

UCLA

Recent Work

Title

Demographics and the Equity Premium

Permalink

<https://escholarship.org/uc/item/0q95p4n0>

Author

Luo, Jiang

Publication Date

2000-12-15

Preliminary and Comments Welcome

Demographics and the Equity Premium*

Jiang Luo[†]

University of California, Los Angeles

December 15, 2000

Abstract

This paper studies the relation between demographics and the equity 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 their labor income. The per capita stock supply is normalized to unity, and the demographic structure is time varying. In equilibrium, the equity 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 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. My interest in this topic was stimulated by reading an unpublished paper by Amit Goyal on demographics and stock market flows. Helpful comments and suggestions were also received from Antonio Bernardo, Amit Goyal, John Riley, Avaniidhar Subrahmanyam and seminar participants at UCLA. 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 premium. The stock market has had a strong upwards movement, while the dividend yield and the dividend payout ratio have been relatively low (See figure 9). These two variables, which are significant in predicting the equity 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. It is possible that this is just a coincidence. But if not, it is important to understand how stock price movements are related to variation in the demographic structure, and to consider whether it is demographic change that has caused the simple predictive relations between the dividend yield and/or the dividend payout ratio and the equity 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 premium can be easily derived from life cycle portfolio theories, which were pioneered by Modigliani and Brumberg (1954), Friedman (1957) and Samuelson (1969). These models assume that the objective of an investor's consumption-investment decision is to smooth consumption over time in order to maximize lifetime utility, and this induces a life cycle pattern in investment behavior. Suppose that the supply of risky assets is constant or inelastic. If an investor's demand for risky assets is affected by his age, one might expect as a first approximation that when the population is dominated by investors of the age at which the demand for risky assets is highest, the stock price would rise to ensure market clearing causing the equity premium to decline.¹ In addition, when the population is dominated by investors of the age at which the demand for risky assets is the most sensitive to certain macroeconomic factors, the stock price will also become sensitive to these factors. However, the situation is complicated by the fact that the demands of today's investors are affected by their forecasts

¹This 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.

of future investors, which will depend on the future demographic structure as well as future expectation about future demand.

In this paper, we develop a dynamic overlapping generations (OLG) equilibrium model, which allows us to study the relation between demographics, stock prices and the equity premium. We assume that the per capita stock supply is constant, and that the demographic structure is deterministic and exogenous but time varying. We are primarily interested in the information for stock returns that is contained in the demand side of the economy, which is represented by the demographic structure.

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 both by the expected returns on stocks and by their ability 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.² 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 is also time varying so that 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.

The equilibrium price function implies a non-linear relation between the equity premium and demographic variables, the real per capita stock price, the dividend yield and the dividend payout ratio. The equity 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 in a nonlinear fashion on the demographic structure, which is time varying. These time varying coefficients are functions of moments of the current and all the future

²Samuelson (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.

demographic structures.

Calibration results show that the observed demographic change in the post World War II period could induce significant variation in the demographic variables in the predictive relation. In addition, these demographic variables are more highly correlated with the change in the share of population in the age range 40-64 than with the level. Intuitively, as shown in the estimates of the labor income process, households in this age range have the highest labor income, and their labor income is the least sensitive to macroeconomic risk, as measured by the unemployment rate. An increase in the share of population in this age range reduces the labor income risk in the whole economy, tending to reduce the importance of stocks as a hedging vehicle. However, since households in this age range have the highest level of idiosyncratic labor income risk, and their risk aversion is relatively high, the increase in this share of population makes the population more risk averse, which tends to make the role of stocks as a hedging vehicle more important. Previous research on stock returns, (for example, Poterba (2000)), uses this share as an explanatory variable in predictive regressions, although Poterba focuses on the predictive power of this share for real stock returns, instead of the equity premium, and finds no significant results.

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 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 premium, whether or not the late 1990s data are included. The inclusion of demographic variables significantly improves on previous predictive regressions, in which the equity excess return 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 excess return 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 Lamont predictive regressions are no longer significant, while our predictive regression is significant with R^2 of 13.6%.

Thus, our model salvages the predictive power of the dividend yield and the dividend payout ratio, but suggests a changing structure for the predictive relation due to demographic change.

Prior research on the predictability of the equity premium either ignores demographic considerations or, as in Poterba (2000), focuses solely on the effects of demographic change on stock returns. These two strands of research on the equity premium have remained totally distinct despite 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.³ The relation between the equity 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. Several explanations have been offered for this failure. Most focus on the statistical biases of empirical tests.⁴

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 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

³See Cochrane (1999) for survey.

⁴For example, Goetzmann 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) and Torous and Yan (1999) on bias arising from near-nonstationarity in regressors; and etc.

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 premium investigates the relation between demographics and stock returns.⁵ 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,⁶ 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 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.

⁵See Poterba (2000) for survey.

⁶For 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 (2000).

Our model is different from Bergantino's 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, while Bergantino simply takes investors' life cycle pattern of investment behavior as exogenous.

Goyal (1999) also adopts a partial equilibrium approach. Like Bergantino (1998), he takes the price as exogeneous, and relates capital flow to change in demographic structure; he also examines the relation between demographic change and the equity premium but not in a context of a rational expectation economy.

Although it is commonly agreed that demographic factors affect stock returns,⁷ 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 excess return, 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 smooth. Not surprisingly, regressions of short-term, for example, quarterly, equity excess returns on demographic variables have low R^2 s. 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 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. They then conclude that demographic change should be correlated with stock prices and returns. In other 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 premium

⁷See Bakshi and Chen (1994), Macunovich (1997), Erb, Harvey and Viskanta (1997), Bergantino (1998), Poterba (2000), Goyal (1999) and etc.

within an OLG general equilibrium model. The model accounts for the information for stock returns that is contained in stock prices, dividends and earnings on the one hand, and for the information in demographics on the other hand, and provides a clear description of the determination of the equity premium. The non-linear predictive relation between the equity 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 calibrates a simplified version of the model; Section 5 derives the non-linear predictive regression for the equity premium and conducts the empirical test; Section 6 concludes.

2 The Model

Consider an 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 , $\tau = 0, 1, \dots, T$, denote the age of an investor who is assumed to live for $T + 1$ periods. For simplicity, T is assumed to be a constant.⁸ The number of investors alive at time t is denoted by $G(t)$. 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 fully anticipate the effects of demographic change on the economy when making their consumption-investment decisions.

⁸We ignore uncertainty about life-span, which has been studied by Hubbard, Skinner and Zeldes (1994). Incorporating uncertain life-span in the model may introduce second order effects, but is unlikely to change our main results.

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

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

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.⁹

2.3 Endowments

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

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

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

⁹We ignore bequests in the model. Including bequests would have 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.

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.¹⁰ The sign of Z may not affect the labor income process because of the quadratic form, but will affect the stock price and the expected excess return because its correlation with the dividend process, which is defined later, makes its effects on the economy asymmetric. 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.¹¹

Z_t is assumed to follow an AR(1) process

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

where a_Z is an $N \times N$ constant matrix whose eigenvalues are all 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).

In addition to his endowment of non-tradeable labor income, each investor is endowed at birth with one share of risky asset whose characteristics are described in the following section.

2.4 Financial assets

There are two publicly traded assets in the economy, a riskless asset and a risky asset (stock). The riskless asset, which is assumed to be in infinitely elastic supply, offers a positive constant

¹⁰The non-trivial quadratic form of $\frac{1}{2}Z_t^T \omega_\tau Z_t$ is essential to our model. It captures the fact that the risk of labor income is itself stochastic. Without this feature, the stock price would not depend on the labor income state variable vector, Z , and the equity premium would depend only on the real per capita stock price.

¹¹See Gollier and Pratt (1996) and Franke, Stapleton and Subrahmanyam (1998) for the effect of this idiosyncratic risk of labor income on investors' investment behavior.

rate of return, r . Its gross rate of return is thus $R = 1 + r$. Given the endowment process, the per capita supply of stocks is one so that the number of shares outstanding at time t is $G(t)$.¹² Each share of stocks pays a dividend D_t at time t , which is governed by the process

$$D_t = \bar{D} + F_t + \epsilon_{D,t} \quad (2.4)$$

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

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

where a_F is a constant with $0 \leq a_F < 1$, and $\epsilon_{F,t}$ is a temporary shock to F_t which 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.

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 serially uncorrelated and independent of the other random variables, $\epsilon_{D,t}$, $\epsilon_{F,t}$ and $\epsilon_{Z,t}$. Define the vector of shocks at time t , $\xi_t \equiv (\epsilon_{D,t}, \epsilon_{F,t}, \epsilon_{Z,t})^T$. 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 their labor income.

¹²This normalization is for simplicity. Allowing for a stochastic stock supply would introduce another state variable, which would affect the stock price and the excess return. But the predictive relation for the equity premium, which is derived later, would still hold.

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 about dividends and labor income is redundant for investors' consumption-investment decisions. Therefore, investors have identical sufficient information sets $\mathfrak{S}_t \equiv \{D_t, F_t, Z_t\}$.

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 tractable.

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

$$J(W_t, \mathfrak{S}_t, \tau; t) \equiv \max_{c_t, X_t} E_t \left[- \sum_{s=t}^{\infty} \beta^{\tau+s-t} \exp(-\alpha c_s) \right] \quad (3.1)$$

subject to

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

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 excess rate of return, which is defined as the excess return on one dollar invested in stocks. The share return must be divided by the share price to get the rate of return.

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

Before analyzing the complete model with a changing demographics, we assume first 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' ages. 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 is same for all periods in an infinite horizon setting.¹³

While the constant demographic structure precludes 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, as in the cross-national study of Erb, Harvey and Viskanta (1999), a comparative analysis of the stationary case is informative. Secondly, intuition about investors' life cycle pattern of 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 for the non-stationary model for which the equilibrium is computed by the backward induction in the numerical example in Section 4.

¹³A constant demographic structure is only a convenient abstraction. In general, only certain age structures, such as the uniform, remain constant over time.

Consider first a further simplified economy in which labor income is deterministic, i.e. $\{\omega_\tau\}_{\tau=0}^T \equiv 0_{N \times N}$ and $\{\sigma_\tau\}_{\tau=0}^T \equiv 0$. The derivation of the equilibrium for this economy is similar to that of the general case, the results of which are stated in proposition 3.2 and theorem 3.3. Here, we just give the equilibrium price function.

Proposition 3.1 When the demographic structure is constant over time, i.e. $g(\tau; t) \equiv g(\tau)$, and labor income is deterministic, i.e. $\{\omega_\tau\}_{\tau=0}^T \equiv 0_{N \times N}$ and $\{\sigma_\tau\}_{\tau=0}^T \equiv 0$, the economy defined in Section 2 has a steady-state equilibrium in which the equilibrium price function is

$$P_t = p_0^* + p_F^* F_t \quad (3.3)$$

where p_0^* is a constant, and $p_F^* = \frac{a_F}{R - a_F}$.

The first term of the price function, p_0^* , can be divided into two parts: the discounted value of the constant level of dividends, $\frac{\bar{D}}{r}$, and the discount in the price to compensate for the risk of dividends, $p_0^* - \frac{\bar{D}}{r}$. The second term, $p_F^* F_t$, is the discounted value of the stochastic part of dividends. As new information about dividends, F , arrives, the stock price adjusts to fully reflect it.

In this economy, there is no labor income risk. Investors trade stocks only for consumption purposes. Investors buy stocks when born and slowly sell them until they die.¹⁴

When there is systemic labor income risk, i.e. $\{\omega_\tau\}_{\tau=0}^T$ are not equivalent to $0_{N \times N}$, investors trade stocks both for consumption purposes and to hedge against the risk of their labor income. The equilibrium price function has to adjust to reflect the role of stocks as a hedging vehicle.

Proposition 3.2 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

$$P_t = p_0 + p_F F_t + p_Z Z_t \quad (3.4)$$

¹⁴See Vayanos (1998) for a similar set up.

where p_0 is a constant; $p_F = \frac{a_F}{R - a_F}$; 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. Besides the discounted value of the constant level of dividends, $\frac{\bar{D}}{r}$, and the discount in the price to compensate for the risk of dividends, $p_0^* - \frac{\bar{D}}{r}$, which are obtained from the economy without labor income risk, it also includes the discount in the price to compensate for the risk of labor income, $p_0 - p_0^*$. The second term, $p_F F_t$, is the discounted value of the stochastic part of dividends. As new information about 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 about labor income, Z , arrives, investors rebalance their portfolios 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 about the future dividend, F , but also on the information about the background labor income risk, Z .

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

$$Q_{t+1} \equiv P_{t+1} + D_{t+1} - RP_t = \Theta \Pi_t + \Phi \xi_{t+1} \quad (3.5)$$

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 . The expected excess return depends only on the labor income variables, Z . When new information about the fundamental, F , arrives, the stock price adjusts to fully reflect it, and the expected excess share return stays the same. However, when new information about labor income, Z , arrives, the stock price and 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 3.3 When the demographic structure is constant over time, problem (3.1)-(3.2) has the following solution

$$J(W_t, \mathfrak{S}_t, \tau; t) = -\beta^\tau \exp(-\gamma_\tau W_t - \mu_\tau t - \frac{1}{2} \Pi_t^T \nu_\tau \Pi_t) \quad (3.6)$$

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

$$X_t = \frac{1}{\gamma_{\tau+1}} \Gamma_{\tau+1} E_t(Q_{t+1}) - \frac{1}{\gamma_{\tau+1}} \kappa_{\tau+1} \Pi_t \quad (3.7)$$

$$c_t = \bar{c}_\tau + \frac{R\gamma_{\tau+1}}{\alpha + R\gamma_{\tau+1}} W_t + \frac{\gamma_{\tau+1} n_{\tau+1} + \mu_{\tau+1}}{\alpha + R\gamma_{\tau+1}} t + \frac{1}{2(\alpha + R\gamma_{\tau+1})} \Pi_t^T m_{\tau+1} \Pi_t \quad (3.8)$$

where $\Gamma_{\tau+1}$ and \bar{c}_τ 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 at time t have higher utility than the investors of the same age at earlier dates. Note that Π is obtained by augmenting Z with 1. Investors trade stocks both for consumption purposes and to hedge against the risk of their labor income. The incentive to trade is much clearer if one considers investors' demand for stocks, X . Similar to Wang (1994), X consists of two components. The first component, $\frac{1}{\gamma_{\tau+1}} \Gamma_{\tau+1} E_t(Q_{t+1})$, is a mean-variance efficient portfolio reflecting the trade off between expected return and risk. Like Vayanos (1998), $\gamma_{\tau+1}$ is the coefficient of absolute risk aversion of investors with respect of their wealth at the age of $\tau + 1$. The term, $\Gamma_{\tau+1}$,

is the inverse of the renormalized covariance matrix of returns for investors of age τ . The second component 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.

Since the coefficients of the value functions and consumption-investment policies are age specific, investment behavior exhibits a life cycle pattern. It is difficult to tell how the demand of an investor of certain age responds to the labor income state variable vector, Z , since the response depends on Γ_τ and κ_τ , and we have no closed form expression for these coefficients. However, the risk aversion of the value function, γ_τ , provides an informative way to study investors' average holdings of stocks. γ_τ , which is positive from its expression derived in Appendix A, 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.¹⁵

[Insert Figure 1 Here]

From equation (A.9) in Appendix A, we have $\gamma_\tau = \frac{\alpha R \gamma_{\tau+1}}{\alpha + R \gamma_{\tau+1}}$. 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 represent calendar age. We have tried other parameter values, and 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 are willing to hold more risky assets than old investors. 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 reflects the fact that the same calendar change shortens an old investor's investment horizon proportionately much more than it does a young investor's investment horizon.

¹⁵From their expressions in Appendix A, Γ_τ and κ_τ are just rearrangements of the coefficients of the mean and the variance of the excess return, Θ and Φ . They should have second order effects on investors' average holdings compared to γ_τ .

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' ages, 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 3.4.

Proposition 3.4 When the demographic structure is time varying, the economy defined in Section 2 has an equilibrium in which the equilibrium price function is

$$P_t = p_{0,t} + p_F F_t + p_{Z,t} Z_t \quad (3.9)$$

where $p_{0,t}$ is a function of t ; $p_F = \frac{a_F}{R - a_F}$; $p_{Z,t}$ 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 $g(\tau; 0) \neq g(\tau; 1)$, 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

$$Q_{t+1} \equiv P_{t+1} + D_{t+1} - RP_t = \Theta_{t+1}\Pi_t + \Phi_{t+1}\xi_{t+1} \quad (3.10)$$

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 now 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 the labor income process, we can derive the investors' optimal consumption-investment policy.

Theorem 3.5 When the demographic structure is time varying, problem (3.1)-(3.2) has the following solution

$$J(W_t, \mathfrak{S}_t, \tau; t) = -\beta^\tau \exp(-\gamma_\tau W_t - \mu_\tau t - \frac{1}{2}\Pi_t^T \nu_{\tau,t} \Pi_t) \quad (3.11)$$

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

$$X_t = \frac{1}{\gamma_{\tau+1}} \Gamma_{\tau+1,t+1} E_t(Q_{t+1}) - \frac{1}{\gamma_{\tau+1}} \kappa_{\tau+1,t+1} \Pi_t \quad (3.12)$$

$$c_t = \bar{c}_{\tau,t} + \frac{R\gamma_{\tau+1}}{\alpha + R\gamma_{\tau+1}} W_t + \frac{\gamma_{\tau+1}n_{\tau+1} + \mu_{\tau+1}}{\alpha + R\gamma_{\tau+1}} t + \frac{1}{2(\alpha + R\gamma_{\tau+1})} \Pi_t^T m_{\tau+1,t+1} \Pi_t \quad (3.13)$$

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 expectations property of our model is in contrast to previous work on the relation between demographics and stock returns, which relies on partial equilibrium models of life cycle portfolio selection.¹⁶

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. By assuming a time varying but deterministic demographic structure, we are able to capture demographic state variables in a single variable, time.

3.3 Derivation of the Predictive Relation for the Equity Risk Premium

The equity premium is obtained by dividing the expected excess share return by the stock price. From equation (3.10), the equity premium, $R_{m,t+1} - R$, can be written as

$$R_{m,t+1} - R = \frac{\Theta_{t+1}\Pi_t}{P_t} = \theta_{0,t+1}\frac{1}{P_t} + \theta_{Z,t+1}\frac{Z_t}{P_t} \quad (3.14)$$

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

The econometrician cannot directly observe the labor income state variable vector, Z . It must be inferred from the available information, such as the stock price, P , and the dividends, D . From the econometrician' point of view, the dividend state variable, F , can

¹⁶See Bergantino (1998).

be interpreted empirically as a value proportional to the demeaned earnings plus a noise,¹⁷ i.e.

$$F_t = a_B(B_t - \bar{B}) + \epsilon_{B,t} \quad (3.15)$$

where B_t is the earnings with mean \bar{B} ; a_B and \bar{B} are constants; $\epsilon_{B,t}$ is a temporary shock, which has an independent and identical normal distribution, $\epsilon_{B,t} \sim N(0, \sigma_B^2)$. $\epsilon_{B,t}$ can be correlated with other innovations to the economy, ξ_t . From the normality assumption of the model, the labor income state variable vector, Z , can be expressed as

$$Z_t = \eta_{0,t} + \eta_{1,t}P_t + \eta_{2,t}D_t + \eta_{3,t}B_t + v_t \quad (3.16)$$

where $\eta_{0,t}$, $\eta_{1,t}$, $\eta_{2,t}$ and $\eta_{3,t}$ are $N \times 1$ matrices dependent on t . v_t is the unexpected component of Z conditional on P , D and B . 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 due to demographic change, and thereafter the correlation matrix of Z , P , D and B is time varying.

Substituting the expression for Z back into equation (3.14), we have the predictive relation for the equity premium

$$R_{m,t+1} - R = \lambda_{0,t} + \lambda_{1,t} \frac{1}{P_t} + \lambda_{2,t} \frac{D_t}{P_t} + \lambda_{3,t} \frac{B_t}{P_t} + \varepsilon_{t+1} \quad (3.17)$$

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 $\varepsilon_t = \frac{\theta_{Z,t+1}v_t}{P_t}$. Equation (3.17) is the one of the main results of this paper. In equilibrium, the equity 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

¹⁷As Lintner (1956) documents, most managers have a target level of dividends equal to a fraction of the earnings. A immediate implication is that the de-meaned dividends will be a fraction of the de-meaned earnings, as we assumed here.

change. These coefficients are highly non-linear functions of all the moments of the current and the future demographic structures. Without further assumptions, we cannot determine their signs and magnitude.

In what follows, we try different strategies to study this predictive relation. In Section 4, we calibrate a simplified version of the model, and obtain the dynamics of the coefficients of the predictive relation. We are primarily interested in whether demographic change can induce significant variation 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 with the results of previous empirical research.

4 Calibration

In this section, we calibrate a simplified version of the model. We want to show that demographic change can induce variation in the coefficients of the linear predictive relation so that simple linear predictive regressions that ignore the structural change may produce misleading results. 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 for 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 population. Members of

this generation 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 and the population average age, and their changes. The share of population in the age range 40-64 is the highest in the middle 1960s, then decreases and obtain its lowest level in 1986, and increases in the 1990s. The population age profile has a 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 $Y_{\tau,t}$.¹⁸ We deflated $Y_{\tau,t}$ by the annual average CPI to obtain real income.

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 polynomials to approximate these estimates. We will use the cubic polynomial approximations as the true parameters in the numerical solution. We regress Z_t on its lagged values to obtain the estimates of $a_Z = 0.74$ and $\sigma_Z = 1.00$.

From this rough estimation of the labor income process, we can see that households in the age range 40-64 have significantly different labor income from others. They have the highest

¹⁸The 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.

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 . The absolute value of the polynomial approximation for ω_τ of investors in the age range 40-64 is about half of that for investors of age 20. 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 age groups.

Prices, P , dividends per share, D , and earnings per share, B , which correspond to the S&P Composite Index, were obtained from the Security Price Index Record Published by Standard & Poor's Statistical Service. We deflated them by CPI and population to obtain the real per capita items.¹⁹ We use the long run average of dividends as $\bar{D} = 8.54$ and the long run average of earnings as $\bar{B} = 17.45$. Note that, after substituting the expression for F into equation (2.2), we have

$$D_t - \bar{D} = a_B(B_t - \bar{B}) + \epsilon_{B,t} + \epsilon_{D,t} \quad (4.1)$$

We regress the de-meaned dividends, $D_t - \bar{D}$, on the demeaned earnings, $B_t - \bar{B}$, to obtain the estimates of $a_B = 0.21$ and $\sigma_B^2 + \sigma_D^2 = 1.00$. Denote $\hat{F}_t = a_B(B_t - \bar{B})$. Note

$$\hat{F}_t = a_F \hat{F}_{t-1} + a_F \epsilon_{B,t-1} + \epsilon_{F,t} - \epsilon_{B,t} \quad (4.2)$$

After regressing \hat{F}_t on a constant and its lagged values, we have the estimates of $a_F = 0.74$ and $(1 + a_F^2)\sigma_B^2 + \sigma_F^2 = 0.27$.

For the correlation between the innovations to 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 $\rho_{FZ} = -0.8$. We assume that the innovations to

¹⁹We use the annual average stock price as that year's stock price, the annual dividends as that year's dividends, and the annual earnings as that year's earnings. Then, we divide them by the annual average of CPI and that year's population to obtain the real per capita items.

earnings, $\epsilon_{B,t}$, are independent of $\epsilon_{D,t}$, $\epsilon_{F,t}$ and $\epsilon_{Z,t}$, and that their volatility is $\sigma_B = 0.32$. Thus, from the estimation results of equation (4.1)-(4.2), we have $\sigma_D = 0.95$ and $\sigma_F = 0.33$. For other parameters, we let the riskfree interest rate $r = 1.2\%$. The preference parameters are set $\alpha = 0.08$ and $\beta = 1.5$.²⁰

[Insert Table 1 Here]

Table 1 summarizes the estimates or the prespecified values of the parameters. In this section, we are primarily interested in whether demographic change can induce significant variation 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, $p_F = 2.79$. 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 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.

[Insert Figure 6 Here]

Figure 6 plots the resulting 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

²⁰A number of capital market studies have obtain estimates of the coefficient of relative risk aversion ranging from 2 to 16. We use 16 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.

The discounted factor, β , is set to be greater than 1 because utility is negative.

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 .

[Insert Figure 7 Here]

Figure 7 plots the time varying coefficients of the predictive relation specified in equation (3.17). The signs of these coefficients are determined by the expectation of Z conditional on the stock price, P , dividends, D , and earnings, B . In equilibrium, $\lambda_{1,t}$ and $\lambda_{2,t}$ are positive, and $\lambda_{3,t}$ is negative, which is consistent to the empirical findings that the equity premium decreases with the stock price, and increases with the dividend yield (Lamont (1998) and Fama (1988)). Lamont also finds that conditional on the dividend yield, the earnings yield is negatively correlated with the future return.

Demographic change causes significant variation in these coefficients. For example, in the dividend yield case, the difference between the highest and the lowest level of $\lambda_{2,t}$ is about 6% of its average level. Therefore, if one regresses the excess return 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 premium, and a negative result can not rule it out, either.

[Insert Figure 8 Here]

The pattern of $\lambda_{0,t}$, $\lambda_{1,t}$, $\lambda_{2,t}$ 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 this period, the stock price was extraordinarily high and dividends were quite low (See figure 9). The traditional predictive models, for example, Fama and French (1988) and Lamont (1998), predict a low equity premium, while the actual equity excess return 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 absolute values of $\lambda_{1,t}$, $\lambda_{2,t}$ and $\lambda_{3,t}$ are historically low, which means that the equity premium is not sensitive to the stock price, the dividend yield and the earnings yield. The high level of the stock price and the low level of the dividend yield do not mean that the equity 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 premium will be higher.

Figure 8 plots the fitted value of the predictive relation where the stock price, the dividend yield and the earnings yield are obtained from the historical data and the coefficients are derived under the calibrated parameter values. The predictive relation forecasts significant variation in the equity premium. The highest level of the equity premium, 6.39%, is obtained in 1954, and the lowest, 0.85%, is obtained in 1977. In particular, in the 1990s, the predictive relation forecasts a high equity premium. The average of the fitted equity premium is 2.9%, while the historical average of realized excess returns for the period 1947-1999 is 7.8%. We are not trying to fit the model to this value.²¹ The calibration is simplified so that we consider only one source of labor income risk, which is measured by the unemployment rate, and the calibration is imprecise.

[Insert Table 2 Here]

Table 2 reports the correlations between moments of the demographic structure and the calibrated coefficients, $\lambda_{0,t}$, $\lambda_{1,t}$, $\lambda_{2,t}$ and $\lambda_{3,t}$. In order to minimize the effects of arbitrarily choosing the future date, which is 2050, as the boundary in the numerical derivation, we take the sample period as 1947-1999. The coefficients of the predictive relation are highly correlated with changes in demographic moments, in particular, the change in population average age and the change in the share of population in the age range 40-64. We conjecture that it is the change in the share of population in the age range 40-64 that causes the variation of the predictive relation. 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 labor income 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

²¹Several researchers have questioned the use of historical average of realized returns as the estimate of the equity premium. For example, Brown, Goetzmann and Ross (1995) obtain an equity premium of 4% after adjusting the survival bias in the historical estimates of the equity premium; Blanchard (1995) claims that the equity premium has gone steadily down since the early 1950s, and the premium appears to be around 2%-3% in the middle 1990s.

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.

In what follows, we will empirically test the predictive relation as specified in equation (3.17). 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 premium. Since the calibration is highly simplified, the results are only indicative of the potential importance of fully anticipated demographic change.

5 Empirical Tests

In this section, we test the predictive relation specified in equation (3.17) using quarterly equity excess returns 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.²²

Therefore, we loglinearize equation (3.17) as

²²Lamont (1998) has shown that, in the predictive relation, using the dividend payout ratio is numerically identical to using the earnings yield. To be consistent to previous research on the equity premium, we use the dividend payout ratio here.

$$r_{m,t+1} - r_{f,t+1} = \lambda_{0,t}^* + \lambda_{1,t}^*(p_t - \varsigma_t) + \lambda_{2,t}^*(d_t - p_t) + \lambda_{3,t}^*(d_t - e_t) + \epsilon_{t+1} \quad (5.3)$$

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 $\ln(y + x) \approx \ln(y + e^{\ln x}) + e^{\ln x}(\bar{y} + e^{\ln x})^{-1}(\ln x - \ln \bar{x})$, which is derived in Appendix B.

We cannot estimate equation (5.3) 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, denoted by ϕ , to proxy the demographic variables in the predictive relation, i.e.

$$\lambda_{0,t}^* = \beta_0 + \beta_1\phi_t, \quad \lambda_{1,t}^* = \beta_2 + \beta_3\phi_t, \quad \lambda_{2,t}^* = \beta_4 + \beta_5\phi_t, \quad \lambda_{3,t}^* = \beta_6 + \beta_7\phi_t \quad (5.4)$$

Bakshi and Chen (1994) have found that the change in average age is significant in predicting the annual equity excess return for the post-1947 period. In fact, as shown in figure 3 and reported below, 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 obtained using linear functions of the change in average age to proxy the demographic variables in the predictive relation are similar to those reported below.

Substituting the expressions (5.4) into equation (5.3), we have the predictive regression

$$\begin{aligned} r_{m,t+1} - r_{f,t+1} = & \beta_0 + \beta_1\phi_t + \beta_2(p_t - \varsigma_t) + \beta_3\phi_t(p_t - \varsigma_t) \\ & + \beta_4(d_t - p_t) + \beta_5\phi_t(d_t - p_t) + \beta_6(d_t - e_t) + \beta_7\phi_t(d_t - e_t) + \epsilon_{t+1} \end{aligned} \quad (5.5)$$

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 underwent 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 9 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.²³ Log earnings, e , are the natural logarithm of a single quarter's nominal earnings per share. The predictive relation specified in equation (5.5) does not require that earnings and dividends 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 9 plots the historical log real per capita stock price, $p - \varsigma$, the log dividend yield, $d - p$, and the log dividend payout ratio, $d - e$.

²³We use the past four quarters of dividends to adjust the seasonality in dividends.

[Insert Table 3 Here]

Summary statistics are reported in table 3. 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, a multicollinearity problem will arise. Secondly, the change in average age, ϑ , and the change in the share of population in the age range 40-64, ϕ , are also highly correlated. 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 excess return 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 4 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 4 reports the results of this quadri-variate error-correction VAR. The null hypothesis of no cointegration is rejected.²⁴

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 specified in equation (5.5). We also try to replicate previous empirical work of 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 5 Here]

Table 5 replicates Fama and French (1988) and Lamont (1998) in predicting the quarterly equity excess return. We regress the equity excess return on the dividend yield and/or 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 excess return will be higher. Row 3 shows that the dividend yield and the dividend payout ratio combined explain 11.7% of the variation in the equity excess return for this period.

²⁴Using 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.

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. Now the dividend yield is not significant in predicting the equity excess return in either the univariate or the 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 excess return for this period.

[Insert Table 6 Here]

Table 6 replicates Bakshi and Chen (1994) in using the change in average age to predict the quarterly and annual equity excess return for the period 1947Q1-1999Q4. We also regress the quarterly and annual equity excess return 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 excess return, on the level of this share, and finds no significant relation. As suggested by the calibration results in Section 4, we expect that the change in this share matters in predicting the equity excess return.

Panel A regresses the annual and quarterly equity excess return 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 excess return is higher. However, the performance of the change in average age in forecasting the short term, for example, quarterly, excess return, is quite poor. As shown in row 2, the R^2 is only 2.4%. Panel B shows that the change in the share of population in the age range 40-64 has the similar problem in the predictive regression.

The empirical evidence supports the claim that demographics matters in the determination of the equity excess return. However, the low regression R^2 , especially in predicting the short-term equity excess return, suggests that considerable amount of information for stock returns is missing.

[Insert Table 7 Here]

Table 7 tests the predictive relation specified in equation (5.5). Panel A reports the regression results for the period 1947Q1-1994Q4. Row 1 reports the coefficient estimates of the predictive relation specified in equation (5.5). 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 causes lack of identification. 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 excess return 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 excess return still increases with the dividend yield and the dividend payout ratio. But the sensitivities of the equity excess return to the dividend yield and the dividend payout ratio are time varying. When the change in this share of population is higher, the equity excess return becomes less sensitive to the dividend yield and the dividend payout ratio.

Comparing row 4, table 7 to row 3, table 5, 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 5 reports R^2 of 11.7%, while row 4, table 7 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 coefficient estimates in row 8 have similar patten to the calibrated coefficients in Section 4. For example, the coefficient of dividends, $(0.076 - 2.215\phi + 0.096 - 9.731\phi = 0.172 - 11.946\phi)$, is positive, which is obtained after averaging ϕ out, and negatively correlated with ϕ , as predicted by $\lambda_{2,t}$; the coefficient of earnings, $-0.096 + 9.731\phi$, is negative, which is obtained after averaging ϕ out, and positively correlated with ϕ , as predicted by $\lambda_{3,t}$. This confirms the quality of our calibration results.

The most significant improvement of our model on previous research on the predictability of the equity premium is shown from the comparison between row 8 of table 7 and the regressions in panel B of table 5, and that between row 8 of table 7 and row 4 of table 6.

Row 8 of table 7 states that the equity excess return 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 5 ignore the changing structure in the linear regression and produce misleading coefficient estimates. Statistically, the regressions in panel B of table 5 can be viewed as misspecified versions of row 8 of table 7. 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 5 fail to detect any significance of the dividend yield and the dividend payout ratio in predicting the equity excess return.

Figure 10 reports the forecasts of the quarterly log excess return on the S&P Composite Index, $r_m - r_f$, from the predictive regression in row 6 of table 5 and from that in row 8 of table 7. The predictive regression that incorporates demographic change forecasts more variation in the equity premium than does the predictive regression that ignores it. The most significant difference between these two forecasts happened in the early 1950s, the 1970s and the 1990s, when the demographic structure changed drastically and the change in the share of population in the age range 40-64 was far from its average level. Demographic change helps to understand the relation between the equity premium and the dividend yield and the dividend payout ratio. For example, in 1974, there were no significant change in these two explanatory variables, but the change in the share of population in the age range 40-64 was

quite low, which made the equity premium sensitive to the explanatory variables. Therefore, a modest change in these two variables produced significant change in the equity premium. In contrast, in the 1990s, demographic change made the equity premium less sensitive to these three variables. Even though the dividend yield and the dividend payout ratio were low, the equity premium was still high. The predictive regression forecasts negative equity premia in the middle 1970s and in 1987, which correspond to the market crash then.²⁵

Row 4 of table 6 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 7. 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 7 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 6. 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 excess return, and use demographic variables to explain the changing structure in this linear relationship.

[Insert Table 8 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 excess return. Table 8 reports the regression results of predicting the equity excess return of 1 up to 4 quarters. As claimed by Fama and French (1988), the predictability of the equity excess return increases with the return horizons. The regression reports R^2 of 13.6% when

²⁵Boudoukh, Richardson and Smith (1993) report that the equity premium is negative in some states of the world.

predicting the quarterly equity excess return, while 35.8% when predicting the annual equity excess return.

Compare row 4 of table 8 to row 2 and 4 of table 6. 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 annual equity excess return respectively, adding the dividend yield and the dividend payout ratio increases this proportion further. Row 4 of table 8 reports R^2 of 35.8%. This further confirms our theoretical analysis and empirical results.

Figure 11 plots the forecast of the annual equity premium obtained from the fitted value of the predictive regression in row 4 of table 8. Compare figure 11 to figure 8. Since the interval between consecutive observations is one quarter in figure 11, and one year in figure 8, the former figure documents much more variation in the equity premium than the latter. Also, the scales of the equity premium are different in these two figures. However, the moving trends of the equity premium are quite similar. This further confirms the quality prediction of our calibration results.

6 Conclusion

In this paper, we have developed a dynamic overlapping generations (OLG) equilibrium model with a constant per capita stock supply, and a deterministic but time varying demographic structure to study the relation between demographics and the equity premium. 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 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 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 observed historical demographic change could induce significant variations in the coefficients of the predictive relation, and that 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 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 excess return 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 excess return 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 3.2 and Theorem 3.3

In what follows, we first treat proposition 3.2 as a conjecture, and prove theorem 3.3. We then use the results from the proof of theorem 3.3 to prove proposition 3.2.

Proof of Theorem 3.3

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

of the Bellman equation

$$0 = \max_{X,c} \left[-\beta^\tau \exp(-\alpha c_t) + \mathbb{E}_t J(W_{t+1}, \mathfrak{S}_{t+1}, \tau + 1; t + 1) - J(W_t, \mathfrak{S}_t, \tau; t) \right] \quad (\text{A.1})$$

subject to

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

where Q_{t+1} is expressed as in equation (3.5), given that proposition 3.2 is true.

Consider the following trial solution for the value function

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

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

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

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

$$-\gamma_{\tau+1}q_{\tau+1}\Pi_t + \gamma_{\tau+1}^2\Gamma_{\tau+1}^{-1}X_t = 0 \quad (\text{A.5})$$

$$\alpha\beta^\tau \exp(-\alpha c_t) + \gamma_{\tau+1}RE_t J(W_{t+1}, \Pi_{t+1}, \tau + 1; t + 1) = 0 \quad (\text{A.6})$$

Denote $m_\tau = \nu_\tau^{aa} + \gamma_\tau \omega_\tau^{aa} - (\gamma_\tau \omega_\tau^{ab} + \nu_\tau^{ab})\Omega_\tau(\gamma_\tau \omega_\tau^{ab} + \nu_\tau^{ab})^T + q_\tau^T \Gamma_\tau q_\tau$ and $\bar{c}_\tau = \frac{1}{\alpha + R\gamma_{\tau+1}} \ln\left(\frac{\alpha}{d_{\tau+1}\beta R\gamma_{\tau+1}}\right)$.

The optimal investment-consumption policy is

$$X_t = \frac{1}{\gamma_{\tau+1}}\Gamma_{\tau+1}q_{\tau+1}\Pi_t \quad (\text{A.7})$$

$$c_t = \bar{c}_\tau + \frac{R\gamma_{\tau+1}}{\alpha + R\gamma_{\tau+1}}W_t + \frac{\gamma_{\tau+1}n_{\tau+1} + \mu_{\tau+1}}{\alpha + R\gamma_{\tau+1}}t + \frac{1}{2(\alpha + R\gamma_{\tau+1})}\Pi_t^T m_{\tau+1}\Pi_t \quad (\text{A.8})$$

The optimal demand for stocks as expressed in theorem 3.2 is immediate if one substitutes the expression for $q_{\tau+1}$ into equation A.7. Substituting the optimal consumption-investment policy back into the Bellman equation (A.1), we obtain

$$\gamma_\tau = \frac{\alpha R\gamma_{\tau+1}}{\alpha + R\gamma_{\tau+1}} \quad (\text{A.9})$$

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

$$\nu_\tau = -2 \ln \exp(-\alpha\bar{c}_\tau) + d_{\tau+1}\beta \exp(R\gamma_{\tau+1}\bar{c}_\tau) \mathring{i}_{11}^{(N+1, N+1)} + \frac{\alpha}{\alpha + R\gamma_{\tau+1}}m_{\tau+1} \quad (\text{A.11})$$

where $\mathring{i}_{11}^{(n, n)}$ is an $n \times n$ index matrix.²⁶ 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 3.2. To ensure that the value function is actually the equilibrium value function, we have to verify proposition 3.2.

²⁶An index matrix $\mathring{i}_{ij}^{(m, n)}$ is an $m \times n$ matrix with the element $\{i, j\}$ being one and all other elements being zero.

Proof of Proposition 3.2

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

$$G(t) = \sum_{\tau=0}^{\infty} g(\tau) \frac{1}{\gamma_{\tau+1}} \kappa_{\tau+1} \Pi_t = G(t) \quad (\text{A.12})$$

Therefore

$$\sum_{\tau=0}^{\infty} g(\tau) \frac{1}{\gamma_{\tau+1}} \kappa_{\tau+1} = 1 \quad (\text{A.13})$$

$$\sum_{\tau=0}^{\infty} g(\tau) \frac{1}{\gamma_{\tau+1}} \kappa_{\tau+1} = 0_{1 \times N} \quad (\text{A.14})$$

where $[\cdot]_{mn}$ 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 3.2, and further confirms the proof of theorem 3.3.

A.2 Proof of Proposition 3.4 and Theorem 3.5

In what follows, we first treat proposition 3.4 as a conjecture, and prove theorem 3.5. We then use the results from the proof of theorem 3.5 to prove proposition 3.4. The derivation is quite similar to that for proposition 3.2 and theorem 3.3. But allowing for demographic change makes the coefficients of the price function and the value function dependent on time t .

Proof of theorem 3.5

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

$$0 = \max_{X,c} \frac{1}{2} \left[-\beta^\tau \exp(-\alpha c_t) + \mathbb{E}_t J(W_{t+1}, \mathfrak{S}_{t+1}, \tau + 1; t + 1) \right] - J(W_t, \mathfrak{S}_t, \tau; t) \quad (\text{A.15})$$

subject to

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

where Q_{t+1} is expressed as in equation 3.10, given proposition 3.4 is true.

Consider the following trial solution for the value function

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

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

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

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

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

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

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

$$X_t = \frac{1}{\gamma_{\tau+1}}\Gamma_{\tau+1,t+1}q_{\tau+1,t+1}\Pi_t \quad (\text{A.21})$$

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

The optimal demand for stocks as expressed in theorem 3.5 is immediate if one substitutes the expression for $q_{\tau+1,t+1}$ into equation (A.21). Substituting the optimal consumption-investment policy back into the Bellman equation (A.15), we obtain

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

The trial value function is confirmed on condition that the price function is as claimed in proposition 3.4. To ensure that the trial indirect utility function is actually the equilibrium value function, we have to verify proposition 3.4.

Proof of theorem 3.5

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

$$G(t) \stackrel{\times}{=} \sum_{\tau=0} g(\tau;t) \frac{1}{\gamma_{\tau+1}}\kappa_{\tau+1,t+1}\Pi_t = G(t) \quad (\text{A.24})$$

Therefore

$$\sum_{\tau=0}^{\infty} g(\tau; t) \frac{1}{\gamma_{\tau+1}} \kappa_{\tau+1, t+1} = 1 \quad (\text{A.25})$$

$$\sum_{\tau=0}^{\infty} g(\tau; t) \frac{1}{\gamma_{\tau+1}} \kappa_{\tau+1, t+1} = 0_{1 \times N} \quad (\text{A.26})$$

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 3.4, and further confirms the proof of theorem 3.5.

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 Taylor expansion around $\bar{\ln} x$. The second approximation is obtained by substituting \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 annual average unemployment rate as that year's unemployment rate.

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 $\ln(\text{CSTIND}_{t+1}/\text{CSTIND}_t) - \ln(\text{USTIND}_{t+1}/\text{USTIND}_t)$, 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, Andrew B., 1999, The effects of a baby boom on stock prices and capital accumulation in the presence of social security, Working Paper, Wharton School, University of Pennsylvania.

Bakshi, Gurdip S., and Zhiwu Chen, 1994, Baby boom, population aging, and capital markets, *Journal of Business* 67, 165-202.

Bergantino, Steven M., 1998, Lifecycle investment behavior, demographics, and asset prices, Doctoral Dissertation, MIT.

Blanchard, Olivier J., 1993, Movements in the equity premium, *Brooking Papers on Economic Activity*, 2, 75-138.

Bodie, Zvi, Robert C. Merton, and William F. Samuelson, 1992, Labor supply flexibility and portfolio choice in a life cycle model, *Journal of Economic Dynamics and Control* 16, 427-449.

Bossaerts, Peter, and Pierre 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 Raymond W. Goldsmith, ed.: *Institutional Investors and Corporate Stock* (NBER, New York, NY.).

Boudoukh, Jacob, Matthew Richardson, and Tom Smith, 1993, Is the ex ante risk premium always positive? *Journal of Financial Economics* 34, 387-408.

Brennan, Michael J., 1998, The role of learning in dynamic portfolio decision, *European Finance Review* 1, 295-306.

Brooks, Robin, 1999, What will happen to financial markets when baby boomers retire? Working Paper, Federal Reserve Board of Governors.

Brown, Stephen J., William N. Goetzmann, and Stephen A. Ross, 1995, Survival, *Journal of Finance* 50, 853-873.

Campbell, John Y., 1996, Understanding risk and return, *Journal of Political Economy* 104, 298-345.

Campbell, John Y., and Albert S. Kyle, 1993, Smart money, noise trading and stock price behavior, *Review of Economic Studies* 60, 1-34.

Campbell, John Y., and Robert J. Shiller, 1988, The dividend-price ratio and expectations of future dividends and discount factors, *Review of Financial Studies* 1, 195-228.

Canner, Niko, N. Gregory Mankiw, and David N. Weil, 1997, An asset allocation puzzle, *American*

Economic Review 87, 181-191.

Cochrane, John H., 1999, New facts in finance, *Economic Perspectives* 23, 36-58.

Erb, Claude B., Campbell R. Harvey, and Tadas E. Viskanta, 1997, Demographics and international investments, *Financial Analysts Journal* 53, 14-28.

Fama, Eugene F., and Kenneth R. French, 1988, Dividend yields and expected stock returns, *Journal of Financial Economics* 22, 3-25.

Franke, Günter, Richard C. Stapleton, and Marti G. Subrahmanyam, 1998, Who buys and who sells options: The role of options in an economy with background risk, *Journal of Economic Theory* 82, 89-109.

Friedman, Milton, 1957, *A Theory of the Consumption Function* (Princeton University Press, Princeton, NY.).

Goetzmann, William N., and Philippe Jorion, 1995, A longer look at dividend yields, *Journal of Business* 68, 483-508.

Goetzmann, William N., and Philippe Jorion, 1993, Testing the predictive power of dividend yields, *Journal of Finance* 48, 663-679.

Gollier, Christian, and John W. Pratt, 1996, Risk vulnerability and the tempering effect of background risk, *Econometrica* 64, 1109-1123.

Gompers, Paul, and Josh Lerner, 2000, Money chasing deals? The impact of fund inflows on private equity valuations, *Journal of Financial Economics* 55, 281-325.

Goyal, Amit, 1999, Demographics and stock market flows, Unpublished Paper, UCLA.

Goyal, Amit, and Ivo Welch, 1999, Predicting the equity premium, Working Paper, UCLA.

Heaton, John, and Deborah Lucas, 2000, Portfolio choice and asset prices: The importance of entrepreneurial risk? *Journal of Finance* 55, 1163-1198.

Hirshleifer, David, and Avanidhar Subrahmanyam, 1993, Futures versus share contracting as means of diversifying output risk, *Economic Journal* 103, 620-638.

Hodrick, Robert J., 1992, Dividend yields and expected stock returns: Alternative procedures for inference and measurement, *Review of Financial Studies* 5, 357-386.

Horvath, Michael T.K., and Mark W. Watson, 1995, Testing for cointegration when some of the cointegrating vectors are prespecified, *Econometric Theory* 11, 984-1014.

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

Kimball, Miles S., 1993, Standard risk aversion, *Econometrica* 61, 589-611.

Lamont, Owen, 1998, Earnings and expected returns, *Journal of Finance* 53, 1563-1587.

Lintner, John, 1956, Distribution of incomes of corporations among dividends, retained earnings and taxes, *American Economic Review* 46, 97-113.

- Mankiw, N. Gregory, and David N. Weil, 1989, The baby boom, the baby bust, and the housing market, *Regional Science and Urban Economics* 19, 235-258.
- Macunovich, Diane J., 1997, Discussion of “social security: How social and secure should it be?” in Steven Sass and Robert Triest, eds., *Social Security Reform: Links to Saving, Investment, and Growth* (Federal Reserve Bank of Boston, Boston, MA.).
- Modigliani, Franco, 1986, Life cycle, individual thrift, and the wealth of nations, *American Economic Review* 76, 297-313.
- Modigliani, Franco, and R. Brumberg, 1954, Utility analysis and the consumption function: An interpretation of cross-section data, in Kenneth K. Kurihara ed., *Post-Keynesian Economics* (Rutgers University Press, New Brunswick, NJ.).
- Nelson, Charles R., and Myung J. Kim, 1993, Predictable stock returns: The role of small sample bias, *Journal of Finance* 48, 641-661.
- Newey, Whitney K., and Kenneth D. West, 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703-708.
- Pesaran, M. Hashem, and Alan Timmermann, 1995, Predictability of stock returns: Robustness and economic significance, *Journal of Finance* 50, 1201-1228.
- Pischke, Jörn-Steffen, 1995, Individual income, incomplete information, and aggregate consumption, *Econometrica* 63, 805-940.
- Poterba, James M., 2000, Population age structure and asset returns: An empirical investigation, Working Paper, NBER.
- Poterba, James M., and Andrew A. Samwick, 1997, Household portfolio allocation over the lifecycle, Working Paper, NBER.
- Samuelson, Paul A., 1969, Lifetime portfolio selection by dynamic stochastic programming, *Review of Economics and Statistics*, 239-246.
- Samuelson, Paul A., 1989, A case at last for age-phased reduction in equity, *Proceedings of the National Academy of Sciences* 86, 9048-9051.
- Samuelson, Paul A., 1991, Long-run risk tolerance when equity returns are mean regressing: Pseudoparadoxes and vindication of “businessman’s risk”, in William C. Brainard, William D. Nordhaus and Harold W. Watts eds.: *Money, Macroeconomics, and Economic Policy: Essays in Honor of James Tobin* (MIT Press, Cambridge, MA.).
- Schieber, Sylvester, and John Shoven, 1997, The consequences of population aging on private pension fund saving and asset markets, in Sylvester Schieber and John Shoven, eds.: *Public Policy Towards Pension* (MIT Press, Cambridge, MA.).
- Shiller, Robert J., 1984, Stock prices and social dynamics, *Brookings Papers on Economic Activity*, 2, 457-498.
- Siegel, Jeremy J., 1998, *Stocks for the Long Run*, 2nd ed. (McGraw Hill, New York, NY.).

- Stambaugh, Robert F., 1999, Predictive regressions, *Journal of Financial Economics* 54, 375-421.
- Torous, Walter, and Shu Yan, 1999, Predictive Regressions Revisited, Working Paper, UCLA.
- Vayanos, Dimitri, 1998, Transaction costs and asset prices: A dynamic equilibrium model, *Review of Financial Studies* 11, 1-58.
- Viceria, Luis M., 1997, Testing for structural change in the predictability of asset returns, Working Paper, Harvard University.
- Wang, Jiang, 1994, A model of competitive stock trading volume, *Journal of Political Economy* 102, 127-168.
- Yoo, Peter S., 1994a, Age dependent portfolio selection, Working Paper, Federal Reserve Bank of St. Louis.
- Yoo, Peter S., 1994b, Age distributions and returns of financial assets, Working Paper, Federal Reserve Bank of St. Louis.
- Xia, Yihong, 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 $\alpha = 0.08$, $r = 1.2\%$ 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.

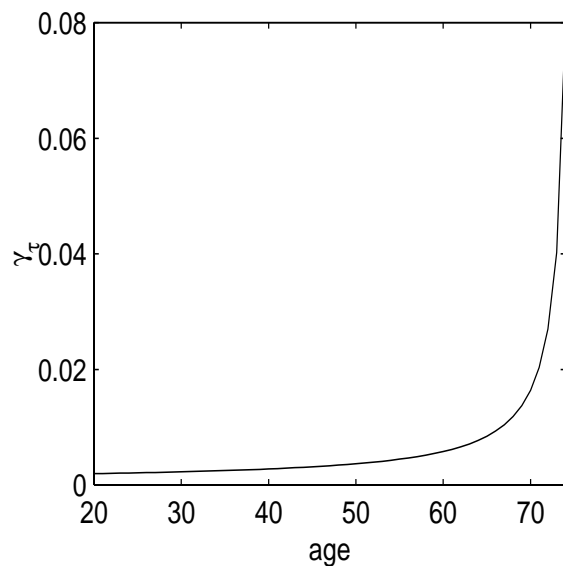


Figure 2: The Demographic Structure of the U.S. Population in the Age Range 20-74 for 1947-2050

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

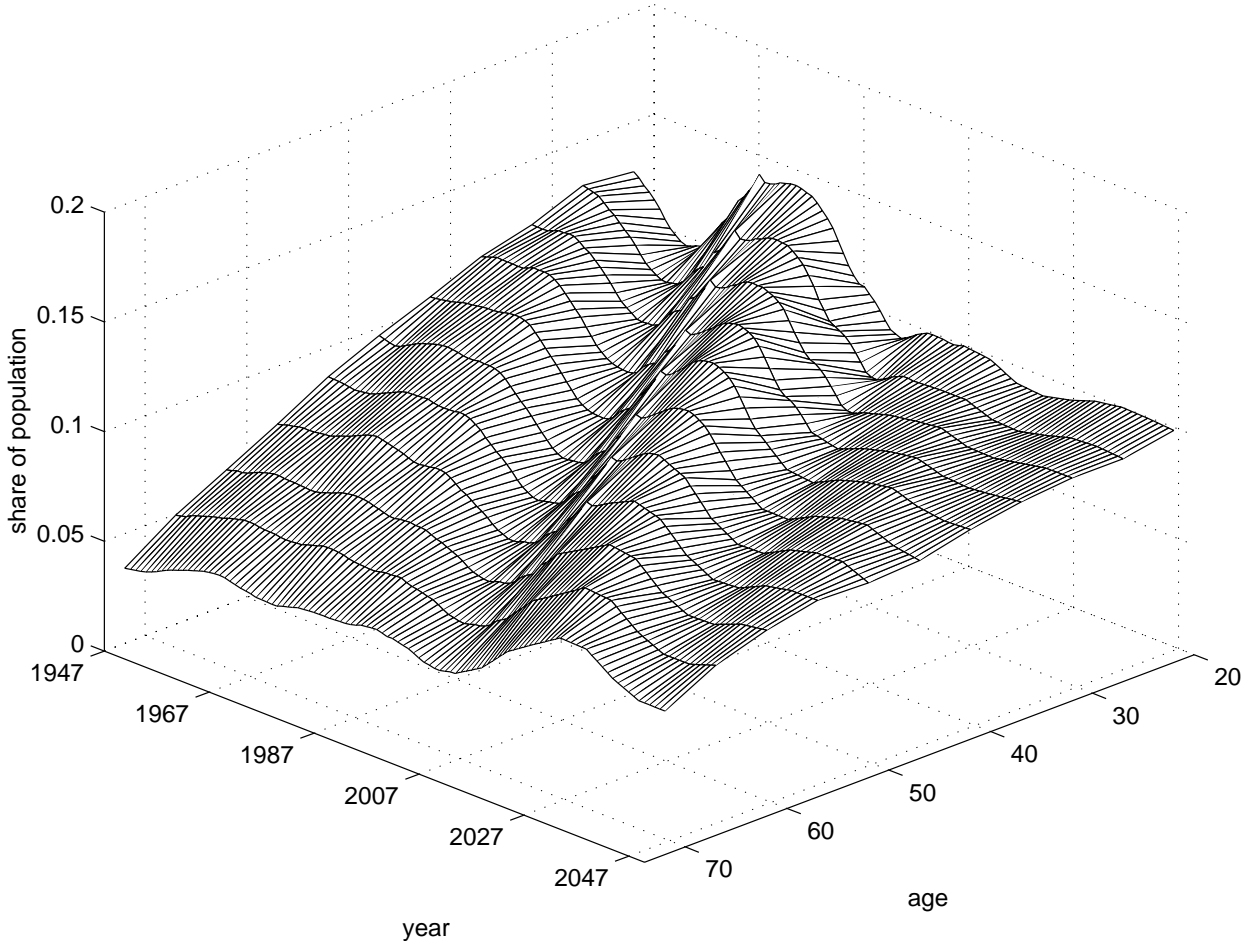


Figure 3: Moments of the U.S. Demographic Structure for 1947-1999

Source: Citibase. The whole population include only people in the age range 20-74.

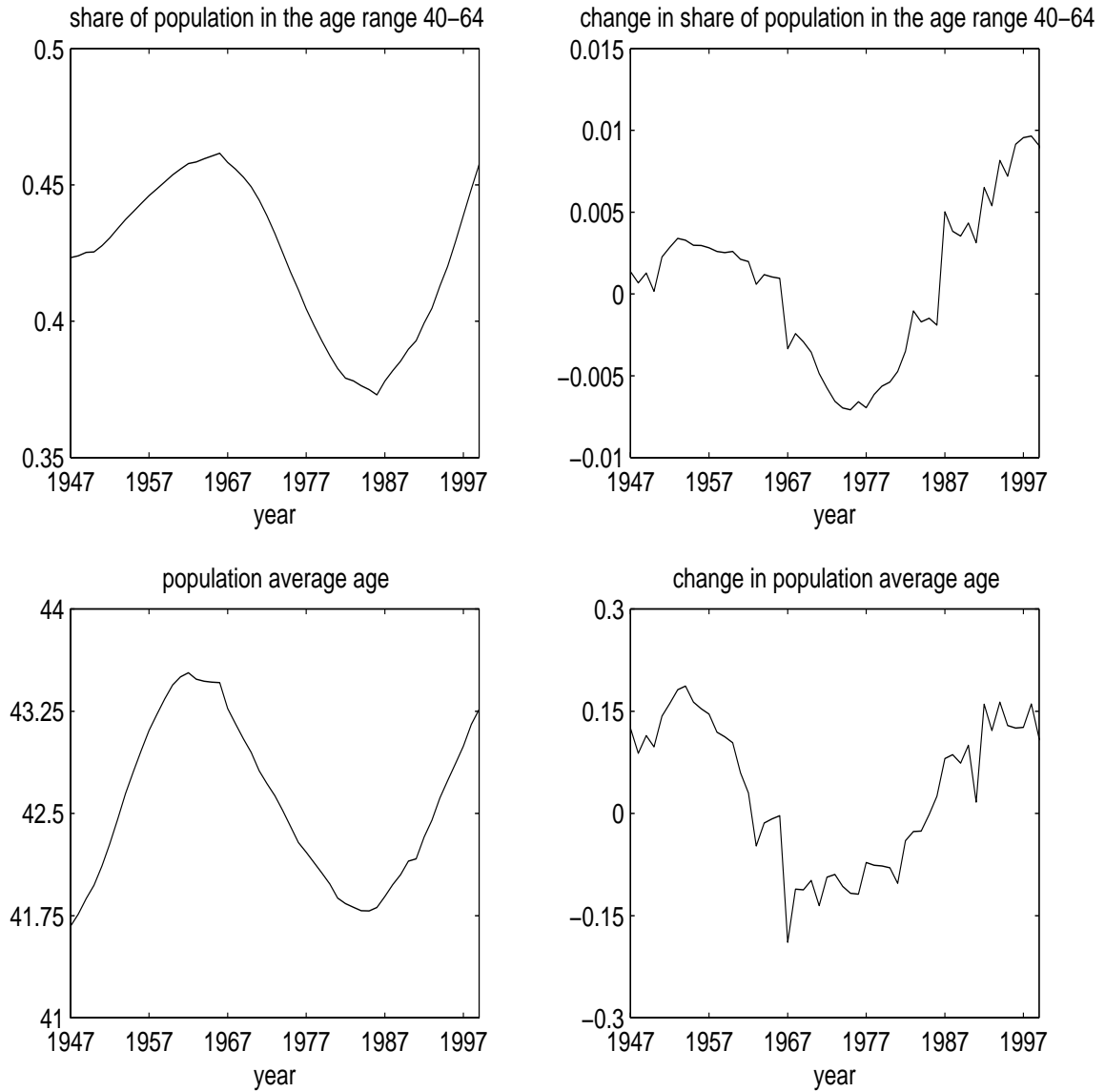


Figure 4: The Estimates of the Labor Income Process

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 dashed lines represent the cubic polynomial approximations of the empirical estimates.

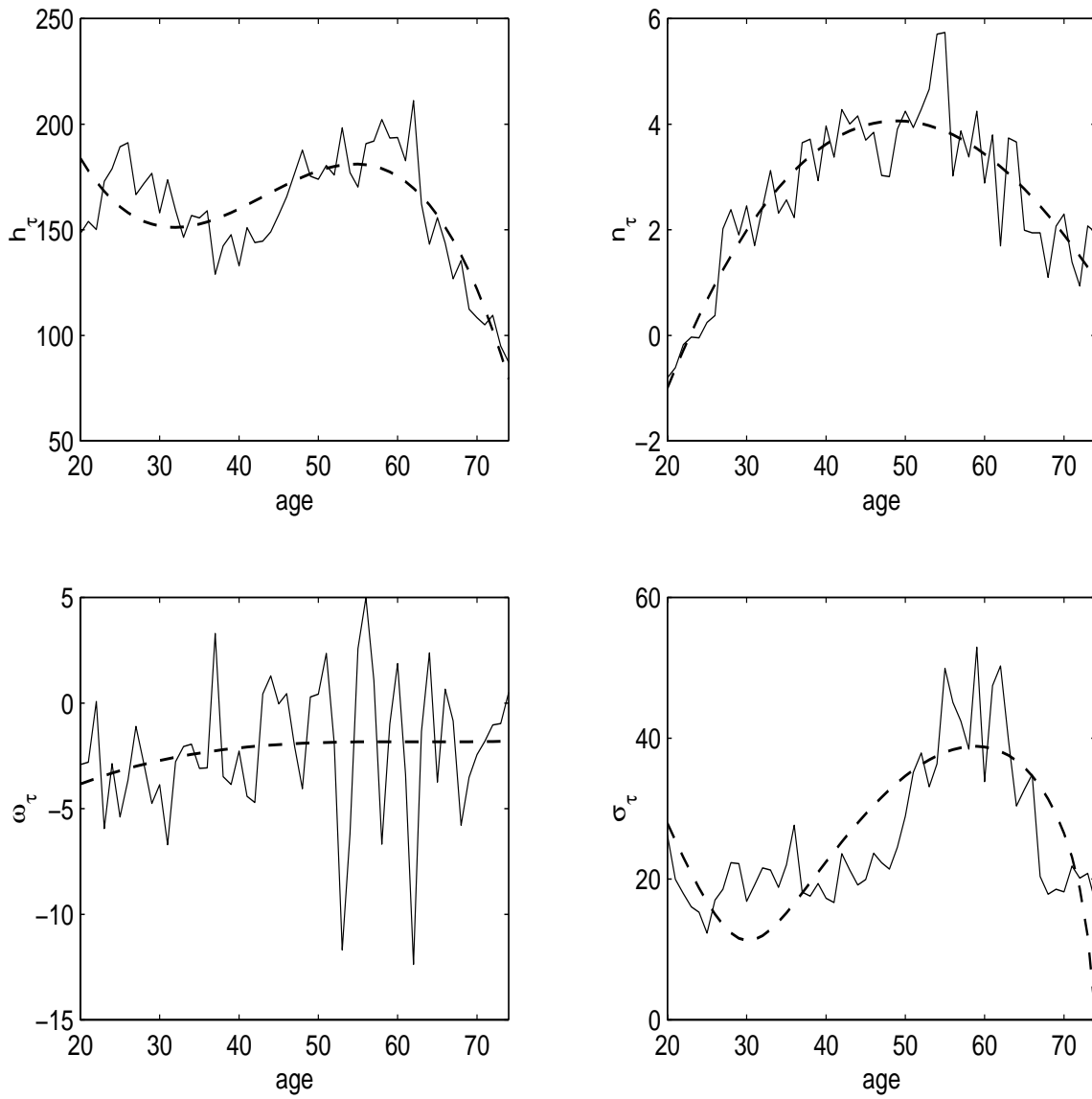


Figure 5: The Time Varying Coefficients of the Price Function

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 2.79 under the calibration.

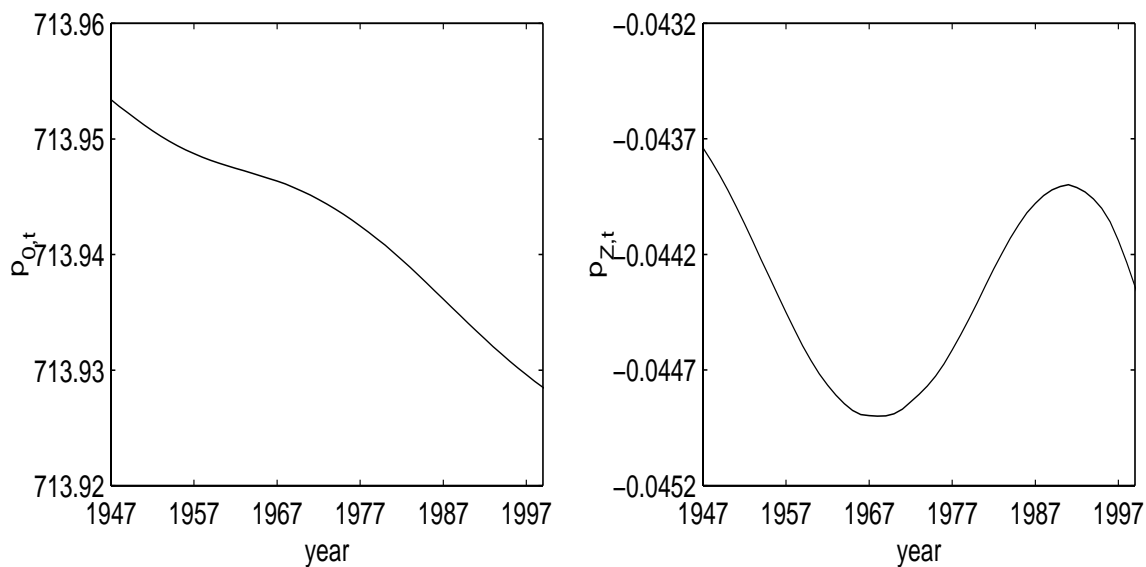


Figure 6: The Time Varying Coefficients of the Expected Excess Return

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 .

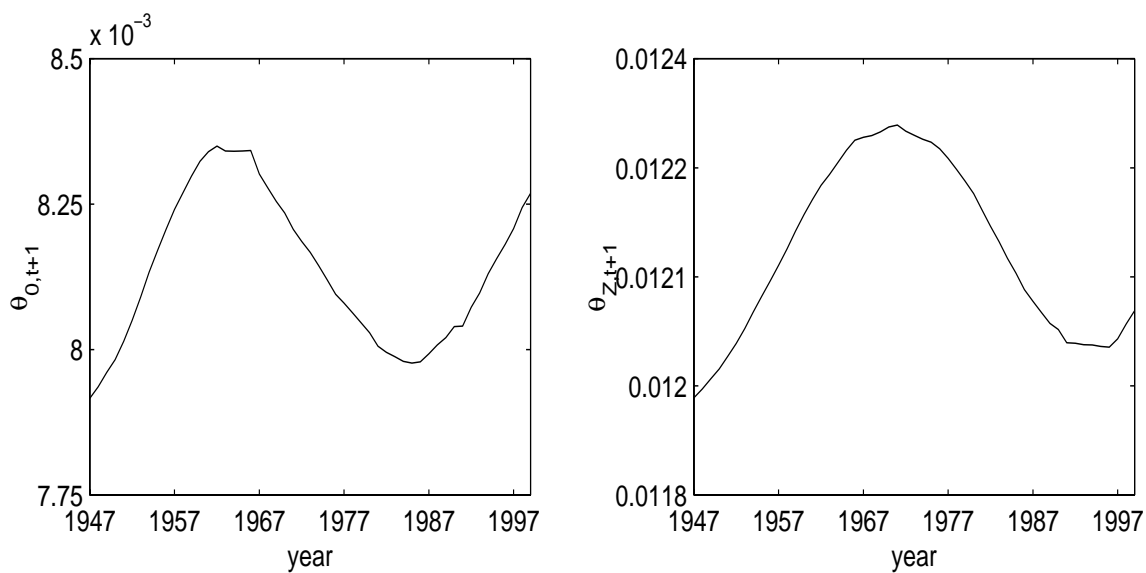


Figure 7: The Time Varying Coefficients of the Predictive Relation

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, $\frac{B_t}{P_t}$.

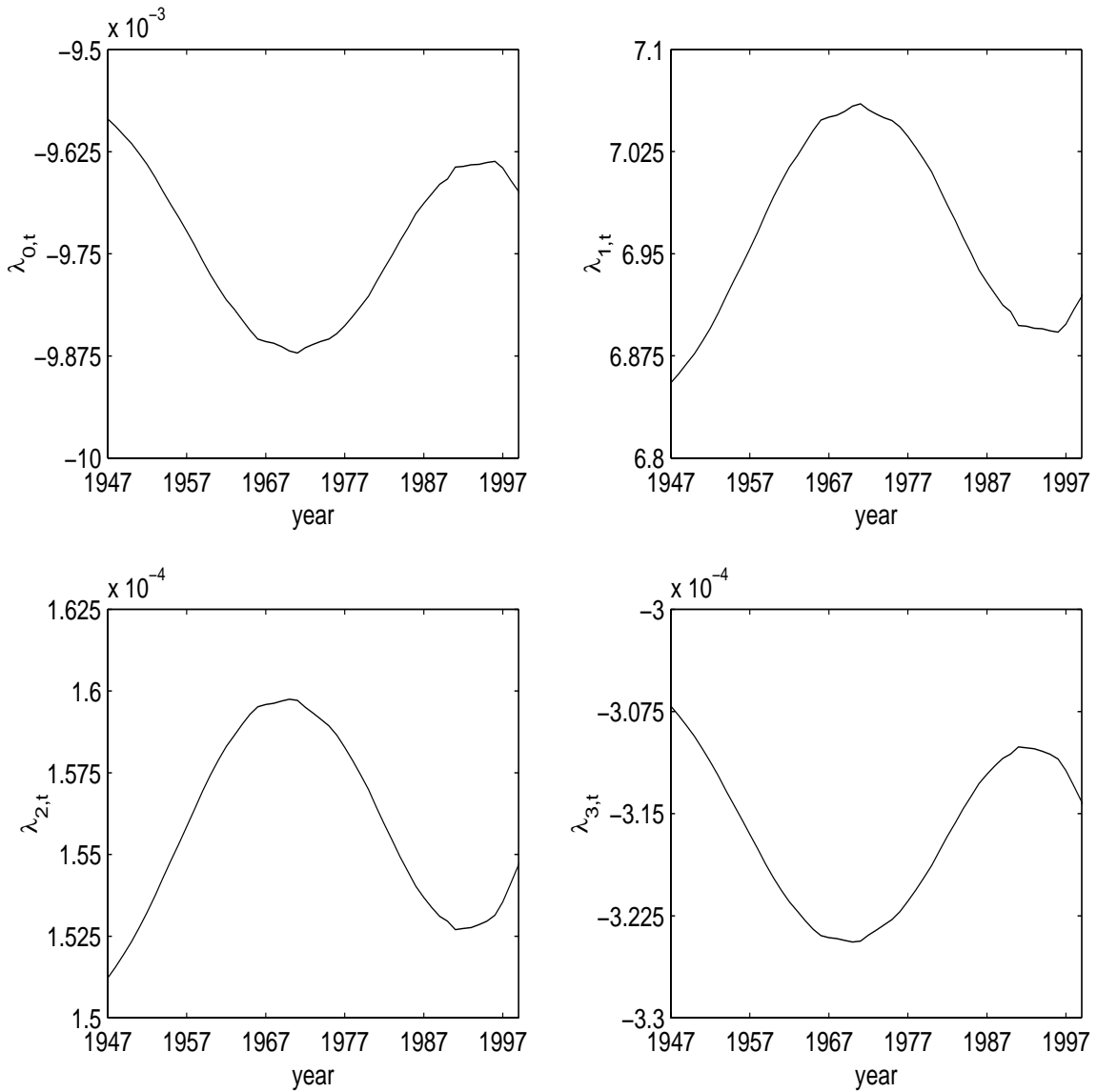


Figure 8: The Fitted Value of the Calibrated Predictive Relation

The figure plots the fitted value 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 P_t , $\frac{D_t}{P_t}$ and $\frac{B_t}{P_t}$ are the historical values of the real per capita stock price, the dividend yield and the earnings yield respectively, which correspond the S&P Composite Index; $\lambda_{0,t}$, $\lambda_{1,t}$, $\lambda_{2,t}$ and $\lambda_{3,t}$ are derived in the numerical solution under the calibrated parameter values.

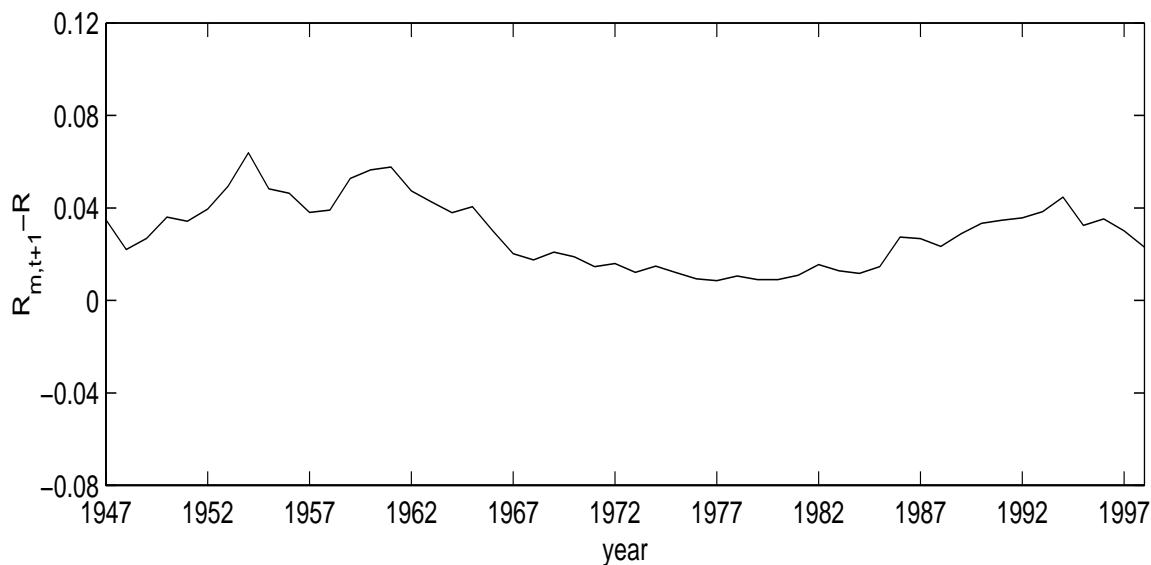


Figure 9: The Real Per Capita Stock Price, the Dividend Yield and the Dividend Payout Ratio 1947Q1-1999Q4

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

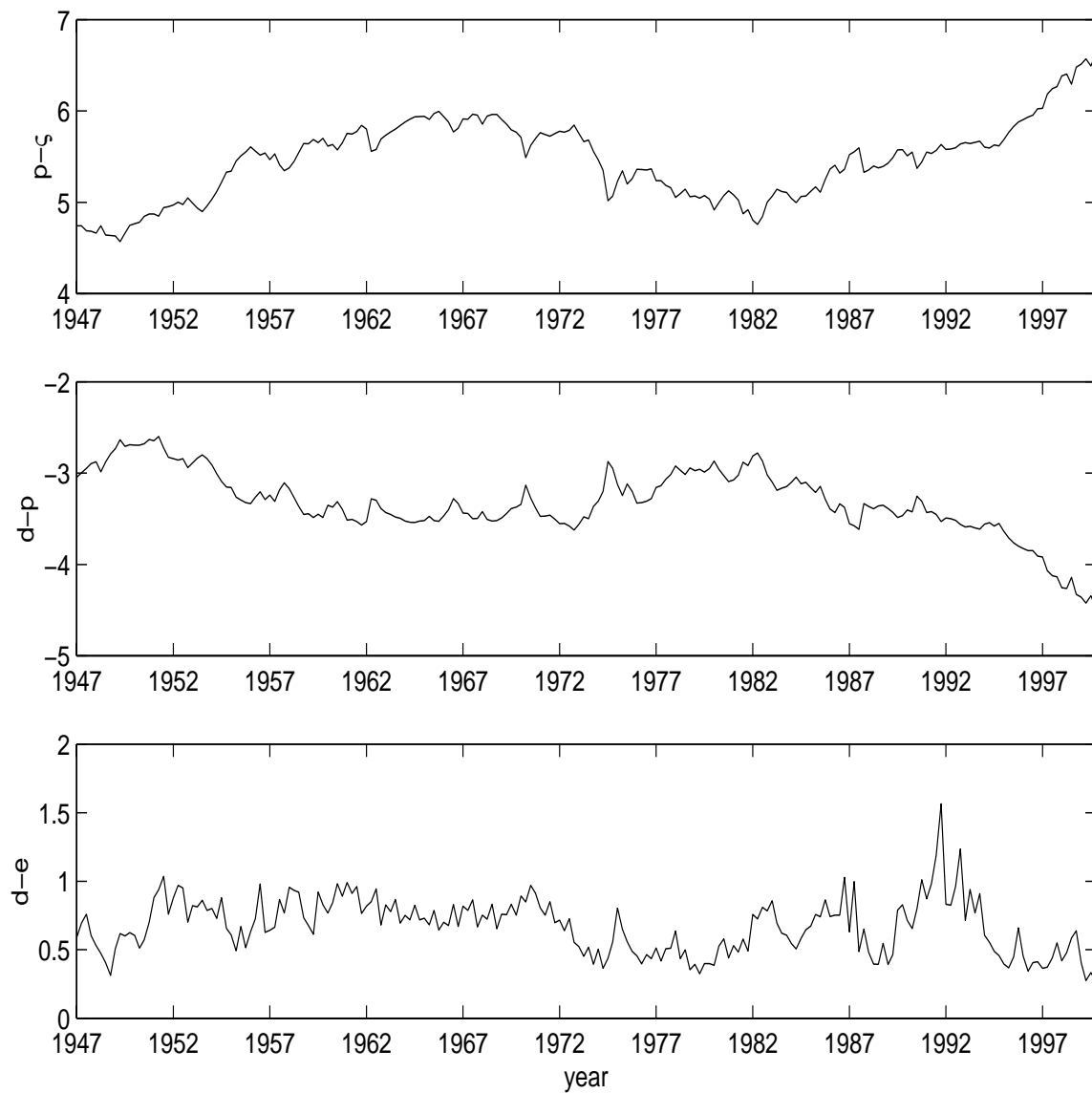


Figure 10: The Fitted Value of the Predictive Regression

The solid line 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 $d - p$ 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. This predictive regression is estimated in row 8 of table 7. The dashed line represents the fitted value of the predictive regression by Lamont (1998), calculated by

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

This predictive regression is estimated in row 6 of table 5.

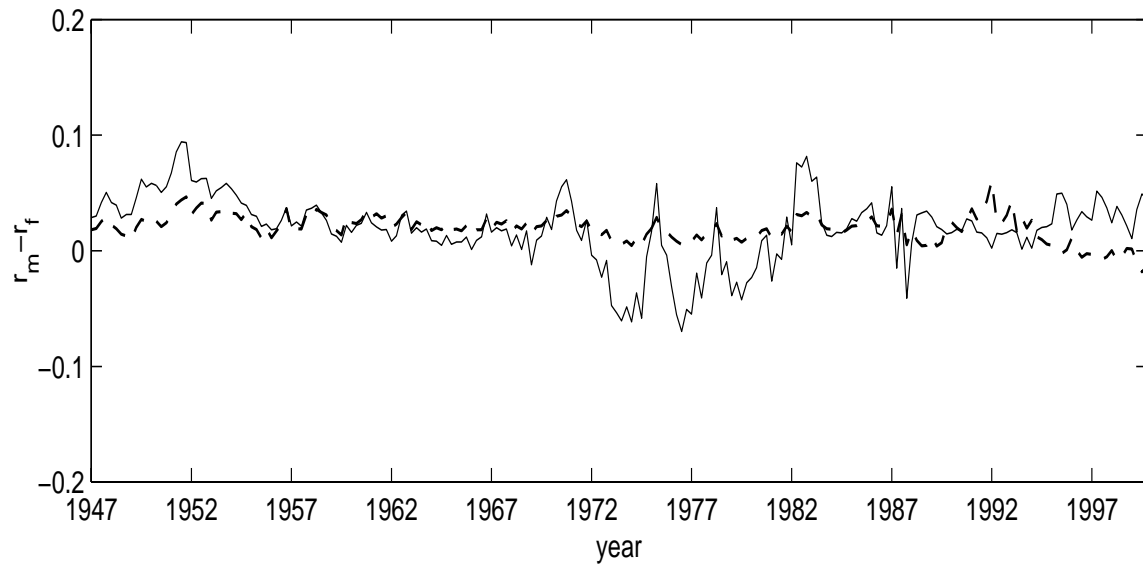


Figure 11: The Fitted Value of the Long-run Predictive Regression

The figure plots the fitted value of the long-run predictive regression, calculated by

$$r_{m,t+1} - r_{f,t+1} = 0.918 + 0.280(d_t - p_t) - 6.332\phi_t(d_t - p_t) + 0.110(d_t - e_t) - 23.064\phi_t(d_t - e_t)$$

where $r_{m,t+1} - r_{f,t+1}$ is the annual equity premium; $d - p$ 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. This predictive regression is estimated in row 4 of table 8.

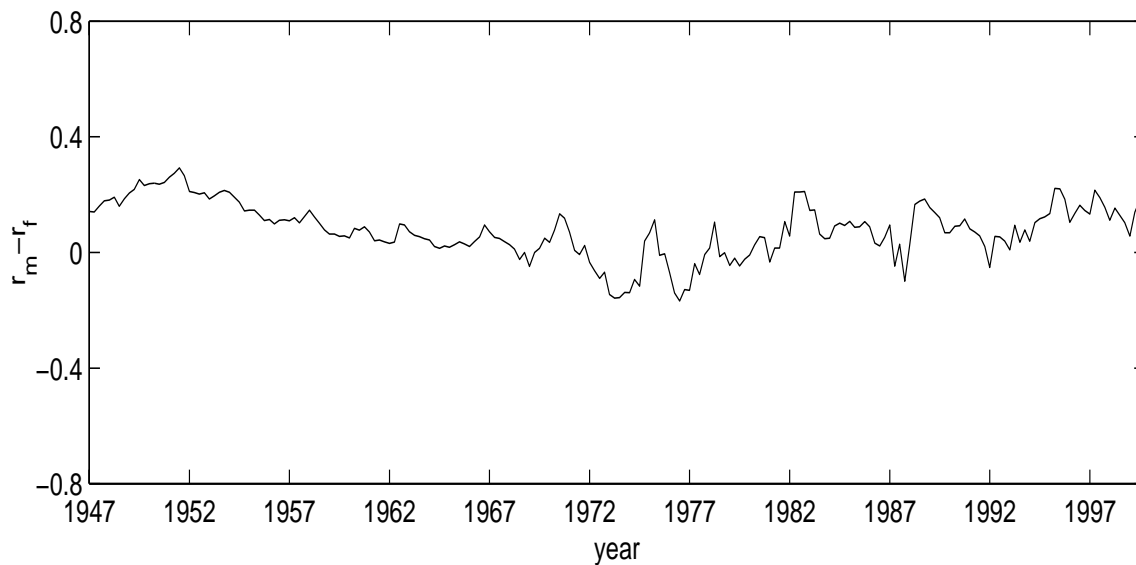


Table 1: Model Calibration I

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}	8.54
2 Volatility of Temporary Shocks to Dividends, $\epsilon_{D,t}$	σ_D	0.95
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	0.33
5 Long-run Level of Earnings, B_t	\bar{B}	17.45
6 Ratio of the Dividend State Variable, F_t , to the De-meaned Earnings, B_t	a_B	0.22
7 Volatility of Temporary Shocks to Earnings, $\epsilon_{B,t}$	σ_B	0.32
The Labor Income Process		
8 AR(1) Coefficient of the Labor Income State Variable, Z_t	a_Z	0.74
9 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		
10 Correlation Coefficients between $\epsilon_{D,t}$ and $\epsilon_{F,t}$	ρ_{DF}	0
11 Correlation Coefficients between $\epsilon_{D,t}$ and $\epsilon_{Z,t}$	ρ_{DZ}	0
12 Correlation Coefficients between $\epsilon_{F,t}$ and $\epsilon_{Z,t}$	ρ_{FZ}	-0.80
Other Parameters		
13 Risk Aversion Coefficient	α	0.08
14 Subjective Discount Factor	β	1.5
15 Real Interest Rate	r	1.2%

Table 2: Model Calibration II

This table reports the correlations between the calibrated coefficients of the predictive relation moments of the demographic structure. The sample period is 1947-1999.

	$\lambda_{0,t}$	$\lambda_{1,t}$	$\lambda_{2,t}$	$\lambda_{3,t}$
Population Average Age	-0.36	0.37	0.51	-0.51
Change in Population Average Age	0.85	-0.85	-0.77	0.77
Share of Population in the Age Range 40-64	-0.28	0.28	0.44	-0.44
Change in Share of Population in the Age Range 40-64	0.75	-0.75	-0.66	0.66

Table 3: 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 4: 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 (0.080)	0.057 (0.069)	-0.801 (0.450)	-0.327 (0.281)	-0.099 (0.037)	-0.036 (0.034)	-0.022 (0.042)	-0.006 (0.032)	0.069
$\Delta \varsigma_{t+1}$	0.015 (0.011)	0.001 (0.010)	0.299 (0.065)	0.084 (0.040)	-0.000 (0.005)	0.007 (0.005)	0.011 (0.006)	-0.010 (0.005)	0.184
Δd_{t+1}	0.067 (0.017)	0.022 (0.015)	-0.023 (0.099)	0.373 (0.062)	-0.008 (0.008)	-0.011 (0.008)	-0.006 (0.009)	-0.028 (0.007)	0.303
Δe_{t+1}	-0.026 (0.142)	0.117 (0.123)	1.261 (0.804)	0.184 (0.502)	-0.324 (0.066)	-0.162 (0.061)	-0.234 (0.076)	0.203 (0.057)	0.214

Horvath-Watson test for cointegration: 61.135

Table 5: Predicting the Quarterly Equity Excess Return 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 excess return 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 (0.054)	0.064 (0.017)		0.046	1.766	10.196
2	-0.042 (0.025)		0.083 (0.033)	0.039	1.802	8.778
3	0.207 (0.057)	0.083 (0.020)	0.112 (0.035)	0.117	1.805	13.599
Panel B: 1947Q1-1999Q4 Obs=212						
4	0.080 (0.051)	0.018 (0.016)		0.002	1.812	1.513
5	-0.016 (0.022)		0.053 (0.029)	0.015	1.821	4.131
6	0.038 (0.056)	0.016 (0.017)	0.050 (0.029)	0.015	1.804	2.639

Table 6: Predicting the Quarterly and Annual Equity Excess Return by Demographic Variables 1947-1999

This table reports the results of the regressions that replicate Bakshi and Chen (1994), in which the equity excess return is regressed on the change in average age, and the regressions that run the equity excess return 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 excess return, 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 (0.018)	21.533 (7.275)	0.113	0.552	209	27.478
2	Quarterly	0.016 (0.005)	5.121 (2.137)	0.024	1.892	212	6.090
Panel B: $r_{t+1} = \beta_0 + \beta_1\phi_t + \epsilon_{t+1}$							
3	Annual	0.074 (0.017)	4.577 (1.532)	0.094	0.533	209	22.634
4	Quarterly	0.018 (0.005)	1.160 (0.463)	0.023	1.886	212	6.012

Table 7: Predicting the Quarterly Equity Excess Return 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 specified in equation (5.5). 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

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

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 (0.066)	-9.182 (15.741)	0.000 (0.028)	1.355 (6.633)	0.086 (0.043)	-2.373 (7.638)	0.117 (0.035)	-8.390 (3.317)	0.137	1.823	5.342
2	0.217 (0.060)	-9.182 (15.801)		1.354 (6.527)	0.086 (0.020)	-2.374 (7.407)	0.117 (0.036)	-8.390 (3.321)	0.142	1.823	6.266
3	0.219 (0.061)	-6.799 (8.711)			0.087 (0.020)	-3.860 (2.942)	0.118 (0.035)	-8.322 (3.422)	0.146	1.820	7.548
4	0.223 (0.063)				0.087 (0.020)	-1.737 (0.674)	0.114 (0.034)	-8.096 (3.351)	0.147	1.825	9.241
Panel B: 1947Q1-1999Q4 Obs=212											
5	0.223 (0.068)	-8.067 (13.267)	-0.002 (0.028)	0.439 (6.502)	0.083 (0.044)	-3.638 (7.175)	0.109 (0.036)	-8.432 (2.660)	0.133	1.905	5.644
6	0.221 (0.061)	-8.084 (13.326)		0.422 (6.394)	0.085 (0.021)	-3.671 (6.964)	0.108 (0.036)	-8.425 (2.668)	0.138	1.904	6.616
7	0.221 (0.062)	-7.327 (4.872)			0.085 (0.021)	-4.128 (1.444)	0.109 (0.035)	-8.398 (2.795)	0.142	1.903	7.977
8	0.202 (0.056)				0.076 (0.018)	-2.215 (0.518)	0.096 (0.035)	-9.731 (2.802)	0.136	1.911	9.338