The Expected Real Return to Equity

Missaka Warusawitharana[*]

Board of Governors of the Federal Reserve System

Abstract

The expected return to equity — typically measured as a historical average — is a key variable in the decision making of investors. A recent literature uses analysts' forecasts, investor surveys or present-value relationships and finds estimates of expected returns that are sometimes much lower than historical averages. This study extends the present-value approach to a dynamic optimizing framework. Given a model that captures this relationship, one can use data on dividends, earnings and valuations to infer the model-implied expected return. Using this method, the estimated expected real return to equity ranges from 4.9 to 5.6 percent. Furthermore, the analysis indicates that expected returns have declined by about 3 percentage points over the past forty years. These results indicate that future returns to equity may be lower than past realized returns.

March 18, 2013

1 Introduction

The expected return to the aggregate stock market is a key variable in the decisions of both individual investors and corporations, as emphasized by Merton (1980). A sample average of realized returns provides the simplest, and most widely used, estimate of expected returns. However, future returns may differ from past values. A recent literature examines this possibility by constructing forward-looking measures of expected returns based on analysts' forecasts, investor surveys or a present-value relationship linking dividends to valuations.[1]

Analysts' forecasts and investor surveys have their own limitations. Easton and Sommers (2007) argue that analysts' forecasts have an upward bias that can significantly affect estimates of expected returns based on their forecasts. Findings from surveys capture the views of only one segment of the market and, in addition, survey-based methods can suffer from many statistical biases (see Foreman (1991)).

Using a present-value relationship avoids the above problems with analysts' forecasts or investor surveys. The present-value approach uses data on dividend yields and dividends (or earnings yields and earnings) and computes expected returns based on the static Gordon growth model for dividends (see Gordon (1959)). However, there exists a much richer relationship between dividends and valuations than given by the present-value relationship. Specifically, dividends and valuations are related through a dynamic optimizing model. The key insight of this study is that given such a model, one could use data on the dividends, earnings, investment and valuations of the US stock market to estimate the model parameters and thereby infer the expected return to equity. This idea differs from the approach in the broader equity premium literature, which fits models to data on fundamentals and asset returns to draw inference on the models. Instead, the novel approach in this study fits a model to data on fundamentals only, and uses the estimates to draw inference on asset returns.

There are many benefits to taking a model-based approach to estimating expected returns. Such an approach avoids potential biases associated with analysts' forecasts. The estimation uses data for the aggregate market, thus taking into account the expectations of all market participants as opposed to only analysts or survey participants. The model features productivity shocks, earnings dynamics and the endogeneity of dividends, all of which are absent from the Gordon growth model. The model also incorporates a time-varying pricing kernel that places a greater valuation on

---

[1] Graham and Harvey (2005) and Fernandez and Baonza (2010) employ investor surveys; Claus and Thomas (2001) and Pastor, Sinha, and Swaminathan (2008) use analysts' forecasts; and Blanchard (1993), Jagannathan, McGrattan, and Scherbina (2001) and Fama and French (2002) base their analysis on present-value relationships.

dividend payouts in recessions. The estimation simultaneously includes data on dividends, earnings, investment and valuations, thereby incorporating more information than that used by the present-value approach. On the other hand, a model-based approach is necessarily predicated on a particular model of equity values. As such, I investigate the robustness of the findings to some changes to the underlying model.

The model underlying the estimation is a variant of the standard production-based asset pricing model employed by Cochrane (1991), Cochrane (1996), Jermann (1998), Gomes, Kogan, and Zhang (2003), Kogan (2004), Zhang (2005), Gomes, Yaron, and Zhang (2006), Liu, Whited, and Zhang (2009), Jermann (2011) and others. The model prices an aggregate equity claim, compared to the per-share claim typically priced in the literature.[2] One deviation from the standard model is to incorporate decreasing returns to scale in the profit function, reflecting firm markups. This feature enables a more general treatment of profits than in a constant returns to scale framework. It also leads firms to vary their optimal scale following persistent shocks to productivity and helps provide an additional source of value compared to models that take dividends as exogenous.

I estimate the parameters of the model using data on a U.S. representative firm constructed by aggregating firm level data from the CRSP/Compustat merged data set. I perform the analysis on two samples: the first uses annual data from 1966 to 2009; the second uses quarterly data from 1984:Q1 to 2009:Q4. Data availability on stockholders' equity and dividends constrain the start dates of the two samples, respectively. Both samples exclude financial firms and regulated utilities.

The parameter estimates from the annual data imply a mean expected annual real return to equity of about 4.9% to 5.6% over the period from 1966 to 2009. These estimates are within the range of values for expected returns obtained by Blanchard (1993), Fama and French (2002), and Graham and Harvey (2005).[3] The findings also support the decision in the limited participation literature to use a lower return to equity than would be indicated by historical returns (see Gomes and Michaelides (2005) and Polkovnichenko (2007)).

Separating the sample into two periods, I find that average expected returns to equity decline sharply from about 6.6% in the period from 1966 to 1987 to about 3.7% in the period from 1988 to 2009. This decline reflects the fact that while earnings and dividends are lower in the latter periods, valuations are higher. The model reconciles this apparent discrepancy with a lower expected return, which raises valuations by lowering the discount rate applied to future dividends. Repeating the

---

[2]Bansal and Yaron (2007) emphasize this distinction and compare the implications of pricing an aggregate equity claim versus a per-share claim. Larrain and Yogo (2008) examine the present value relationship between asset prices and payouts using data on the aggregate stock market.

[3]The estimated expected returns are higher than the estimates of Claus and Thomas (2001).

2

estimation using the quarterly data from 1984:Q1 to 2009:Q4 generates an average real return to equity of 3.6%, similar to that obtained using the annual data from 1988 to 2009. This finding is also consistent with the theoretical models of Lettau, Ludvigson, and Wachter (2008) and Cogley and Sargent (2008), which emphasize a potential reduction of risk in the economy.

The model performs fairly well on a range of diagnostic tests. Comparing moments not used in the estimation with the corresponding data reveals that the model generates similar properties for the value-to-dividend ratio as the data. The standard deviation of returns implied by the estimates ranges from 7.0% to 11.6%, compared to about 18.3% in the data. The model also fares well on an external validation test, with the model-implied consumption series having similar moments as the actual data. In addition, one could use the parameter estimates and data on earnings to generate a model-implied conditional expected return series. Consistent with economic intuition, this series is countercyclical, with annual expected returns around 8 percent or higher in recessions. In comparison, the conditional expected return series reported in the literature exhibits less volatility (see Campbell and Thompson (2008) and Rapach, Strauss, and Zhou (2010)).

One limitation of the model is that it also generates a relatively high risk-free rate. As I argue later, this reflects the fact that standard production-based models cannot generate risk-free rates much below the economic growth rate. Thus, the approach used in the study cannot be used to provide an estimate of the equity premium. However, this does not invalidate using the model to estimate the expected real return to equity.

The findings have sharp implications for the investment decisions of investors. Current asset allocation advice is mostly based on properties of historical returns. A lower expected return to equity implies that individuals need to save more to fund retirement expenses. They may also need to reduce their allocation to equities in their portfolios. A lower expected return also impacts the actuarial calculation of pension funds and insurance firms, who base their decisions on annual expected nominal returns around 8%. In particular, it would further exacerbate the under-funding problems of state pension funds, as discussed by Novy-Marx and Rauh (2009).

This study is organized as follows. Section 2 presents the model relating earnings, dividends, and valuations. Section 3 discusses the data and the identification of the model parameters. Section 4 presents the results. Section 5 examines other implications of the model and Section 6 concludes.

2   Model

Consider an economy with a representative agent that provides a fixed supply of labor, L, at a wage rate, w_t, determined by market clearing conditions in the labor market. The economy consists of otherwise identical firms that produce differentiated goods $y_{j,t}$, where j indexes each good. The differentiated goods are combined to form an aggregate good, Y_t, using the Dixit-Stiglitz aggregator (see Dixit and Stiglitz (1977)) with inverse elasticity of substitution, ν,

$$Y_t = \left(\int y_{j,t}^{(1-\nu)} dj \right)^{\frac{1}{1-\nu}}.$$

Let $p_{j,t}$ denote the price of the jth good. The aggregate price index is given by

$$P_t = \left(\int p_{j,t}^{\frac{\nu-1}{\nu}} dj \right)^{\frac{\nu}{\nu-1}}.$$

Given this specification, one can derive the following demand function for each good:

$$p_{j,t} = y_{j,t}^{-\nu} Y_t^{\nu} P_t. \tag{1}$$

For simplicity, we subsequently normalize the aggregate price index P_t to be equal to 1, so that valuations and profits are in real terms.

2.1   Firm profits, dividends, and value

Each firm uses assets, $k_{j,t}$, and labor, $l_{j,t}$, to produce output. Let X_t denote the aggregate level of labor augmenting technology and ψ_t denote trend deviations of aggregate productivity. The output of each firm is given by:

$$y_{j,t} = \psi_t k_{j,t}^{\alpha} (X_t l_{j,t})^{1-\alpha}, \tag{2}$$

where α denotes the elasticity of output with respect to assets. The profits of each firm is given by maximizing with respect to the labor input,

$$\pi_{j,t} = \max_{l_{j,t}} p_{j,t} y_{j,t} - w_t l_{j,t}. \tag{3}$$

Using the equations for the price and output of each firm given in (1) and (2), Appendix A solves for the optimal labor choice and derives the following expression for the profits of each firm as a function of its assets and aggregate conditions:

$$\pi_{j,t} = a\psi_t^{\frac{1-\nu}{1-(1-\alpha)(1-\nu)}} X_t^{\frac{\nu}{1-(1-\alpha)(1-\nu)}} k_{j,t}^{\frac{\alpha(1-\nu)}{1-(1-\alpha)(1-\nu)}}, \tag{4}$$

where a is a constant term. The wage w_t will be set such that the labor market clears, i.e., $\int l_{j,t} dj = L$, where the labor demand is given in (A.4).

Each firm is financed through debt and equity. Denote the level of debt by $b_{j,t}$. Following much of the literature, the book leverage is assumed to be a constant, ϕ, implying that the level of debt $b_{j,t} = \phi k_{j,t}$. This debt is assumed to be riskless with an interest rate r. In unreported results, I examine the effect of allowing the interest rate to vary with aggregate conditions and find that this has very little effect on the estimates.

Each firm funds physical investment, $i_{j,t}$, from its cash flow. The asset accumulation equation is given by

$$k_{j,t+1} = k_{j,t}(1 - \delta) + i_{j,t},$$

where δ equals the depreciation rate. Each firm also faces a quadratic adjustment cost of investment given by $\lambda \frac{i_{j,t}^2}{2k_{j,t}}$, as in Hayashi (1982). Firms do not face any additional costs of disinvestment as in Abel and Eberly (1996).

Each firm uses its cash flow to fund investment, pay interest to creditors, pay taxes to the government, and pay dividends to shareholders. The taxes are paid on profits net of depreciation and interest expenses, as in the U.S. tax code. I consider a linear tax code with a tax rate, τ. Firms also face a fixed cost of operations $f_t = X_t f$ each period. The scaling with aggregate technology ensures that the fixed costs do not vanish over time. Given these assumptions, the dividend payout of the firm, $d_{j,t}$, is given by

$$d_{j,t} = (\pi_{j,t} - fX_t)(1 - \tau) - b_{j,t}(1 + r(1 - \tau)) + b_{j,t+1} + \delta k_{j,t}\tau - i_{j,t} - \frac{\lambda i_{j,t}^2}{2k_{j,t}}. \tag{5}$$

Let $M_{t,t+1}$ denote the pricing kernel of the economy, which we will parametrize subsequently. The value of each firm, $v(k_{j,t}, X_t, \psi_t)$, can be expressed as the solution to the following Bellman

equation:

$$v(k_{j,t}, X_t, \psi_t) = \max_{k_{j,t+1}, b_{j,t+1}, i_{j,t}} d_{j,t} + E\left[M_{t,t+1}v(k_{j,t+1}, X_{t+1}, z_{t+1})\right], \tag{6}$$

$$k_{j,t+1} = k_{j,t}(1-\delta) + i_{j,t},$$

$$X_{t+1} = X_t(1+\gamma),$$

where γ denotes the growth rate of aggregate technology and the expression for dividends is given above in (5).

In effect, each firm chooses the optimal dividend and investment policies to maximize firm value. The value of the firm depends on the endogenous dividend process, which itself is influenced by the pricing kernel, $M_{t,t+1}$. For instance, an increase in expected returns will lower the average pricing kernel, leading to both an increase in dividends and a decrease in firm value.

2.2 Aggregate firm value

As all the firms are otherwise identical and face only aggregate shocks, they will use the same inputs for $k_{i,t}$ and $l_{i,t}$.[4] Further, they will have the same investment and dividend policies. As such, one can rewrite the model in terms of aggregate variables. Let K_t, Π_t, B_t, I_t, D_t, V_t denote aggregate assets, profits, debt, investment, dividends and firm value, respectively. Then, aggregate profits can be written as

$$\Pi_t = A\psi_t^{\frac{1-\nu}{1-(1-\alpha)(1-\nu)}} X_t^{\frac{\nu}{1-(1-\alpha)(1-\nu)}} K_t^{\frac{\alpha(1-\nu)}{1-(1-\alpha)(1-\nu)}},$$

where A is another constant that incorporates integration constants. We can simplify the above expression by noting that the exponents of K_t and X_t sum to one. Let $\theta = \frac{\alpha(1-\nu)}{1-(1-\alpha)(1-\nu)}$ be the exponent on K_t. Then, some algebra yields that the exponent on X_t equals $1-\theta$. In addition, simplify further by renormalizing the trend deviation of productivity term, ψ, such that $z_t = \psi_t^{\frac{1-\nu}{1-(1-\alpha)(1-\nu)}}$. Thus, one can write aggregate profits as

$$\Pi_t = Az_t X_t^{1-\theta} K_t^{\theta}. \tag{7}$$

---

[4]Bilbiie, Ghironi, and Melitz (2012) examine business cycle fluctuations in a similar setting with symmetric firms producing heterogenous goods.

One important point to note here is that the exponent θ on K_t is not equal to α. Instead, θ is a function of both the asset share, α, and the inverse elasticity of substitution across goods, ν.

Using a similar aggregation argument, aggregate dividends are given by the following

$$D_t = \Pi_t(1-\tau) - fX_t(1-\tau) - B_t(1 + r(1-\tau)) + B_{t+1} + \delta K_t \tau - I_t - \frac{\lambda I_t^2}{2K_t}. \tag{8}$$

The value function for the aggregate firm is then given by

$$
\begin{aligned}
V(K_t, X_t, z_t) &= \max_{K_{t+1}, B_{t+1}, I_t} D_t + E\left[M_{t,t+1} V(K_{t+1}, X_{t+1}, z_{t+1})\right], \tag{9} \\
K_{t+1} &= K_t(1-\delta) + I_t, \\
X_{t+1} &= X_t(1+\gamma),
\end{aligned}
$$

where dividends are given in (8).

The assumption of differentiated products with a Dixit-Stiglitz aggregator implies that the value function for the aggregate firm exhibits decreasing returns to scale in K_t. This departs from standard models that assume constant returns to scale in K_t.[5] In addition to the empirical evidence in favor of downward sloping demand curves for firms' profits, this departure is helpful in the context of this study. It implies that firms will change their optimal scale following productivity shocks. This optimal response acts as an added source of firm value that would be absent otherwise. In addition, a constant returns to scale specification implies that valuation ratios are mainly driven by investment adjustment costs, an unappealing restriction for understanding variation in valuations. A decreasing returns to scale specification also gives the model additional flexibility in matching the averages of earnings and valuations, which can be challenging for constant returns to scale models. As I estimate θ, the degree of decreasing returns to scale required to best fit the data becomes a question to be resolved by the estimation.

---

[5]While a C.R.S. specification may be more suited for understanding variation in labor and output, the objective of this study to understand variation in profits and valuations.

2.3 Parametrization of the pricing kernel

The productivity term, z_t, measures deviations from trend growth levels and is assumed to follow an auto-regressive process with

$$\log(z_{t+1}) \;=\; \rho \log(z_t) + \epsilon_t, \tag{10}$$
$$\epsilon_t \;\sim\; N(0, \sigma),$$

where ϵ_t denotes shocks to aggregate productivity.

I parametrize the log pricing kernel as a log-linear function of current and future aggregate productivity:

$$\log(M_{t,t+1}) = -(1 + b_0)\log(1 + \gamma) - b_1(\log(z_{t+1}) - \log(z_t)) - b_2 \log(z_t). \tag{11}$$

The above parametrization is similar to the ones employed by Berk, Green, and Naik (1999) and Zhang (2005). This specification enables one to capture relevant features of pricing kernels in a parsimonious manner. The b_1 term captures the impact of changes in aggregate conditions on the pricing kernel. For instance, a negative value for b_1 implies that the agent places low valuations on transitions from high to low aggregate productivity. In addition, b_2 enables average expected returns to vary across current productivity states. These are economically meaningful features of the pricing kernel that also arise in consumption-based models. The above specification also nests a constant discount rate when $b_1 = b_2 = 0$. Note that as the long-run growth rate of the economy, γ, will be fixed in the estimation, the presence of the $\log(1 + \gamma)$ term in the constant term merely functions to rescale b_0.[6] Eliminating the $\log(1 + \gamma)$ term would have no effect on the results other than to change the estimate of b_0.

2.4 Mapping the model to a stationary economy

The value function given in equation (9) incorporates X_t, the trend productivity term. In order to simulate and estimate the model, the value function needs to be detrended into a stationary form.

---

[6]This particular specification helps the estimation by ensuring that the model converges for all values of b_0.

Denote the detrended variables using tildes. Thus,

$$\tilde{K}_t = \frac{K_t}{X_t}, \ \tilde{I}_t = \frac{I_t}{X_t}, \ \tilde{\Pi}_t = \frac{\Pi_t}{X_t}, \ \tilde{B}_t = \frac{B_t}{X_t}, \ \tilde{D}_t = \frac{D_t}{X_t}$$

and

$$\tilde{V}(\tilde{K}_t, z_t) = \frac{V(K_t, X_t, z_t)}{X_t}.$$

It is fairly straightforward to derive that the detrended value of the firm is given by the following Bellman equation:

$$
\begin{aligned}
\tilde{V}(\tilde{K}_t, z_t) &= \max_{\tilde{K}_{t+1}, B_{t+1}, \tilde{I}_t} \tilde{D}_t + E\left[M_{t,t+1}(1+\gamma)\tilde{V}(\tilde{K}_{t+1}, z_{t+1}) \right] && (12) \\
\tilde{D}_t &= \tilde{\Pi}(1-\tau) - f(1-\tau) - \tilde{B}_t(1 + r(1-\tau)) + \tilde{B}_{t+1}(1+\gamma) + \delta \tilde{K}_t \tau - \tilde{I}_t - \frac{\lambda \tilde{I}_t^2}{2\tilde{K}_t}. \\
(1+\gamma)\tilde{K}_{t+1} &= \tilde{K}_t(1-\delta) + \tilde{I}_t.
\end{aligned}
$$

Observe that the terms involving $t+1$ are now multiplied by $(1+\gamma)$. This arises due to the fact that when detrending $t+1$ variables by X_t, one needs to include a $\frac{X_{t+1}}{X_t} = 1 + \gamma$ term.

More importantly, this implies that the transformation of the economy into a stationary form involves an adjustment in the pricing kernel to take into account economic growth. In the absence of this adjustment, one would derive incorrect valuation ratios for a given set of parameter values. The intuition for this result is that observed data on valuations incorporate a component for economic growth. As such, when attempting to understand the relationship between earnings, dividends and valuations using a stationary model, one needs to incorporate the adjustment for growth. This adjustment for economic growth also implies that the model cannot generate low risk-free rates below the economic growth rate as it would violate the contraction mapping necessary for convergence.

3 Data and estimation

I employ the above model given in (12) to provide an estimate of the expected real return to equity. Fundamentally, data on income accruing to shareholders, dividends, investment and firm value can be used to estimate the structural parameters of the model. The expected return to equity can

then be derived from the estimated parameters for the pricing kernel. The intuition behind the estimation is that the expected return maps dividends onto valuations. Therefore, given data, one can infer the expected return.

This analysis provides an alternate perspective on the expected real return to equity that may differ from historical averages. While the hypothesis that future returns should simply reflect past returns is a compelling null hypothesis, there are robust arguments to the contrary. Cogley and Sargent (2008) argue that the Great Depression led to an increase in the market price of risk that has slowly dissipated over time. Lettau, Ludvigson, and Wachter (2008) argue that declining macroeconomic risk has led to a decline in the expected equity premium. These arguments motivate using the above model to provide a forward looking estimate of the expected return to equity.

That said, the recent recession may have resulted in a reexamination of the riskiness of equities with an associated change in expected returns. Unfortunately, the approach in this study cannot examine this question without data on a sufficient number of periods following the recession. However, much of the recent market turmoil was driven mainly by financial firms, and it is not necessarily given that this turmoil would have affected expected returns to nonfinancial firms, the focus of this study. These firms exhibit a quick rebound in earnings from 2009:Q2 onwards.

3.1 Data

The data for the estimation is obtained from the CRSP/Compustat merged data set. I estimate the model on both an annual data set and a quarterly data set. The sample periods for the annual and quarterly data sets extend from 1966 to 2009 and 1984:Q1 to 2009:Q4, respectively. The use of the quarterly data functions as a robustness check for the estimations with the annual data, particularly when the model is estimated annually over split samples. The start dates are determined by the availability of data on shareholders equity and dividends for the annual and quarterly data sets, respectively. The sample excludes financial firms and regulated utilities as the model would not be appropriate to use for such firms. The data sets for the representative firm are constructed by aggregating firm level data on total assets, capital expenditures, common dividends, income accruing to shareholders, total liabilities, corporate taxes and the book and market values of equity from the respective samples. Any firm with missing values for the market value of equity is excluded from the aggregation.

An alternative data set that could be used for such an estimation is the corporate sector income

statement and balance sheet from the U.S. Flow of Funds accounts.[7] However, the Flow of Funds data have one limitation that makes them less attractive for the purpose of this study. The Flow of Funds data is designed on a territorial basis — that is, only designed to reflect the U.S. operations of firms — and indeed reports fixed assets for all firms and financial assets for manufacturing firms on this basis using Census data. Due to data limitations, it reports financial assets for non-manufacturing firms on a consolidated basis using IRS data, which includes financial assets held abroad. As such, the Flow of Funds data does not provide the necessary like-for-like measure of the moments necessary for the estimation. Further, this gap in the data has likely widened over time as U.S. corporations have become increasingly global.

Table 1 reports the summary statistics of interest for the aggregate firm: Panels A and B, respectively, report values at annual and quarterly frequencies. All variables except the market-to-book ratio, leverage and interest expenses are constructed after scaling by lagged total assets.[8] The market-to-book ratio equals the market value of common equity divided by book equity. The mean value of earnings, measured as income accruing to shareholders, indicates that the aggregate earnings of firms are noticeably higher over the annual data period than over the quarterly data period. Aggregate dividends exhibit a similar pattern. In contrast, the aggregate valuation of firms is lower over the annual data period than over the quarterly data period.

Figure 1 plots the time series of earnings and valuations of the aggregate firm. This confirms the observation that, over time, the aggregate earnings of the selected firms have declined while their aggregate valuation has increased. Further, this effect does not arise from any particular outliers. While the previous literature has noted the increase in valuations over time, the finding of a decline in earnings has not been highlighted before, to the best of my knowledge.[9] One could reconcile higher valuations with lower earnings and dividends with a lower expected return, which, *ceteris paribus*, increases firm value by lowering the discount rate applied to future cash flows. The subsequent estimation examines this possibility more formally.

3.2   Calibrated parameters

The above model includes many auxiliary parameters, such as the corporate tax rate. One approach would be to include relevant information and estimate all of these parameters. Another approach

---

[7]McGrattan and Prescott (2005) use the aggregate data from Flow of Funds accounts for their analysis.

[8]This scaling transforms all the data into real terms under the assumption that the same deflator applies to all the data series.

[9]The Flow of Funds accounts also exhibit a similar decline in earnings over time.

would be to calibrate some of the model parameters to match the data, and estimate the rest. I take the second approach, which has the benefit of focusing the estimation on the parameters of interest, thereby improving the identification.

In the calibration, the depreciation rate, δ, is set equal to the mean ratio of aggregate depreciation to total assets over time. This ensures that the depreciation rate in the simulations matches that in the data. In a similar vein, the interest rate, r, equals the mean ratio of interest payable to total liabilities observed in the data. The linear tax rate equals mean aggregate taxes to taxable income. The calibrated tax rate, τ, equals 40.2%, close to the federal tax rate of 35% plus the average state tax rate of 4%. The leverage level, ϕ, is set to match the ratio of total liabilities to total assets. Finally, the fixed cost of operations, f, is set to match the observed ratio of selling, general and administration expenses to total assets.

The economic growth rate used in the estimation, γ, equals 1.99%, the per capita consumption growth rate over the sample period. The quarterly estimates assume a quarterly growth rate of 0.5%.

3.3   Estimation details

The remaining parameters of the model are estimated using the simulated method of moments (SMM) method introduced by Lee and Ingram (1991) and Duffie and Singleton (1993).[10] As detailed in Appendix C, for each parameter vector, the value function given in equation (12) is solved numerically and the solution is used to generate a set of simulated moments. The program then searches for the parameter vector that minimizes a quadratic form of the distance between the simulated moments and the data moments.

The estimation is carried out using a one-step GMM estimator, with the optimal weighting matrix given by the inverse of the covariance matrix of the data moments. For robustness, I examine the effect of using a two-step GMM estimator, where the weighting matrix for the second step is derived using the first step estimate. The two-step estimation leads to similar results as the one-step estimation.[11]

---

[10]Other studies that use this estimation method include Hennessy and Whited (2005), Cooper and Haltiwanger (2006), Hennessy and Whited (2007), Eberly, Rebelo, and Vincent (2008), Kogan, Livdan, and Yaron (2009), and Bloom (2009).

[11]Comparing the one-step and two-step estimators reveals that the first stage of the two-step estimator, which uses an identity weight matrix, matches each individual moment more closely than either the second stage of the two-step estimator or the one-step estimator, both of which use non-diagonal weighting matrices. In these cases, divergences between data and model moments are sometimes offset by off-diagonal terms in the weighting matrix.

3.4 Identification of model parameters

The matched moments include selected means, variances, autocorrelations and regression coefficients. Given the black box nature of the estimation, it is helpful to provide some intuition as to which moment helps identify which parameter.

The average earnings level helps pin down the curvature of the profit function, θ. This reflects the link between θ and the markup charged by the firm. The autocorrelation of dividends and earnings helps pin down the autocorrelation parameter, ρ. Similarly, the variances of earnings and earnings differences inform the volatility of the productivity shock process, σ. Due to the curvature of the value function, the volatility of the productivity measure also influences the mean market-to-book ratio. The mean dividend level helps pin down the adjustment cost parameter, λ. The resource constraint for the firm implies that dividends equals earnings minus the cost of investment. Given levels of earnings and investment, an increase in adjustment costs lowers the dividend payout.

Once the model has pinned down the above parameters, the average market-to-book ratio helps pin down the mean pricing kernel. Effectively, given earnings and dividends, a lower market-to-book ratio requires a lower mean pricing kernel. Finally, the expected return to equity equals the average realized return obtained from simulating the model given the estimated parameter values. The expected return is reported after transforming back into the economy with trend growth, so as to be comparable with data on real returns.

Although the above discussion focuses on selected mappings between the moments and the parameters, the estimation employs data on all the moments to pin down all the parameters. In the model simulations, a parameter change will directly or indirectly affect all the moments. For example, an increase in ρ leads to a higher average market-to-book ratio, in addition to its direct effect on the autocorrelation moments.

4 Results

This section presents the results from the estimation. The first subsection estimates the model using annual data from 1966 to 2009. The next two subsections, respectively, estimate the model on the annual data split into two subperiods and on quarterly data from 1984 onwards. This analysis helps examine whether average expected returns have changed over time. The last subsection examines an extension of the model that incorporates long-run changes in economic growth, motivated by

the long-run risks literature (see Bansal and Yaron (2004)).

4.1 Annual data

Panel A of Table 2 presents the parameter estimates from estimating the model on annual data from 1966 to 2009. The expected mean return to equity from the estimation equals 4.9%, lower than the long-run historical average real return of 6.5% used in the literature. The estimates indicate that the expected return to equity is lower than historical values, consistent with the findings of Fama and French (2002), Siegel (1999), Blanchard (1993), Claus and Thomas (2001), and Jagannathan, McGrattan, and Scherbina (2001). This suggests that historical realized returns may be higher than expected returns in the future. This could be due to either decreasing macroeconomic risk (see Lettau, Ludvigson, and Wachter (2008)), decreased market price of risk (see Cogley and Sargent (2008)) or changes in taxes (see McGrattan and Prescott (2005)).

The intuition for the estimate is that it equals the expected return at which data on earnings, dividends, and associated growth options are best reconciled with data on valuations. A higher expected return would lower the discounted present value of future dividend streams, resulting in a smaller aggregate market-to-book ratio. The precision of the estimate arises from the sensitivity of the model to changes in the expected return. Thus, the method employed in the study helps filter some of the noise in observed returns.

The point estimate for θ indicates significant curvature of the profit function.[12] The estimates for ρ and σ are mainly determined by the data on the earnings process, while the estimate for λ enables the model to match the data on dividends. Note that the investment adjustment cost is the only friction in the model, and the estimate reflects the gap between earnings, dividends and investment observed in the data.

Panel B of Table 2 presents the moments used in the estimation. The model moments are constructed using the parameter values reported in Panel A. The model successfully matches the first moments of earnings, dividends and the market-to-book ratio in the data. However, the model fails to match the average earnings to value ratio in the data, indicating that data exhibit a somewhat weaker relationship between firm value and earnings than in the model.[13]

---

[12]The estimate for θ is very close to the value of 0.57 obtained by setting $\alpha = 1/3$ and $\nu = 1/5$ into the definition of θ. This suggests that the θ estimate is consistent with accepted values of α and ν.

[13]The $\mathcal{J}$-statistic for a formal over-identification test of the model equals 60.6, indicating rejection of the model. This is a common occurrence with SMM. For example, the SMM estimations in seminal studies such as Cooper and Haltiwanger (2006) and Bloom (2009) also yield $\mathcal{J}$-statistics that imply rejection. The reason for this is that the moments used in SMM are typically identified very precisely in the data, and models cannot match these moments

The model matches the volatility of earnings differences, but not earnings themselves. This reflects the fact that, given a single shock process, the model has difficulty matching multiple variances.[14] The model matches the persistence of earnings and dividends, with dividends being smoother than earnings, as in the data. Further, the model generates a similar relationship between dividends and earnings as in the data, but fails to generate the relationship between investment and earnings in the data. The model finds investment to be sensitive to current earnings, while investment responds more to lagged earnings in the data, perhaps reflecting lead times associated with planning investment projects.

4.2   Annual data - split sample

Estimating the model after separating the sample into two periods enables one to examine whether the expected real return to equity has declined further over time. Table 3 presents the results from estimating the model after splitting the annual sample at its midpoint, 1987. Panel A reports the parameter estimates from the two samples, and Panel B reports the matched moments.

The split sample results provide clear evidence of a decline in the expected real return to equity, which declines from 6.6% in the twenty two years from 1966-87 to 3.7% from 1988-2009. This decline is both economically and statistically significant. This indicates that the decline in average earnings and the increase in valuations documented in Table 1 and Figure 1 translate to a noticeable decline in the expected return to equity.

One criticism that the reader may pose is whether the lower estimated expected returns reflect an inability of the model to generate high returns to equity. However, note that the estimated expected return for the period from 1966-87 is similar to the long-run historical average. This indicates that the model can generate high real returns to equity, and the lower estimates for the overall sample (and the latter sample period) are not an artifact of the model, but instead an estimate that reflects the data over these periods.

The gap between the estimated expected return to equity and realized returns may reflect a structural break in the economy or a slow moving change in expected returns.[15] Lettau, Lud-

---

with the precision necessary to generate low $\mathcal{J}$-statistics.

[14]For instance, the model also fails to match the observed volatility of the market-to-book ratio in the data. However, adding the volatility of the market-to-book ratio as a matched moment has little effect on the parameter estimates, as the covariance matrix used to weight the moments gives it little weight.

[15]See Pastor and Stambaugh (2001) and Kim, Morley, and Nelson (2005) for a Bayesian approach to structural breaks in equity returns. Lettau and Nieuwerburgh (2008) provide evidence of a structural break in the price-dividend ratio at about the same time.

vigson, and Wachter (2008) argue that the decline in equity premium arose from a reduction in macroeconomic volatility. Cogley and Sargent (2008) attribute the decline in the equity premium to slow moving changes in the market price of risk following the Great Depression. Bansal and Yaron (2004) argue that slow moving changes in consumption growth drives asset prices. Garleanu, Panageas, and Yu (2012) present a model in which technological change drives slow moving changes in expected returns over long horizons. Increased participation in the stock market may also lower the expected returns to equity. In contrast, Pastor and Veronesi (2009) provide an alternate view that the increase in equity values was a temporary phenomenon associated with the introduction and adoption of the Internet. In a related study, Pastor and Veronesi (2006) argue that the high observed valuations for the NASDAQ could be rationalized given uncertainty about future profitability. Ritter and Warr (2002) argue that the bull market in the 1990s may have been partly due to a decrease in inflation. While the results in this study cannot distinguish between these explanations, they suggest that investors should not simply rely on historical returns to equity when forming portfolio and savings choices.

The parameter estimates from the split samples are comparable to those obtained with the full sample. The model manages to fit the chosen moments quite well over the first period.[16] It is less successful at matching all the moments over the second period, indicating that the model has more difficulty generating the relationship between earnings, dividends and valuations observed over this period.

4.3   Quarterly data

The previous analysis suggests that expected returns to equity have declined noticeably over the past forty years. Another way to verify this result would be to examine whether the model estimated using quarterly data, available only from 1984:Q1 onwards, generates similar results to that obtained with annual data over the later sample period. As such, this functions as a robustness check for the annual results with the split samples.

Panels A and B of Table 4 present the estimated parameter values and the matched moments, respectively, from the estimation of the model using quarterly data from 1984:Q1 to 2009:Q4. The estimated expected real return to equity equals 3.6%, similar to the estimate obtained using the annual data for the period from 1988 onwards. The finding provides further evidence of a decline

---

[16]The one exception being the regression coefficients of investment on earnings and lagged earnings, where the data find a much stronger relationship for lagged earnings than for contemporaneous earnings, unlike the model.

in average expected returns to equity.

The estimates for the other parameters are driven by the matched moments and are, for the most part, comparable to the annual estimates. The estimate for the adjustment cost parameter, λ, jumps, mainly as a result of the quadratic specification. That is, λ has to be much higher in order to generate the same resource cost of investment with much lower average investment rates, since investment is now measured quarterly, not annually. As before, the mean expected return is fairly precisely estimated, while the model has difficulty pinning down some of the other pricing kernel parameters.

4.4   Long-run changes in economic growth

Bansal and Yaron (2004) emphasize long-run changes in consumption growth as a possible explanation for the equity premium. One could examine the effect of such long-run economic changes within the context of this study by incorporating an additional long-run growth term into the value function.

Let x_t denote a long-run growth term that takes values of either $\exp(-lt)$ or $\exp(lt)$, where lt denotes the deviation from trend growth. Let p_l and p_h denote the transition probabilities of the economy remaining in the low and high long-run growth states, respectively. The long-run growth term also impacts the pricing kernel. Denote the pricing kernel of the augmented model as

$$
\begin{aligned}
\log(M_{t,t+1}) \quad = \quad & -(1+b_0)\log(1+\gamma) - b_1(\log(z_{t+1}) - \log(z_t)) - b_2\log(z_t) \\
& -b_3(\log(x_t) - \log(x_{t+1})) - b_4(\log(x_t) - \mu_x),
\end{aligned}
\tag{13}
$$

where b_3, b_4 denote the parameters associated with the long-run growth terms.

Table 5 presents the results obtained from estimating the augmented model. Panel A presents the parameters of interest, while Panel B presents the matched moments. Consistent with the long-run risks literature, this model generates a higher expected return than the baseline model. However, the estimation fails to identify any of the augmented model parameters, leading to a large standard error for the expected return.

5 Model implications

This section evaluates the model along some dimensions not used in the estimation and presents the conditional expected returns implied by the model.

5.1 Model evaluation

The model presented in this study is primarily aimed towards estimating the expected return to equity. The reader may be interested in understanding how the model fares with regards to features of the data that were not employed in the estimation.[17] Such a comparison serves as an out-of-sample evaluation of the estimates.

Table 6 presents some statistics of interest obtained from the data and the model simulations. Panel A presents the results using the estimates obtained with the annual data reported in Table 2. Panel B presents the corresponding statistics using the results reported in Table 3. The data statistics are obtained from the corresponding data sets used in the estimations. The reported statistics relate to the key variables employed in the estimation: aggregate dividends, investment and market value.

The models generate an annualized equity return volatility of 7.0% to 11.6%, about half the value observed in the data. This indicates that although the model has a flexible pricing kernel, it has some difficulty generating the high equity return volatility observed in the data.

The log market value to dividend ratio from the model matches the data well with the annual estimates, but less so for the split sample periods. On the other hand, the model comes close to capturing the autocorrelation of the log market value to dividend ratio over all periods. As Bansal and Yaron (2007) find, the log value to dividend ratio exhibits less persistence than the per share price-dividend ratio.

The model matches the volatility of dividend growth over the longer period, but overshoots the data values for the two sub-periods. Reflecting the mean-reverting nature of the productivity function, the model generates a small negative autocorrelation for dividend growth. However, this autocorrelation takes a small positive value in the data. Finally, the model generates a somewhat smaller correlation between dividends and investment than in the data, which is perhaps surprising given the fact that the model has a stronger relationship between investment and contemporaneous earnings than in the data.

---

[17]I thank João Gomes for this suggestion.

A further external validity test of the model would be to compare the model-implied data with aggregate data from other sources. Panel C of Table 6 reports the results of comparing the model-implied consumption and investment series with the corresponding data from the National Income and Product Accounts (NIPA). The model-implied consumption series is derived as the sum of labor income from operations, fixed costs, interest income and dividend income. The consumption series from the NIPA is obtained as consumption on nondurables and services from NIPA Table 7.1. The investment series from the NIPA is obtained as private nonresidential fixed investment from the fixed asset tables.

As the table indicates the model matches the consumption growth moments quite well. Average consumption growth matches the data by construction, as the growth rate in the model is calibrated using consumption growth data. The fact that the model matches the volatility and persistence of consumption growth is noteworthy, as neither of these moments were used at all in the estimation. In addition, the model-implied investment series also matches the data fairly well. Overall, the ability of the model to match these external data moments helps validate the model.

One dimension along which the model fails to match the data is the risk-free rate. The implied risk-free rates are always above the assumed economic growth rates. As discussed in Section 2.4, production-based models with time-separable utility functions have substantial difficulty generating low risk-free rates once one accounts for economic growth. One natural criticism that follows is whether a model that does not match the low risk-free rate is helpful for understanding equity returns. However, Weil (1989) indicates that one can think of the low risk-free rate as a distinct puzzle separate from the equity premium puzzle. Further, Krishnamurthy and Vissing-Jorgensen (2012) argue that the low risk-free rates observed in the data partially reflect investor demand for the safety and liquidity of U.S. Treasury bonds.

5.2 Time variation in expected returns

The previous discussion focused on the mean expected return to equity. However, the model also implies that expected returns vary with the aggregate earnings of firms.[18] An examination of the model-implied conditional expected return provides another criterion for evaluating the estimates.

Figure 2 plots the model-implied conditional expected return to equity. Panel A plots the series constructed using the estimates from the annual data reported in Table 2; Panel B plots the series

---

[18]Chen, Novy-Marx, and Zhang (2010) present a three-factor cross-sectional asset pricing model that employs ROA as one factor.

constructed using the estimates from the split sample data reported in Table 3. I construct the expected return to equity for each of the earnings grid values used in the estimation. Interpolating the observed earnings data onto the conditional expected returns to equity at these grid points generates the model-implied expected return at each data point. In order to offset the decline in earnings observed in Figure 1, the expected returns using the full sample estimates are calculated after removing the difference in earnings over the two subperiods.[19] For the split sample series using annual data, I splice together the implied expected return from the sample from 1966 to 1987 with that from the sample from 1988 to 2009.

The figure demonstrates that the model generates noticeable variation in expected returns. Consistent with economic reasoning, the model-implied conditional expected returns reach their peak values during the recessions. In comparison, studies that use predictive regressions with scaled price ratios, such as Campbell and Thompson (2008) and Wachter and Warusawitharana (2011), find that conditional expected returns vary slowly over time, reflecting the slow movement in scaled price ratios. Rapach, Strauss, and Zhou (2010) combine together predicted equity premia using regressions on scaled price ratios and economic variables and find that conditional expected returns vary with the business cycle, as shown above. The conditional expected returns shown above are more volatile than those reported in the above studies, possibly reflecting the fact that expected returns are a function of only aggregate earnings, which can vary substantially with the business cycle. That said, the model-implied conditional expected returns take mostly positive values, satisfying the non-negativity criterion emphasized by Campbell and Thompson (2008).

The model-implied conditional expected return values are driven by the parameters b_1 and b_2. The negative estimate for b_1 implies that increases in productivity have high returns associated with them. As the productivity process mean reverts, productivity changes are expected to be positive when current productivity is low, leading to higher expected returns in these periods. A lower value for b_1 dampens the valuation variation associated with productivity changes, and therefore dampens the observed time variation in expected returns. Further, b_2 changes average expected returns across current states. Holding b_1 fixed, the positive estimate for b_2 implies that expected returns are higher in high productivity states, which mitigates the variation in conditional expected returns implied by the estimate for b_1.

Figure 3 examines the sensitivity of the reported time-variation in expected returns to changes

---

[19]Eliminating this difference makes the conditional expected returns reported in Panels A and B comparable. Constructing this figure using the unadjusted data generates similar cyclical variation. However, since this estimate assumes that expected return are, on average, unchanged over the entire period, it shows an upward trend in conditional expected returns as a result of the decline in earnings over this period.

in b_1 and b_2. The baseline parameter values are taken from the estimates reported in Table 2 that are used to construct Panel A of Figure 2. Panel A of Figure 3 plots the time-variation in expected returns obtained using the baseline estimates except with b_1 equal to half of its estimated value. This change substantially lowers the model-implied time-variation in expected returns, as the lower absolute value for b_1 implies that changes in productivity are associated with much smaller changes in returns. Panel B of Figure 3 plots the corresponding time-variation in expected returns obtained using the baseline estimates with b_2 equal to half of its estimated value. This change reduces the dampening effect of b_2, thereby increasing the time-variation in expected returns a bit.

6   Conclusion

This study estimates the expected real return to equity using a production-based asset pricing approach. The intuition underlying the approach is that the expected return provides a mapping from data on earnings, dividends and investment onto the valuation of firms. Given data on these variables and a dynamic optimizing model for their evolution over time, one can estimate the expected return to equity that generates the best fit for the model.

Using this approach, I obtain estimates for the expected real return to equity of about 4.9%. These results are lower than mean historical returns to equity, indicating that investors expect future returns to equity to be lower than past returns. This finding is consistent with other studies in the literature. In addition, the results also indicate a further drop in average expected returns over time.

One interesting question that this study cannot answer is whether the decline in expected returns reflects a permanent change or a temporary regime shift. It is quite possible that, in light of the recent financial crisis, expected returns to equity have increased again as investors revise upward their views on the riskiness of equities. That said, some commentators have suggested that stock returns may remain low for a long period following the financial crisis. The method in this study cannot evaluate this question without data on a sufficient sample period following the crisis.

Appendix

A   Labor demand and firm profits

Substitute the demand function for the firm's products (1), after normalizing $P_t = 1$, into its profit function (3) to obtain

$$\pi_{j,t} = \max_{l_{j,t}} y_{j,t}^{1-\nu} Y_t^{\nu} - w_t l_{j,t}. \tag{A.1}$$

The first order conditions with respect to $l_{j,t}$ imply that

$$w_t = (1-\nu) y_{j,t}^{-\nu} Y_t^{\nu} \frac{\partial y_{j,t}}{\partial l_{j,t}}.$$

From (2), the derivative of output with respect to labor equals

$$\frac{\partial y_{j,t}}{\partial l_{j,t}} = (1-\alpha) \frac{y_{j,t}}{l_{j,t}}.$$

Substituting the above two expression, one can rewrite the first order condition for labor and the implied profits of the firm, respectively, as:

$$w_t l_{j,t} = (1-\alpha)(1-\nu) y_{j,t}^{1-\nu} Y_t^{\nu}, \tag{A.2}$$

and

$$\pi_{j,t} = \frac{(1 - (1-\nu)(1-\alpha))}{(1-\nu)(1-\alpha)} w_t l_{j,t}. \tag{A.3}$$

Substituting the expression for output (2) into (A.2) and rearranging terms, one can write the optimal labor choice as

$$l_{j,t} = \left[\frac{(1-\alpha)(1-\nu)}{w_t} Y_t^{\nu} \psi_t^{(1-\nu)} k_{j,t}^{\alpha(1-\nu)} X_t^{(1-\alpha)(1-\nu)} \right]^{\frac{1}{1-(1-\alpha)(1-\nu)}} \tag{A.4}$$

Substituting the above expression into (A.3) one can write the profits of the firm as

$$\pi_{j,t} = (1-(1-\nu)(1-\alpha))((1-\nu)(1-\alpha))^{\frac{(1-\nu)(1-\alpha)}{1-(1-\nu)(1-\alpha)}} \left[\left(\frac{Y_t}{X_t}\right)^{\nu} X_t^{\nu} \psi_t^{(1-\nu)} k_{j,t}^{\alpha(1-\nu)} \left(\frac{X_t}{w_t}\right)^{(1-\nu)(1-\alpha)} \right]^{\frac{1}{1-(1-\nu)(1-\alpha)}}.$$

Note that along a balanced growth path, the ratios of aggregate technology to real wages and aggregate output to technology are constants. Thus, one can gather the constants into a single term a, and rewrite the above profit function as

$$\pi_{j,t} = a\psi_t^{\frac{1-\nu}{1-(1-\alpha)(1-\nu)}} X_t^{\frac{\nu}{1-(1-\nu)(1-\alpha)}} k_{j,t}^{\frac{\alpha(1-\nu)}{1-(1-\nu)(1-\alpha)}}, \qquad (A.5)$$

where the power coefficients on X_t and $k_{j,t}$ sum to 1.

B   Data appendix

This appendix details the firm-level data used in the construction of the aggregate nonfinancial firm. It also provides the respective Compustat Fundamentals Annual data set variable names in parenthesis.

Assets - Total assets (AT)

Liabilities - Total liabilities (LT) plus preferred equity capital (PSTK)

Shareholders equity - Common and preferred equity capital (SEQ) minus preferred equity capital (PSTK)

Earnings - Income before extraordinary items available for common shareholders (IBCOM)

Investment - Funds used for property, plant and equipment, excluding acquisitions (CAPX)

Market value - Number of common shares outstanding (CSHO) times share price at the close of the fiscal year (PRCC_F)

Interest expenses - Interest expenses and financing charges on debt (XINT)

Corporate taxes - All income taxes imposed by federal, state and foreign governments (TXT)

Depreciation - Depreciation and amortization (DP)

Fixed costs - Selling, general and administrative expense (XSGA).

C   Numerical solution

The simulations require a numerical solution of the value function for the aggregate firm. The capital grid has 150 points and the productivity grid has 10 points. The capital grid is centered around an approximation of the median size of the firm given the parameters. Let β denote the risk-free rate implied by the pricing kernel. The approximate value of the steady state capital stock

$(\hat{k})$ is given by the following:

$$
\begin{aligned}
\hat{k} &= \left(\frac{\theta \mu_z (1-\tau)}{\text{usr}}\right)^{1/(1-\theta)}, \\
\mu_z &= \exp\left(0.5\sigma^2/(1-\rho^2)\right), \\
\text{usr} &= 1/\beta - 1 + \delta - \tau\delta + \lambda(\delta+\gamma)(1/\beta - 1 + \delta - .5(\delta+\gamma)),
\end{aligned}
$$

where μ_z and usr denote the mean productivity level and the steady state Jorgensonian user cost of capital, respectively.[20] A check reveals that the steady state firm size obtained from the simulations lies very close to this approximation. The profit grid is formed using the quadrature method of Tauchen and Hussey (1991).

The simulated sample is generated using the value and policy functions for the aggregate firm. The law of motion for productivity is generated directly using the transition equations (10). The simulation is run for 20,000 years, with the initial 10,000 discarded as a burn-in sample. The value of the quadratic form of the distance between the data moments and the simulated moments is computed for each simulation. The program searches for the parameters that minimize this distance using the simulated annealing algorithm. Each estimation involved evaluating more than 100,000 candidate parameter sets and took a day or two of computing time.

---

[20]The details of this calculation are available from the author.

References

Abel, Andrew, and Janice C. Eberly, 1996, Optimal investment with costly reversibility, *Review of Economic Studies* 63, 581–593.

Bansal, Ravi, and Amir Yaron, 2004, Risks for the long-run: A potential resolution of asset pricing puzzles, *Journal of Finance* 59, 1481–1509.

Bansal, Ravi, and Amir Yaron, 2007, The asset pricing-macro nexus and return-cash flow predictability, Working paper, Duke University.

Berk, Jonathan B., Richard C. Green, and Vasant Naik, 1999, Optimal investment, growth options, and security returns, *Journal of Finance* 54, 1553–1607.

Bilbiie, Florin O., Fabio Ghironi, and Marc J. Melitz, 2012, Endogenous entry, product variety, and business cycles, *Journal of Political Economy* 120, 304–345.

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

Bloom, Nick, 2009, The impact of uncertainty shocks, *Econometrica* 77, 623–685.

Campbell, John Y., and Samuel Thompson, 2008, Predicting excess stock returns out of sample: Can anything beat the historical average?, *Review of Financial Studies* 21, 1509–1531.

Chen, Long, Robert Novy-Marx, and Lu Zhang, 2010, An alternative three-factor model, Working paper, Washington University in St. Louis.

Claus, James, and Jacob Thomas, 2001, Equity premia as low as three percent? Evidence from analysts' earnings forecasts for domestic and international stock markets, *Journal of Finance* 56, 1629–1666.

Cochrane, John H., 1991, Production-based asset pricing and the link between stock returns and economic fluctuations, *Journal of Finance* 46, 209–237.

Cochrane, John H., 1996, A cross-sectional test of an investment-based asset pricing model, *Journal of Political Economy* 104, 572–621.

Cogley, Timothy, and Thomas J. Sargent, 2008, The market price of risk and the equity premium: A legacy of the Great Depression?, *Journal of Monetary Economics* 55, 454–476.

Cooper, Russell W., and John Haltiwanger, 2006, On the nature of capital adjustment costs, *Review of Economic Studies* 73, 611–633.

Dixit, Avinash K., and Joseph E. Stiglitz, 1977, Monopolistic competition and optimum product diversity, *Americal Economic Review* 67, 297–308.

Duffie, Darrell, and Kenneth J. Singleton, 1993, Simulated moments estimation of Markov models of asset prices, *Econometrica* 61, 929–952.

Easton, Peter D., and Gregory A. Sommers, 2007, Effect of analysts' optimism on estimates of the expected return implied by earnings forecasts, *Journal of Accounting Research* 45, 983–1016.

Eberly, Janice, Sergio Rebelo, and Nicolas Vincent, 2008, Investment and value: A neoclassical benchmark, NBER working paper #13866.

Fama, Eugene F., and Kenneth R. French, 2002, The equity premium, *Journal of Finance* 57, 637–659.

Fernandez, Pablo, and Javier Del Campo Baonza, 2010, Market risk premium used in 2010 by analysts and companies: A survey with 2400 answers, Working paper.

Foreman, E. K., 1991, *Survey Sampling Principles*. (Marcel Dekker New York, NY).

Garleanu, Nicolae, Stavros Panageas, and Jianfeng Yu, 2012, Technological growth and asset pricing, *Journal of Finance* 67, 1265–1292.

Gomes, Francisco, and Alexander Michaelides, 2005, Optimal life-cycle asset allocation: Understanding the empirical evidence, *Journal of Finance* 60, 869–904.

Gomes, Joao, Leonid Kogan, and Lu Zhang, 2003, Equilibrium cross section of returns, *Journal of Political Economy* 111, 693–732.

Gomes, Joao, Amir Yaron, and Lu Zhang, 2006, Asset pricing implications of firms' financing constraints, *Review of Financial Studies* 19, 1321–1356.

Gordon, Myron J., 1959, Dividends, earnings and stock prices, *Review of Economics and Statistics* 41, 99–105.

Graham, John R., and Campbell Harvey, 2005, The long-run equity premium, *Finance Research Letters* 2, 185–194.

Hayashi, Fumio, 1982, Tobin's marginal q and average q: A neoclassical interpretation, *Econometrica* 50, 213–224.

Hennessy, Christopher A., and Toni M. Whited, 2005, Debt dynamics, *Journal of Finance* 60, 1129–1165.

Hennessy, Christopher A., and Toni M. Whited, 2007, How costly is external financing? Evidence from a structural estimation, *Journal of Finance* 62, 1705–1745.

Jagannathan, Ravi, Ellen R. McGrattan, and Anna Scherbina, 2001, The declining U.S. equity premium, *Federal Reserve Bank of Minneapolis Quarterly Review* 24, 3–19.

Jermann, Urban, 1998, Asset pricing in production economies, *Journal of Monetary Economics* 41, 257–275.

Jermann, Urban, 2011, A production-based model for the term structure, Working paper, University of Pennsylvania.

Kim, Chang-Jin, James C. Morley, and Charles R. Nelson, 2005, The structural break in the equity premium, *Journal of Business and Economic Statistics* 23, 181–191.

Kogan, Leonid, 2004, Asset prices and real investment, *Journal of Financial Economics* 73, 411–431.

Kogan, Leonid, Dmitry Livdan, and Amir Yaron, 2009, Oil price futures in a production economy with investment constraints, *Journal of Finance* 64, 1345–1375.

Krishnamurthy, Arvind, and Annette Vissing-Jorgensen, 2012, The aggregate demand for Treasury debt, *Journal of Political Economy* 120, 233–267.

Larrain, Borja, and Motohiro Yogo, 2008, Does firm value move too much to be justified by subsequent changes in cash flow?, *Journal of Financial Economics* 87, 200–226.

Lee, Bong-Soo, and Beth F. Ingram, 1991, Simulation estimation of time-series models, *Journal of Econometrics* 47, 197–205.

Lettau, Martin, Sydney C. Ludvigson, and Jessica A. Wachter, 2008, The declining equity premium: What role does macroeconomic risk play?, *Review of Financial Studies* 21, 1653–1687.

Lettau, Martin, and Stijn Van Nieuwerburgh, 2008, Reconciling the return predictability evidence, *Review of Financial Studies* 21, 1607–1652.

Liu, Laura Xiaolei, Toni Whited, and Lu Zhang, 2009, Investment-based expected stock returns, *Journal of Political Economy* 117, 1105–1139.

McGrattan, Ellen R., and Edward Prescott, 2005, Taxes, regulations, and the value of U.S. and U.K. corporations, *Review of Economic Studies* 72, 767–796.

Merton, Robert C., 1980, On estimating the expected return on the market: An exploratory investigation, *Journal of Financial Economics* 8, 323–361.

Novy-Marx, Robert, and Joshua D. Rauh, 2009, The liabilities and risks of state-sponsored pension plans, *Journal of Economic Perspectives* 23, 191–210.

Pastor, Lubos, Meenakshi Sinha, and Bhaskaran Swaminathan, 2008, Estimating the intertemporal risk-return tradeoff using the implied cost of capital, *Journal of Finance* 63, 2859–2897.

Pastor, Lubos, and Robert F. Stambaugh, 2001, The equity premium and structural breaks, *Journal of Finance* 56, 1207–1239.

Pastor, Lubos, and Pietro Veronesi, 2006, Was there a NASDAQ bubble in the late 1990s?, *Journal of Financial Economics* 81, 61–100.

Pastor, Lubos, and Pietro Veronesi, 2009, Technological revolutions and stock prices, *American Economic Review* 99, 1451–83.

Polkovnichenko, Valery, 2007, Life-cycle portfolio choice with additive habit formation preferences and uninsurable labor income risk, *Review of Financial Studies* 20, 83–124.

Rapach, David E., Jack K. Strauss, and Guofu Zhou, 2010, Out-of-sample equity premium prediction: Combination forecasts and links to the real economy, *Review of Financial Studies* 23, 821–862.

Ritter, Jay R., and Richard S. Warr, 2002, The decline of inflation and the bull market of 1982-1999, *Journal of Financial and Quantitative Analysis* 37, 29–61.

Siegel, Jeremy J., 1999, The shrinking equity premium, *Journal of Portfolio Management* 26, 10–17.

Tauchen, George, and Robert Hussey, 1991, Quadrature-based methods for obtaining approximate solutions to nonlinear asset pricing models, *Econometrica* 59, 371–396.

Wachter, Jessica A., and Missaka Warusawitharana, 2011, What is the chance that the equity premium varies over time? Evidence from predictive regressions, NBER Working paper #17334.

Weil, Philippe, 1989, The equity premium puzzle and the risk-free rate puzzle, *Journal of Monetary Economics* 24, 402–421.

Zhang, Lu, 2005, The value premium, *Journal of Finance* 60, 67–103.

Table 1: Summary statistics

The table reports the summary statistics for the aggregate firm constructed by aggregating firm level data from the CRSP/Compustat merged data set at annual and quarterly frequencies. The sample periods for the annual and quarterly date are from 1966 to 2009 and 1984:Q1 to 2009:Q4, respectively. 'St. dev.' denotes standard deviation. The earnings variable measures income accruing to common shareholders scaled by lagged total assets. Market-to-book equals the aggregate market value of equity scaled by the book value of equity. Leverage equals total liabilities divided by total book assets. Interest costs are scaled by total liabilities. Earnings, dividends, investment, corporate taxes and depreciation are all scaled by lagged total assets and reported as percentages.

	Annual data		Quarterly data	
Variable	Mean	St. dev.	Mean	St. dev.
Earnings	5.42	1.49	1.15	0.35
Market-to-book	2.19	0.82	2.87	0.89
Dividends	2.32	0.61	0.47	0.10
Investment	8.79	2.03	1.73	0.36
Leverage	0.59	0.07	0.64	0.03
Interest costs	4.03	1.02	0.86	0.23
Corporate taxes	4.05	1.65	0.70	0.19
Depreciation	5.38	0.39	1.19	0.09

Table 2: Annual data

Panel A reports the parameters values obtained from estimating the model on the aggregate firm using annual data. Panel B reports the corresponding moment values from the data and the model. The data moments are obtained using data on the aggregate firm constructed by summing the variables across all firms in a given year. Earnings, dividends and investment are all reported as percentages. The sample period is from 1966 to 2009. The annual economic growth rate, γ, is assumed to be 1.99%. The expected real return to equity, $E[r]$, equals the average equity return in the simulated data set constructed using the estimated parameter values. The standard error for the expected real return to equity is computed using the delta method.

Panel A: Parameters

Parameter	θ	ρ	σ	λ	b_0	b_1	b_2	$E[r]$
Estimate	0.59	0.60	0.044	5.40	1.34	-1.65	0.25	4.94
Standard error	(0.02)	(0.03)	(0.005)	(0.38)	(0.10)	(0.71)	(0.37)	(0.20)

Panel B: Moments

Moment	Data	Model
Average of		
Earnings	5.42	5.29
Market-to-book	2.19	2.32
Dividends	2.32	2.38
Earnings to value	7.03	5.54
Variance of		
Earnings	2.17	1.01
Earnings difference	0.97	0.80
Autocorrelation of		
Earnings	0.75	0.60
Dividends	0.85	0.81
Earnings difference	0.16	-0.19
Regression coefficient of		
Investment on earnings	0.24	0.83
Investment on lagged earnings	0.81	-0.16
Dividends on earnings	0.15	0.26
Dividends on lagged earnings	0.19	0.14

Table 3: Annual data - split sample

Panel A reports the parameters values obtained from estimating the model on the aggregate firm for two annual sub-samples. Panel B reports the corresponding moment values from the data and the model. The data moments are obtained using data on the aggregate firm constructed by summing the variables across all firms in a given year. Earnings, dividends and investment are all reported as percentages. The sub-sample periods are from 1966 to 1987 and 1988 to 2009, respectively. The annual economic growth rate, γ, is assumed to be 1.99%. The expected real return to equity, $E[r]$, equals the average equity return in the simulated data set constructed using the estimated parameter values. The standard error for the expected real return to equity is computed using the delta method.

Panel A: Parameters

Sample	Parameter	θ	ρ	σ	λ	b_0	b_1	b_2	$E[r]$
1966 - 1987	Estimate	0.68	0.71	0.050	9.10	1.95	-1.97	0.16	6.62
	Standard error	(0.02)	(0.04)	(0.003)	(0.62)	(0.08)	(0.68)	(0.26)	(0.23)
1988 - 2009	Estimate	0.55	0.66	0.042	10.22	0.73	-1.68	0.17	3.69
	Standard error	(0.01)	(0.02)	(0.009)	(0.66)	(0.14)	(2.48)	(0.94)	(0.16)

Panel B: Moments

Moment	1966 - 1987		1988 - 2009	
	Data	Model	Data	Model
Average of				
Earnings	6.22	6.26	4.61	5.15
Market-to-book	1.58	1.62	2.79	3.78
Dividends	2.71	2.80	1.92	1.82
Earnings to value	9.12	8.22	4.94	3.84
Variance of				
Earnings	1.34	2.08	1.72	0.91
Earnings difference	0.83	1.12	1.11	0.62
Autocorrelation of				
Earnings	0.66	0.73	0.68	0.66
Dividends	0.81	0.85	0.76	0.82
Earnings difference	-0.03	-0.14	0.25	-0.16
Regression coefficient of				
Investment on earnings	0.27	0.54	-0.11	0.68
Investment on lagged earnings	0.71	-0.22	0.38	-0.11
Dividends on earnings	0.15	0.39	0.10	0.27
Dividends on lagged earnings	0.20	0.25	0.11	0.12

Table 4: Quarterly data

Panel A reports the parameters values obtained from estimating the model on the aggregate firm. Panel B reports the corresponding moment values from the data and the model. The data moments are obtained using data on the aggregate firm constructed by summing the variables across all firms in a given quarter. Earnings, dividends and investment are all reported as percentages. The sample period is from 1984:Q1 to 2009:Q4. The quarterly economic growth rate, γ, is assumed to be 0.50%. The annualized expected real return to equity, $E[r]$, equals the average equity return in the simulated data set constructed using the estimated parameter values. The standard error for the expected real return to equity is computed using the delta method.

Panel A: Parameters

Parameter	θ	ρ	σ	λ	b_0	b_1	b_2	$E[r]$
Estimate	0.47	0.85	0.026	23.50	0.67	-0.73	-0.09	3.56
Standard error	(0.01)	(0.06)	(0.004)	(0.83)	(0.06)	(1.23)	(0.13)	(0.23)

Panel B: Moments

Moment	Data	Model
Average of		
Earnings	1.15	1.00
Market-to-book	2.87	4.20
Dividends	0.47	0.47
Earnings to value	0.50	0.66
Variance of		
Earnings	0.12	0.04
Earnings difference	0.05	0.01
Autocorrelation of		
Earnings	0.80	0.84
Dividends	0.64	0.89
Earnings difference	-0.08	-0.08
Regression coefficient of		
Investment on earnings	-0.25	0.55
Investment on lagged earnings	0.42	-0.20
Dividends on earnings	0.11	0.58
Dividends on lagged earnings	0.00	0.15

Table 5: Long-run changes in growth

Panel A reports the parameters values obtained from estimating the model with long-run changes in growth on the aggregate firm. Panel B reports the corresponding moment values from the data and the model. The data moments are obtained using data on the aggregate firm constructed by summing the variables across all firms in a given year. Earnings, dividends and investment are all reported as percentages. The sample period is from 1966 to 2009. The mean annual economic growth rate, γ, is assumed to be 1.99%. The expected real return to equity, $E[r]$, equals the average equity return in the simulated data set constructed using the estimated parameter values. The standard error for the expected real return to equity is computed using the delta method.

Panel A: Parameters

Parameter	θ	ρ	σ	λ	b_0	b_1	b_2	$E[r]$
Estimate	0.50	0.62	0.031	5.28	0.90	-1.72	0.09	5.64
Standard error	(3.47)	(0.61)	(0.192)	(17.00)	(19.29)	(52.44)	(89.16)	(44.62)
Additional parameters								
	lt	p_l	p_h	b_3	b_4			
Estimate	0.04	0.97	0.33	-0.49	-0.39			
Standard error	(0.84)	(2.31)	(3.360)	(68.29)	(11.05)			

Panel B: Moments

Moment	Data	Model
Average of		
Earnings	5.42	5.61
Market-to-book	2.19	2.03
Dividends	2.32	2.50
Earnings to value	7.03	6.73
Variance of		
Earnings	2.17	0.95
Earnings difference	0.97	0.78
Autocorrelation of		
Earnings	0.75	0.59
Dividends	0.85	0.68
Earnings difference	0.16	-0.21
Regression coefficient of		
Investment on earnings	0.24	0.92
Investment on lagged earnings	0.81	-0.18
Dividends on earnings	0.15	-0.31
Dividends on lagged earnings	0.19	0.24

Table 6: Model evaluation

The table reports statistics of interest from the data and the model simulations. Panel A reports values from the estimation using annual data from 1966 to 2009 reported in Table 2. Panel B reports values from the estimation using the split sample results reported in Table 3. Panel C presents the findings from an external validation exercise using the estimates from Table 2, where the moments of the model-implied consumption and investment series are compared with the corresponding moments for data on consumption and investment from the National Income and Product Accounts. The data values represent the corresponding statistics from the actual data sets, while the model values are obtained from the simulated data sets given the estimated parameter values. The value-dividend ratio equals the log of the aggregate market value divided by aggregate dividends.

Panel A: Full sample results

| Statistic	Data	Model
Standard deviation of market return	18.3	7.0
Average of value-dividend ratio	3.687	3.656
Standard deviation of value-dividend ratio	0.43	0.15
Autocorrelation of value-dividend ratio	0.91	0.85
Standard deviation of dividend growth	0.08	0.12
Autocorrelation of dividend growth	0.10	-0.23
Correlation between dividends and investment	0.75	0.38

Panel B: Split sample results

	1966-1987		1988-2009	
Statistic	Data	Model	Data	Model
Standard deviation of market return	17.7	11.6	19.1	7.80
Average of value-dividend ratio	3.379	3.292	3.994	4.29
Standard deviation of value-dividend ratio	0.195	0.27	0.37	0.44
Autocorrelation of value-dividend ratio	0.65	0.86	0.79	0.80
Standard deviation of dividend growth	0.05	0.18	0.10	0.31
Autocorrelation of dividend growth	0.19	-0.13	0.18	-0.26
Correlation between dividends and investment	0.52	0.10	0.43	0.38

Panel C: External validation

Statistic	NIPA data	Model
Average of consumption growth	1.87	1.92
Standard deviation of consumption growth	1.36	1.51
Autocorrelation of consumption growth	0.49	0.60
Average of investment	10.53	8.87
Standard deviation of investment	1.05	0.91
Autocorrelation of investment	0.76	0.60

Figure 1: Aggregate earnings and valuations

Panel A: Earnings

Panel B: Valuations
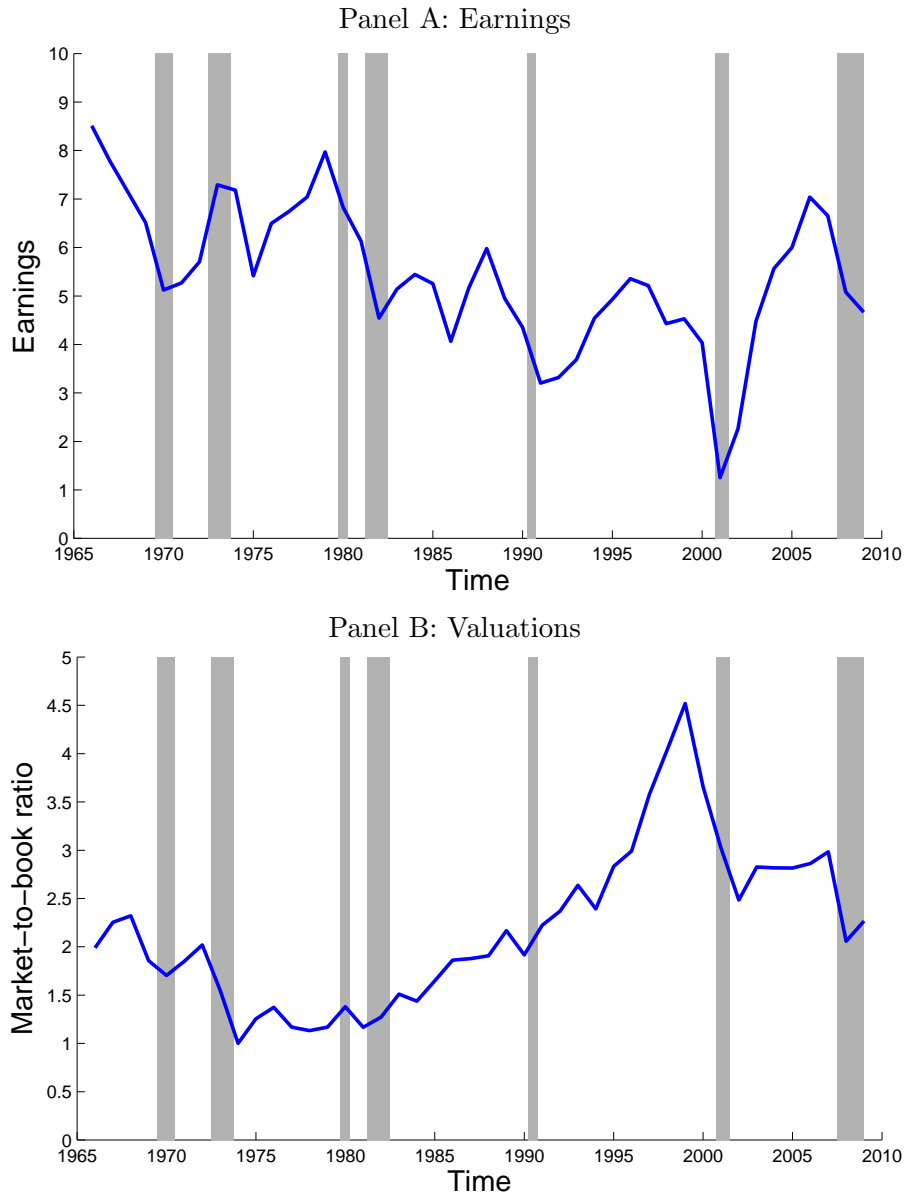
Panel A plots the aggregate earnings of nonfinancial corporations as constructed from the CRSP/Compustat merged data set. Earnings are measured as aggregate income accruing to shareholders divided by aggregate lagged total assets. Panel B plots the aggregate valuation of nonfinancial corporations, where the aggregate valuation is measured as the aggregate market-to-book ratio. The sample period is from 1966 to 2009 and the data are annual. The shaded areas correspond to NBER recession periods.
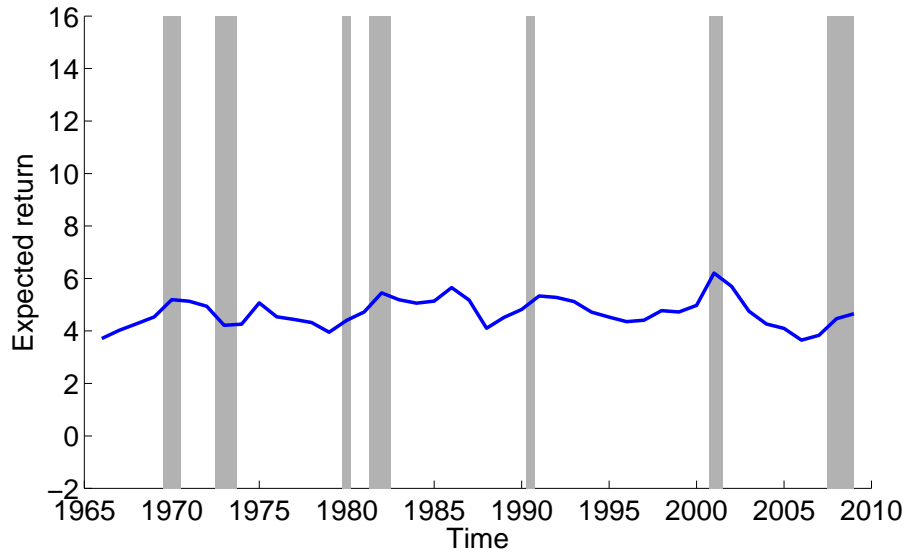
Figure 2: Time variation in expected returns
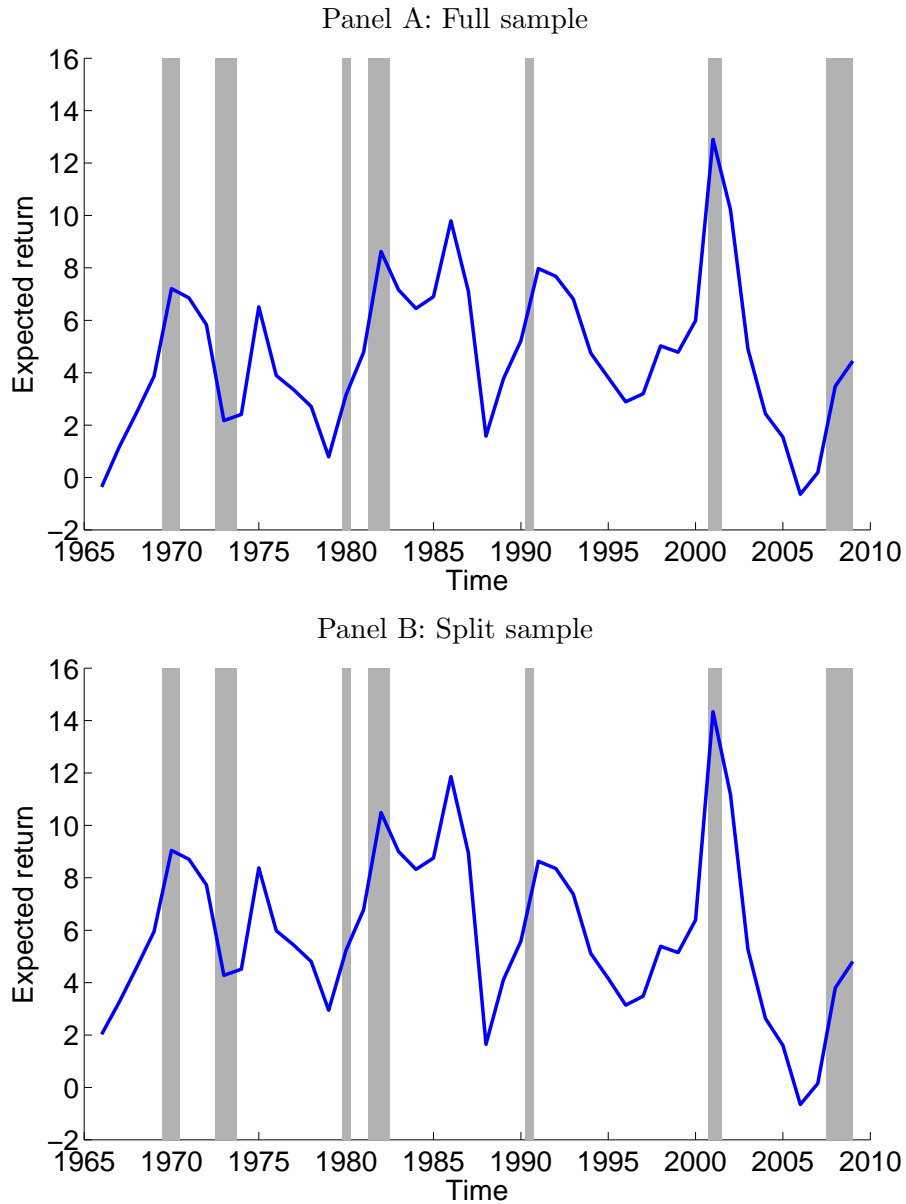
Panel A: Full sample
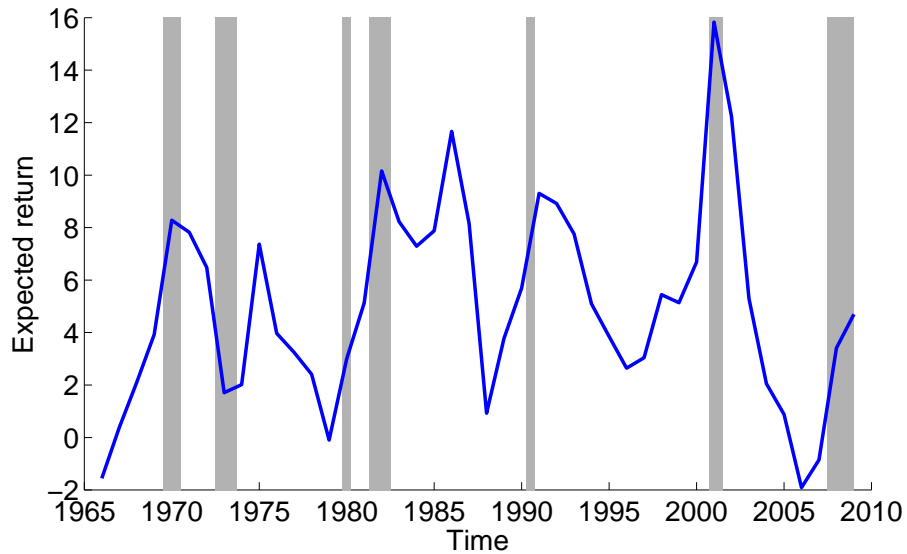

Panel B: Split sample

The figure plots the model-implied annualized conditional expected return to equity. Panel A plots the expected return from the estimates obtained using annual data reported in Table 2. Panel B plots the corresponding series from the split sample results reported in Table 3. The expected return to equity varies with the current value for income accruing to shareholders scaled by lagged total assets. The conditional expected returns in Panel A are calculated after adjusting for changes in average earnings over the two periods. The shaded areas correspond to NBER recession periods.

Figure 3: Sensitivity of conditional expected returns

Panel A: Lower absolute value for b_1 compared to Panel A of Figure 2

Panel B: Lower absolute value for b_2 compared to Panel A of Figure 2

The figure plots the model-implied annualized conditional expected return to equity for various parameter values. Panel A plots the conditional expected return from the estimates reported in Table 2 except for $b_1 = -0.82$, which equals half the point estimate. Panel B plots the conditional expected return from the estimates reported in Table 2 except for $b_2 = 0.12$, also half the point estimate. The shaded areas correspond to NBER recession periods.

38