations as the true parameters in the numerical solution. We also regress Z_t on its lagged values to obtain the estimates of a_Z and σ_Z.

From this rough estimation of the labor income process, we can tell that households in the age range 40-64 have significantly different labor income from others. They have the highest level of income from the estimates of h_τ and n_τ, and the highest level of idiosyncratic risk from the estimate of σ_τ. But their labor income has the lowest sensitivity to the labor income state variable from the estimate of ω_τ. Combined with the dynamics of the risk aversion coefficient as a function of age, these characteristics make the population in the age range 40-64 have significantly different effects on the economy from other group of people.

Prices, P, dividends per share, D, and earnings per share, B, all correspond to S&P Composite Index. We deflated them by CPI and population to obtain the real per capita items. We use the long run average of dividends as $\bar{D}$ and the long run average of earnings as $\bar{B}$. Note that, after substituting the expression of F into equation (2.2), we have $D_t - \bar{D} = a_B(B_t - \bar{B}) + \epsilon_{D,t}$. We regress the de-meaned dividends, $D_t - \bar{D}$, on the demeaned earnings, $B_t - \bar{B}$, to obtain the estimates of a_B and σ_D. We recover the time series of F from the earnings by $F_t = a_B(B_t - \bar{B})$, and then regress F on its own lags to obtain the estimates of

[15]The average per capita market capitalization of one point of the S&P Composite Index is about 60 dollars. So we divide the family labor income per family member by 60 to adjust it to the S&P Composite Index. The average real per capita labor income from 1967 to 1992 is 11420 dollars, which is 192.19 index points.

a_F and σ_F.

For the correlation among the innovations to the dividends, $\epsilon_{D,t}$, to the dividend state variable, $\epsilon_{F,t}$, and to the labor income state variable, $\epsilon_{Z,t}$, we assume that only $\epsilon_{F,t}$ and $\epsilon_{Z,t}$ are correlated with a correlation coefficient $\rho_{FZ} = -0.8$. Thus, the information in dividends, D, for the labor income state variable, Z is contained in that in earnings, B. The conditional distribution of Z on P, D and B is equivalent to that on P and B. $\lambda_{2,t}$ equals zero. The equity risk premium depends only on the inverse of the real per capita stock price and the earnings yield with time varying coefficients. For other parameters, we let the riskfree interest rate $r = 1.2\%$. The preference parameters are set $\alpha = 0.05$ and $\beta = 1.2$.[16]

[Insert Table 1 Here]

Table 1 summarizes the estimates or the prespecified values of the parameters. We are not trying to fit our prediction of the equity risk premium to the historical data. In this section, we are primarily interested in whether demographic change can induce significant variations in the coefficients of the predictive relation. Other parameter values have been tried. They give similar qualitative results.

[Insert Figure 5 Here]

Figure 5 plots the time varying coefficients of the price function, $p_{0,t}$ and $p_{Z,t}$. The coefficients of the dividend state variable is a constant, $p_F = 2.75$. We focus our analysis on the pattern of $p_{Z,t}$.

$p_{Z,t}$ is negative. Intuitively, the common component of investors' labor income, $\frac{1}{2}\omega_\tau Z^2$, can be viewed as a non-tradable security, whose next period payoff is Z_{t+1}, and investors of age τ have to hold $\frac{1}{2}\omega_\tau Z_t$ share of this security. Since the dividend state variable, F, and the labor income state variable, Z, are negatively correlated, dividends and labor income are substitutes for each other. When Z_t increases, everything else equal, an average investor

[16] A number of capital market studies have obtain estimates of the coefficient of relative risk aversion ranging from 2 to 16. We use 10 as an intermediate value of the coefficient of relative risk aversion and scale it down by an income level of 200, adjusted to the S&P Composite Index, to estimate the coefficient of absolute risk aversion.

will reduce his demand for stocks, though he can not increase his position of labor income. The stock price has to decrease to clear the market. Demographic change induce significant variations in $p_{Z,t}$. The difference between the highest and the lowest level of $p_{Z,t}$ is about 8% of its average level.

[Insert Figure 6 Here]

Figure 6 plots the time varying coefficients of the expected excess return, $\theta_{0,t+1}$ and $\theta_{Z,t+1}$. We are primarily interested in the pattern of $\theta_{Z,t+1}$. $\theta_{Z,t+1}$ is positive. As shown before, the stock price decreases with Z. The expected excess return increases as the stock price decreases. Therefore, the expected excess return increases with Z. Demographic change induces significant variations in both $\theta_{0,t+1}$ and $\theta_{Z,t+1}$. In particular, the difference between the highest and the lowest level of $\theta_{Z,t+1}$ is about 8% of its average level.

[Insert Figure 7 Here]

Figure 7 plots the time varying coefficients of the predictive relation as specified in equation (3.16). The signs of these coefficients are determined by the expectation of Z conditional on P and F. They depend not only on the covariance between Z and P, and Z and F, which are both negative, but also on that between P and F. In equilibrium, $\lambda_{1,t}$ is positive and $\lambda_{3,t}$ is positive, which is consistent to the empirical findings that the equity risk premium decreases with the stock price, and increases with the earnings yield.[17]

Demographic change causes significant variations in these coefficients. For example, the differences between the highest and the lowest levels of $\lambda_{0,t}$, $\lambda_{1,t}$ and $\lambda_{3,t}$ are around 0.004, 3 and 0.002. $\lambda_{2,t}$ is zero from the model specification. Therefore, if one regresses the equity risk premium on the stock price, the dividend yield and/or the earnings yield without considering the variation in the linear coefficients, he is likely to obtain misleading results. A positive result can not conclude the predictability of the equity risk premium, and a negative result can not rule it out, either.

The pattern of $\lambda_{0,t}$, $\lambda_{1,t}$ and $\lambda_{3,t}$ in the late 1990s coincides with the stock market performance at the same time, which the traditional predictive models fail to explain. During

[17] For example, Campbell and Shiller (1988) and Lamont (1998).

this period, the stock price was extraordinarily high and the earnings was quite low. The traditional predictive models predict a low equity risk premium, while the actual equity risk premium was quite high. Our model, which incorporates demographic changes in the predictive relation, can be used to explain this. As shown in figure 7, during this period, the sensitivity of the equity risk premium to the inverse of the stock price, $\lambda_{1,t}$, and that to the earnings yield, $\lambda_{3,t}$, are relatively low. The equity risk premium is less sensitive to the stock price and the earnings yield. The high level of the stock price and the low level of the earnings yield do not mean that the equity risk premium is necessarily low. In addition, the constant term in the predictive relation, $\lambda_{0,t}$, is relatively high. Everything else equal, this means that the equity risk premium will be higher.

Simple comparison of figures 3 and 7 suggests that $\lambda_{0,t}$, $\lambda_{1,t}$ and $\lambda_{3,t}$ are correlated with the change in the share of population in the age range 40-64. For example, in the mid 1970s, $\lambda_{0,t}$ is low, and the change in this share of population is low; in the late 1990s, $\lambda_{1,t}$ is high, and the change in this share of population is high. $\lambda_{0,t}$ is positively correlated with the change in this share. Similarly, $\lambda_{1,t}$ and $\lambda_{3,t}$ are negatively correlated with the change in this share. Intuitively, since households in this age range have the least sensitive labor income to the macroeconomic risk, represented by the unemployment rate here, the increase in this share of population introduces less macroeconomic risk to the whole economy. The role of stocks as a hedging vehicle should be less important. However, they have the highest level of idiosyncratic labor income risk, and their risk aversion is relatively high as shown in figure 1, the increase in this share of population also makes the population more risk averse, which makes the role of stocks as a hedging vehicle more important. The net effect is different from zero. This motivates the later empirical tests. From the comparison of the patters of $\lambda_{1,t}$ and $\lambda_{3,t}$ and that of the change in the share of population in the age range 40-64, we expect that the second effect dominates.

In what follows, we will empirically test the predictive relation as specified in equation (3.16). We use the change in the share of population in the age range 40-64 to summarize the demographic information in the predictive relation for the equity risk premium. Since the calibration is simplified, the empirical results are not necessarily same as predicted here.

5 Empirical Tests

In this section, we test the predictive relation as specified in equation (3.16) in predicting the short-term, in particular the quarterly, equity risk premium for the post-1947 period. As suggested in Section 4, we use the change in the share of population in the age range 40-64 to summarize the demographic information in the predictive relation.

5.1 The Derivation of the Predictive Regression

We follow the tradition of using log variables. Define ς as the sum of log CPI and log population. Write the log excess return as $r_m - r_f$, the log real per capita stock price as $\ln(P) = p - \varsigma$, where log price, p, is the natural logarithm of the nominal stock price, the log real per capita dividends as $\ln(D) = d - \varsigma$, where log dividends, d, are the natural logarithm of nominal dividends, and the log real per capita earnings as $\ln(B) = e - \varsigma$, where log earnings, e, are the natural logarithm of nominal earnings. In addition, as in Lamont (1998), we use the dividend payout ratio instead of the earnings yield in the predictive regression.[18] Therefore, we loglinearize equation (3.16) as

$$r_{m,t+1} - r_{f,t+1} = \lambda^*_{0,t} + \lambda^*_{1,t}(p_t - \varsigma_t) + \lambda^*_{2,t}(d_t - p_t) + \lambda^*_{3,t}(d_t - e_t) + \epsilon_{t+1} \tag{5.1}$$

where $\lambda^*_{0,t}$, $\lambda^*_{1,t}$, $\lambda^*_{2,t}$ and $\lambda^*_{3,t}$ depend on $\lambda_{0,t}$, $\lambda_{1,t}$, $\lambda_{2,t}$, $\lambda_{3,t}$ and the means of $p - \varsigma$, $d - p$ and $d - e$. ϵ_{t+1} is the residual term, which is, in general, different from ε_{t+1}. The derivation uses the fact $\ln(y + x) \approx \ln(y + e^{\bar{\ln x}}) + e^{\bar{\ln x}}(\bar{y} + e^{\bar{\ln x}})^{-1}(\ln x - \bar{\ln x})$, which is derived in Appendix B.

We can not estimate equation (5.1) directly because the coefficients are time varying due to demographic change. As suggested in the calibration results in Section 4, we use linear functions of the change in the share of population in the age range 40-64 to proxy the demographic variables in the predictive relation, i.e.

[18] Lamont (1998) has shown that, in the predictive relation, using the dividend payout ratio is numerically identical to using the earnings yield. To be compatible to previous research for the equity risk premium, we use the dividend payout ratio here.

$$\lambda^*_{0,t} = \beta_0 + \beta_1\phi_t, \quad \lambda^*_{1,t} = \beta_2 + \beta_3\phi_t, \quad \lambda^*_{2,t} = \beta_4 + \beta_5\phi_t, \quad \lambda^*_{3,t} = \beta_6 + \beta_7\phi_t \tag{5.2}$$

Bakshi and Chen (1994) have found that the change in average age, denoted by ϑ here, is significant in predicting the annual equity risk premium for the post-1947 period. In fact, as shown in figure 3 and the correlation coefficient reported later, the change in the share of population in the age range 40-64 is highly positively correlated with the change in average age. The information in the change in this share of population for stock returns is similar to that in the change in average age. The regression results of using linear functions of the change in average age to proxy the demographic variables in the predictive relation are similar.

Substituting the expressions (5.2) into equation (5.1), we have the predictive regression

$$\begin{aligned} r_{m,t+1} - r_{f,t+1} &= \beta_0 + \beta_1\phi_t + \beta_2(p_t - \varsigma_t) + \beta_3\phi_t(p_t - \varsigma_t) \\ &\quad +\beta_4(d_t - p_t) + \beta_5\phi_t(d_t - p_t) + \beta_6(d_t - e_t) + \beta_7\phi_t(d_t - e_t) + \epsilon_{t+1} \end{aligned} \tag{5.3}$$

ϵ_{t+1} can be interpreted as the sum of two parts. The first is the difference between the ex ante excess return, i.e. the equity risk premium, and the ex post excess return, which is $\frac{\Phi_t\xi_{t+1}}{P_t}$ from equation (3.9). The second is the noise due to the population growth rate in the historical excess return. Note that by assumption the per capita stock supply is normalized to unity and thereafter that the number of outstanding shares of stocks at time t is proportional to the population at time t, $G(t)$. We should have adjusted the historical excess return to the per capita level. However, we only have annual population data. The adjustment for short-term, for example quarterly, returns is not feasible. In addition, in order to be directly compared to other research, we would rather keep the dependent variable on the left hand side of the predictive regression unadjusted.

5.2 Data

The sample period is 1947Q1-1999Q4. This period includes the late 1990s, in which the baby boomers entered their middle age and the demographic structure had drastic change as shown in figure 2. We also consider the subsample period 1947Q1-1994Q4, which has been studied before, so that we can compare the relative performance of our model and previous predictive regressions.

Stock returns, prices, dividends per share, and quarterly earnings per share all correspond to the S&P Composite Index, because historical quarterly earnings data for the index are available. Additional information on the variables is given in the Appendix C.

[Insert Figure 8 Here]

Excess returns, $r_m - r_f$, are total stock returns (continuously compounded including reinvested dividends) minus returns on a portfolio of treasury bills. Log price, p, is the natural logarithm of the nominal S&P Composite Index. Log dividends, d, are the natural logarithm of the sum of the past four quarters of nominal dividends per share.[19] Log earnings, e, are the natural logarithm of a single quarter's nominal earnings per share. The predictive relation as specified in equation (5.3) does not require that earnings and dividends have to be contemporary. The derivation of the predictive relation just requires that earnings and dividends of whatever period(s) are correlated with the labor income state variable vector, Z, and thereafter that the expectation of Z is linear in the stock price, dividends and earnings. Log sum of CPI and population, ς, is the sum of the natural logarithm of CPI and that of the previous year's population. The change in average age, ϑ, is the difference between the natural logarithm of the population average age in the previous year and that in the year before. The change in the share of population in the age range 40-64, ϕ, is the difference between the natural logarithm of this share of population in the previous year and that in the year before. Thus, all the explanatory variables on the right hand side of the predictive regression are predetermined. Figure 8 plots the historical log real per capita stock price, $p - \varsigma$, the log dividend yield, $d - p$, and the log dividend payout ratio, $d - e$.

[19] We use the past four quarters of dividends to adjust the seasonality in dividends.

[**Insert Table 2 Here**]

Summary statistics are reported in table 2. Several observations require special attention. First, the real per capita stock price, $p-\varsigma$, and the dividend yield, $d-p$, are highly correlated. The correlation coefficient for the sample period 1947Q1-1999Q4 is -0.932. Therefore, if a predictive regression has both variables on the right hand side, the multicollinearity problem in explanatory variables is expected. The inference may be biased. Adjustment is necessary. Secondly, the change in average age, ϑ, and the change in the share of population in the age range 40-64, ϕ, are highly correlated as well. The correlation coefficient for the sample period 1947Q1-1999Q4 is 0.841. These two variables contain similar information, and can be viewed as substitutes for each other. Therefore, in the following empirical tests, we just use one of them, the change in this share of population, to summarize the demographic information for stock returns.

Finally and most importantly, the real per capita stock price, $p-\varsigma$, the dividend yield, $d-p$, and the dividend payout ratio, $d-e$, are highly autocorrelated. The autocorrelation coefficients are 0.984, 0.976 and 0.725 respectively. If these three time series are non-stationary, the regressions that run the equity risk premium on these variables will be spurious. Therefore, we have to clear up the stationarity issue, before we go further to the empirical tests. The annual change in average age, ϑ, and the annual change in the share of population in the age range 40-64, ϕ, are also highly autocorrelated. The autocorrelation coefficients are 0.908 and 0.939 respectively. But there is no reason to believe that these two time series are non-stationary.

[**Insert Table 3 Here**]

We follow Horvath and Watson (1995) in testing the stationarity of the real per capita stock price, $p-\varsigma$, the dividend yield, $d-p$, and the dividend payout ratio, $d-e$. We just need to show that any pair of log price, p, log dividends, d, log earnings, e, and log sum of CPI and population, ς, which are non-stationary because of nominal and real growth in the economy and population, are cointegrated and that the cointegration vectors are (1, -1). The Horvath and Watson test provides an informative way to test the joint hypothesis that all

four variables are cointegrated with unitary coefficients. The procedure tests the alternative of known cointegrating vectors against the null of no cointegration. The test statistic is identical to a Wald test for whether the error correction terms, $p-\varsigma$, $d-p$ and $d-e$, belong on the right hand side of a vector autoregression (VAR) of Δp, $\Delta\varsigma$, Δd and Δe. Table 3 reports the results of this quadri-variate error-correction VAR. The null hypothesis of no cointegration is rejected.[20]

In summary, the statistical tests show that log price, p, log dividends, d, and log earnings, e, and log sum of CPI and population, ς, all share a common trend. Thus, the difference of any two of these variables are stationary. This finding also confirms our assumption that the (real per capita) dividend process is stationary in our theoretical model.

5.3 Regression Results

In this section, we test the predictive relation as specified in equation (5.3). We also try to replicate previous empirical work, for example, Fama and French (1988) and Lamont (1998) who focus on the information in dividends and/or earnings for stock returns, and Bakshi and Chen (1994) who focus on the demographic information for stock returns, in order to compare the performance of our model to theirs.

[Insert Table 4 Here]

Table 4 replicates Fama and French (1988) and Lamont (1998) in predicting the quarterly equity risk premium. We regress the equity risk premium on the dividend yield and the dividend payout ratio. Panel A reports the regression results for the period 1947Q1-1994Q4, which has been studied by Lamont (1998). The dividend yield and the dividend payout ratio are significant in both the univariate and the multivariate regressions. When the dividend yield and the dividend payout ratio are higher, the equity risk premium will be higher. Row

[20] Using quarterly data, the test statistics are 61.135 (and 88.989 with trend) for a VAR with one lag, and 63.208 for a VAR with four lags. All these are well above the 1 percent critical value of 41.08 (or 56.17) for a system with four variables, three known cointegrating relationships with (or without) an unknown cointegrating relationship, and a null hypothesis of no cointegration.

3 shows that the dividend yield and the dividend payout ratio combined explain 11.7% of the variation in the equity risk premium for this period.

Panel B reports the regression results for the period 1947Q1-1999Q4. This period includes the late 1990s in which the baby boomers began to enter their middle age. The calibration results in Section 4 suggest a significant changing structure in the predictive relation for this period. Not surprisingly, the dividend yield is not significant in predicting the equity risk premium in both the univariate and multivariate regressions for this period. Only the dividend payout ratio is marginally significant. Row 6 shows that the dividend yield and the dividend payout ratio combined explain only 1.5% of the variation in the equity risk premium for this period.

[Insert Table 5 Here]

Table 5 replicates Bakshi and Chen (1994) in using the change in average age to predict the quarterly and annual equity risk premium for the period 1947Q1-1999Q4. We also regress the quarterly and annual equity risk premium on the change in the share of population in the age range 40-64 for the same period. Poterba (2000) regresses the real stock return, instead of the equity risk premium, on the level of this share, and finds no significant returns. As suggested by the calibration results in Section 4, we expect that the *change* in this share matters in predicting the equity risk premium.

Panel A regresses the annual and quarterly equity risk premium on the change in average age. As in Bakshi and Chen (1994), the change in average age is significant. When the change in average age is higher, the equity risk premium is higher. However, the predictive power of the change in average age is quite sensitive to the return horizons. When the return horizon decreases from one year to one quarter, the R^2 decreases sharply from 11.3% to 2.4%. The change in the share of population in the age range 40-64 has the similar problem in the predictive regression. Panel B shows that the change in this share of population is significant in predicting the equity risk premium. It explains 9.4% of the variation in the annual equity risk premium, but only 2.3% for the quarterly equity risk premium.

The empirical evidence supports the claim that demographics matters in the determination of the equity risk premium. However, the low regression R^2, especially in predicting the short-term equity risk premium, suggests that considerable amount of information for stock returns is missing. This makes that demographics is virtually of no practical use in reality.

[Insert Table 6 Here]

Table 6 tests the predictive relation as specified in equation (5.3). Panel A reports the regression results for the period 1947Q1-1994Q4. Row 1 reports the coefficient estimates of the predictive relation exactly as specified in equation (5.3). Although the 13.7% R^2 is high, only two explanatory variables, the dividend payout ratio, $d - e$, and the product of the dividend payout ratio and the change in the share of population in the age range 40-64, $\phi(d - e)$, have significant t-statistics. Remember that the real per capita stock price, $p - \varsigma$, and the dividend yield, $d - p$, are highly correlated. The multicollinearity problem may contaminate the estimation. Therefore, in rows 2-4, we take away the explanatory variables that have the lowest t-statistics one by one, and conduct the regressions again until all the coefficients are significant. The fact that there is no significant change in the R^2 during this process confirms our suspicion of the multicollinearity problem in the regressors in row 1.

Row 4 is the main result of our empirical tests. The equity risk premium can be explained by the dividend yield, $d - p$, the dividend payout ratio, $d - e$, and their products with the change in the share of population in the age range 40-64, $\phi(d - p)$ and $\phi(d - e)$. The implicit coefficients with the dividend yield and the dividend payout ratio, calculated by averaging the change in this share of population out, are positive. Therefore, the equity risk premium still increases with the dividend yield and the dividend payout ratio. But the sensitivities of the equity risk premium to the dividend yield and the dividend payout ratio are time varying. When the change in this share of population is higher, the equity risk premium becomes less sensitive to the dividend yield and the dividend payout ratio.

Comparing row 4, table 6 to row 3, table 4, we find that the coefficient estimates of the dividend yield, $d - p$, and the dividend payout ratio, $d - e$, of these two regressions have the same signs and similar magnitude. However, the inclusion of demographic information

significantly improves the R^2. Row 3, table 4 reports R^2 of 11.7%, while row 4, table 6 reports 14.7%.

Panel B reports the regression results for the period 1947Q1-1999Q4. We follow the same procedure as in panel A, and end up with row 8. Row 8 has the same explanatory variables as row 4 does. In addition, the coefficient estimates in row 8 have the same signs and similar magnitude as in row 4. The late 1990s data did not damage the predictive power of our model. Row 8 reports R^2 of 13.6%, which is only slightly lower than 14.7% of row 4. The F-statistic is well about the 95 percent critical value of 2.37.

The most significant improvement of our model on previous research on the predictability of the equity risk premium is shown from the comparison between row 8 of table 6 and the regressions in panel B of table 4, and that between row 8 of table 6 and row 4 of table 5.

Row 8 of table 6 states that the equity risk premium can be explained by the dividend yield and the dividend payout ratio, but the coefficients of the linear regression are time varying because of demographic change. The regressions in panel B of table 4 ignore the changing structure in the linear regression. When structural change is relatively small, for example, in the period 1947Q1-1994Q4, ignoring its effects may still produce significant t-statistics, though the R^2 is lower. But when structural change is relatively large, for example, in the late 1990s, ignoring its effects will produce misleading coefficient estimates. Statistically, the regressions in panel B of table 4 can be viewed as misspecified versions of row 8 of table 6. The two missing variables, $\phi(d-p)$ and $\phi(d-e)$, which are obviously correlated with the dividend yield, $d-p$, and the dividend payout ratio, $d-e$, make classic inference biased. The regressions in panel B of table 4 fail to detect any significance of the dividend yield and the dividend payout ratio in predicting the equity risk premium.

Figure 9 plots the quarterly log excess return on the S&P Composite Index, $r_m - r_f$, the fitted value of the predictive regression as in row 6 of table 4 and the fitted value of the predictive regression as in row 8 of table 6. From the figure, the predictive regression that incorporates demographic change fits the historical equity risk premium better than the predictive regression that ignores it. In particular, for the late 1990s, the former regression

forecasts a higher equity risk premium than the latter.[21]

Row 4 of table 5 does not consider the information in dividends and earnings for stock returns. It can be viewed as an unconditional version of row 8 of table 6. Actually, if one calculates the implicit constant term and the coefficient of the change in the share of population in the age range 40-64, ϕ, in row 8 of table 6 by substituting the means of the dividend yield, $d-p$, and the dividend payout ratio, $d-e$, into the regression, he will obtain the similar estimates as in row 4 of table 5. Failure to consider the information in dividends and earnings still produces a significant result about the predictive power of demographic variables, but the R^2 is quite low.

In summary, our predictive relation that incorporates not only the information in dividends and earnings, but also the demographic information significantly improves on the previous predictive regressions that only consider either of them. The sources of predictive power are different. We use a linear combination of the dividend yield and the dividend payout ratio to predict the equity risk premium, and use demographic variables to explain the changing structure in this linear relationship.

[Insert Table 7 Here]

We also test the significance of the dividend yield, the dividend payout ratio and the change in share of population in the age range 40-64 in predicting the long-run equity risk premium. Table 7 reports the regression results of predicting the equity risk premium of 1 up to 4 quarters. As claimed by Fama and French (1988), the predictability of the equity risk premium increases with the return horizons. The regression reports R^2 of 13.6% when predicting the quarterly equity risk premium, while 35.8% when predicting the annual equity risk premium.

Compare row 4 of table 7 to row 2 and 4 of table 5. Even though the demographic variables, the change in average age and the change in the share of population in the age range 40-64 in this case, explain a significant proportion, around 10%, of the variation in the

[21] We also examined the fitted value of the predictive regression using the estimates in row 3 of table 4. It gives the similar forecasts of the equity risk premium as our mode does for the period 1947-1994. But it predicts even lower equity risk premium than row 6 of table 4.

annual equity risk premium respectively, adding the dividend yield and the dividend payout ratio increases this proportion further. Row 4 of table 7 reports R^2 of 35.8%. This further confirms our theoretical analysis and empirical results.

6 Conclusion

In this paper, we develop a dynamic overlapping generations (OLG) equilibrium model to study the relation between demographics and the equity risk premium. We let the per capita stock supply be constant, and the demographic structure be deterministic, exogenously given and time varying. Investors' income derives from two sources: labor and capital investments. The labor income process and the dividend process are correlated. Investors trade stocks both for consumption purposes and to hedge against the risk of labor income. Since the equilibrium excess return on stocks follows an AR(1) process (with deterministic time varying coefficients), investors with exponential utility over consumption are more risk taking and, everything else equal, hold more risky assets on average when young than when old.

A non-linear predictive relation between the equity risk premium and demographic variables, the real per capita stock price, the dividend yield and the dividend payout ratio is derived from the equilibrium price function. In particular, the equity risk premium is linear in the real per capita stock price, the dividend yield and the dividend payout ratio, but the coefficients of this linear relation are time varying because of demographic change. These coefficients depend on the current and all the future demographic structures.

Calibration results suggest that the historical demographic change could induce significant variations in the coefficients of the predictive relation. Also, these coefficients are correlated with the change in the share of population in the age range 40-64. Writing these coefficients as linear functions of the change in this share of population, we express the equity risk premium as a linear function of the change in this share of population, the real per capita stock price, the dividend yield, the dividend payout ratio and the products of the change in this share of population with the real per capita stock price, the dividend yield and the dividend payout ratio.

The inclusion of demographic variables significantly improves on the previous predictive regressions, in which the equity risk premium is regressed on the dividend yield and/or the dividend payout ratio. In particular, replication of the regression by Lamont (1998) for the period 1947-1994, in which the quarterly equity risk premium is regressed on the dividend yield and the dividend payout ratio, yields R^2 of 11.7%, while adding the change in the share of population in the age range 40-64 to the regression raises R^2 to 14.7%. For the period 1947-1999, the previous predictive regressions are no longer significant, while our predictive regression is significant with R^2 of 13.6%.

Appendix A

A.1 Proof of Proposition 3.1 and Theorem 3.2

In what follows, we first treat proposition 3.1 as a conjecture, and prove theorem 3.2. Then, we use the results from the proof of theorem 3.2 to prove proposition 3.1.

Proof of Theorem 3.2

The investors' optimization problem, equations (3.1)-(3.2), can be expressed in the form of the Bellman equation

$$0 = \max_{X,c}\left\{-\beta^{\tau}\exp(-\alpha c_t) + \mathrm{E}_t\left[J(W_{t+1}, \Im_{t+1}, \tau+1; t+1)\right] - J(W_t, \Im_t, \tau; t)\right\} \tag{A.1}$$

subject to

$$W_{t+1} = (W_t - c_t)R + X_t Q_{t+1} + Y_{\tau+1,t+1} \tag{A.2}$$

where Q_{t+1} is expressed as in equation (3.4), given that proposition 3.1 is true.

Consider the following trial solution for the value function

$$J(W_t, \Im_t, \tau; t) = -\beta^{\tau}\exp(-\gamma_\tau W_t - \mu_\tau t - \frac{1}{2}\Pi_t^T \nu_\tau \Pi_t) \tag{A.3}$$

where γ_τ and μ_τ are functions of τ. ν_τ is a $(N+1)\times(N+1)$ symmetric matrix dependent on τ.

Let $\Pi_{t+1} = a\Pi_t + b\epsilon_{t+1}$, where $a = \begin{pmatrix} 1 & 0_{1\times N} \\ 0_{N\times 1} & a_Z \end{pmatrix}$ and $b = \begin{pmatrix} 0 & 0 & 0_{1\times N} \\ 0_{N\times 1} & 0_{N\times 1} & I_{N\times N} \end{pmatrix}$. Denote $\omega_\tau^* = \begin{pmatrix} 0 & 0_{1\times N} \\ 0_{N\times 1} & \omega_\tau \end{pmatrix}$, $\omega_\tau^{aa} = a^T\omega_\tau^* a$, $\omega_\tau^{ab} = a^T\omega_\tau^* b$, $\omega_\tau^{bb} = b^T\omega_\tau^* b$, $\nu_\tau^{aa} = a^T\nu_\tau a$, $\nu_\tau^{ab} = a^T\nu_\tau b$, $\nu_\tau^{bb} = b^T\nu_\tau b$, $\Sigma_\tau = (\Sigma^{-1}+\nu_\tau^{bb}+\gamma_\tau\omega_\tau^{bb})^{-1}$, $\Gamma_\tau = (\Phi\Sigma_\tau\Phi^T)^{-1}$, $q_\tau = \Theta-\Phi\Sigma_\tau(\gamma_\tau\omega_\tau^{ab}+\nu_\tau^{ab})^T$ and $d_\tau = |\Sigma_\tau^{-1}\Sigma|^{-\frac{1}{2}}\exp\left\{\frac{1}{2}\gamma_\tau^2\sigma_\tau^2 - \left[\gamma_\tau(h_\tau+n_\tau)+\mu_\tau\right]\right\}$. It is straightforward to show that

$$
\begin{aligned}
&\mathrm{E}_\tau\left[J(W_{t+1},\Pi_{t+1},\tau+1;t+1)\right] \\
= \; & -d_{\tau+1}\beta^{\tau+1}\exp\Big\{-\gamma_{\tau+1}(W_t-c_t)R-(\gamma_{\tau+1}n_{\tau+1}+\mu_{\tau+1})t \\
& -\gamma_{\tau+1}X_t q_{\tau+1}\Pi_t + \frac{1}{2}\gamma_{\tau+1}^2 X_t^2\Gamma_{\tau+1}^{-1} \\
& -\frac{1}{2}\Pi_t^T\left[\gamma_{\tau+1}\omega_{\tau+1}^{aa}+\nu_{\tau+1}^{aa}-(\gamma_{\tau+1}\omega_{\tau+1}^{ab}+\nu_{\tau+1}^{ab})\Sigma_{\tau+1}(\gamma_{\tau+1}\omega_{\tau+1}^{ab}+\nu_{\tau+1}^{ab})^T\right]\Pi_t\Big\} \qquad \text{(A.4)}
\end{aligned}
$$

Substitute expression (A.4) into equation (A.1) and take derivatives with respect to X_t and c_t to obtain

$$
\begin{aligned}
-\gamma_{\tau+1}q_{\tau+1}\Pi_t + \gamma_{\tau+1}^2\Gamma_{\tau+1}^{-1}X_t &= 0 \qquad \text{(A.5)} \\
\alpha\beta^\tau\exp(-\alpha c_t) + \gamma_{\tau+1}R\mathrm{E}_t\left[J(W_{t+1},\Pi_{t+1},\tau+1;t+1)\right] &= 0 \qquad \text{(A.6)}
\end{aligned}
$$

Denote $m_\tau = \nu_\tau^{aa}+\gamma_\tau\omega_\tau^{aa}-(\gamma_\tau\omega_\tau^{ab}+\nu_\tau^{ab})\Sigma_\tau(\gamma_\tau\omega_\tau^{ab}+\nu_\tau^{ab})^T+q_\tau^T\Gamma_\tau q_\tau$ and $\bar{c}_\tau = \dfrac{1}{\alpha+R\gamma_{\tau+1}}\ln(\dfrac{\alpha}{d_{\tau+1}\beta R\gamma_{\tau+1}})$. The optimal investment-consumption policy is

$$
\begin{aligned}
X_t &= \frac{1}{\gamma_{\tau+1}}\Gamma_{\tau+1}q_{\tau+1}\Pi_t \qquad \text{(A.7)} \\
c_t &= \bar{c}_\tau + \frac{R\gamma_{\tau+1}}{\alpha+R\gamma_{\tau+1}}W_t + \frac{\gamma_{\tau+1}n_{\tau+1}+\mu_{\tau+1}}{\alpha+R\gamma_{\tau+1}}t + \frac{1}{2(\alpha+R\gamma_{\tau+1})}\Pi_t^T m_{\tau+1}\Pi_t \qquad \text{(A.8)}
\end{aligned}
$$

The optimal demand for stocks as expressed in theorem 3.1 is immediate if one substitutes

the expression of $q_{\tau+1}$ into equation A.7. Substituting the optimal consumption-investment policy back into the Bellman equation (A.1), we obtain

$$\gamma_\tau = \frac{\alpha R \gamma_{\tau+1}}{\alpha + R\gamma_{\tau+1}} \tag{A.9}$$

$$\mu_\tau = \frac{\alpha(\gamma_{\tau+1} n_{\tau+1} + \mu_{\tau+1})}{\alpha + R\gamma_{\tau+1}} \tag{A.10}$$

$$\nu_\tau = -2\ln\left[\exp(-\alpha\bar{c}_\tau) + d_{\tau+1}\beta\exp(R\gamma_{\tau+1}\bar{c}_\tau)\right]\mathrm{i}_{11}^{(N+1,N+1)} + \frac{\alpha}{\alpha + R\gamma_{\tau+1}} m_{\tau+1} \tag{A.11}$$

where $\mathrm{i}_{11}^{(n,n)}$ is an $n \times n$ index matrix.[22] Equations (A.9)-(A.11), combined with the initial values $\gamma_T = \alpha$, $\mu_T = 0$ and $\nu_T = 0_{(N+1)\times(N+1)}$, recursively solve γ_τ, μ_τ and ν_τ. The solution determines γ_τ, μ_τ and ν_τ in the trial value function, equation (A.3), and thereafter fully specifies the investors' optimal consumption-investment policy, equations (A.7)-(A.8).

The trial value function is confirmed on condition that the price function is as claimed in proposition 3.1. To ensure that the value function is actually the equilibrium value function, we have to verify proposition 3.1.

Proof of Proposition 3.1

From equation (A.7), at time t, each investor of age τ has demand for stocks, $\frac{1}{\gamma_{\tau+1}}\kappa_{\tau+1}\Pi_t$. Market clearing requires

$$G(t)\sum_{\tau=0}^{T-1} g(\tau)\frac{1}{\gamma_{\tau+1}}\kappa_{\tau+1}\Pi_t = G(t) \tag{A.12}$$

Therefore

[22] An index matrix $\mathrm{i}_{ij}^{(m,n)}$ is an $m \times n$ matrix with the element $\{i, j\}$ being one and all other elements being zero.

$$\left[\sum_{\tau=0}^{T-1} g(\tau)\frac{1}{\gamma_{\tau+1}}\kappa_{\tau+1}\right]_{11} = 1 \tag{A.13}$$

$$\left[\sum_{\tau=0}^{T-1} g(\tau)\frac{1}{\gamma_{\tau+1}}\kappa_{\tau+1}\right]_{12} = 0_{1\times N} \tag{A.14}$$

where $[.]_{mn}$ is the element $\{m, n\}$ of the matrix. Equations (A.13)-(A.14) are a set of algebraic equations of p_0 and p_Z. The solution completely specifies the proposed equilibrium price function. We are not able to express the roots in analytical form. A numerical method can be used to solve equations (A.13)-(A.14). This verifies proposition 3.1, and further confirms the proof of theorem 3.2.

A.2 Proof of Proposition 3.3 and Theorem 3.4

In what follows, we first treat proposition 3.3 as a conjecture, and prove theorem 3.4. Then, we use the results from the proof of theorem 3.4 to prove proposition 3.3. The derivation is quite similar to that for proposition 3.1 and theorem 3.2. But to allow for demographic change makes the coefficients of the price function and the value function dependent on time t.

Proof of theorem 3.4

The investors' optimization problem, equations (3.1)-(3.2), can be expressed in the form of the Bellman equation

$$0 = \max_{X,c}\left\{-\beta^{\tau}\exp(-\alpha c_t) + \mathrm{E}_t\Big[J(W_{t+1}, \Im_{t+1}, \tau+1; t+1)\Big] - J(W_t, \Im_t, \tau; t)\right\} \tag{A.15}$$

subject to

$$W_{t+1} = (W_t - c_t)R + X_t Q_{t+1} + Y_{\tau+1,t+1} \tag{A.16}$$

where Q_{t+1} is expressed as in equation 3.9, given proposition 3.3 is true.

Consider the following trial solution for the value function

$$J(W_t, \Im_t, \tau; t) = -\beta^\tau \exp\left[-\gamma_\tau W_t - \mu_\tau t - \frac{1}{2}\Pi_t^T \nu_{\tau,t}\Pi_t\right] \tag{A.17}$$

where γ_τ and μ_τ are defined as in equation (A.9)-(A.10). $\nu_{\tau,t}$ is a $(N+1)\times(N+1)$ symmetric matrix dependent on τ and t.

Denote $\nu_{\tau,t}^{aa} = a^T\nu_{\tau,t}a$, $\nu_{\tau,t}^{ab} = a^T\nu_{\tau,t}b$, $\nu_{\tau,t}^{bb} = b^T\nu_{\tau,t}b$, $\Sigma_{\tau,t} = (\Sigma^{-1} + \nu_{\tau,t}^{bb} + \gamma_\tau\omega_\tau^{bb})^{-1}$, $\Gamma_{\tau,t} = (\Phi_t\Sigma_{\tau,t}\Phi_t^T)^{-1}$, $q_{\tau,t} = \Theta_t - \Phi_t\Sigma_{\tau,t}(\gamma_\tau\omega_\tau^{ab} + \nu_{\tau,t}^{ab})^T$ and $d_{\tau,t} = |\Sigma_{\tau,t}^{-1}\Sigma|^{-\frac{1}{2}}\exp\left\{\frac{1}{2}\gamma_\tau^2\sigma_\tau^2 - \left[\gamma_\tau(h_\tau + n_\tau) + \mu_\tau\right]\right\}$. It is straightforward to show that

$$\begin{aligned}
&\mathrm{E}_\tau\left[J(W_{t+1}, \Pi_{t+1}, \tau+1; t+1)\right] \\
= \; & -d_{\tau+1,t+1}\beta^{\tau+1}\exp\Big\{-\gamma_{\tau+1}(W_t - c_t)R - (\gamma_{\tau+1}n_{\tau+1} + \mu_{\tau+1})t \\
& -\gamma_{\tau+1}X_t q_{\tau+1,t+1}\Pi_t + \frac{1}{2}\gamma_{\tau+1}^2 X_t^2\Gamma_{\tau+1,t+1}^{-1} - \frac{1}{2}\Pi_t^T\Big[\gamma_{\tau+1}\omega_{\tau+1}^{aa} + \nu_{\tau+1,t+1}^{aa} \\
& -(\gamma_{\tau+1}\omega_{\tau+1}^{ab} + \nu_{\tau+1,t+1}^{ab})\Sigma_{\tau+1,t+1}(\gamma_{\tau+1}\omega_{\tau+1}^{ab} + \nu_{\tau+1,t+1}^{ab})^T\Big]\Pi_t\Big\}
\end{aligned} \tag{A.18}$$

Substitute expression (A.18) into equation (A.15) and take the derivatives with respect to X_t and c_t to obtain

$$-\gamma_{\tau+1}q_{\tau+1,t+1}\Pi_t + \gamma_{\tau+1}^2\Gamma_{\tau+1,t+1}^{-1}X_t = 0 \tag{A.19}$$

$$\alpha\beta^\tau\exp(-\alpha c_t) + \gamma_{\tau+1}R\mathrm{E}_t\left[J(W_{t+1}, \Pi_{t+1}, \tau+1; t+1)\right] = 0 \tag{A.20}$$

Denote $\bar{c}_{\tau,t} = \dfrac{1}{\alpha + R\gamma_{\tau+1}}\ln\left(\dfrac{\alpha}{d_{\tau+1,t+1}\beta R\gamma_{\tau+1}}\right)$, and $m_{\tau,t} = \nu_{\tau,t}^{aa} + \gamma_\tau\omega_\tau^{aa} - (\gamma_\tau\omega_\tau^{ab} + \nu_{\tau,t}^{ab})\Sigma_{\tau,t}(\gamma_\tau\omega_\tau^{ab} + \nu_{\tau,t}^{ab})^T + q_{\tau,t}^T\Gamma_{\tau,t}q_{\tau,t}$. The optimal investment-consumption policy is

$$
\begin{aligned}
X_t &= \frac{1}{\gamma_{\tau+1}}\Gamma_{\tau+1,t+1}q_{\tau+1,t+1}\Pi_t && \text{(A.21)}\\
c_t &= \bar{c}_{\tau,t} + \frac{R\gamma_{\tau+1}}{\alpha + R\gamma_{\tau+1}}W_t + \frac{\gamma_{\tau+1}n_{\tau+1} + \mu_{\tau+1}}{\alpha + R\gamma_{\tau+1}}t \\
&\quad + \frac{1}{2(\alpha + R\gamma_{\tau+1})}\Pi_t^T m_{\tau+1,t+1}\Pi_t && \text{(A.22)}
\end{aligned}
$$

The optimal demand for stocks as expressed in theorem 3.4 is immediate if one substitutes the expression of $q_{\tau+1,t+1}$ into equation (A.21). Substituting the optimal consumption-investment policy back into the Bellman equation (A.15), we obtain

$$
\begin{aligned}
\nu_{\tau,t} &= -2\ln\left[\exp(-\alpha\bar{c}_{\tau,t}) + d_{\tau+1,t+1}\beta\exp(R\gamma_{\tau+1}\bar{c}_{\tau,t})\right]\mathrm{i}_{11}^{(N+1,N+1)} \\
&\quad + \frac{\alpha}{\alpha + R\gamma_{\tau+1}}m_{\tau+1,t+1} && \text{(A.23)}
\end{aligned}
$$

The trial value function is confirmed on condition that the price function is as claimed in proposition 3.3. To ensure that the trial indirect utility function is actually the equilibrium value function, we have to verify proposition 3.3.

Proof of theorem 3.4

From equation (A.21), at time t, each investor of age τ has demand for stocks, $\frac{1}{\gamma_{\tau+1}}\kappa_{\tau+1,t}\Pi_t$. Market clearing requires

$$
G(t)\sum_{\tau=0}^{T-1} g(\tau;t)\frac{1}{\gamma_{\tau+1}}\kappa_{\tau+1,t+1}\Pi_t = G(t) \qquad \text{(A.24)}
$$

Therefore

$$
\begin{aligned}
\left[\sum_{\tau=0}^{T-1} g(\tau;t)\frac{1}{\gamma_{\tau+1}}\kappa_{\tau+1,t+1}\right]_{11} &= 1 && \text{(A.25)}\\
\left[\sum_{\tau=0}^{T-1} g(\tau;t)\frac{1}{\gamma_{\tau+1}}\kappa_{\tau+1,t+1}\right]_{12} &= 0_{1\times N} && \text{(A.26)}
\end{aligned}
$$

Equations (A.25)-(A.26) are a set of algebraic equations of p_0 and p_Z. The solution completely specifies the proposed equilibrium price function. We are not able to express the roots in analytical form. A numerical method can be used to solve equations (A.25)-(A.26). This verifies proposition 3.3, and further confirms the proof of theorem 3.4.

Appendix B

Derivation of Approximation Formula

$$\begin{aligned}
\ln(y+x) &= \ln(y+e^{\ln x}) \\
&\approx \ln(y+e^{\bar{\ln} x}) + e^{\bar{\ln} x}(y+e^{\bar{\ln} x})^{-1}(\ln x - \bar{\ln} x) \\
&\approx \ln(y+e^{\bar{\ln} x}) + e^{\bar{\ln} x}(\bar{y}+e^{\bar{\ln} x})^{-1}(\ln x - \bar{\ln} x)
\end{aligned}$$

The first approximation is obtained by viewing $\ln(y+x)$ as a function of $\ln x$ and taking Tailor expansion around $\bar{\ln} x$. The second approximation is obtained by substitute $\bar{y}$ for y.

Appendix C

Data

The Consumer Price index (CPI) data, 1982-1984=100, are from the Department of Labor, Bureau of Labor Statistics. We use the CPI of the last month of the quarter as the CPI for that quarter.

The annual demographic data are from Citibase. Citibase provides the annual demographic data from 1946 to 1997 in 5 years interval in people's age. It also provides the demographic data after 1997, projected by the Bureau of Census. We view the projection as true historical data. We only consider the population in the age range 20-74, with the average population from 1982 to 1984 normalized to unity. A simple interpolation is used to calculate the population of any age in each year.

The Panel Study of Income Dynamics (PSID) provides a panel of annual observations of individual and family income and other variables from 1967 to 1992. The PSID oversamples poorer members of the U.S. population by including a sample of poor families from the Survey of Economic Opportunity (SEO). We dropped the families that were originally part of the SEO to obtain a random sample. Only families with a male head from age 20 to 74 are used. We take a broad definition of labor income. Total family labor income includes total labor income of the head of the family and his wife along with total transfers to the family. The transfers include unemployment compensation, workers' compensation, pension income, child support, social security, and so on. Observations that still report zero for this broad income category are dropped. Labor income defined this way is adjusted to the S&P Composite Index.

The unemployment rate data are from the Department of Labor, Bureau of Labor Statistics. We use the average unemployment rate in a quarter as the unemployment rate of that quarter.

All data on stock and bill returns come from Ibbotson Associates. Excess stock returns are $r_{m,t+1} - r_{f,t+1}$, defined as $\ln(\text{CSTIND}_{t+1}/\text{CSTIND}_t) - \ln(\text{CSTIND}_{t+1}/\text{CSTIND}_t)$, where CSTIND is an index of total return (including reinvested dividends) on the S&P Composite Index and USTIND is an index of total return on T-bills, as of the last day of quarter t.

The basic earnings and dividends data are from the *Security Price Index Record* published by Standard & Poor's Statistical Service. EPS is quarterly earnings per share, Adjusted to Index, Composite. DPS is 12-month moving total dividends per share, Adjusted to Index, Composite. The *Index Record* report dividends and earnings indexed to their composite price index, SPLEVEL. We define log price as $p \equiv \ln(\text{SPLEVEL})$, log dividends as $d \equiv \ln(\text{DPS})$ and log earnings as $e \equiv \ln(\text{EPS})$ and . Log sum of CPI and population is defined as $\varsigma \equiv \ln(\text{CPI}/100) + \ln(\text{Pop})$, where CPI is the price level of the last month in the quarter, and Pop is the amount of population in the previous year.

Reference

Abel, A. B., 1999, The Effects of a Baby Boom on Stock Prices and Capital Accumulation in the Presence of Social Security, mimeo, Wharton School, University of Pennsylvania.

Bakshi, G., and Z. Chen, 1994, Baby Boom, Population Aging, and Capital Markets, *Journal of Business* 67, 165-202.

Bergantino, S., 1998, Lifecycle Investment Behavior, Demographics, and Asset Prices, Doctoral Dissertation, MIT.

Bodie, Z., R. C. Merton, and W. Samuelson, 1992, Labor Supply Flexibility and Portfolio choice in a Life Cycle Model, *Journal of Economic Dynamics and Control* 16, 427-449.

Bossaerts, P., and P. Hillion, 1999, Implementing Statistical Criteria to Select Return Forecasting Models: What Do We Learn? *Review of Financial Studies* 12, 405-428.

Bossons, J., 1973, The Distribution of Assets among Individuals of Different Age and Wealth. In R. Goldsmith (ed.), *Institutional Investors and Corporate Stock*, New York: NBER.

Brennan, M. J., 1998, The Role of Learning in Dynamic Portfolio Decision, *European Finance Review* 1, 295-306.

Brooks, R. J., 1999, What will Happen to Financial Markets When Baby Boomers Retire? mimeo, Federal Reserve Board of Governors.

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

Campbell, J. Y., and A. S. Kyle, 1993, Smart Money, Noise Trading and Stock Price Behavior, *Review of Economic Studies* 60, 1-34.

Campbell, J. Y., and R. J. Shiller, 1988, The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors, *Review of Financial Studies* 1, 195-228.

Canner, N., N. G. Mankiw, and D. N. Weil, 1997, An Asset Allocation Puzzle, *American Economic Review* 87, 181-191.

Cochrane, J. H., 1997, Where is the Market Going? Uncertain Facts and Novel theories, *Economic Perspectives* 21, 3-37.

Erb, C. B., C. R. Harvey and T. E. Viskanta, 1997, Demographics and International Investments, *Financial Analysts Journal* 53, 14-28.

Fama, E. F., and K. R. French, 1988, Dividend Yields and Expected Stock Returns, *Journal of Financial Economics* 22, 3-25.

Friedman, M, 1957, A Theory of Consumption Function, a Study by the National Bureau of Economic Research, New York, Princeton, Princeton University Press.

Goetzman, W. M., and P. Jorion, 1995, A Longer Look at Dividend Yields, *Journal of Business* 68, 483-508.

Goetzman, W. M., and P. Jorion, 1993, Testing the Predictive Power of Dividend Yields, *Journal of Finance* 48, 663-679.

Gompers, P., and J. Lerner, 2000, Money Chasing Deals? the Impact of Fund Inflows on Private Equity Valuations, *Journal of Financial Economics* 55, 281-325.

Goyal, A., and I. Welch, 1999, Predicting the Equity Premium, Working Paper, UCLA.

Heaton, J., and D. J. Lucas, 1999, Saving Behavior and Portfolio Choice: Which Risks Matter? *Journal of Finance*, Forthcoming.

Hirshleifer, D., and A. Subrahmanyam, 1993, futures versus Share Contracting as Means of Diversifying Output Risk, *Economic Journal* 103, 620-638.

Hodrick, R. J., 1992, Dividend Yields and Expected Stock Returns: Alternative Procedures for Inference and Measurement, *Review of Financial Studies* 5, 357-386.

Horvath, M. T. K., and M. W. Watson, 1995, Testing for Cointegration when Some of the Cointegrating Vectors are Prespecified, *Econometric Theory* 11, 984-1014.

Hubbard, R. G., and J. Skinner, and S. P. Zeldes, Expanding the life-cycle model: Precautionary saving and public policy, *American Economic Review* 84, 174-179.

Kimball, M. S., 1993, Standard Risk Aversion, *Econometrica* 61, 589-611.

Lamont, O., 1998, Earnings and Expected Returns, *Journal of Finance* 53, 1563-1587.

Mankiw, N. G., and D. Weil, 1989, The Baby Boom, the Baby Bust, and the Housing Market, *Regional Science and Urban Economics* 19, 235-258.

Macunovich, D., 1997, Discussion of 'Social Security: How Social and Secure should It be? in S. Sass and R. Triest, eds., *Social Security Reform: Links to Saving, Investment, and Growth* (Boston: Federal Reserve Bank of Boston), 64-76.

Modigliani, F., 1986, Life Cycle, Individual thrift, and the Wealth of the Nations, *American Economic Review* 76, 297-312.

Modigliani, F., and R. Brumberg, 1954, Utility Analysis and the Consumption Function: and Interpretation of Cross-section Data, in K. Kurihara eds., *Post-Keynesian Economics*, New Brunswich, N.J.: Rutgers University Press.

Nelson, C. R., and M. J. Kim, 1993, Predictable Stock Returns: Robustness and Economic Significance, *Journal of Finance* 48, 641-661.

Newey, W., and K. West, 1987, A Simple, Positive Semi-definite, Heteroskedastically and Autocorrelation Consistent Covariance Matrix, *Econometrica* 55, 703-708.

Pesaran, M. H., and A. Timmermann, 1995, Predictability of Stock Returns: Robustness and Economics Significance, *Journal of Finance* 50, 1201-1228.

Pischke, J., 1995, Individual Income, Incomplete Information, and Aggregate Consumption, *Econometrica* 63, 805-940.

Poterba, J. M., 2000, Population Age Structure and Asset Returns: an Empirical Investigation, NBER Working Paper.

Poterba, J. M., and A. A. Samwick, 1997, Household Portfolio Allocation over the Lifecycle, NBER Working Paper.

Samuelson, P. A., 1969, Lifetime Portfolio Selection by Dynamic Stochastic Programming, *Review of Economics and Statistics*, 239-246.

Samuelson, P. A., 1989, A Case at Last for Age-phased Reduction in Equity, *Proceedings of the National Academy of Sciences* 86, 9048-9051.

Samuelson, P. A., 1991, Long-Run Risk Tolerance when Equity Returns are Mean Regressing: Pseudoparadoxes and Vindication of a 'Businessman's Risk', in William C. Brainard, W. D. Nordhaus and H. W. Watts eds., *Money, Macroeconomics, and Economic Policy: Essays in Honor of James Tobin.* Cambridge and London: MIT Press, 181-200.

Schieber, S., and J. Shoven, 1997, The Consequences of Population Aging on Private Pension Fund Saving and Asset Markets, in S. Schieber and J. Shoven, eds., *Public Policy Towards Pension*, Cambridge: MIT Press, 219-245..

Siegel, J., 1998, *Stocks for the Long Run, 2nd*, New York: McGraw Hill.

Shiller, R. J., 1984, Stock Prices and Social Dynamics, *Brookings Papers on Economic Activity*, 457-498.

Stambaugh, R., 1999, Predictive Regressions, *Journal of Financial Economics* 54, 375-421.

Viceria, L. M., 1997, Testing for Structural Changes in the Predictability of Asset Returns, Working Paper, Harvard University.

Yoo, P. S., 1994a, Age Dependent Portfolio Selection, Working Paper 94-003, Federal Reserve Bank of St. Louis.

Yoo, P. S., 1994b, Age Distributions and Returns of Financial Assets, Working Paper 94-002b, Federal Reserve Bank of St. Louis.

Xia, Y., 2000, Learning about Predictability: the Effect of Parameter Uncertainty on Optimal Dynamic Consumption and Asset Allocation, *Journal of Finance*, Forthcoming.

Figure 1: The Risk Aversion Coefficient of the Value Function

The figure plots the risk aversion coefficient of the value function, γ_τ. The parameter values are $\alpha = 0.05$, $r = 1.2\%$ and $T = 54$. We consider the population in the age range 20-74, so that the economic age of an investor of calendar age 20 is $\tau = 0$. The x-axis is of calendar age.

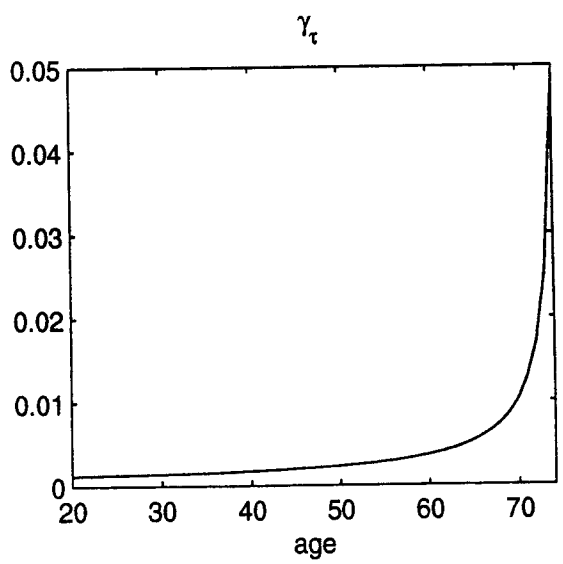


Figure 2: The Demographic Structure of the U.S. Population in the Age Range 20-74 in 1947-2050

Source: Citibase. The demographic data of U.S. population after 1997 are based on projections by Bureau of Census.

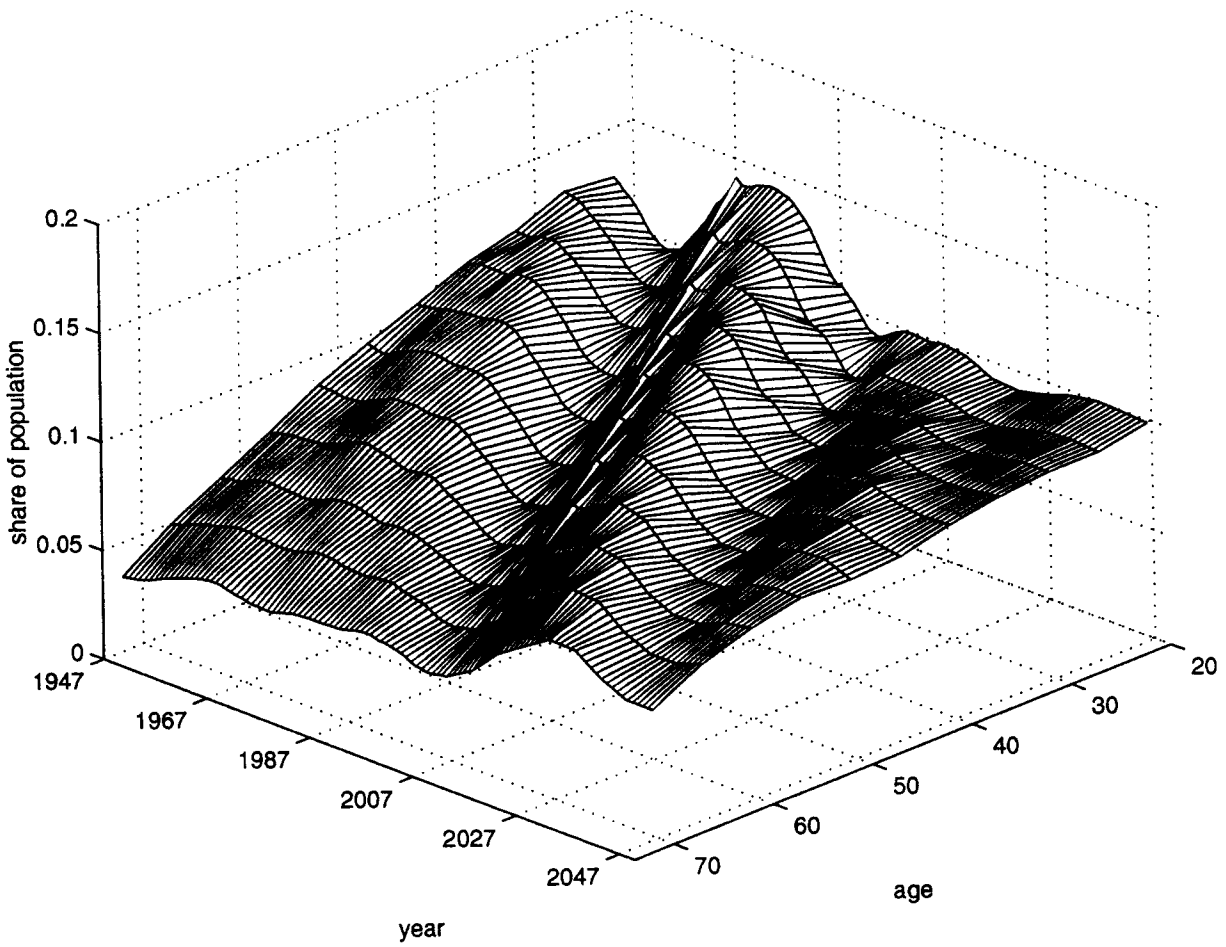


Figure 3: Moments of the U.S. Demographic Structure in 1947-2050

Source: Citibase. The whole population include only people in the age range 20-74. The demographic data of U.S. population after 1997 are based on projections by Bureau of Census.

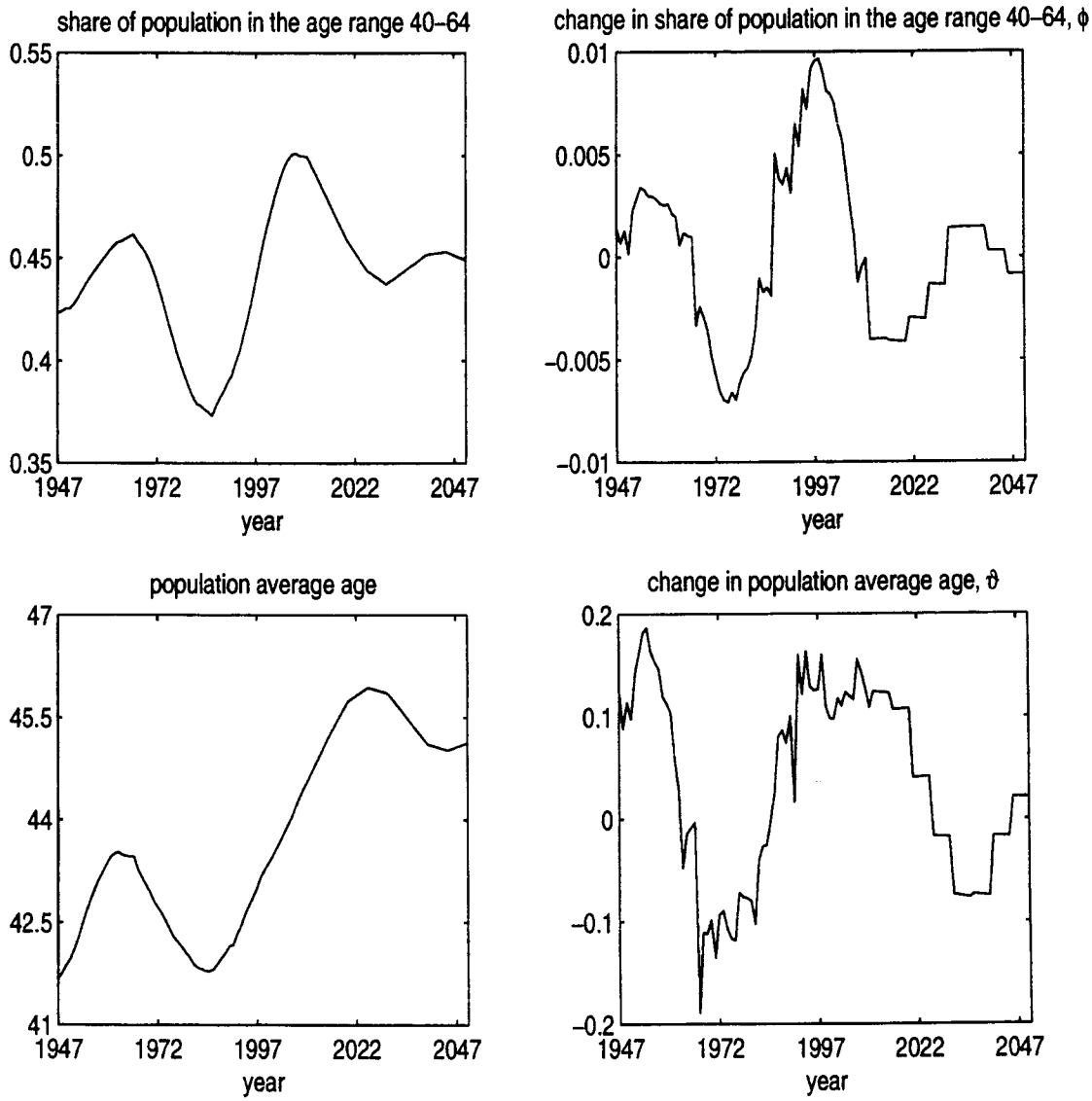


Figure 4: The Estimates of the Labor Income Process

At time t, an investor of age τ has labor income

$$Y_{\tau,t} = h_\tau + n_\tau t + \frac{1}{2} Z_t^T \omega_\tau Z_t + \epsilon_{\tau,t}$$

The variance of $\epsilon_{\tau,t}$ is σ_τ^2. The solid lines represent the empirical estimates of the labor income process, h_τ, n_τ, ω_τ and σ_τ; the dark dashed lines represent the cubic polynomial approximations of the empirical estimates.

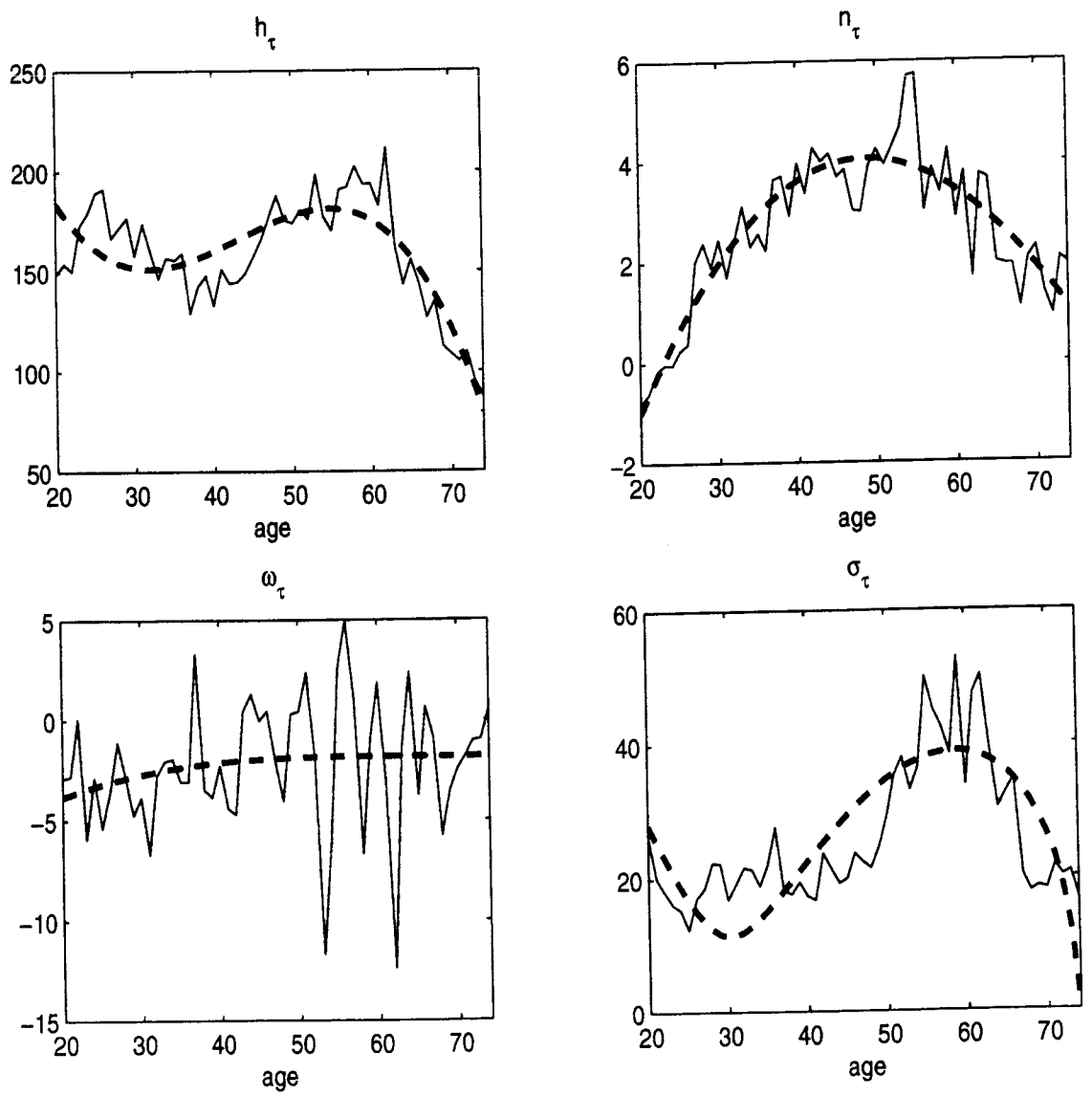


Figure 5: The Time Varying Coefficients of the Price Function

The figures plot the coefficients of the price function

$$P_t = p_{0,t} + p_F F_t + p_{Z,t} Z_t$$

where $p_{0,t}$ is the constant term; p_F is the coefficient of the dividend state variable, F; $p_{Z,t}$ is that of the labor income state variable, Z. $p_{0,t}$ and $p_{Z,t}$ are time varying because of demographic change. p_F is a constant, which equals 2.757 under the calibration.

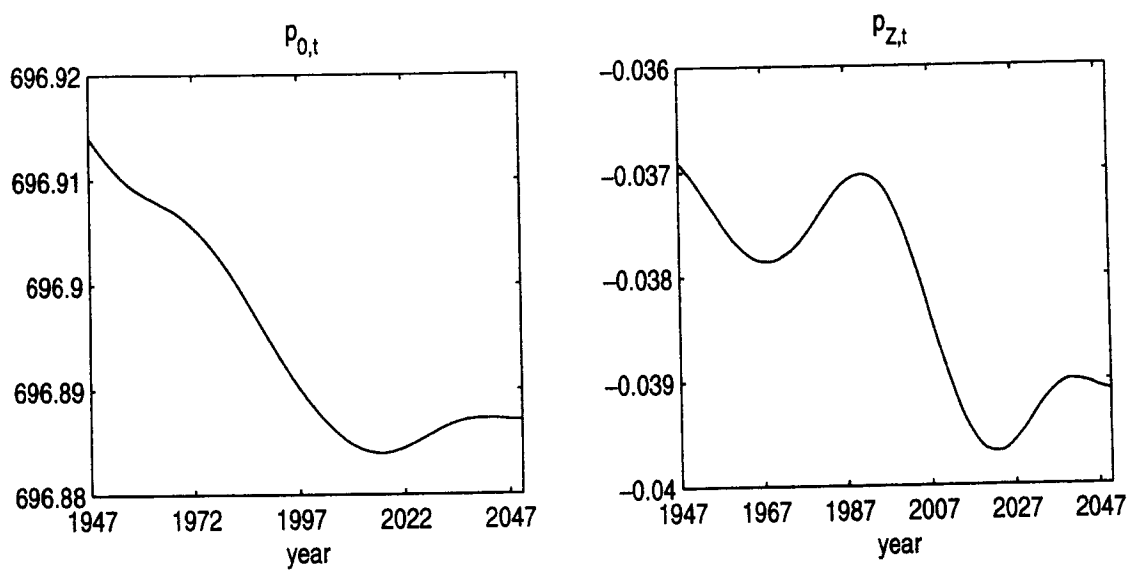


Figure 6: The Time Varying Coefficients of the Expected Excess Return

The figures plot the coefficients of the expected excess return

$$E_t(Q_{t+1}) = \theta_{0,t+1} + \theta_{Z,t+1} Z_t$$

where $\theta_{0,t+1}$ is the constant term; $\theta_{Z,t+1}$ is the coefficient of the labor income state variable, Z. These coefficients are time varying because of demographic change.

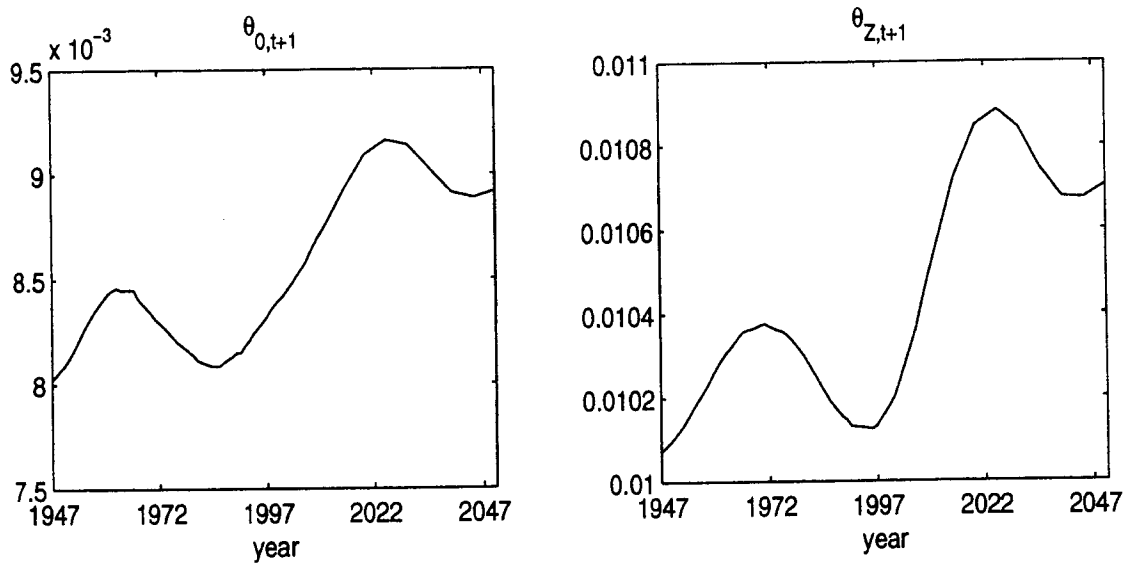


Figure 7: The Time Varying Coefficients of the Predictive Relation

The figures plot the coefficients of the predictive relation

$$R_{m,t+1} - R = \lambda_{0,t} + \lambda_{1,t}\frac{1}{P_t} + \lambda_{2,t}\frac{D_t}{P_t} + \lambda_{3,t}\frac{B_t}{P_t} + \varepsilon_{t+1}$$

where $\lambda_{0,t}$ is the constant term; $\lambda_{1,t}$ is the coefficient of the inverse of the real per capita stock price, $\frac{1}{P_t}$; $\lambda_{2,t}$ is that of the dividend yield, $\frac{D_t}{P_t}$; $\lambda_{3,t}$ is that of the earnings yield, $\frac{B_t}{P_t}$. In general, these coefficients are time varying because of demographic change. $\lambda_{2,t}$ is zero under the calibration.

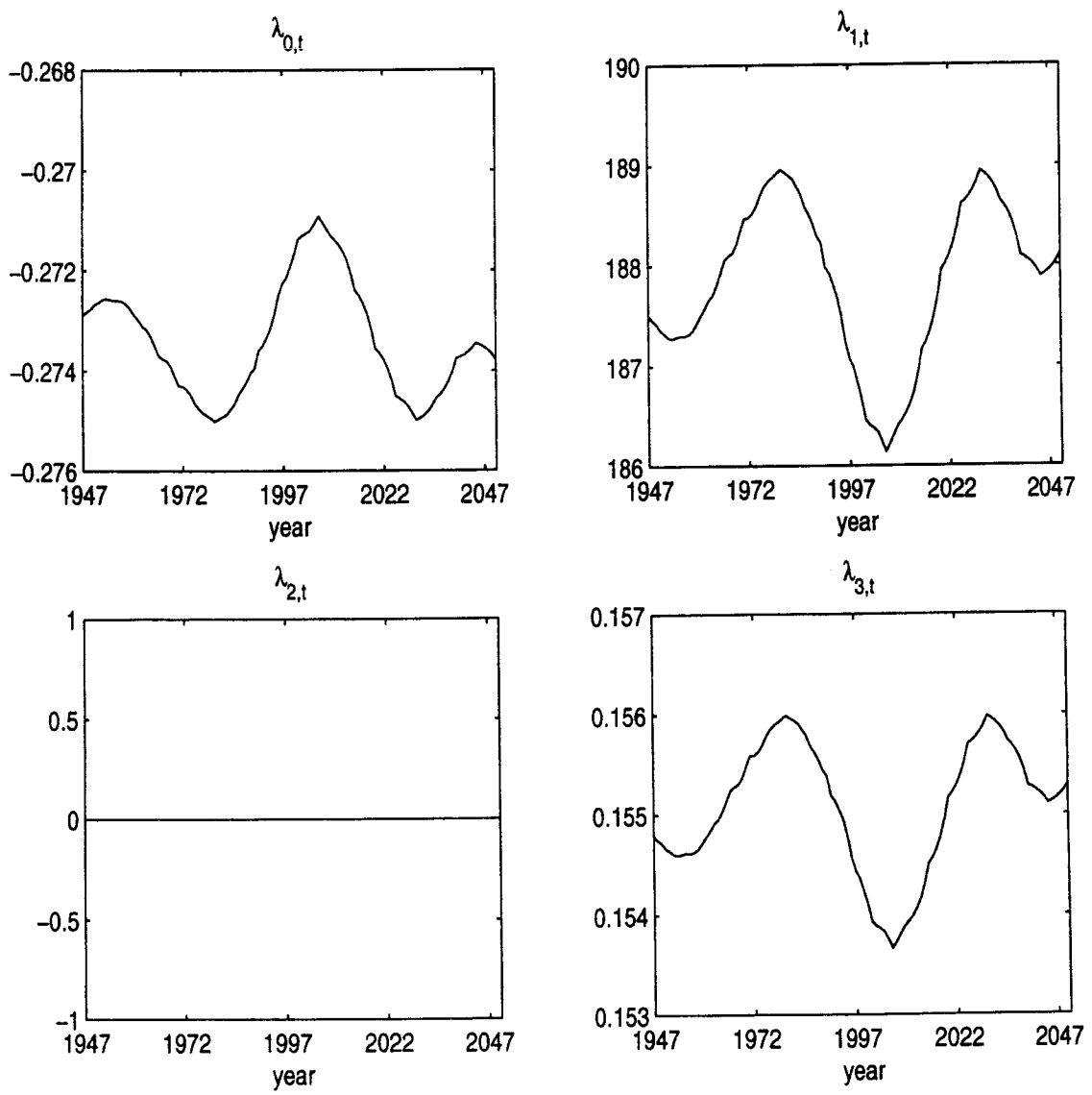


Figure 9: The Historical Equity Risk Premium and the Fitted Value of the Predictive Regression

The solid line represents the quarterly log excess return on the S&P Composite Index, $r_m - r_f$. The dashed line represents the fitted value of the predictive regression by Lamont (1994), calculated by

$$r_{m,t+1} - r_{f,t+1} = 0.038 + 0.016(d_t - p_t) + 0.050(d_t - e_t)$$

where $d - p$ is the log dividend yield; $d - e$ is the log dividend payout ratio. This predictive regression is estimated in row 6 of table 4. The asterisk represents the fitted value of the predictive regression, calculated by

$$r_{m,t+1} - r_{f,t+1} = 0.202 + 0.076(d_t - p_t) - 2.215\phi_t(d_t - p_t) + 0.096(d_t - e_t) - 9.731\phi_t(d_t - e_t)$$

where ϕ is the previous year's change in the log share of population in the age range 40-64. This predictive regression is estimated in row 8 of table 6.

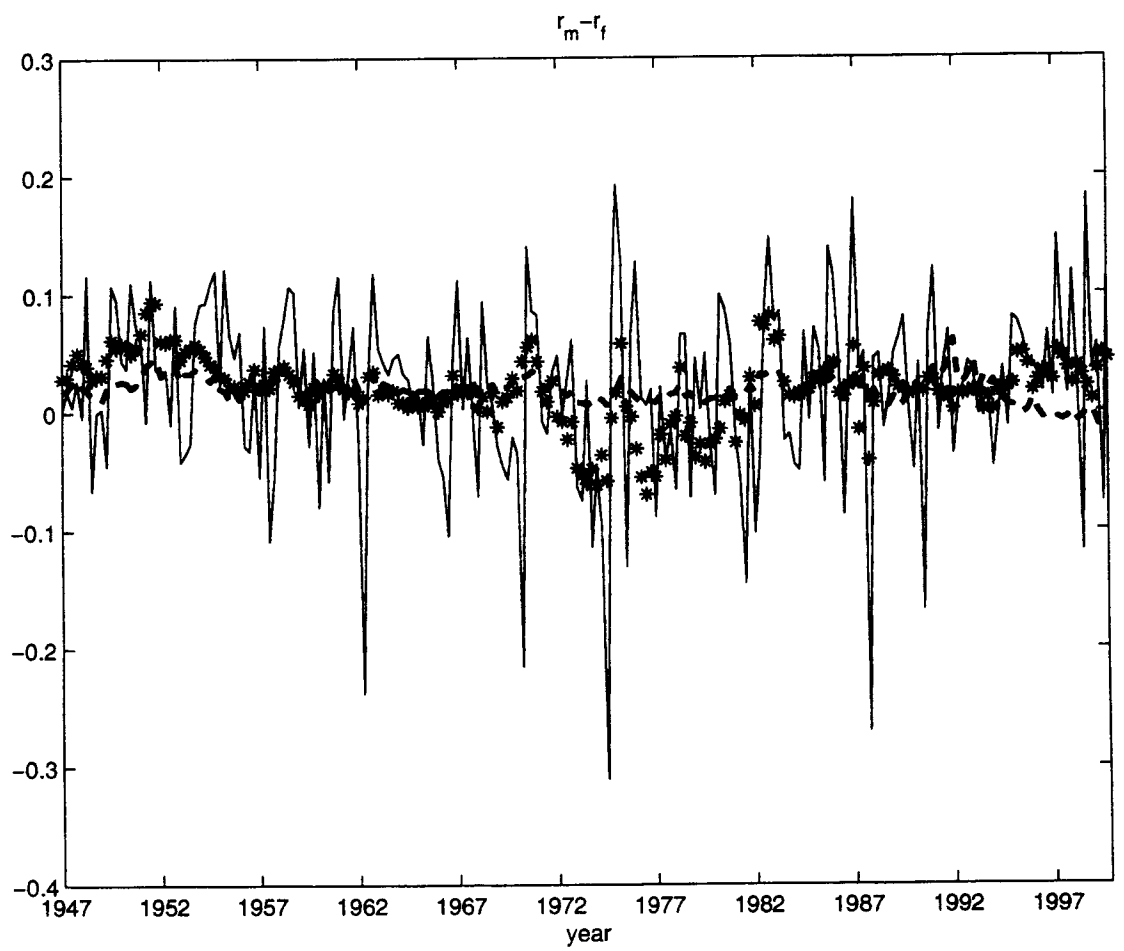


Figure 8: The Real Per Capita Stock Price, the Dividend Yield and the Dividend Payout Ratio

The figures plot the historical log real per capita stock price, $p - \varsigma$, the log dividend yield, $d - e$, and the dividend payout ratio, $d - e$.

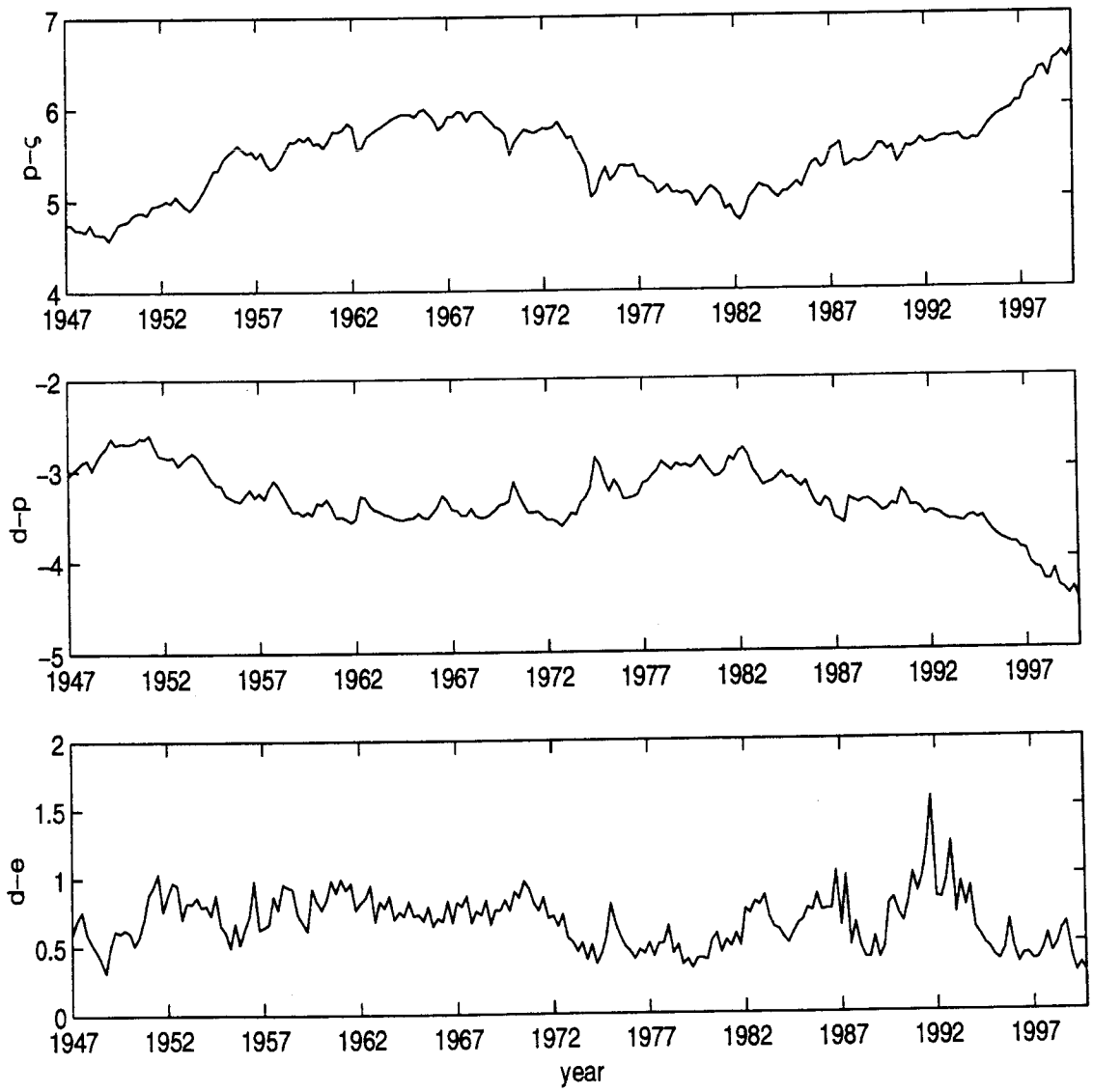


Table 1: Model Calibration

This table lists the estimates or the prespecified values of the parameters used in the numerical analysis.

	Parameter Descriptions	Notation	Parameter Values
	The Dividend Process		
1	Long-run Level of Dividends, D_t	$\bar{D}$	8.33
2	Volatility of Temporary Shocks to Dividends, $\epsilon_{D,t}$	σ_D	1.01
3	AR(1) Coefficient of the Dividend State Variable, F_t	a_F	0.74
4	Volatility of Temporary Shocks to the Dividend State Variable, $\epsilon_{F,t}$	σ_F	0.47
5	Long-run Level of Earnings, B_t	$\bar{B}$	17.36
6	Ratio of the Dividend State Variable, F_t, to the De-meaned Earnings, B_t	a_B	0.21
	The Labor Income Process		
7	AR(1) Coefficient of the Labor Income State Variable, Z_t	a_Z	0.74
8	Volatility of Temporary Shocks to the Labor Income State Variable, $\epsilon_{Z,t}$	σ_Z	1.00
	Correlations between the Dividend Process and the Labor Income Process		
9	Correlation Coefficients between $\epsilon_{D,t}$ and $\epsilon_{F,t}$	ρ_{DF}	-0.26
10	Correlation Coefficients between $\epsilon_{D,t}$ and $\epsilon_{Z,t}$	ρ_{DZ}	-0.04
11	Correlation Coefficients between $\epsilon_{F,t}$ and $\epsilon_{Z,t}$	ρ_{FZ}	-0.70
	Other Parameters		
12	Risk Aversion Coefficient	α	0.05
13	Subjective Discount Factor	β	1.2
14	Real Interest Rate	r	1.2%

Table 2: Summary Statistics 1947Q1-1999Q4

$r_m - r_f$ are quarterly log excess returns, calculated as total returns on the S&P composite Index minus total returns on T-bills. $p - \varsigma$ is the log real per capita stock price. $d - p$ is the log dividend yield. $d - e$ is the log dividend payout ratio. Log price, p, is the natural logarithm of the nominal S&P Composite Index. Log dividends, d, are the natural logarithm of the sum of the past four quarters of nominal dividends per share. Log earnings, e, are the natural logarithm of a single quarter's nominal earnings per share. Log sum of CPI and population, ς, is the sum of the natural logarithm of CPI and that of the previous year's population. The change in average age, ϑ, is the previous year's change in the natural logarithm of the population average age. The change in the share of population in the age range 40-64, ϕ, is the previous year's change in the natural logarithm of this share of population.

Correlation Matrix						
	$r_{m,t+1} - r_{f,t+1}$	$p_t - \varsigma_t$	$d_t - p_t$	$d_t - e_t$	ϑ_t	ϕ_t
$r_{m,t+1} - r_{f,t+1}$	1.000					
$p_t - \varsigma_t$	-0.092	1.000				
$d_t - p_t$	0.085	-0.932	1.000			
$d_t - e_t$	0.139	0.044	0.086	1.000		
ϑ_t	0.168	0.034	-0.125	0.103	1.000	
ϕ_t	0.167	0.441	-0.522	0.104	0.841	1.000
Univariate Summary Statistics						
Mean	0.019	5.462	-3.318	0.673	0.760^a	1.180^a
Standard deviation	0.075	0.431	0.359	0.201	2.471^a	10.910^a
Min	-0.311	4.569	-4.474	0.268	-4.362^a	-17.064^a
Max	0.193	6.624	-2.597	1.565	4.395^a	-22.006^a
Autocorrelation	0.079	0.984	0.976	0.725	0.908^b	0.939^b

[a] We multiply the value by 10^3.
[b] The value is with respect to annual observations.

Table 3: Quadri-variate Cointegration Tests 1947Q1-1999Q4

This table shows a first-order error-correction vector autoregression of the changes in log price, Δp, log dividends, Δd, log earnings, Δe, and log sum of CPI and population, $\Delta \varsigma$, on their own lags and the lagged level of the real per capita stock price, $p - \varsigma$, the dividend yield, $d - p$, and the dividend payout ratio, $d - e$. The Horvath-Watson statistic tests the alternative hypothesis that log price, p, log dividends, d, log earnings, e and log sum of CPI and population, ς, are cointegrated with unitary coefficients, against the null hypothesis of no cointegration. The test statistic is an exclusion test for the real per capita stock price $p - \varsigma$, the dividend yield, $d - p$, and the dividend payout ratio, $d - e$, in this vector autoregression. OLS standard errors are in parentheses below the coefficient estimates.

Dependent variable	Const	Δp_t	$\Delta \varsigma_t$	Δd_t	Δe_t	$p_t - \varsigma_t$	$d_t - p_t$	$d_t - e_t$	R^2
Δp_{t+1}	0.164	0.057	-0.801	-0.327	-0.099	-0.036	-0.022	-0.006	0.069
	(0.080)	(0.069)	(0.450)	(0.281)	(0.037)	(0.034)	(0.042)	(0.032)	
$\Delta \varsigma_{t+1}$	0.015	0.001	0.299	0.084	-0.000	0.007	0.011	-0.010	0.184
	(0.011)	(0.010)	(0.065)	(0.040)	(0.005)	(0.005)	(0.006)	(0.005)	
Δd_{t+1}	0.067	0.022	-0.023	0.373	-0.008	-0.011	-0.006	-0.028	0.303
	(0.017)	(0.015)	(0.099)	(0.062)	(0.008)	(0.008)	(0.009)	(0.007)	
Δe_{t+1}	-0.026	0.117	1.261	0.184	-0.324	-0.162	-0.234	0.203	0.214
	(0.142)	(0.123)	(0.804)	(0.502)	(0.066)	(0.061)	(0.076)	(0.057)	

Horvath-Watson test for cointegration: 61.135

Table 4: Predicting the Quarterly Equity Risk Premium by the Dividend Yield and/or the Dividend Payout Ratio

This table reports the results of the regressions that replicate Fama and French (1988) and Lamont (1998), in which the equity risk premium is regressed on the dividend yield and/or the dividend payout ratio, for different periods. The predictive regression is

$$r_{m,t+1} - r_{f,t+1} = \beta_0 + \beta_1(d_t - p_t) + \beta_2(d_t - e_t) + \epsilon_{t+1}$$

where $r_m - r_f$ is the quarterly log excess return on the S&P Composite Index; $d - p$ is the log dividend yield; $d - e$ is the log dividend payout ratio. Standard errors, calculated using the method outlined in Newey and West (1987) with a lag length of 4, are in parentheses below the coefficient estimates. The reported R^2 is the adjusted R^2 statistic. D-W is the Durbin-Watson statistic for the error term. Obs is the number of observations. F is the F statistic.

No.	Const	$d-p$	$d-e$	R^2	D-W	F
			Panel A: 1947Q1-1994Q4 Obs=192			
1	0.222	0.064		0.046	1.766	10.196
	(0.054)	(0.017)				
2	-0.042		0.083	0.039	1.802	8.778
	(0.025)		(0.033)			
3	0.207	0.083	0.112	0.117	1.805	13.599
	(0.057)	(0.020)	(0.035)			
			Panel B: 1947Q1-1999Q4 Obs=212			
4	0.080	0.018		0.002	1.812	1.513
	(0.051)	(0.016)				
5	-0.016		0.053	0.015	1.821	4.131
	(0.022)		(0.029)			
6	0.038	0.016	0.050	0.015	1.804	2.639
	(0.056)	(0.017)	(0.029)			

Table 5: Predicting the Quarterly and Annual Equity Risk Premium by Demographic Variables 1947-1999

This table reports the results of the regressions that replicate Bakshi and Chen (1994), in which the equity risk premium is regressed on the change in average age, and the regressions that run the equity risk premium on the change in the share of population in the age range 40-64. The predictive regression is

$$r_{m,t+1} - r_{f,t+1} = \beta_0 + \beta_1(\text{Demographic Variables})_t + \epsilon_{t+1}$$

where $r_m - r_f$ is the quarterly or annual log excess return on the S&P Composite Index. In calculating the annual equity risk premium, over-lapping data are used. ϑ is the previous year's change in the natural logarithm of the population average age. ϕ is the previous year's change in the log share of population in the age range 40-64. Standard errors, calculated using the method outlined in Newey and West (1987) with a lag length of 4, are in parentheses below the coefficient estimates. The reported R^2 is the adjusted R^2 statistic. D-W is the Durbin-Watson statistic for the error term. Obs is the number of observations. F is the F statistic.

No.	Return Horizons	Const	Demographic Variables	R^2	D-W	Obs	F
		Panel A: $r_{t+1} = \beta_0 + \beta_1\vartheta_t + \epsilon_{t+1}$					
1	Annual	0.062	21.533	0.113	0.552	209	27.478
		(0.018)	(7.275)				
2	Quarterly	0.016	5.121	0.024	1.892	212	6.090
		(0.005)	(2.137)				
		Panel B: $r_{t+1} = \beta_0 + \beta_1\phi_t + \epsilon_{t+1}$					
3	Annual	0.074	4.577	0.094	0.533	209	22.634
		(0.017)	(1.532)				
4	Quarterly	0.018	1.160	0.023	1.886	212	6.012
		(0.005)	(0.463)				

Table 6: Predicting the Quarterly Equity Risk Premium by Using the Change in the Share of Population in the Age Range 40-64 to Proxy Demographic Change

This table reports the results of the regressions as specified in equation (5.3). We use linear functions of the change in the share of population in the age range 40-64 to proxy the time varying coefficients. The predictive regression is

$$\begin{aligned} r_{m,t+1} - r_{f,t+1} &= \beta_0 + \beta_1\phi_t + \beta_2(p_t - \varsigma_t) + \beta_3\phi_t(p_t - \varsigma_t) \\ &\quad + \beta_4(d_t - p_t) + \beta_5\phi_t(d_t - p_t) + \beta_6(d_t - e_t) + \beta_7\phi_t(d_t - e_t) + \epsilon_{t+1} \end{aligned}$$

where $r_m - r_f$ is the quarterly log excess return on the S&P Composite Index; $p - \varsigma$ is the log real per capita stock price; $d - p$ is the log dividend yield; $d - e$ is the log dividend payout ratio; ϕ is the previous year's change in the log share of population in the age range 40-64. Standard errors, calculated using the method outlined in Newey and West (1987) with a lag length of 4, are in parentheses below the coefficient estimates. The reported R^2 is the adjusted R^2 statistic. D-W is the Durbin-Watson statistic for the error term. Obs is the number of observations. F is the F statistic.

No.	Const	ϕ_t	$p_t - \varsigma_t$	$\phi_t(p_t - \varsigma_t)$	$d_t - p_t$	$\phi_t(d_t - p_t)$	$d_t - e_t$	$\phi_t(d_t - e_t)$	R^2	D-W	F
					Panel A: 1947Q1-1994Q4 Obs=192						
1	0.217	-9.182	0.000	1.355	0.086	-2.373	0.117	-8.390	0.137	1.823	5.342
	(0.066)	(15.741)	(0.028)	(6.633)	(0.043)	(7.638)	(0.035)	(3.317)			
2	0.217	-9.182		1.354	0.086	-2.374	0.117	-8.390	0.142	1.823	6.266
	(0.060)	(15.801)		(6.527)	(0.020)	(7.407)	(0.036)	(3.321)			
3	0.219	-6.799			0.087	-3.860	0.118	-8.322	0.146	1.820	7.548
	(0.061)	(8.711)			(0.020)	(2.942)	(0.035)	(3.422)			
4	0.223				0.087	-1.737	0.114	-8.096	0.147	1.825	9.241
	(0.063)				(0.020)	(0.674)	(0.034)	(3.351)			
					Panel B: 1947Q1-1999Q4 Obs=212						
5	0.223	-8.067	-0.002	0.439	0.083	-3.638	0.109	-8.432	0.133	1.905	5.644
	(0.068)	(13.267)	(0.028)	(6.502)	(0.044)	(7.175)	(0.036)	(2.660)			
6	0.221	-8.084		0.422	0.085	-3.671	0.108	-8.425	0.138	1.904	6.616
	(0.061)	(13.326)		(6.394)	(0.021)	(6.964)	(0.036)	(2.668)			
7	0.221	-7.327			0.085	-4.128	0.109	-8.398	0.142	1.903	7.977
	(0.062)	(4.872)			(0.021)	(1.444)	(0.035)	(2.795)			
8	0.202				0.076	-2.215	0.096	-9.731	0.136	1.911	9.338
	(0.056)				(0.018)	(0.518)	(0.035)	(2.802)			

Table 7: Predicting the Long-run Equity Risk Premium by the Dividend Yield, the Dividend Payout Ratio and the Change in the Share of Population in the Age Range 40-64. 1947-1999

This table reports the regression results of predicting the long-run equity risk premium. We uses linear functions of the change in the share of population in the age range 40-64 to proxy demographic change. The predictive regression is

$$r_{m,t+1} - r_{f,t+1} = \beta_0 + \beta_1(d_t - p_t) + \beta_2\phi_t(d_t - p_t) + \beta_3(d_t - e_t) + \beta_4\phi_t(d_t - e_t) + \epsilon_{t+1}$$

where $r_{m,t+1} - r_{f,t+1}$ is the log excess return of 1 up to 4 quarters on the S&P Composite Index; $d - p$ is the log dividend yield; $d - e$ is the log dividend payout ratio; ϕ is the previous year's change in the log share of population in the age range 40-64. In calculating the long-run equity risk premium, overlapping data are used. Standard errors, calculated using the method outlined in Newey and West (1987) with a lag length of 4, are in parentheses below the coefficient estimates. The reported R^2 is the adjusted R^2 statistic. D-W is the Durbin-Watson statistic for the error term. Obs is the number of observations. F is the F statistic.

No.	Return Horizons	Const	$d_t - p_t$	$\phi_t(d_t - p_t)$	$d_t - e_t$	$\phi_t(d_t - e_t)$	R^2	D-W	Obs	F
1	1 Quarter	0.202	0.076	-2.215	0.096	-9.731	0.136	1.911	212	9.338
		(0.056)	(0.018)	(0.518)	(0.035)	(2.802)				
2	2 Quarter	0.458	0.151	-3.651	0.113	-14.489	0.204	1.134	211	14.447
		(0.106)	(0.032)	(0.683)	(0.046)	(3.471)				
3	3 Quarter	0.689	0.221	-5.386	0.136	-21.118	0.301	0.743	210	23.489
		(0.140)	(0.042)	(0.879)	(0.061)	(4.469)				
4	4 Quarter	0.918	0.280	-6.332	0.110	-23.064	0.358	0.631	219	29.966
		(0.167)	(0.051)	(1.004)	(0.067)	(5.259)				

THE ANDERSON SCHOOL AT UCLA

An electronic version of this paper can be found on The Anderson School at UCLA's Finance Working Papers site at the following url:

http://www.anderson.ucla.edu/acad_unit/finance/workpaper2.htm

Check this page for related papers that might be of interest.

All 1999 & 2000 working papers are now available there as pdf files.

THE JOHN E. ANDERSON
GRADUATE
SCHOOL OF MANAGEMENT
AT UCLA

110 WESTWOOD PLAZA
BOX 951481
LOS ANGELES, CA 90095-1481
WEB: http://www.anderson.ucla.edu/