Online Appendix for
Price-Earnings Ratios:
Growth and Discount Rates

Andrew Ang*
Columbia University and NBER

Xiaoyan Zhang†
Purdue University

This Version: 19 May 2011

*Columbia Business School, 3022 Broadway 413 Uris, New York NY 10027, ph: (212) 854-9154; email: aa610@columbia.edu; WWW: http://www.columbia.edu/~aa610.
†Krannert School of Management, Purdue University, West Lafayette, IN 47907. Ph: (765) 496-7674; email: zhang654@purdue.edu WWW: http://web.ics.purdue.edu/~zhang654/
1 Model

1.1 Price-Earnings Ratios

We start with a standard present value model where the stock price, \( P_t \), is given by the discounted value of future dividends:

\[
P_t = E_t \left[ \sum_{i=1}^{\infty} \exp \left( -\sum_{j=0}^{i-1} \delta_{t+j} \right) D_{t+i} \right],
\]

(1)

where \( D_t \) is the dividend paid at time \( t \), and the time interval \( t \) to \( t+1 \) represents one year. The discount rate, \( \delta_t \), discounts cashflows at \( t+1 \) back to time \( t \), \( \delta_t = \ln E_t \left[ (P_{t+1} + D_{t+1})/P_t \right] \).

We can transform equation (1) into a statement about earnings by introducing the log payout ratio, \( po_t \):

\[
po_t = \ln \left( \frac{D_t}{EA_t} \right),
\]

(2)

where \( EA_t \) are aggregate earnings reported at time \( t \) representing earnings from \( t-1 \) to time \( t \). Multiplying both sides of equation (1) by \( 1/EA_t \), we obtain a relation for the aggregate PE ratio, \( PE_t \), in terms of discount rates, growth rates, and payout ratios:

\[
PE_t \equiv \frac{P_t}{EA_t} = E_t \left[ \sum_{i=1}^{\infty} \exp \left( -\sum_{j=0}^{i-1} \delta_{t+j} \right) \frac{EA_{t+i} D_{t+i}}{EA_t EA_{t+i}} \right]
\]

(3)

\[
= E_t \left[ \sum_{i=1}^{\infty} \exp \left( -\sum_{j=0}^{i-1} \delta_{t+j} + \sum_{j=1}^{i} g_{t+j} + po_{t+i} \right) \right],
\]

where \( g_t \) is log earnings growth,

\[
g_t = \ln \left( \frac{EA_t}{EA_{t-1}} \right).
\]

(4)

In equation (5), the PE ratio is not simply the discounted sum of future earnings growth. The PE ratio involves the payout ratio because we value the portion of earnings at each future time to which equity holders have claim.\(^1\)

The price-earnings ratio at time \( t \), \( PE_t \), is given by

\[
PE_t = E_t \left[ \sum_{i=1}^{\infty} \exp \left( -\sum_{j=0}^{i-1} \delta_{t+j} + \sum_{j=1}^{i} g_{t+j} + po_{t+i} \right) \right],
\]

where the discount rate \( \delta_t \), discounts cashflows at \( t+1 \) back to time \( t \):

\[
\delta_t = \ln E_t \left[ \frac{P_{t+1} + D_{t+1}}{P_t} \right],
\]

(5)

\(^1\) Simply discounting the entire future earnings stream yields neither the value of the firm nor the value of equity (see Miller and Modigliani, 1961).
$g_t$ is log earnings growth,

$$g_t = \ln \left( \frac{EA_t}{EA_{t-1}} \right),$$

and $p_{ot}$ is the log payout ratio:

$$p_{ot} = \ln \left( \frac{D_t}{EA_t} \right).$$

The discount rate is a linear function of state variables, $X_t$:

$$\delta_t = \delta_0 + \delta'_1 X_t.$$  \hfill (8)

We specify the first three factors in $X_t$ to be the log risk-free rate $r^f_t$, log aggregate earnings growth $g_t$, and the log payout ratio, $p_{ot}$. We also include other variables in $X_t$. The final factor in $X_t$ is a latent factor, $f_t$. The state variables $X_t$ follow

$$X_{t+1} = \mu + \Phi X_t + \Sigma \epsilon_{t+1},$$

where $\epsilon_t \sim \text{iid } \mathcal{N}(0, I)$. We specify the latent factor $f_t$ to be orthogonal to the other factors at all leads and lags.

Under assumptions (8) and (9), the PE ratio can be written as

$$P/E_t = \sum_{i=1}^{\infty} \exp(a_i + b'_i X_t),$$

The coefficients $a_i$ and $b_i$ follow the recursions

$$a_{i+1} = -\delta_0 + a_i + (e_2 + b_i)'\mu + \frac{1}{2} (e_2 + b_i)'\Sigma \Sigma'(e_2 + b_i)$$

$$b_{i+1} = -\delta_1 + \Phi'(e_2 + b_i),$$

where $e_n$ is a vector of zero’s with a 1 in the $n$th position. The initial conditions are $a_1 = -\delta_0 + (e_2 + e_3)'\mu + \frac{1}{2} (e_2 + e_3)'\Sigma \Sigma'(e_2 + e_3)$ and $b_1 = -\delta_1 + \Phi'(e_2 + e_3)$.

To derive the initial condition, consider the first term in the sum:

$$E_t[\exp(-\delta_t + g_{t+1} + p_{ot+1})] = E_t[\exp(-\delta_0 - \delta'_1 X_t + (e_2 + e_3)'X_{t+1})]$$

$$= \exp(a_1 + b'_1 X_t).$$

We derive the recursion in (10) by induction. Assume that the $i$th term of the PE in takes the form $\exp(a_i + b'_i X_t)$. Then consider the $(i+1)$th term:

$$E_t \left[ \exp(-\delta_t + g_{t+1}) E_{t+1} \left[ \exp \left( -\sum_{j=0}^{i-1} \delta_{t+1+j} + \sum_{j=1}^{i} g_{t+1+j} + p_{ot+1+i} \right) \right] \right]$$

$$= E_t[\exp(-\delta_t + g_{t+1} + a_i + b'_i X_{t+1})].$$

2
Substituting $g_{t+1} = e'_2 X_{t+1}$ and taking the expectation yields $\exp(a_{i+1} + b'_{i+1} X_t)$, where $a_{i+1}$ and $b_{i+1}$ are given in equation (5).

To obtain some intuition on the recursions in equation (10), note that $\delta_1$ enters with a negative sign in the $b_i$ loadings. Higher discount rate loadings on the $X_t$ factors decrease PE ratios as discount rates appear in the denominator of the present-value sum. An increase in unexpected earnings growth increases the current cashflow, through $b_1$, and hence the PE, but the higher growth today also translates to higher growth at other horizons, captured through $b_i$, because $g_t$ is persistent.

### 1.2 No-Growth Price-Earnings Ratio

We define the no-growth PE ratio, $PE_{t}^{ng}$, where earnings growth is everywhere zero and the payout ratio is equal to one:

$$
PE_{t}^{ng} = \frac{P_{t}^{ng}}{E_{t}^{A}} = E_t \left[ \sum_{i=1}^{\infty} \exp \left( \delta_{t+j} \right) \right].
$$

(13)

This can be evaluated as

$$
\frac{P_{t}^{ng}}{E_{t}^{A}} = E_t \left[ \sum_{i=1}^{\infty} \exp \left( \delta_{t+j} \right) \right] = \sum_{i=1}^{\infty} \exp(a_i^* + b_i^* X_t),
$$

(14)

where $a_i^*$ and $b_i^*$ have initial values $a_1^* = -\delta_0$ and $b_1^* = -\delta_1$ and follow the recursions

$$
a_{i+1}^* = -\delta_0 + a_i^* + b_i^* \mu + \frac{1}{2} b_i^* \Sigma \Sigma' b_i^*
$$

$$
b_{i+1}^* = -\delta_1 + \Phi' b_i^*.
$$

(15)

In the no-growth PE, Jensen’s terms enter the $a_i^*$ recursions. As Pastor and Veronesi (2006) note, uncertainty in discount rates and cashflows increase price-earnings ratios because the PE ratio is a convex function of discount rates and cashflow growth. These Jensen’s terms play important roles in determining the PE ratio and are large especially during times when PE ratios are very high and expected returns are low, such as during the late 1990s.\(^2\)

\(^2\) The convexity and non-linearity in the PE ratio is an important component and is preserved in our full valuation model in equation (10) and the no-growth values in equation (13). An alternative approach taken by Gabaix (2007) is to write down an exponential affine function for the PE ratio and then estimate non-linear processes for the factors which yield the linear form for the log PE ratio. This shuts down much of the inherent convexity in a standard valuation formula.
2 Estimation

We estimate the PE model using a two-step procedure. First, we estimate all parameters of the processes associated with observable factors. Then, we estimate the parameters of the latent $f_t$ process. This separation is possible because the latent factor is specified to represent the remainder of variation in PE ratios not accounted for by the other observable factors of the model. We label $z_t$ the observable factors, where $X_t = (z_t', f_t)'$.

We work at a quarterly frequency, but specify $t$ to $t + 1$ to denote one year. While the estimates of the parameters are consistent in the presence of the moving average terms induced by the overlapping observations, the standard errors of the estimates must be adjusted. We compute standard errors by GMM that take into account the two stage process as well as being robust to heteroskedasticity and moving average effects by using four lags in the GMM covariance weighting matrix.

2.1 Identification

The discount rate $\delta_t$ is defined by

$$\exp(\delta_t) = E_t \left[ \frac{P_{t+1} + D_{t+1}}{P_t} \right] = E_t \exp(r_{t+1}),$$

where $r_{t+1}$ is the log total return over $t$ to $t + 1$. We write the discount rate in equation (8) as

$$\delta_t = \delta_0 + \delta'_z z_t + f_t.$$  

(17)

We run the regression

$$r_{t+1} = \alpha + \beta' z_t + \sigma_{rt} \varepsilon_{t+1},$$  

(18)

where $\varepsilon_{t+1}$ is a mean-zero error following Fama and French (1988) and many others. The PE model in equation (5) generates heteroskedasticity in returns, which we denote by $\sigma_{rt}$, but this cannot be computed in closed form, at least in discrete time.\(^3\) Consider $E \delta_t = E \ln E_t \exp(r_{t+1})$. From the regression (18), we can write

$$E_t \exp(r_{t+1}) = \exp \left( \alpha + \beta' z_t + \frac{1}{2} \sigma^2_{rt} \right)$$

$$E \ln E_t \exp(r_{t+1}) = \alpha + \beta' E z_t + \frac{1}{2} E \sigma^2_{rt}.$$  

(19)

\(^3\) In the classification of Ang and Liu (2007), the PE model specifies cashflows and a discount rate, which endogenously determines predictability of conditional expected returns and heteroskedasticity of returns.
From the definition of the discount rate (17) we have

\[ E \delta_t = \delta_0 + \delta'_1 E z_t + E f_t. \]  

(20)

Thus, equating (19) and (20) we can identify \( \delta_{1z} = \beta \) by OLS.

The constant term \( \delta \) in the discount rate equation (17) is related to the constant term \( \alpha \) in the OLS regression (18) by

\[ \alpha + \frac{1}{2} E \sigma^2_{zt} = \delta_0 + E f_t \]  

(21)

and thus \( \delta_0 \) is not separately identified from \( \mu_f \). Intuitively, shifting up or down the unconditional mean of \( f_t \) is equivalent to changing the level of \( \delta_t \). For identification we always normalize the latent factors to be zero mean and report \( f_t - E f_t \). This allows us to directly compute the conditional forecast

\[ E_t(r_{t+1}) = \alpha + \beta' z_t + (f_t - E f_t) \]  

(22)

where \( \beta = \delta_{1z} \) and the unconditional effect of the latent factors is absorbed into the residual term of the regression (18).

### 2.2 Estimating the Model

We estimate the model in two stages. In the first stage, we estimate the parameters of the VAR in equation (9) corresponding to the observable factors \( z_t \). In this first stage we also estimate the parameters \( \delta_1 \) which identify the discount rate process by running the regression in equation (18) of log total market returns onto \( z_t \). Because \( f_t \) is orthogonal to \( z_t \), the OLS coefficients \( \beta \) are consistent estimates for \( \delta_1 \). We set \( \delta_0 \) to be the constant \( \alpha \) in the regression (18). This is to facilitate our estimate of the latent factor parameters, but the value of \( \delta_0 \) by itself is not used in inference as it is not separately identifiable from \( \mu_f \) or the heteroskedasticity of returns. Thus, we do not include a moment condition for \( \delta_0 \) in estimating the standard errors of the system.

We denote the set of first-stage parameters as \( \theta_1 = (\mu_z \Phi_z \Sigma_z \delta_0 \delta_1) \). We denote the set of moments of the first-stage estimation as \( g_1(\theta_1) \) which is exactly identified. The first-stage parameter estimates, \( \hat{\theta}_1 \) satisfy the system

\[ \sqrt{T} g_1(\hat{\theta}_1) = 0. \]  

(23)

In the second stage we estimate the parameters \( \theta_2 = (\mu_f \Phi_f \Sigma_f) \) corresponding to the process \( f_t \) taking the parameters from the first stage as given. We extract estimates of \( f_t \) from PE ratios in data. Equation (10) links PE ratios and factors, given the parameters employed in the coefficients \( a_i \) and \( b_i \). For a fixed set of parameters, we can invert equation (10) to solve for \( f_t \)
at each time \( t \) to match the data PE. We truncate the sums at \( i = 500 \) years, where the terms are extremely small. Note that by directly working with the non-linear PE in equation (10), our approach is very different from log-linearized models of the PE ratio such as Campbell and Shiller (1988) which ignore the large convexity present when discount rates are small (see comments by Pastor and Veronesi, 2006).

Our method of estimating components of expected returns from prices is similar to the exercises done by Lee, Myers and Swaminathan (1999), Claus and Thomas (2001), and Pastor, Sinha and Swaminathan (2008). However, these studies do not allow the expected return to vary over time and proxy expected cashflows by analysts’ forecasts with a residual terminal value. In contrast, we endogenously take into account the time variation of expected returns and by matching PE ratios in data, we estimate a discount rate process consistent with observed prices. However, we do not use analysts’ forecast information in these studies and estimate the \( f_t \) process only using realized PE ratios. Analysts’ forecasts are not available over the full sample and few analysts produce forecasts of year-ahead aggregate market PE ratios compared to future earnings of individual firms.

To solve for \( f_t \) at each time \( t \), we find \( \hat{f}_t \) which equates the PE to the RHS of equation (10). The series \( \{\hat{f}_t\} \) matches the PE in data exactly for a given set of parameter values \( \theta = (\theta_1, \theta_2) \). To estimate \( \theta_2 \) taking \( \theta_1 \) as given in the second stage, we match moments from the sample \( \{\hat{f}_t\} \) with moments of \( f_t \) implied by the VAR from equation (9). Specifically, we match the mean, variance, and autocovariance \( f_t \) implied by the VAR with the sample mean, variance, and autocovariance of the implied extracted series \( \{\hat{f}_t\} \). Formally, we denote the second stage estimation as

\[
\min_{\theta_2} \quad g_2(\theta_2, \hat{\theta}_1)'g_2(\theta_2, \hat{\theta}_1),
\]

where \( g_2 \) contains the autocovariance and variance moments taking the first-stage estimates \( \hat{\theta}_1 \) as given. The first-order conditions for this second stage are:

\[
G_{2,2}'\sqrt{T}g_2(\hat{\theta}_2, \hat{\theta}_1) = 0,
\]

for the second-stage estimates \( \hat{\theta}_2 \) and \( G_{2,2} = \partial g_2/\partial \theta_2' \). The motivation in the second-stage minimization (24) is that the sample of extracted factors \( \{f_t\} \) has its own associated moments. We ensure that the model’s implied moments for \( f_t \) match the sample moments of \( \{f_t\} \) as close as possible. This procedure estimates a value for \( \mu_f \) given a first-stage value \( \delta_0 \). Both are not separately identified from the previous section. Thus, we report the demeaned latent factor series \( \{\hat{f}_t - \hat{E}f_t\} \) in our empirical results.
To obtain standard errors, we follow Newey and MacFadden (1994) and Bekaert and Hodrick (2001) and set up the joint system

\[ \sqrt{T} g_1(\hat{\theta}_1) = 0 \]
\[ \sqrt{T} g_2(\hat{\theta}_2, \hat{\theta}_1) = 0. \]

Using a first-order Taylor expansion around the true parameters \( \theta_1 \) and \( \theta_2 \), we can write

\[ \sqrt{T} g_1(\hat{\theta}_1) = \sqrt{T} g_1(\theta_1) + G_{1,1} \sqrt{T}(\hat{\theta}_1 - \theta_1) \]
\[ \sqrt{T} g_2(\hat{\theta}_1, \hat{\theta}_2) = \sqrt{T} g_2(\theta_2, \theta_1) + G_{2,1} \sqrt{T}(\hat{\theta}_1 - \theta_1) + G_{2,2} \sqrt{T}(\hat{\theta}_2 - \theta_2), \] (26)

where \( G_{i,j} = \partial g_i / \partial \theta'_j \), for \( i, j = 1, 2 \).

Rearranging and taking standard Central Limit Theorems allows us to derive the asymptotic variance of \( \sqrt{T}(\hat{\theta}_1 - \theta_1) \) and \( \sqrt{T}(\hat{\theta}_2 - \theta_2) \) as:

\[ \text{var}(\sqrt{T}(\hat{\theta}_1 - \theta_1)) = G_{1,1}^{-1} \Sigma_{1,1} G_{1,1}^{-1} \]
\[ \text{var}(\sqrt{T}(\hat{\theta}_2 - \theta_2)) = \left[ -(G'_{2,2} G_{2,2})^{-1} (G'_{2,2} G_{2,2})^{-1} G'_{2,2} \right] \Omega \left[ -(G'_{2,2} G_{2,2})^{-1} (G'_{2,2} G_{2,2})^{-1} G'_{2,2} \right]' , \] (27)

where

\[ \Omega = \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix} \]

is the covariance of the moments \( g = \{ g_1, g_2 \} \):

\[ \Omega = \sum_{j=-\infty}^{\infty} \text{E}[g_{t-j} g'_t]. \]

To take account of the overlapping quarterly frequency observations with an annual horizon we estimate \( \Omega \) using a Hansen and Hodrick (1980) estimate with four lags.

### 2.3 Goodness of Fit

To judge the fit of the model we conduct several residual and stability tests. We construct standardized residuals of the latent factor process, defined as

\[ \eta'_t = \frac{\hat{f}_t - \mu_f - \Phi_f \hat{f}_{t-1}}{\sigma_f}, \] (28)

where \( \hat{f} \) is the estimated latent factor, and \( \mu_f, \Phi_f, \text{ and } \sigma_f \) are the parameters of the \( f_t \) process in the VAR. We fail to reject the hypothesis that \( \text{E}[\eta'_t \eta'_{t-1}] = 0 \) with a p-value of 0.21. We
also fail to reject that there is zero serial correlation in the squared residuals, with a p-value of 0.13 for the GMM test $E[((\eta^f_t)^2 - 1)((\eta^f_{t-1})^2 - 1)] = 0$. Second, we fail to reject that the correlations of $\eta^f_t$ with all the other factor residuals are jointly equal to zero with a p-value of 0.08. Finally, we fail to reject that the AR(1) specification for the latent factor has structural breaks. The Andrews (1993) supLM test fails to reject the null that the coefficients in the VAR corresponding to the latent factor dynamics are stable both individually and jointly. In particular, the $\chi^2$ statistic for jointly testing that $(\mu_f, \Phi_f, \sigma_f^2)$ are stable has a value of 7.2 compared to a critical 95% value of 13.6. In summary, the latent factor model comfortably passes residual and stability specification tests.
References


