This document serves as the Appendix and describes in detail key technical components of *Cross-Sectional Factor Dynamics and Momentum Returns*. The first section solves the asset-pricing model, the second section describes how the return and dividend data are constructed for the different momentum portfolios, the third section provides the details of the MCMC approach in estimating model parameters and state-variables, the fourth section provides the parameter estimates and asset-pricing results of the one-factor (Base) model, and the fifth section provides comparison of moments from the data on dividend growth and the models of dividend growth for Small, Low/Small and High/Small Stocks.

1 Model

1.1 The Value function:

The representative agent is endowed with Duffie-Epstein preferences with unit elasticity of intertemporal substitution which is given by

$$f(C,J) = \beta(1-\gamma)J\left[\log C - \frac{\log(1-\gamma)J}{1-\gamma}\right] \tag{1}$$

where C is aggregate consumption, J is the value function, β is the time-discount parameter, and γ is risk-aversion. Let aggregate consumption C_t follow

$$\frac{dC_t}{C_t} = (\mu_C + X_t)dt + \sigma_C dW_C \tag{2}$$

$$dX_t = -\kappa_X X_t dt + \sigma_X dW_X \tag{3}$$

where the Brownian innovations dW_C and dW_X are independent.

The solution of the value function J corresponding to preferences given by (1) and consumption dynamics given by (2)-(3) is given by

$$E_t[\mathcal{D}J] + f(C,J) = 0$$

where $\mathcal{D}$ is the differential operator applied to J. Applying, Ito's lemma this implies

$$J_C C(\mu_C + X_t) - J_X \kappa_X X_t + \frac{1}{2}\left[J_{CC}C^2\sigma_C^2 + J_{XX}\sigma_X^2\right] + f(C,J) = 0$$

The above PDE is homogeneous of order $1-\gamma$ in C. Guess a solution of the form

$J = \frac{C^{1-\gamma}}{1-\gamma} g(X_t)$ and plug it into the above PDE to obtain another ODE in g

$$\mu_C + X_t - \frac{\gamma}{2}\sigma_C^2 - \frac{\beta \ln g}{1-\gamma} - \frac{\left[\kappa_X X_t \frac{g_X}{g} - \frac{\sigma_X^2}{2}\frac{g_{XX}}{g}\right]}{1-\gamma} = 0$$

which has a solution given by $g(X_t) = e^{\bar{A}+\bar{B}X_t}$ where

$$\begin{aligned} \bar{A} &= \frac{(1-\gamma)}{\beta}\left[\mu_C - \frac{\gamma}{2}\sigma_C^2 + \frac{\sigma_X^2}{2(1-\gamma)}\bar{B}^2\right] \\ \bar{B} &= \frac{1-\gamma}{\kappa_X + \beta} \end{aligned}$$

1.2 The Pricing Kernel:

Duffie and Epstein (1992) shows that the pricing kernel for preferences given by (1) is given by $\Lambda_t = e^{\int_0^t f_J ds} f_c$. Given the solution of the value function from above, the marginal utilities are given by $f_c = \beta C^{-\gamma} g(X_t)$ and $f_J = -\beta(1 + \ln g)$. Applying Ito's lemma to Λ_t, we can write it in differential form as

$$\frac{d\Lambda}{\Lambda} = -r_t^f dt - \gamma\sigma_C dW_C + \bar{B}\sigma_X dW_X \tag{4}$$

where $r_t^f = \beta + \mu_C - \gamma\sigma_C^2 + X_t$.

1.3 Asset Return Dynamics:

The dividend process of portfolio i follows

$$\frac{dD^i}{D^i} = \left[\mu_D^i + \phi_1^i X_t + Y_t^i\right] dt + \sigma_D^i dW_D^i \tag{5}$$

$$dY_t^i = -\kappa_Y^i Y_t^i dt + \sigma_Y^i dW_{Y^i} \tag{6}$$

where $Corr(dW_X, dW_{Y^i}) = \rho_{XY^i}^i$. The value of portfolio i can be derived from discounting future cash-flows using the discount rate derived above. Using (5)-(6) and (4)

$$\begin{aligned} P_t^i &= \frac{1}{\Lambda_t} E_t \int_t^\infty \Lambda_s D_s^i ds \\ &= \frac{1}{\Lambda_t} \int_t^\infty E_t \left[\Lambda_s D_s^i\right] ds \end{aligned} \tag{7}$$

To compute the inner expectation, first let $h_s = \Lambda_s D_s^i$. Applying Ito's Lemma, it follows a process given by

$$\begin{aligned}\frac{dh}{h} &= \left[\mu_D^i - \beta - \mu_C + \gamma\sigma_C^2 + (\phi_1^i - 1)X_t + Y_t^i\right] dt \\ &\quad -\gamma\sigma_C dW_C + \bar{B}\sigma_X dW_X + \sigma_D^i dW_D^i\end{aligned}$$

Now we are left with evaluating $E_t[h_s]$. Applying Feynman-Kac, $E_t[h_s] = f(h_t, X_t, Y_t^i, \tau)$, where $\tau = s - t$. Consider h_s to be an asset that pays dividends D_s^i at time $s > t$ and f represents the expected discounted value of this asset at time t. Since f has to be a martingale, that implies $E_t[df] = 0$. Applying Ito's lemma to $f(h_t, X_t, Y_t^i, \tau)$ and using the fact that it is a martingale gives us the PDE for f

$$\begin{aligned}f_\tau &= f_h h\left[\mu_D^i - \beta - \mu_C + \gamma\sigma_C^2 + (\phi_1^i - 1)X_t + Y_t^i\right] \\ &\quad -\kappa_X X_t f_X - \kappa_Y^i Y_t^i f_{Y^i} + \frac{1}{2}\left[f_{hh} dh' dh + f_{XX}\sigma_X^2 + f_{Y^i Y^i}{\sigma_Y^i}^2\right] \\ &\quad + f_{Xh} h\bar{B}\sigma_X^2 + \left(f_{hY^i} h\bar{B} + f_{XY^i}\right)\sigma_X\sigma_Y^i\rho_{XY^i} \qquad (8)\end{aligned}$$

which has a solution $f(h_t, X_t, Y_t^i, \tau) = h_t m(X_t, Y_t^i, \tau)$. Plugging the solution into (8) we get

$$\begin{aligned}\frac{m_\tau}{m} &= \left[\mu_D^i - \beta - \mu_C + \gamma\sigma_C^2 + (\phi_1^i - 1)X_t + Y_t^i\right] - \kappa_X X_t \frac{m_X}{m} - \kappa_Y^i Y_t^i \frac{m_{Y^i}}{m} \\ &\quad + \frac{1}{2}\left[\frac{m_{XX}}{m}\sigma_X^2 + \frac{m_{Y^iY^i}}{m}{\sigma_Y^i}^2\right] + \frac{m_X}{m}\bar{B}\sigma_X^2 + \left(\frac{m_{Y^i}}{m}\bar{B} + \frac{m_{XY^i}}{m}\right)\sigma_X\sigma_Y^i\rho_{XY^i}\end{aligned}$$

whose solution is given by $m(X_t, Y_t^i, \tau) = e^{P_1^i(\tau) + P_2^i(\tau)X_t + P_3^i(\tau)Y_t^i}$ where, $P_1^i(\tau)$, $P_2^i(\tau)$ and $P_3^i(\tau)$ and follow a system of ODEs

$$\begin{aligned}P_3^{i'} &= 1 - \kappa_Y^i P_3^i \\ P_2^{i'} &= (\phi_1^i - 1) - \kappa_X P_2^i \\ P_1^{i'} &= \left[\mu_D^i - \beta - \mu_C + \gamma\sigma_C^2\right] + \frac{1}{2}\left(P_2^i\sigma_X^2\left[P_2^i + 2\bar{B}\right] + P_3^{i2}{\sigma_Y^i}^2\right) + P_3^i(\bar{B} + P_2^i)\sigma_X\sigma_Y^i\rho_{XY^i}\end{aligned}$$

with initial condition $P_1^i(0) = P_2^i(0) = P_3^i(0) = 0$. P_2^i and P_3^i are solved in closed form - $P_2^i(\tau) = \frac{\phi_1^i - 1}{\kappa_X}\left(1 - e^{-\kappa_X\tau}\right)$ and $P_3^i(\tau) = \frac{1}{\kappa_Y^i}\left(1 - e^{-\kappa_Y^i\tau}\right)$. P_1^i is solved numerically using Matlab's ODE45 algorithm.

Now, we can represent the price-dividend ratio of portfolio i as

$$\begin{aligned} P_t^i &= \frac{1}{\Lambda_t}\int_t^\infty E_t\left[\Lambda_s D_s^i\right] ds \\ &= \frac{1}{\Lambda_t}\int_t^\infty \Lambda_t D_t^i m(X_t, Y_t^i, \tau) ds \qquad (9) \\ &= D_t^i M_t^i \qquad (10) \end{aligned}$$

where $M_t^i = \int_t^\infty m(X_t, Y_t^i, \tau)ds$ where $\tau = s - t$.

Let cumulative excess return be given by $dR_t^i = \frac{dP^i + D_t^i dt}{P_t^i} - r_t^f dt$. The return dynamics follow

$$\begin{aligned} dR_t^i &= \mu_t^i dt + \sigma_t^i \cdot dW \qquad (11) \\ \mu_t^i &= -Cov\left(\frac{d\Lambda}{\Lambda}, \frac{dP^i}{P^i}\right) \\ &= -\left(\frac{M_X^i}{M^i}\bar{B}\sigma_X^2 + \frac{M_{Y^i}^i}{M^i}\bar{B}\sigma_X\sigma_Y^i\rho_{XY^i}^i\right) \qquad (12) \\ \sigma_t^i &= \begin{bmatrix}\sigma_D^i & \frac{M_X^i}{M^i}\sigma_X & \frac{M_{Y^i}^i}{M^i}\sigma_Y^i\end{bmatrix} \qquad (13) \\ dW &= [dW_D^i \quad dW_X \quad dW_{Y^i}]' \end{aligned}$$

2 Portfolio Construction

To construct the portfolios, we start by pulling the entire universe of firms in the CRSP monthly stock database by PERMNO. We calculate book value of stocks using data from COMPUSTAT, and merge it with CRSP by pulling the entire universe of firms in the CRSP/COMPUSTAT merged file by "LPERMNO" (restricting to link types LC and LU). We cleanse the data by first removing firms with missing SIC codes. We further restrict the sample by only focussing on share codes 10-12 (ordinary common shares) which are traded in NYSE, AMEX or NASDAQ. Our data period is from 1963-2013.

For each firm, we obtain two different sets of return series - total returns cumulative of dividends (RET) and return without dividends (RETX). We adjust RETX for repurchases by the following algorithm. First, shares outstanding are adjusted for splits/reverse splits using the cumulative share adjustment factor. If the number of adjusted shares outstanding at time t is smaller than the number of adjusted shares outstanding at $t-1$, then let $adj = \min\left(\frac{Shares_t}{Shares_{t-1}}, 1\right)$. The only exception to this is when the number of shares outstanding is reduced by a tender offer, and in this case adj is set to 1, since CRSP automatically adjusts for tender offers in its RET calculation. At the end, we adjust returns without dividends as follows:

$$(1 + \text{RETX}^*) = (1 + \text{RETX}) \times adj.$$

2.1 Momentum Portfolios:

At the beginning of month t, we take the entire universe of stocks selected based on the above criteria and compute their geometric returns from months $t-12$ to $t-2$. We further restrict the sample by dropping firm/month observations that have missing/stale prices at the end of month $t-12$, and missing return or market cap for month $t-1$. Firms are then sorted into 10 portfolios based on Ken French's Prior Return Breakpoints. Each portfolio return is constructed by weighing individual stocks by its prior month's market cap. The portfolio where the highest return stocks were placed is designated Winner, the lowest return stocks designated Loser, We hold each portfolio for one month. At the beginning of the next month, we repeat this process. Monthly returns data are then compounded to form quarterly returns which are converted into real returns using the PCE deflator. Excess returns are excess over the 90 day Treasury Bill yield.

We reconstructed the portfolios by implementing our own break-points, and by dropping stocks in the bottom 10 percent of average market cap over the previous year in line with Zurek (2007). None of the findings change significantly under these additional criteria.

2.2 Size-Momentum Portfolios:

At the beginning of each month t, we look at the prior month's market cap for every stock. We discard stocks which do not have number of shares outstanding or the stock price at the end of month $t-1$. Then we sort the firms into 3 portfolios based on Ken French's ME breakpoints. Firms below the 30th percentile of market cap are put in the first portfolio (Small), while firms above the 70th percentile are put in the third portfolio(Big). The rest of the firms are put in the second portfolio (Medium).

We further sort the stocks in each of the size portfolios into three momentum portfolios based on the methodology described above. The only difference is that the breakpoint for prior returns within each size portfolio is calculated such that each momentum portfolio has the same number of firms. For each size portfolio this process produces three momentum portfolios for a total of nine portfolios. We hold these portfolios for one month. At the beginning of the next month, we repeat this process. For each of the size portfolios we have three momentum portfolios, which gives us 9 portfolios in this cross-section. Monthly returns data are then compounded to form quarterly returns which are converted into real returns using the PCE deflator. Excess returns are excess over the 90 day Treasury Bill yield.

2.3 BM-Size-Momentum Portfolios:

Book-to-Market (BM) of a stock is its book value divided by its market cap. Book value of a stock is computed as the book value of common equity + balance sheet deferred taxes and investment tax credits - book value of preferred shares. We remove firms which are missing book value of common equity, negative book value and market capitalization. In line with Ken French, we form portfolios based on the BM ratio at the end of December of year t. Firms are sorted into three portfolios based on Ken French's BM breakpoints. Firms below the 30th percentile in BM ratio are grouped into the first portfolio (Low) and above the 70th percentile in the third portfolio (High). The rest of the firms are in the second portfolio (Moderate). Even though the firms are sorted into portfolios at the end of December of year t, these portfolio assignments aren't used until July of year $t+1$. Therefore, the portfolio composition of each of the three portfolios stay the same from July of year $t+1$ to the end of June of year $t+2$.

Within each BM portfolio, we sort stocks into three size portfolios using our own breakpoints. Even though the portfolio composition of each of the BM portfolios stays the same from July to June of next year, each month we recreate the size portfolios because we have data on market cap on a monthly frequency. Now, within each BM and Size sort, we take these firms and sort them into three momentum portfolios using our own breakpoints. For each BM and for each Size portfolio, we have three momentum portfolios for a total of 27 different portfolios. The returns on each of these 27 portfolios is available on a monthly frequency. We discard the first few years of data because we simply didn't have enough firms for each of the 27 portfolios. We restrict our analysis using these portfolios from 1971-2013. Monthly returns data are then compounded to form quarterly returns which are converted into real returns using the PCE deflator. Excess returns are excess over the 90 day Treasury Bill yield.

2.4 Dividend Growth data:

Dividend growth is computed in the same way as BDL (2005). To compute dividend growth for each of the portfolios described above, a simulated portfolio is created at the beginning of the monthly return series. The value of this portfolio in every period is the value in the previous period multiplied by (1+RETX*) from the current period. Dividend yield is computed as the difference between RET and RETX*; dividends paid are calculated as the dividend yield times the value of the simulated portfolio at the end of the previous period. Dividends paid for a month are smoothed by adding up all dividends paid from months t to $t-11$. They are converted into real dividends by using the PCE deflator. To form quarterly dividend growths, first we isolate dividends paid at the end of every quarter (March, June, September and December of each year). Quarterly dividend growth is then taken as the percentage change of dividends between the end of successive quarters.

3 Empirical Analysis using MCMC:

We estimate the parameters and latent state variables of the model using a Bayesian methodology. Let the full parameter space and state variables for consumption growth and dividend growth of every asset i be given by $\Theta = \{\mu_C, \sigma_C, \kappa_X, \sigma_X, \mu_D^i, \sigma_D^i, \kappa_Y^i, \sigma_Y^i, \phi^i, \rho^i\}$ and (X, Y^i). Here X and Y^i denote the full time-series $\{X_1, \cdots, X_T\}$ and $\{Y_1^i, \cdots, Y_T^i\}$. We follow a Markov Chain Monte Carlo (MCMC) algorithm to get joint estimates of the parameters and state variables conditional on the time-series data on aggregate consumption and dividend growths of every asset i, i.e. $p(\Theta, X, Y^i|\text{data})$. First, we break up the full parameter space into the consumption parameters and dividend parameters. Let $\Theta_C = \{\mu_C, \sigma_C, \kappa_X, \sigma_X\}$ be the parameters driving the aggregate consumption process and $\Theta_D^i = \{\mu_D^i, \sigma_D^i, \kappa_Y^i, \sigma_Y^i, \phi^i, \rho^i\}$ be the parameters of the dividend process for each asset i. The MCMC algorithm first draws the consumption parameters and expected consumption growth rate X conditional on the consumption data, i.e. $p(\Theta_C, X|\text{consumption growth})$, and then draws the parameters of the dividend process and dividend state variable Y^i of each asset i, i.e. $p(\Theta_D^i, Y^i|\text{dividend growth of asset i}, X)$.

First, let's consider the aggregate consumption growth dynamics. Discretizing (2)-(3) and setting $dt = 1$, we get

$$\Delta c_{t+1} = (\mu_C + X_t) + \sigma_C Z_{t+1}^1 \quad (14)$$

$$X_{t+1} = \delta_X X_t + \sigma_X Z_{t+1}^2 \quad (15)$$

where $\Delta c_{t+1} = \ln C_{t+1}/C_t$ is a time-series of real consumption growth, $\delta_X = e^{-\kappa_X}$ and Z_{t+1}^1 and Z_{t+1}^2 are iid N(0,1) shocks. To draw the aggregate parameters Θ_C and state variable X jointly, we draw them one at a time conditioning on the other, i.e. $p(\Theta_C|X, \Delta c_t)$ and $p(X|\Theta_C, \Delta c_t)$. First, we draw the time-series of the expected consumption growth rate X conditional on the parameter space Θ_C and the full time-series of consumption growth. We follow a Bayesian version of Kalman filter called Forward Filtering Backward Sampling (FFBS) as introduced by Carter and Kohn (1994) and Frühwirth-Schnatter (1994). In the next step, we draw the parameter set Θ_C conditional on the full time-series of consumption growth and X. Here we break up the parameter space even further and draw each component of Θ_C one at a time conditional on the others - $p(\Theta_j|\Theta_{-j}, \Delta c_t, X)$, where Θ_{-j} is the rest of the parameters modulo the $j-$th one. In summary, the consumption parameters and state variable X are drawn the following way:

- $p(X|\Theta_C, \Delta c_t)$ is a Bayesian Kalman filter implemented with FFBS.
- $p(\mu_C|\Theta_{-\mu_C}, X, \Delta c_t)$ is posterior mean of Δc_t conditional on X.
- $p(\sigma_C|\Theta_{-\sigma_C}, X, \Delta c_t)$ is a posterior standard deviation of Δc_t conditional on X.
- $p(\delta_X, \sigma_X|\Theta_{-\delta_X, -\sigma_X}, X)$ obtained by a time-series regression of X_{t+1} on X_t.

In the next step we draw the parameters of the dividend process of each asset i, conditional on expected consumption growth rate, X, i.e. $p(\Theta_D^i, Y^i|\text{dividend data of asset } i, X)$. Discretizing (5)-(6), using the correlation structure between (3)-(6) and setting $dt = 1$, we get

$$\Delta d_{t+1}^i = \left(\mu_D^i + \phi_1^i X_t + Y_t^i\right) + \sigma_D^i Z_{t+1}^3 \tag{16}$$

$$Y_{t+1}^i = \left(\frac{\rho^i \sigma_Y^i}{\sigma_X}\left(X_{t+1} - \delta_X X_t\right) + \delta_Y^i Y_t^i\right) + \sigma_Y^i \sqrt{1 - {\rho^i}^2} Z_{t+1}^4 \tag{17}$$

where $\Delta d_{t+1}^i = \ln D_{t+1}^i / D_t^i$ is a time-series of real dividend growth of asset i, $\delta_Y^i = e^{-\kappa_Y^i}$ and Z_{t+1}^3 and Z_{t+1}^4 are iid N(0,1) shocks. Since X is known in this step, this is an univariate state-space similar to (14)-(15). As before, we break up the joint distribution of $p(\Theta_D^i, Y^i|\text{dividend data}, X)$ into $p(\Theta_D^i|\text{dividend data}, X, Y^i)$ and $p(Y^i|\text{dividend data}, X, \Theta_D^i)$. Like expected consumption growth rate, we obtain $p(Y^i|\text{dividend data}, X, \Theta_D^i)$ by FFBS using the state-space (16)-(17). We break up the dividend growth parameter space even further and draw each of component of Θ_D^i one at a time conditonal on the others. In summary, the dividend growth parameters and state variable Y^i of each asset i are drawn the following way:

- $p(Y^i|\Theta_D^i, X, \Delta d_t^i)$ is a Bayesian Kalman filter implemented with FFBS.
- $p(\mu_D^i|\Theta_{-\mu_D^i}^i, X, Y^i, \Delta d_t^i)$ is posterior mean of Δd_t^i conditional on X and Y^i.
- $p(\phi_1^i, \sigma_D^i|\Theta_{-\phi_1^i, -\sigma_D^i}, Y^i, X)$ obtained by a time-series regression of Δd_{t+1}^i on Y^i and X.
- $p(\delta_Y^i, \sigma_Y^i|\Theta_{-\delta_Y^i, -\sigma_Y^i}, Y^i, X)$ obtained by a time-series regression of Y_{t+1}^i on Y_t^i after adjusting for $\frac{\rho^i \sigma_Y^i}{\sigma_X}(X_{t+1} - \delta_X X_t)$ as in (17).
- $p(\rho^i|\Theta_{-\rho^i}^i, X, Y^i)$ obtained from the posterior variance-covariance matrix of ΔX and ΔY^i.

The exact formulation of the posterior distribution of each parameter above is available in elementary conjugate form. For details on the exact specification, please consult Allenby, Rossi and McCulloch (2005). We use a Gibbs sampler to draw iteratively from each posterior distribution. In each iteration, we first draw the consumption parameters and the state variable X. Then, for each asset i, we draw the dividend growth parameters and Y^i. Thus, after one iteration of the MCMC, we have one draw of the aggregate parameters and one draw of all the cross-sectional parameters for all assets we considered. We iterate 25,000 times and discard the first 10,000 for burn-in.

4 The Base Model

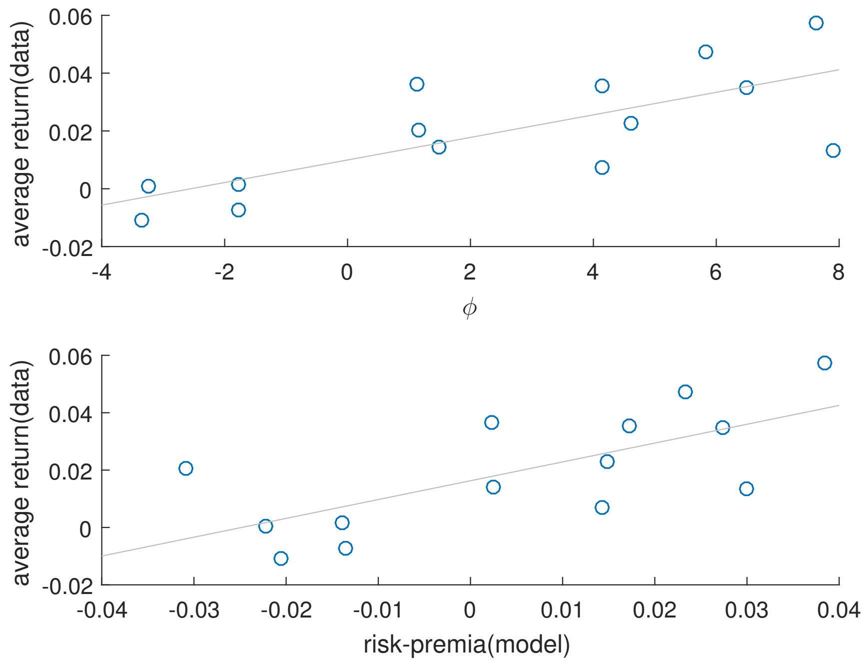
In this section, we provide parameter estimates and asset-pricing results of the one-factor (Base) model. First, we provide the parameter estimates of the model from Bayesian MCMC. The parameters of the one factor model are estimated the same way as the parameters of the two factor model. Both are estimated by Bayesian MCMC, and the prior settings for parameters which are common across the two models are exactly the same. The aggregate parameters of the Base model are identical to the Alternate model, and reported in Table 4 in the main paper.

Table 1: Cross-Sectional "Leverage" and Volatility parameters from the MCMC algorithm obtained from the Base model (Equations 1-2) in the main paper. For each parameter, we show the median, as well as the 5th and 95th percentiles.

	Portfolios	ϕ	σ_D
	Loser	-3.35	0.19
		(-9.24, 2.31)	(0.14, 0.27)
	8	-3.25	0.17
		(-8.12, 2.06)	(0.13, 0.26)
All	5	-1.77	0.22
Stocks		(-5.27, 2.04)	(0.15, 0.33)
	3	4.13	0.16
		(-2.79, 11.02)	(0.10, 0.27)
	Winner	7.91	0.13
		(1.66, 14.26)	(0.08, 0.26)
	Loser	-5.17	0.14
		(-12.34, 1.95)	(0.08, 0.22)
Small	Middle	1.12	0.15
Stocks		(-5.46, 7.91)	(0.05, 0.22)
	Winner	5.83	0.13
		(-2.75, 13.81)	(0.10, 0.19)
Low	Loser	-1.77	0.17
Small		(-9.23, 5.54)	(0.09, 0.28)
Stocks	Middle	1.49	0.16
		(-6.68, 9.95)	(0.11, 0.24)
	Winner	4.13	0.14
		(-2.08, 11.14)	(0.09, 0.24)
High	Loser	4.61	0.18
Small		(-4.12, 14.06)	(0.12, 0.30)
Stocks	Middle	6.49	0.20
		(-6.11, 18.73)	(0.12, 0.32)
	Winner	7.62	0.24
		(0.05, 14.69)	(0.16, 0.34)

In the first panel of the following figure, first we show that the "leverage" parameter of the Base model, ϕ, lines up quite nicely with average return in the data. Taking the median parameter values and the median draw of the X_t variable, we compute a time-series of risk-premia from (12) by shutting off all the other parameters relevant to the two-factor

Figure 1: The figure shows the asset pricing properties of the one-factor (Base) model. The first panel shows the relationship between the leverage parameter, ϕ, and the average expected return in the data. ϕ is estimated by Bayesian MCMC and reported in Table 1 of this document. The second panel shows the relationship between risk-premia of different assets from the one-factor model, relative to the average return in the data. The data sample period covers from 1963-2013 for momentum assets constructed out of All Stocks and Small Stocks. The data sample period covers from 1971-2013 for momentum assets constructed out of Low/Small and High/Small Stocks.



model. We take time-series averages to create average risk-premia of all assets, and plot them in the second panel against observed average return.

5 Moments of dividend growth for other momentum portfolios

In this section, we show how our models perform in matching key moments of dividend growth in Winner and Loser portfolios in Small, Low/Small and High/Small stocks. We calculate the moments of dividend growth from simulated data using the Base and Alternate models discussed in Section 2 of the main paper. The Base is the one-factor model a la BDL

(2005) and Zurek (2007). The Alternate is the two factor model used in the paper. We use Bayesian MCMC to estimate the parameters and state-variables of the two models. The posterior distributions of the cross-sectional parameters of the Base model are in Table 1 in this document, and the ones from the Alternate model are in Table 5 of the main paper. The aggregate parameters for both models are taken from Table 4 in the main paper.

Taking the posterior median of the parameters, we simulate dividend growth under the two models. Below are the sample moments of the data and from simulated data under the two models.

Table 2: Moments: data versus simulation. The first simulation uses parameters from the two-factor model, where the posterior distribution of aggregate parameters is taken from Table 4 in the main paper and cross-sectional parameters taken from Table 5 in the main paper. The second simulation uses the one-factor (Base) model, where the posterior distribution of aggregate parameters is taken from Table 4 in the main paper and cross-sectional parameters are reported in Table 1 in this document. Parameters for both models are obtained from Bayesian MCMC which is described in detail in Section 3 in this document. The data come from dividend growth of momentum portfolios formed using Small, High/Small and Low/Small Stocks. Data are quarterly and covers the sample period 1963-2013 for Small, and 1971-2013 for Low/Small and High/Small Stocks.

		Data		Simulation		Simulation (Base)	
		Loser	Winner	Loser	Winner	Loser	Winner
Small Stocks	mean	-0.01	0.02	-0.01	0.02	-0.01	0.02
				(-0.04,0.02)	(-0.03,0.06)	(-0.02,0.00)	(0.00,0.04)
	Std. Dev.	0.11	0.14	0.08	0.09	0.13	0.13
				(0.07,0.18)	(0.03,0.17)	(0.10,0.20)	(0.07,0.23)
	AC(1)	0.58	0.41	0.50	0.37	0.00	0.03
				(0.34,0.66)	(0.20,0.54)	(-0.11,0.10)	(-0.08,0.13)
	AC(2)	0.26	0.16	0.37	0.24	-0.01	-0.01
				(0.21,0.52)	(0.07,0.41)	(-0.11,0.09)	(-0.11,0.12)
	AC(3)	0.02	-0.03	0.15	0.12	0.00	0.03
				(-0.05,0.34)	(-0.05,0.30)	(-0.10,0.10)	(-0.07,0.13)
		Data		Simulation		Simulation (Base)	
		Loser	Winner	Loser	Winner	Loser	Winner
Low Small Stocks	mean	-0.03	0.02	-0.03	0.02	-0.03	0.02
				(-0.07,0.01)	(-0.02,0.07)	(-0.05,-0.01)	(0.01,0.04)
	Std. Dev.	0.23	0.19	0.11	0.11	0.16	0.15
				(0.06,0.19)	(0.04,0.17)	(0.06,0.16)	(0.09,0.25)
	AC(1)	0.26	0.24	0.22	0.22	0.04	0.02
				(0.09,0.36)	(0.10,0.34)	(-0.04,0.12)	(-0.08,0.11)
	AC(2)	0.15	0.16	0.18	0.13	0.01	0.02
				(0.05,0.32)	(0.03,0.24)	(-0.08,0.10)	(-0.12,0.16)
	AC(3)	0.02	0.13	0.09	0.06	0.03	0.03
				(-0.05,0.23)	(-0.12,0.23)	(-0.06,0.12)	(-0.09,0.14)
		Data		Simulation		Simulation (Base)	
		Loser	Winner	Loser	Winner	Loser	Winner
High Small Stocks	mean	-0.01	0.05	-0.01	0.04	-0.01	0.04
				(-0.04,0.02)	(0.00,0.08)	(-0.03,0.01)	(0.02,0.07)
	Std. Dev.	0.20	0.24	0.19	0.28	0.19	0.26
				(0.14,0.25)	(0.21,0.37)	(0.13,0.28)	(0.18,0.40)
	AC(1)	0.21	0.16	0.15	0.19	0.00	0.04
				(0.03,0.28)	(0.08,0.31)	(-0.10,0.10)	(-0.07,0.15)
	AC(2)	0.09	0.03	0.12	0.11	-0.01	0.07
				(0.00,0.24)	(-0.04,0.25)	(-0.11,0.10)	(-0.04,0.18)
	AC(3)	0.09	0.02	0.05	-0.05	-0.01	0.03
				(-0.10,0.20)	(-0.25,0.15)	(-0.11,0.10)	(-0.08,0.15)

References

[1] Allenby, G.M., Rossi, P.E., McCulloch R., 2005. Bayesian Statistics and Marketing, Wiley.

[2] Bansal, R., Dittmar, R.F., Lundblad, C.T., 2005. Consumption, Dividends and the Cross Section of Equity Returns. Journal of Finance 60, 1639-1672.

[3] Carter, C.K., Kohn R., 1994. On Gibbs Sampling for State Space Models. Biometrika 81, 541-553.

[4] Duffie, D., Epstein, L.G., 1992. Asset Pricing with Stochastic Differential Utility. Review of Financial Studies 5, 411-436.

[5] Frühwirth-Schnatter, S., 1994. Data Augmentation and Dynamic Linear Models. Journal of Time Series Analysis 15, 183-202.

[6] Zurek, P., 2007. Momentum and Long-Run Risks. Working Paper, University of Pennsylvania.