

Habit, Production, and the Cross-Section of Stock Returns

Andrew Y. Chen

Federal Reserve Board

andrew.y.chen@frb.gov

January 2014*

Abstract

Solutions to the equity premium puzzle should inform us about the cross-section of stock returns. An external habit model with heterogeneous firms reproduces numerous stylized facts about both the equity premium and the value premium. The equity premium is large, time-varying, and linked with consumption volatility. The cross-section of expected returns is log-linear in B/M, and the slope matches the data. The explanation for the value premium lies in the interaction between the cross-section of cash flows and the time-varying risk premium. Value firms are temporarily low productivity firms, which will eventually experience high cash flows. The present value of these temporally distant cash flows is sensitive to risk premium movements. The value premium is the reward for bearing this sensitivity. Empirical evidence verifies that value firms have higher cash-flow growth. The data also show that value stock returns are more sensitive to risk premium movements, as measured by consumption volatility shocks.

*I would like to thank Steve Sharpe, René Stulz, and Lu Zhang for helpful comments, as well as Julia Thomas and Aubhik Khan for the training that made this work possible. I would also like to thank Lars Kuehn and Oliver Boguth for making their data available. The views expressed herein are those of the author and do not necessarily reflect the position of the Board of Governors of the Federal Reserve or the Federal Reserve System.

1. Introduction

In the decades since the publication of Mehra and Prescott (1985)'s equity premium puzzle, economics and finance have produced a handful of models that rationalize the high average returns and volatility of the aggregate stock market. Whether these solutions also help us understand the cross-section of stock returns has received less attention. Exploring this question is important because what we ultimately want is not merely a solution to the equity premium puzzle but a framework for understanding asset prices in general. A model that explains the return on the aggregate stock market but cannot address any other asset leaves much to be desired.

In this paper, I show that one solution to the equity premium puzzle, habit formation, provides insight into the cross-section of stock returns. I embed a variant of Campbell and Cochrane (1999) external habit preferences in a model with heterogeneous firms. The model reproduces numerous stylized facts about the value premium. Both in the model and in the data, expected returns are linear in the log of book-to-market (B/M). Moreover, the model matches the slope of the relationship. The slope on $\log(B/M)$ is approximately five, indicating that a 20% higher B/M implies a 100 b.p. increase in expected returns over the next year. These cross-sectional predictions come with predictions about the equity premium and business cycle that are consistent with the data. The model matches the first two moments of aggregate excess returns, the risk-free rate, consumption, output, and investment, as well as excess return and dividend predictability regressions.

The model's equity premium dynamics are critical to understanding the cross-section of stock returns. The equity premium is time-varying and persistent, resulting in a cross-sectional equity term premium: stocks with temporally distant cash flows earn higher returns. To understand this, note that when the equity premium rises, all stock prices fall, since cash flows are discounted more aggressively. Stocks with temporally distant cash flows are hit harder, however. This is because the rise in discount rates is persistent. Thus, the effects of the rise are compounded for distant cash flows. Since the equity premium rises in bad times, investors demand high returns in exchange for these negative price reactions.

This link between time-varying equity premia and an equity term premium

is found in a number of models (Santos and Veronesi (2010), Chen (2012)). This model differs in that value firms (high B/M) endogenously have temporally *distant* cash flows, and thus higher expected returns. Cash flows are the result of investment and production but, for the basic intuition, one can simply examine productivity. Value firms have low productivity today and, as a result, produce little cash flow. Productivity eventually recovers, however, as will the cash flows of value firms. Thus, value firms have temporally distant cash flows. Growth firms (low B/M) follow the opposite pattern: high productivity and cash flow today, normal productivity and cash flow later, and thus temporally *close* cash flows. The existence of an equity term premium then leads to a value premium.

This explanation of the value premium is consistent with three empirical facts about value firms: compared to growth firms, value firms (1) have lower return on equity (Fama and French (1995)), (2) higher cash-flow growth (Lakonishok, Shleifer, and Vishny (1994), Chen (2012)), and (3) more negative discount rate betas (Campbell and Vuolteenaho (2004), Campbell, Giglio, Polk, and Turley (2012)). The first empirical fact is well established. The second and third are not, and so I present additional evidence showing their robustness.

The second empirical fact, that value firms have higher cash-flow growth, is controversial. This fact conflicts with the basic intuition that “growth” firms should grow, and thus value firms should not. In the value premium literature, however, the terms “value” and “growth” refer to high and low B/M. Low B/M stocks may not grow much at all, and indeed the empirical evidence for low B/M stocks having high cash-flow growth comes entirely from two studies of equity duration: Dechow, Sloan, and Soliman (2004) and Da (2009). The trouble with duration studies is that for equities, duration is very difficult to measure. Measuring duration requires estimating both a discount rate and terminal value, both of which are very difficult to identify at the firm-level.

Measuring cash-flow growth, on the other hand, is straightforward. As recently emphasized by Chen (2012), by a number of measures, value firms have higher cash-flow growth, and thus temporally distant cash flows.¹ I present additional empirical evidence that value firms have high cash-flow growth. Existing papers focus on dividends, but there are other, arguably more appropriate, measures of cash flows. I also consider earnings before extraordinary income, earnings after extraordinary income, earnings plus depreciation, and cash flows from

¹See also Lakonishok, Shleifer, and Vishny (1994), Bansal, Dittmar, and Lundblad (2005), Hansen, Heaton, and Li (2008), Chen, Petkova, and Zhang (2008).

operating and investing activities. I examine buy-and-hold portfolios, which arguably are a better match for the model's mechanics. I find that, regardless of the definition of cash flow, value firms have higher cash-flow growth.

The remaining empirical fact is that value firms have more negative discount rate betas. Campbell and Vuolteenaho (2004) and Campbell, Giglio, Polk, and Turley (2012) find evidence for this view, but they use vector autoregression frameworks which may be sensitive to specification (Chen and Zhao (2009)). Since discount rates are not directly observable, one cannot entirely avoid this critique, but one can provide supporting evidence using an entirely different framework. To this end, I examine consumption volatility betas of value and growth firms. Unlike standard external habit models, the discount rate shocks of this model are closely linked with time-varying consumption volatility.² To measure consumption volatility, I use Boguth and Kuehn (2013)'s Markov chain of the components of consumption growth. I find that value firms have more negative betas with respect to movements in consumption volatility. This relationship is seen both in the consumption volatility betas of book-to-market portfolios, as well as in firm-level regressions.

The paper proceeds as follows. In the remainder of this section I discuss related literature. Because the relationship between value and cash-flow growth is controversial, I then present empirical evidence on value and cash-flow growth in Section 2. Section 3 presents the model, solution method, and calibration. Section 4 shows that the model addresses numerous equity premium facts as well as facts about the value premium. Section 5 inspects the mechanism for generating the value premium and provides empirical evidence about value and discount rate shocks. Section 6 concludes.

Related Literature The model builds off of Santos and Veronesi (2010), who also study the cross-section of stock returns in an external habit model. This paper can be considered an extension of their model into a production economy. Adding production reverses the value – expected return relationship. Without production, value stocks are characterized by a high dividend price ratio. Mean reversion implies low dividend growth and low exposure to discount rate shocks. In contrast, a model with production characterizes value with book-to-market. Value firms are then low productivity firms, and mean reversion implies high

²This result is due to the presence of production and is analyzed in Chen (2014).

cash-flow growth and high exposure to discount rate shocks.

In addition to Santos and Veronesi (2010), this paper is closely related to Chen (2012). Chen (2012) shows that any model with time-varying risk premia will also have a cash-flow growth premium, and that this cash flow premium is empirically related to the value premium. While Chen (2012)'s theoretical results are reduced form and qualitative, the model of this paper is structural and quantitative. Thus, this paper can be considered a formal general equilibrium foundation for the earlier results.

A number of papers also generate a value premium in a Q-theoretical framework (Zhang (2005), Carlson, Fisher, and Giammarino (2005), Cooper (2006)). In these papers, the cash flows of value firms are more cyclical as a result of operating leverage or costly reversibility. These technological features are absent from the model of this paper. Indeed, in this model, the relationship between cash flows and aggregate state variables does depend significantly on B/M. Thus, the mechanism underlying the value premium in this paper is distinct from existing Q-theory models. Empirically, the cash flow risk channel does find some support (Cohen, Polk, and Vuolteenaho (2009), Santos and Veronesi (2010)). This channel can be added to the model with the inclusion of operating leverage or costly reversibility but is beyond the scope of this paper.

A handful of papers model the equity premium and the cross-section in a long-run risk setting. These papers find that the long-run risk framework is consistent with several facts about the cross-section. Avramov, Cederburg, and Hore (2011) find size, value, and momentum effects in an endowment economy. Favilukis and Lin (2011) and Ai and Kiku (2012) find value effects in production economies. While these papers successfully generate a large equity premium and volatile excess returns, they use the version of the long-run risk model which does not produce time-varying risk premia. This paper overcomes this issue by using the external habit framework, which does produce time-varying risk premia. Gabaix (2012) time-varying disaster model produces time-varying risk premia and is qualitatively consistent with the value premium, but Gabaix (2012) does not provide a quantitative analysis.

2. Empirical Evidence: Value Firms Have High Cash-Flow Growth

The model's mechanism relies on value firms having higher cash-flow growth than growth firms. In this section, I discuss the existing empirical evidence regarding the mechanism, and present new evidence in support of the mechanism.

The conventional view is that value firms have low cash-flow growth. Evidence for this view comes exclusively from equity duration studies, namely Dechow, Sloan, and Soliman (2004) and Da (2009). Both papers find that value firms have short durations, and thus, low cash-flow growth. Equity duration, however, is very difficult to measure. Generally speaking, the duration of an asset is something like

$$\text{Duration} = \sum_{t=0}^{\infty} \frac{\text{PV}(\text{CF}_t)}{P_0} t.$$

Note that measuring duration requires both a discount rate and a terminal value. While these two are both directly observable for bonds, they are extremely difficult to measure for equities.

Dechow, Sloan, and Soliman (2004) and Da (2009) rely on identifying assumptions which bias them toward finding that value firms have low durations. In particular, Dechow, Sloan, and Soliman (2004) assume that the terminal value is equal to the market value, less some present value of the next 10 years of cash flow. Since value firms have low market value compared to current cash flows, this leads to a low terminal value and, thus, a low duration. Da (2009) assumes that the terminal ROE is equal to the mean of ROE for the first seven years after portfolio formation. Since value firms have low ROE at portfolio formation (Fama and French (1995)), this assumption leads to a low terminal value and, thus, low duration. These biases are pointed out by Chen (2012).

Measuring cash-flow growth is much more straightforward. Indeed, a number of papers focused on other issues happen to provide summary statistics regarding value and cash-flow growth (Lakonishok, Shleifer, and Vishny (1994), Bansal, Dittmar, and Lundblad (2005), Hansen, Heaton, and Li (2008)). These papers uniformly find that value portfolios have higher cash-flow growth.

Unfortunately, even cash-flow growth offers multiple methods of measurement. For example, the previously mentioned papers use portfolios that are re-

balanced periodically using various methods. Chen (2012) provides a detailed examination of cash-flow growth and value in buy-and-hold portfolios. For the majority of his methods of analysis, he finds that value firms have high cash-flow growth.

This section presents additional empirical evidence regarding value and cash-flow growth. In particular, Chen (2012) focuses on dividends and earnings before extraordinary income. Other definitions of cash flow are arguably more appropriate empirical targets for models. Existing models, including the one in this paper, abstract from dividend policy (Miller and Modigliani (1961)). As a result, model dividends are equal to net income plus depreciation less net investment (cash flow from operating and investing activities). Both sides of the equation can be considered as cash flow, and the right-hand side is arguably a more appropriate empirical target since it captures the real (non financial) activities of the firm.

I look at four different notions of cash flow. The first is a common notion of cash flow, earnings before extraordinary income (ib). This measure is stable and reflects the ongoing activities of the firm, but investors must face the consequences of extraordinary income, and thus the stock price should reflect these items. Thus, I also look at earnings (ni), which includes extraordinary income. Earnings, however, reflect depreciation charges, which are not represented in cash flows in the model. I thus also examine earnings plus depreciation (ni + dp). Lastly, cash flows in the model can also come from selling / buying capital. Thus I also examine earnings plus depreciation less net investment (ni + dp - capx + sppe). This last measure is closest in spirit to the cash flows of the model.

I use tercile book-to-market sorted portfolios and CRSP and COMPUSTAT data from 1971–2011. The portfolios are buy-and-hold portfolios. Cash flows from delisted stocks are reinvested in the remaining stocks, following Chen (2012)'s procedure. The choice of terciles is due to the use of net investment. Net investment is quite volatile, and the use of large portfolios averages out much of this volatility and paints a clearer picture of the typical cash flow dynamics. The relatively short post-1971 sample is due to the limited availability of sales of plant, property, and equipment (spp) data. The relatively modern sample period is useful, however, in that one of the only measurements where Chen (2012) finds that value firms have low cash-flow growth is when he looks at dividend growth for buy-and-hold portfolios in the post-1963 sample. I will show that, during a similar sample period, many other definitions of cash flow present the

opposite picture for buy-and-hold portfolios.

Table 1 shows cash flow levels. It shows the cash flow from a \$1 investment in value or growth portfolios, averaged across portfolio formation years. The first thing that jumps out from the table is that value stocks do not have significantly lower cash flows than growth stocks. In the first year after portfolio formation, value stocks pay 7 cents per dollar invested while growth stocks pay 6 cents, with respect to earnings before extraordinary income. In fact, net of extraordinary income and investment, value firms pay much less. Using this definition, in the first year value pays half a cent per dollar while growth pays an order of magnitude more. The second pattern which emerges from the table is that value has higher cash-flow growth. There is little action in the cash flows of growth firms, but the value cash flows exhibit apparent growth.

[Table 1 about here]

Table 2 shows growth rates of the cash flows from the previous figure. It also considers two additional definitions of cash flow: earnings (after extraordinary income) and earnings plus depreciation. By all definitions of cash flow, value portfolios have much higher cash-flow growth than growth portfolios in year two. Indeed, using the definition closest in spirit to the model (earnings plus depreciation less net investment), value experiences a huge 544% growth in cash flow between years one and two, while growth gets a meager 5% growth. Cash-flow growth is also monotonically increasing in B/M using all definitions. Cash-flow growth of value exceeds that of neutral which exceeds that of growth. An additional pattern which is seen in Table 2 is that the growth rates mean revert. Value begins with strikingly high cash-flow growth in year two, but growth slows down quickly. Growth portfolios follow the opposite pattern. Both the high cash-flow growth of value portfolios and its subsequent mean reversion will be seen in the model.

[Table 2 about here]

Analyzing cash-flow growth for the longer term faces data limitations. Forty years of data provides only 10 non overlapping four-year periods. Thus, it is probably best to focus on the year two and year three growth rates. Nevertheless, the fact that the cash flow patterns are common across multiple definitions of cash flow is reassuring.

3. A General Equilibrium Model with Heterogeneous Firms

Having established that the key element of the mechanism is consistent with the data, I now present the model. The model is an real business cycle model with external habit formation, capital adjustment costs, and idiosyncratic firm productivity. It is designed to have the minimal features for both an equity premium and an endogenous cross-section of firms. It is essentially a heterogeneous firm extension of Chen (2014), and thus differs from traditional habit models by featuring a time-varying consumption volatility channel.

Markets are complete, and time is discrete and infinite. For the remainder of the paper, lowercase denotes logs—that is, $c_t \equiv \log C_t$.

3.1. Representative Household

A unit measure of identical households $j \in [0, 1]$ chooses asset holdings to maximize lifetime utility

$$\mathbb{E}_0 \left\{ \sum_{t=0}^{\infty} \beta^t \frac{(C_{j,t} - H_t)^{1-\gamma}}{1-\gamma} \right\}, \quad (1)$$

where H_t , the level of habit, is determined by aggregate consumption and is taken as external by the household.

I specify the evolution of habit using surplus consumption, rather than the level of habit itself. That is, let

$$s_t \equiv \frac{C_t - H_t}{C_t}, \quad (2)$$

be the surplus consumption ratio, where C_t is aggregate consumption. Then surplus consumption follows an AR1-process in logs

$$s_{t+1} \equiv (1 - \rho_s) \bar{s} + \rho_s s_t + \lambda(c_{t+1} - c_t). \quad (3)$$

This approach leads to a simple stochastic discount factor and makes for ease of comparison with the existing literature on external habit (Campbell and Cochrane (1999), and Wachter (2006), among others).

The habit process differs in that the conditional volatility λ is a constant. This is a substantial deviation from the literature on external habit models in endowment economies, which specify this conditional volatility as as time-varying and countercyclical (e.g. Campbell and Cochrane (1999), Menzly, Santos, and Veronesi (2004)). In Chen (2014), I show that the introduction of production results in countercyclical consumption volatility, which is quantitatively very similar to the assumed countercyclical volatility of surplus consumption typical of endowment economy models. For comparability with Campbell and Cochrane (1999), I fix λ at their steady state value

$$\lambda = \frac{1}{\bar{S}} - 1. \quad (4)$$

The fact that markets are complete and external habit specification mean that the household side of the model boils down to a simple stochastic discount factor

$$M_{t,t+1} = \beta \left(\frac{C_{t+1}}{C_t} \frac{S_{t+1}}{S_t} \right)^{-\gamma}. \quad (5)$$

3.2. Heterogeneous Firms

There is a unit measure of heterogeneous firms, indexed by $i \in [0, 1]$. The firms produce consumption using capital $K_{i,t}$

$$\Pi(K_{i,t}, B_{i,t}, A_t) = A_t B_{i,t} K_{i,t}^\alpha, \quad (6)$$

where aggregate productivity A_t and idiosyncratic productivity $B_{i,t}$ are both AR1 processes in logs:

$$a_{t+1} = \rho_a a_t + \sigma_a \epsilon_{a,t+1} \quad (7)$$

$$b_{i,t+1} = \rho_b b_{i,t} + \sigma_b \epsilon_{b,i,t+1}, \quad (8)$$

where $\epsilon_{a,t+1}$ and $\epsilon_{b,i,t+1}$ are independent standard normal random variables.

All heterogeneity in the models originates from the idiosyncratic productivity process (8). This approach is used for three reasons. The first is that it is a very simple way of introducing a cross-section of firms. The second is that a large literature documents substantial heterogeneity in productivity (Syverson (2011)). The third is that this approach is the standard way of modeling firm heterogene-

ity in both macroeconomics and finance (Hennessy and Whited (2005), Zhang (2005), Khan and Thomas (2008), Bloom (2009)). We will see, however, that this approach has difficulties matching the tremendous heterogeneity in asset valuations that is seen in the data. Matching the heterogeneity in the data with additional sources of heterogeneity is an interesting question for future research—however, it is beyond the scope of this paper.

Capital accumulates according to the usual capital accumulation rule,

$$K_{i,t+1} = I_{i,t} + (1 - \delta)K_{i,t}, \quad (9)$$

and firms face a convex capital adjustment cost

$$\Phi(I_{i,t}, K_{i,t}) = \frac{\phi}{2} \left(\frac{I_{i,t}}{K_{i,t}} - \delta \right)^2 K_{i,t}. \quad (10)$$

I assume that the adjustment costs are a pure loss. They do not represent payments to labor. Adjustment costs are included because production economies produce a counterfactually smooth Tobin's Q unless one includes an investment friction. Quadratic costs are chosen for simplicity, but a richer model would incorporate investment frictions by modeling an investment goods sector, as in Boldrin, Christiano, and Fisher (2001), or would feature heterogeneous plants with non-convex costs of adjustment, as in Khan and Thomas (2008).

Because of complete markets, the firm's objective is standard:

$$\max_{\{I_{i,t}, K_{i,t+1}\}} \mathbb{E}_0 \left\{ \sum_{t=0}^{\infty} M_{0,t} [A_t B_{i,t} K_{i,t}^\alpha - I_{i,t} - \Phi(I_{i,t}, K_{i,t})] \right\}. \quad (11)$$

It chooses investment and capital to maximize future dividends, discounted with the household's SDF.

3.3. Recursive Competitive Equilibrium

Equilibrium is defined recursively. Thus, in the remainder of this section I drop the time subscripts and represent the next period with primes. Let μ represent the distribution of firms over capital K_i and idiosyncratic productivity B_i . The aggregate state is the triple of the distribution of firms μ , surplus consumption S , and aggregate productivity A .

The recursive competitive equilibrium is laws of motion for the distribution

of firms $\Gamma(\mu, S, A)$ and aggregate consumption $C(\mu, S, A)$, a capital policy for the firm $G(K_i, B_i; \mu, S, A)$ and value function for the firm $V(K_i, B_i; \mu, S, A)$ such that

1. Firm optimality holds: $G(K_i, B_i; \mu, S, A)$ and $V(K_i, B_i; \mu, S, A)$ solve

$$V(K_i, B_i; \mu, S, A) = \max_{\{I, K'_i\}} \left\{ \Pi(K_i, A, B_i) - \Phi(I, K_i) + \int_{-\infty}^{\infty} dF(\epsilon'_a) \int_{-\infty}^{\infty} dF(\epsilon'_b) M(A'; \mu, S, A) \times V(K'_i, B'_i; \mu', S', A') \right\}, \quad (12)$$

where the productivity processes are given by (7) and (8), capital accumulation is given by (9), the SDF is the household's intertemporal marginal rate of substitution

$$M(A'; \mu, S, A) = \beta \left(\frac{C(\mu', S', A') S'}{C(\mu, S, A) S} \right)^{-\gamma}, \quad (13)$$

S' evolves according to (3), μ' is given by $\Gamma(\mu, S, A)$, and $F(\epsilon'_a)$ is the standard normal cumulative distribution function.

2. Firm decisions are consistent with the law of motion for consumption:

$$C(\mu, S, A) = \int d\mu(K_i, B_i) \{ \Pi(K_i, B_i, Z) - \Phi(I(K_i, B_i; \mu, S, A), K_i) \} \quad (14)$$

where $I(K_i, B_i; \mu, S, A) = G(K_i, B_i; \mu, S, A) - (1 - \delta)K_i$.

3. Firm decisions are consistent with the law of motion for the distribution of firms—that is, let \mathbb{B} be the Borel algebra for \mathbb{R}_+^2 . Then $\mu' = \Gamma(\mu, S, A)$ is given by

$$\begin{aligned} \forall (K_1, B_1) \in \mathbb{B}, \\ \mu'(K_1, B_1) = \int_{\{(K_i, B_i) \mid G(K_i, B_i; \mu, S, A) \in K_1\}} d\mu(K_i, B_i) \int_{\{\epsilon'_b \mid \exp(\rho_b b + \sigma_b \epsilon'_b) \in B_1\}} dF(\epsilon'_b). \end{aligned} \quad (15)$$

3.4. Krusell-Smith Solution Method

I solve the model with a variant of the Krusell and Smith (1998) algorithm, similar to Khan and Thomas (2008). As in Khan and Thomas (2008), I approxi-

mate the distribution of firms μ with the aggregate capital stock K . Thus, the approximate aggregate state is a triple of aggregate capital, surplus consumption, and aggregate productivity: (K, S, A) .

I discretize the aggregate and idiosyncratic productivity processes (7) and (8) using the Rouwenhorst (1995) method. I then conjecture that the laws of motion for aggregate consumption and capital follow the following log-linear forms:

$$\begin{aligned} c' &= \log \tilde{C}(K, S, A_j) = \theta_{0,j}^C + \theta_{1,j}^C k + \theta_{2,j}^C \hat{s} \\ k' &= \log \tilde{\Gamma}(K, S, A_j) = \theta_{0,j}^\Gamma + \theta_{1,j}^\Gamma k + \theta_{2,j}^\Gamma \hat{s}, \end{aligned} \quad (16)$$

where $j \in \{A_1, \dots, A_{N_A}\}$ represents the aggregate productivity state. Note that the use of aggregate productivity dependent coefficients allows for a non-linear relationship between consumption and the aggregate state.

The goal of the Krusell-Smith method is to find the coefficients $\theta_{i,j}^C, \theta_{i,j}^\Gamma$ such that

1. Firms maximize value given the laws of motion $\theta_{i,j}^C, \theta_{i,j}^\Gamma$
2. Estimates of (16) on simulated data using policies from step 1 produce coefficients close to $\theta_{i,j}^C, \theta_{i,j}^\Gamma$, and R^2 's close to one.

The most straightforward application of Krusell-Smith searches for this approximate equilibrium by doing a fixed-point iteration using the firm's problem defined in (12) and a simulation of a distribution of firms. However, there is no theorem that suggests that this fixed-point iteration will converge, and indeed I find that it typically does not.

To aid in finding equilibrium, I apply the 'equilibrium-in-simulation' method (Krusell and Smith (1997))—that is, I first solve solve the approximate equilibrium version of (12). I then plug the resulting value function into the following problem:

$$\begin{aligned} G(K_i, B_i; K, S, A; C) &= \arg \max_{\{I, K'_i\}} \left\{ \Pi(K_i, A, B_i) - \Phi(I, K_i) \right. \\ &\quad \left. + \int_{-\infty}^{\infty} dF(\epsilon'_a) \int_{-\infty}^{\infty} dF(\epsilon'_b) M^*(A'; K, S, A; C) \right. \\ &\quad \left. \times V(K'_i, B'_i; K', S', A') \right\} \quad (17) \end{aligned}$$

$$M^*(A'; K, S, A; C) = \beta \left(\frac{\tilde{C}(K', S', A') S'}{CS} \right)^{-\gamma}. \quad (18)$$

This procedure introduces today's aggregate consumption as an additional state variable and solves for a new investment policy which accounts for aggregate consumption. I then use this augmented investment policy $G(K_i, B_i; K, S, A; C)$ in the simulation step. This allows me to find a 'market-clearing' C at each date in the simulation. That is, at each date, I use a root finder to find the C that solves equation (14). Note that once the equilibrium is found, aggregate consumption from the simulation of the firms and that produced by the law of motion are equal, and so problem (17) with market clearing (14) and problem (12) produce identical choices.

The presence of external habit significantly complicates the computationally demanding Krusell-Smith algorithm. External habit preferences introduce an additional aggregate state variable, surplus consumption, which is completely absent from the standard RBC economy. As a result, using the RBC equilibrium as an initial guess for the Krusell-Smith algorithm will cause the algorithm to fail. To address this problem, I apply a homotopy method. I solve a series of models with the following altered SDF

$$M' = \beta \left(\frac{C'}{C} \left(\frac{S'}{S} \right)^\chi \right)^{-\gamma}. \quad (19)$$

I begin by solving the model with $\chi = 0$. Here the RBC model serves as a good initial guess. Once the program is fairly close to equilibrium, I increase χ by 0.1 and use the previous laws of motion as an initial guess for the new model. I repeat this process until $\chi = 1.0$, which is equivalent to the model presented in (3).

Surplus consumption also adds the difficulty that it is an endogenous state variable that is not predetermined. As a result, the habit process equation (3) must be solved at every date in the simulation step of the algorithm. Note that the simulation step involves simulating an entire distribution of firms, and so an entire distribution of decision rules must be accounted for in solving equation (3). This also significantly increases the computational burden of the algorithm.

3.5. Calibration to Post-War U.S. Data

The model is calibrated to post-war (post 1947) U.S. data. This sample period is chosen because the World Wars introduce structural changes that may not be captured by the model. In particular, over the long sample (post-1929) HP-filtered output and investment are essentially uncorrelated.

Aggregate quantities are taken from the BEA. Firm-level data are taken from CRSP/COMPUSTAT. B/M-sorted portfolio are taken from Ken French's website. Aggregate asset-price moments are taken from Beeler and Campbell (2009).

Table 3 shows the calibration. Preference parameters are chosen as much as possible to fit unconditional moments of asset prices. Since time preference β is reflected in risk-free assets, I choose it to fit the mean 30-day T-bill return. The model and data T-bill returns match nicely at about 1% per year. The persistence of surplus consumption ρ_s affects the persistence of asset prices. Thus I choose ρ_s to approximately match the annual persistence of the CRSP price/dividend ratio of 0.87. The two remaining preference parameters, the steady state surplus consumption \bar{S} and utility curvature γ , jointly control risk aversion. Thus, it is difficult to identify these parameters separately. For ease of comparison with the literature on external habit, I choose $\gamma = 2$ to match Campbell and Cochrane (1999)'s value, and then choose \bar{S} to fit the mean Sharpe ratio of the CRSP index. The model does a good job matching the data here: both Sharpe ratios are roughly 0.40.

[Table 3 about here]

I choose aggregate technology parameters to fit moments of the real economy. The production curvature α and depreciation rate δ are chosen to fit the capital-output ratio and the mean growth-adjusted investment rate. These moments match nicely at 0.40 and 0.08 respectively. The volatility of aggregate productivity σ_a is chosen to fit the volatility of HP-filtered log GDP (I use a smoothing parameter of 6.25, as argued by Ravn and Uhlig (2002)). The data and model match well in this dimension, producing a volatility of about 1.5%. The persistence of aggregate productivity is chosen to fit the persistence of the Solow residual with constant labor. A critical parameter of the production technology is the quadratic adjustment cost parameter ϕ . I choose this parameter value to hit the volatility of aggregate consumption growth (non durables and services). The data and model match well here, producing a volatility of about 1.5% per year.

Firm-level technology parameters are chosen to fit the cross-sectional means of time-series moments. I target non-micro-cap firms (firms with market equity of more than 600 million). The persistence of idiosyncratic productivity ρ_b is chosen to fit the persistence of firm-level ROE. The volatility of idiosyncratic productivity σ_b is chosen to match the volatility of individual stock returns.

4. Quantitative Results

I begin by showing that the model addresses equity premium puzzles (Section 4.1). This section mostly verifies the results of Chen (2014), so the discussion will be brief. Section 4.2 contains the main quantitative results. It shows that the model reproduces the value premium described by firm-level regressions.

4.1. Matching the Data on Equity Premium Puzzles

Table 4 shows aggregate asset-price moments. The model produces a large and volatile equity premium, and a low and smooth risk-free rate. The large and volatile equity premium is typical of habit models (Jermann (1998), Campbell and Cochrane (1999)). The smooth risk-free rate comes from time-varying consumption volatility that counterbalances the powerful intertemporal substitution effect of habit models. Time-varying volatility, in turn, comes from the “precautionary volatility” channel of Chen (2014). In bad times, the household is unsure of how much precautionary savings it needs next period. This motive and production lead to countercyclical consumption volatility.

[Table 4 about here]

Table 5 shows regressions of future dividend growth and excess returns on the price-dividend ratio. Because the price-dividend ratio must predict dividend growth or returns, these regressions form a nice characterization of the drivers of asset-price fluctuations (Campbell and Shiller (1988), Cochrane (2011) and Kojien and Van Nieuwerburgh (2010)). The table shows that, as in the data, the price dividend ratio has little predictive power for future dividend growth, and has strong predictive power for future excess returns. These results show that the model captures the nature of stock market fluctuations. These time-varying excess returns also come from endogenous time-varying consumption volatility (see Chen (2014)).

[Table 5 about here]

Table 6 shows basic business cycle moments. As intended by the calibration, the model produces low consumption volatility. As in the data, investment is much more volatile than output and consumption is much less volatile. The model also replicates the co-movement of consumption, investment, and GDP.

[Table 6 about here]

Taken together, Tables 4, 5, and 6, show that the model does a good job of addressing the equity premium puzzles. The model generates a large and volatile equity premium, low and smooth risk-free rate, and asset valuations that are tied to expected excess returns. These data-like asset-price dynamics come with good predictions about basic business cycle variables.

4.2. Matching the Data on the Cross-Section of Stock Returns

We have seen that the model is able to address the equity premium puzzles. External habit combined with production produces a large and volatile equity premium, a low and smooth risk-free rate, and asset price fluctuations that are linked to excess returns rather than the price-dividend ratio. This brings us to the main question of the paper. Is external habit consistent with the cross-section of stock returns?

Table 7 examines this question and shows the main result of the paper. It shows regressions of next year's returns on today's log B/M. The regressions are firm-level and follow the Fama and MacBeth (1973) method. In both the model and data, the $\log(B/M)_{i,t}$ coefficient is positive and highly statistically significant. Moreover, the slopes are large and similar in magnitude, with a value of about five. This slope means that in both the model and data, a 20% higher B/M implies a roughly 100 b.p. higher expected return.

[Table 7 about here]

Though firm-level regressions provide the most statistical power and offer the simplest quantitative description of the value premium, the literature often examines value-weighted portfolio sorts. Table 8 shows summary statistics on decile B/M-sorted portfolios. The expected returns columns show that, in both the model and data, expected returns are monotonically increasing in B/M. In the model, expected returns culminate to an economically significant decile 10-1 return of about 2.5% per year.

[Table 8 about here]

The decile value premium is smaller than that of the data, but it comes with a much smaller dispersion in B/M. Recall that all heterogeneity in the model originates from a single AR1 idiosyncratic productivity process. For parsimony and

to maintain clarity of the mechanism, the model abstracts from other sources of heterogeneity such as differences in financial frictions or life-cycle effects. As a result, the model generates a dispersion in B/M that is significantly less than the data. In the model, $\log B/M$ differs by 0.4 between the high and low deciles. In the data, this difference is 1.4. This difference in spreads means that the traditional high-low portfolio returns of the model are not comparable to the data. Reproducing the enormous B/M dispersion in the data is an interesting question, but is beyond the scope of this paper.

A more effective way to illustrate the portfolio sort results is to interpret them as a nonparametric regression (Cochrane (2011)). Figure 1 provides this interpretation. It plots the average returns of the 10 B/M-sorted portfolios against $\log B/M$. Returns are equal-weighted in this figure because the functional form in the data reflects equal-weighting, that is, value-weighted returns do not result in a log linear pattern in the data. The figure shows that the model captures this log-linear form. The match is not only qualitative but quantitative too. The slope of the the relationship between expected returns and $\log B/M$ is similar in both model and data, consistent with the results of the firm-level regressions (Table 7).

[Figure 1 about here]

5. Inspecting the Mechanism

In this section I explain how this value premium works. I begin by showing that value firms have high cash-flow growth (Section 5.1), consistent with the empirical results of Section 2. I go on to illustrate how high cash-flow growth leads to high expected returns through exposure to discount rate shocks (Section 5.2). I then provide empirical evidence for this high exposure (Section 5.3).

The last two subsections provide robustness. Section 5.4 rules out other potential mechanisms. Section 5.5 illustrates the role of general equilibrium.

5.1. Value and Cash-Flow Growth

To gain a picture of the mechanism, it helps to start with the meaning of value in the model. Figure 2 plots B/M and expected returns against the two firm state variables, idiosyncratic productivity and capital. The left panel shows

that value firms are low productivity firms with high capital. Capital, however, is slow-moving. As a result, value is primarily characterized by low productivity. The right panel shows how value is connected to expected returns. Expected returns decline strongly in idiosyncratic productivity. Overall, we see that value firms are low productivity firms that have high expected returns, consistent with the quantitative results of Section 4.

[Figure 2 about here]

The low productivity of value firms leads to cash-flow growth. This relationship is illustrated in Figure 3 which plots the cash-flow growth of portfolios sorted on B/M. The darkest lines show value portfolios. Soon after portfolio formation, value portfolios have high dividend growth, but this growth slows down quickly and eventually reaches the average growth rate of zero (the model abstracts from balanced growth). Intuitively, value firms have low productivity, but this low productivity is temporary. Mean reversion implies that productivity grows, and so cash flows grow. This contrasts with growth portfolios (the lightest lines), which show the reverse pattern. When growth firms are declared as growth, they have very high productivity. Mean reversion then means that their productivity will fall, leading to low cash-flow growth. Both this initial spread in cash-flow growth and its subsequent mean reversion are consistent with the empirical evidence of Section 2. In terms of magnitudes, the dispersion in cash-flow growth is similar to the cash flow measure that includes investment in Table 2, though the measurement of cash-flow growth in the data is quite noisy.

[Figure 3 about here]

While mean reversion provides a simple story of the relationship between value and cash-flow growth, the exact relationship is endogenous. Investment also affects cash flows and is not discussed in the simple story above. Indeed, low productivity tends to encourage disinvestment, increasing cash flows today and decreasing cash-flow growth. Whether the productivity channel or the investment channel dominates depends on the parameter choices.

In fact, the model without habit produces the counterfactual prediction that value firms have low cash-flow growth. This is seen in Figure 4, which plots cash flows for B/M portfolios in a model without habit. As with the habit model, the adjustment costs in this model are chosen to match the volatility of consumption

growth. The figure shows exactly the opposite pattern of figure 3, with growth firms (dotted line) beginning with high cash-flow growth and eventually mean reverting to the steady state of no growth. Here, the investment channel dominates the production channel. Intuitively, removing habit increases the elasticity of intertemporal substitution of the household, and requires that the adjustment costs be lowered in order to fit the volatility of consumption. With low costs of adjusting capital, that is, low costs of investing, the investment channel becomes dominant.

[Figure 4 about here]

5.2. Cash-Flow Growth and Expected Returns

So far, I have shown that value is characterized by low productivity and high cash-flow growth. To finish the story, I need to show that high cash-flow growth leads to high expected returns. This last link is due to the large discount rate shocks that drive the external habit model. High cash-flow growth means that cash flows are distributed far into the future, and thus are more exposed to large discount rate shocks. Investors then demand high returns in exchange for bearing this higher exposure. A number of previous papers discuss this link (Cornell (1999), Lettau and Wachter (2007), Santos and Veronesi (2010), Chen (2012)), but here I provide a new sketch of the intuition. This sketch uses simple Investments 101 formulas and so steps outside of the general equilibrium model.

Consider a growing perpetuity

$$P_0 = \frac{D_1}{\kappa_0 - g}, \quad (20)$$

where κ_0 is the discount rate and g is the growth rate of cash flows. Now suppose that the discount rate gets hit by a completely unexpected shock $\Delta\kappa$. The price next period is then

$$P_1 = \frac{D_1(1+g)}{\kappa_0 + \Delta\kappa - g}. \quad (21)$$

Taking a first-order Taylor approximation of the definition of the return gives us

$$R_1 \equiv \frac{D_1 + P_1}{P_0} \approx (1 + \kappa_0) - \left(\frac{1+g}{\kappa_0 - g} \right) \Delta\kappa. \quad (22)$$

If discount rates suddenly go up, the stock price takes a hit, and so we have a negative sign on the second term. Notice also that the second term is increasing in the cash-flow growth rate g . Discount rate shocks hit high cash-flow growth assets particularly hard. Intuitively, high cash-flow growth means that most of the cash flows will occur in the distant future, and these distant cash flows are hit multiple times by a persistent shock to discount rates.

Informally, the law of one price implies that ³

$$\mathbb{E}_0[R_1 - R_f] \approx \left(\frac{1+g}{\kappa_0 - g} \right) \text{Cov}_t(-\Delta\kappa, -M_1) \frac{\sigma_0(M_1)}{\mathbb{E}_0(M_1)} \quad (23)$$

Provided that the discount rate shock is positively correlated with the SDF, this higher exposure to discount rate shocks commands a risk premium. In the model, this correlation is indeed positive. As in most habit models, a negative shock both increases discount rates and increases the marginal utility of consumption (Campbell and Cochrane (1999)). This paper differs though, in that this simultaneous increase is endogenous and the result of precautionary savings dynamics (Chen (2014)).

This theoretical link between cash-flow growth and expected returns is very similar to that which comes from the standard duration formula

$$\% \Delta P_1 \approx -[\text{Duration}] \Delta \kappa. \quad (24)$$

High cash-flow growth firms have long durations and thus are more sensitive to discount rate shocks. However, I choose to focus on cash-flow growth because it is easier to define and measure.

5.3. Empirical Evidence: Value and Discount Rate Shocks

We've already seen some evidence in support of the mechanism in Section 2. Consistent with the mechanism, value firms have high cash-flow growth by a number of definitions of cash flow. But to complete the empirical verification, we should also see that value firms are more exposed to discount rate shocks. The model's discount rate shocks are due to time-varying consumption volatility (Chen (2014)). This section shows that both in the model and the data, value

³Formally, there is no uncertainty at date 0 and so the expected return should be $1/\mathbb{E}_0[M_1]$. A more formal illustration would involve shocks to the volatility of the SDF but precludes the use of simple formulas.

returns are bad when consumption volatility rises.

To show this I need an empirical measure of consumption volatility. I use Boguth and Kuehn (2013)'s measure, which comes from an estimation of a Markov chain model for the first and second moments of consumption growth. An advantage of Boguth and Kuehn's measure is that they take advantage of the information in the components of consumption, which helps alleviate problems regarding identifying persistent volatility in the short post-war quarterly consumption data. Since the model is annual and lacks the consumption components used in Boguth and Kuehn (2013), I compute consumption volatility from the model's laws of motion for the simulated results.

Figure 5 shows consumption volatility betas for 10 book-to-market sorted portfolios. Consumption volatility betas are constructed by regressing excess returns on changes in consumption volatility. The left panel shows that, with a couple exceptions, consumption volatility betas decline monotonically in B/M. The right panel shows betas from the model. Here, consumption volatility is precisely measured using the laws of motion of the model and the betas are averaged over numerous simulations. As a result, we get a cleanly declining relationship between consumption volatility betas and B/M.

[Figure 5 about here]

The zig-zagging in the data panel is not surprising considering the short sample of quarterly post-war consumption and the high volatility of portfolio returns. Firm level results provide more statistical power, and are shown in Table 9. The table shows Fama-Macbeth regressions of consumption volatility betas on log B/M. Betas are constructed by regressing returns on changes in consumption volatility in rolling windows. The table shows that in both model and data, the relationship is negative, statistically significant, and similar in magnitude. The table uses forward-looking betas, that is, the windows for date t run from date t to 40 quarters after date t . I use forward-looking betas because theory predicts that it's the future return covariance that matters. Using the more traditional backward looking windows does not materially affect the data columns, but it does affect the model columns. This result is likely because firms in the model are characterized by stationary state variables and cannot display permanent differences as in the data. The window is long because the model only contains 3 aggregate technology states in order to maintain tractability. Shorter windows

show a stronger relationship between consumption volatility betas and book-to-market in the data.

[Table 9 about here]

A weakness of the results is that the overall level of the consumption volatility betas differs significantly between the model and data. This deviation is due to the precise measurement of consumption volatility in the model, as well as the single shock nature of the model. These two features mean that TFP, consumption volatility, and stock prices of portfolios move in lock-step, leading to highly negative consumption volatility betas. Softening the level of the betas would involve introducing additional sources of aggregate risk and is an interesting path for future research.

5.4. Other Potential Mechanisms

The model does not have fixed operating costs, irreversible investment, asymmetric adjustment costs on capital, or fixed costs of investment. Thus, the mechanism is distinct from the operating leverage channel of Carlson, Fisher, and Giammarino (2005), the inflexibility channel of Zhang (2005) (see also Gala (2010)), as well as the real option channel of Ai and Kiku (2012).

There is still one important channel to exclude: the cyclicalities of cash flows. It could be that value firms have higher returns due to the fact that their cash flows are more ‘procyclical.’ Figure 6 shows that this is not the case.

[Figure 6 about here]

The left panels of Figure 6 plot the cash flows of value and growth firms against the two state variables that represent the business cycle in this model: surplus consumption and aggregate productivity. For both value and growth firms, cash flows decline in surplus consumption. Since high surplus consumption represents a good state, in this respect both value and growth firms have countercyclical cash flows. Regarding the magnitude of the countercyclicity, there is no apparent difference. On the other hand, in terms of aggregate productivity, growth firms are clearly more procyclical. Value firm cash flows are generally invariant to aggregate productivity, but, growth firm cash flows clearly

increase. The cyclical nature of cash flows itself would then lead to a value discount, not a value premium.

The right panels of Figure 6 show that this result is intuitive. These panels show net investment (investment net of depreciation) for value and growth firms. In bad times, that is, in states with low surplus consumption or low aggregate productivity, value firms are disinvesting. These are times when the household really values consumption, and since value firms are unproductive, it is efficient for the value firms to discard their capital and provide cash flows to the household. This behavior leads to countercyclical cash flows for value firms. On the other hand, growth firms are investing in bad states. The household wants consumption, but since growth firms are so productive, it is efficient for the firm to give the household less consumption so that it can invest for the future. This behavior leads to procyclical and riskier cash flows for growth firms.

Of course, the risk of holding a stock is not just the risk of its cash flow next period. Every cash flow into the infinite future affects the risk of the stock. Both the temporal distribution and the short-term cyclical nature of a firm's cash flows affect its risk and return. In net, the high cash-flow growth of value outweighs the lower cyclical nature of its cash flows.

5.5. The Role of General Equilibrium

The model is general equilibrium (GE), and GE has many important implications for the results. One important role is that it pins down difficult-to-observe investment frictions. These investment frictions have a significant effect on the model's cross-sectional asset pricing results.

To show this, I conduct a partial equilibrium (PE) experiment. First I take the laws of motion for consumption and aggregate capital (16) and apply parameter values from the calibration (Table 3). Note that these parameter values are calibrated using a GE model. I then plug these laws of motion into the firm's problem (12) and solve for firm investment policies, but I change the adjustment costs for the firm's problem to be 1/20th of their calibrated value. These lower adjustment costs are in line with partial equilibrium estimates which use a constant SDF (for example, Whited (1992)). Lastly, I simulate a panel of firms using these PE investment policies (updating aggregates using the GE laws of motion).

This procedure mimics that used in the large literature on partial equilib-

rium dynamic firm models (for example Zhang (2005), Carlson, Fisher, and Giannarino (2005), Hennessy and Whited (2005)). I am conjecturing an SDF, and then solving for optimal firm behavior given this SDF, but I do not go on to check that the SDF is consistent with firm behavior. Note that in this example markets will not clear, that is, equation (14) does not hold. Indeed, consumption is not clearly defined since I can calculate consumption either from the conjectured law of motion or by aggregating in the panel simulation.

Table 10 shows that in this PE model, the value premium disappears. It shows Fama-Macbeth regressions of next year's returns on today's log B/M ratio. While the GE model matches the data quite nicely, in the PE model, the slope on log B/M becomes tiny and statistically insignificant.

[Table 10 about here]

Figure 7 explains why the value premium goes away. It shows the cash-flow growth of value and growth firms, comparing the GE model to the PE model. In the GE model, there is a large spread in cash-flow growth, but in PE, the spread is tiny. Intuitively, a firm does not want to have high cash-flow growth because temporally distant cash flows raise its discount rate and lower its value. The firm tries to reduce its discount rate by shifting its cash flows from the future to the present, that is, by disinvesting. The low adjustment costs of the PE model reduce the costs of this disinvestment, and thus result in a lower value premium.

[Figure 7 about here]

Note that the low elasticity of intertemporal substitution (EIS) implied by external habit preferences and the need to match aggregate consumption volatility are critical to the quantitative effects in this discussion. This low EIS means that the household has a strong desire to smooth consumption across time, and, through the SDF, the firm has a strong incentive to smooth cash flows. This strong smoothing motive combined with the volatility of consumption growth seen in U.S. data then imply large investment frictions. This stands in contrast to long-run risk and disaster models, which typically imply a large EIS, and therefore small investment frictions.

6. Conclusion

I show that external habit formation is consistent with one important aspect of the cross-section of stock returns. A real business cycle model, extended to include external habit preferences and idiosyncratic productivity generates a value premium that is quantitatively consistent with the data. The value premium arises as a result of the temporal distribution of cash flows. Value firms are temporarily low productivity firms, but they tend to have higher cash-flow growth in the future. These temporally distant cash flows are more exposed to the discount rate shocks that originate from time-varying consumption volatility. Empirical evidence confirms that value firms have higher cash-flow growth and are more sensitive to consumption volatility movements.

References

- Ai, Hengjie and Dana Kiku. "Growth to value: Option exercise and the cross section of equity returns". *Journal of Financial Economics* (2012).
- Avramov, Doron, Scott Cederburg, and Satadru Hore. "Cross-sectional asset pricing puzzles: An equilibrium perspective". *Unpublished manuscript* (2011).
- Bansal, Ravi, Robert F Dittmar, and Christian T Lundblad. "Consumption, dividends, and the cross section of equity returns". *The Journal of Finance* 60.4 (2005), pp. 1639–1672.
- Beeler, J. and J.Y. Campbell. *The long-run risks model and aggregate asset prices: an empirical assessment*. Tech. rep. National Bureau of Economic Research, 2009.
- Bloom, N. "The Impact of Uncertainty Shocks". *Econometrica* 77.3 (2009), pp. 623–685.
- Boguth, Oliver and Lars-Alexander Kuehn. "Consumption volatility risk". *The Journal of Finance* (2013).
- Boldrin, M., L.J. Christiano, and J.D.M. Fisher. "Habit persistence, asset returns, and the business cycle". *American Economic Review* (2001), pp. 149–166.
- Campbell, John Y and Tuomo Vuolteenaho. "Bad Beta, Good Beta". *American Economic Review* (2004), pp. 1249–1275.
- Campbell, John Y et al. *An intertemporal capm with stochastic volatility*. Tech. rep. National Bureau of Economic Research, 2012.
- Campbell, J.Y. and J.H. Cochrane. "By Force of Habit: A Consumption Based Explanation of Aggregate Stock Market Behavior". *Journal of Political Economy* 107.2 (1999), pp. 205–251.
- Campbell, J.Y. and R.J. Shiller. "The dividend-price ratio and expectations of future dividends and discount factors". *Review of financial studies* 1.3 (1988), pp. 195–228.
- Carlson, Murray, Adlai Fisher, and Ron Giammarino. "Corporate Investment and Asset Price Dynamics: Implications for the Cross-section of Returns". *The Journal of Finance* 59.6 (2005), pp. 2577–2603.
- Chen, Andrew. "Precautionary Volatility and Asset Prices". *Mimeo* (2014).
- Chen, Huafeng Jason. "Do Cash Flows of Growth Stocks Really Grow Faster?" *Available at SSRN 1903904* (2012).
- Chen, Long, Ralitsa Petkova, and Lu Zhang. "The expected value premium". *Journal of Financial Economics* 87.2 (2008), pp. 269–280.

- Chen, Long and Xinlei Zhao. "Return decomposition". *Review of Financial Studies* (2009), hhp017.
- Cochrane, J.H. "Presidential address: Discount rates". *The Journal of Finance* 66.4 (2011), pp. 1047–1108.
- Cohen, Randolph B, Christopher Polk, and Tuomo Vuolteenaho. "The price is (almost) right". *The Journal of Finance* 64.6 (2009), pp. 2739–2782.
- Cooper, Ilan. "Asset pricing implications of nonconvex adjustment costs and irreversibility of investment". *The Journal of Finance* 61.1 (2006), pp. 139–170.
- Cornell, Bradford. "Risk, Duration, and Capital Budgeting: New Evidence on Some Old Questions*". *The Journal of Business* 72.2 (1999), pp. 183–200.
- Da, Zhi. "Cash Flow, Consumption Risk, and the Cross-section of Stock Returns". *The Journal of Finance* 64.2 (2009), pp. 923–956.
- Dechow, Patricia M, Richard G Sloan, and Mark T Soliman. "Implied equity duration: A new measure of equity risk". *Review of Accounting Studies* 9.2-3 (2004), pp. 197–228.
- Fama, E.F. and K.R. French. "Size and book-to-market factors in earnings and returns". *Journal of finance* (1995), pp. 131–155.
- Fama, Eugene F and James D MacBeth. "Risk, return, and equilibrium: Empirical tests". *The Journal of Political Economy* (1973), pp. 607–636.
- Favilukis, J. and X. Lin. *Infrequent Renegotiation of Wages: A Solution to Several Asset Pricing Puzzles*. Tech. rep. Working paper London School of Economics and Ohio State University, 2011.
- Gabaix, Xavier. "Variable rare disasters: An exactly solved framework for ten puzzles in macro-finance". *The Quarterly Journal of Economics* 127.2 (2012), pp. 645–700.
- Gala, V. "Irreversible Investment and the Cross-Section of Stock Returns in General Equilibrium". *Manuscript, London Business School* (2010).
- Hansen, Lars Peter, John C Heaton, and Nan Li. "Consumption Strikes Back? Measuring Long-Run Risk". *Journal of Political Economy* 116.2 (2008), pp. 260–302.
- Hennessy, C.A. and T.M. Whited. "Debt dynamics". *The Journal of Finance* 60.3 (2005), pp. 1129–1165.
- Jermann, U.J. "Asset pricing in production economies". *Journal of Monetary Economics* 41.2 (1998), pp. 257–275.

- Khan, A. and J.K. Thomas. “Idiosyncratic Shocks and the Role of Nonconvexities in Plant and Aggregate Investment Dynamics”. *Econometrica* 76.2 (2008), pp. 395–436. ISSN: 1468-0262.
- Koijen, R.S.J. and S. Van Nieuwerburgh. *Predictability of Returns and Cash Flows*. Tech. rep. National Bureau of Economic Research, 2010.
- Krusell, P. and A.A. Smith. “Income and wealth heterogeneity in the macroeconomy”. *Journal of Political Economy* (1998), pp. 867–896.
- “Income and wealth heterogeneity, portfolio choice, and equilibrium asset returns”. *Macroeconomic dynamics* 1.02 (1997), pp. 387–422.
- Lakonishok, J., A. Shleifer, and R.W. Vishny. “Contrarian Investment, Extrapolation, and Risk”. *Journal of Finance* (1994), pp. 1541–1578.
- Lettau, M. and J.A. Wachter. “Why Is Long-Horizon Equity Less Risky? A Duration-Based Explanation of the Value Premium”. *The Journal of Finance* 62.1 (2007), pp. 55–92.
- Mehra, R. and E.C. Prescott. “The equity premium: A puzzle”. *Journal of monetary Economics* 15.2 (1985), pp. 145–161.
- Menzly, L., T. Santos, and P. Veronesi. “Understanding predictability”. *Journal of Political Economy* 112.1 (2004), pp. 1–47.
- Miller, Merton H and Franco Modigliani. “Dividend policy, growth, and the valuation of shares”. *the Journal of Business* 34.4 (1961), pp. 411–433.
- Ravn, Morten O and Harald Uhlig. “On adjusting the Hodrick-Prescott filter for the frequency of observations”. *Review of Economics and Statistics* 84.2 (2002), pp. 371–376.
- Rouwenhorst, KG. *Asset pricing implications of equilibrium business cycle models, chapter 10*, 294–330. 1995.
- Santos, T. and P. Veronesi. “Habit formation, the cross section of stock returns and the cash-flow risk puzzle”. *Journal of Financial Economics* 98.2 (2010), pp. 385–413.
- Syverson, C. “What Determines Productivity?” *Journal of Economic Literature* 49.2 (2011), pp. 326–365.
- Wachter, Jessica A. “A consumption-based model of the term structure of interest rates”. *Journal of Financial economics* 79.2 (2006), pp. 365–399.
- Whited, Toni M. “Debt, liquidity constraints, and corporate investment: Evidence from panel data”. *The Journal of Finance* 47.4 (1992), pp. 1425–1460.
- Zhang, L. “The value premium”. *The Journal of Finance* 60.1 (2005), pp. 67–103.

7. Tables and Figures

Table 1: Mean Cash Flow of Value and Growth

'E before Extraordinary' is earnings before extraordinary income (ib). 'E + Dep - Net Inv' is earnings plus depreciation less capital expenditures plus sales of plant, property, and equipment ($ni + dp - capx + sppe$). 'Year' is year after portfolio formation. 'Growth,' 'Neutral,' and 'Value' are buy-and-hold value-weighted tercile portfolios sorted on B/M. Cash flow is averaged across portfolio formation years. Sample is 1971-2011.

Mean Cash Flow per \$ Invested in Year 0						
Year	E Before Extraordinary			E + Dep - Net Inv		
	Growth	Neutral	Value	Growth	Neutral	Value
1	0.063	0.084	0.073	0.029	0.029	0.005
2	0.067	0.093	0.098	0.031	0.034	0.030
3	0.072	0.102	0.113	0.035	0.037	0.034

Table 2: Growth Rates of Mean Cash Flow for Value and Growth

'E before Extraordinary' is earnings before extraordinary income (ib). 'E' is earnings (ni), 'Dep' is depreciation (dp), 'Net Inv' is capital expenditures less sales of plant, property, and equipment (capx-sppe). 'Year' is year after portfolio formation. 'Growth,' 'Neutral,' and 'Value' are buy-and-hold value-weighted tercile portfolios sorted on B/M. The growth rate in year t is the growth rate of mean cash flows between years t and $t - 1$. Sample is 1971-2011.

Growth of Mean Cash Flow (% per year)						
Year	E Before Extraordinary			E		
	Growth	Neutral	Value	Growth	Neutral	Value
2	6.8	10.2	34.5	6.4	10.1	50.2
3	6.6	9.4	14.6	6.6	7.0	16.4
4	8.4	7.7	6.9	7.4	11.0	9.8
Year	E + Dep			E + Dep - Net Inv		
	Growth	Neutral	Value	Growth	Neutral	Value
2	8.9	9.3	17.3	5.1	20.0	544.3
3	8.7	8.3	9.4	12.0	6.4	11.8
4	9.1	10.3	7.2	14.4	22.6	-1.7

Table 3: Calibration

The model is annual, and all parameter values and empirical moments are annual. Consumption is real non-durable goods and services consumption. The volatility of GDP and relative volatility of consumption are logged and HP-filtered with a smoothing parameter of 6.25.

Parameter		Value	Target	Data	Model
Unconditional Asset Price Moments					
β	Time Preference	0.89	Mean 30-Day T-bill Return	0.89	0.96
ρ_s	Persistence of Surplus Consumption	0.86	Persistence of CRSP Price/Div	0.87	0.88
\bar{S}	Steady-State Surplus Consumption	0.06	Mean Sharpe Ratio of CRSP Index	0.44	0.39
Long-Run Growth Moments					
α	Production Curvature	0.35	Mean Output/Capital	0.41	0.42
δ	Depreciation Rate	0.08	Mean Investment Rate	0.07	0.08
Unconditional Business Cycle Moments					
σ_a	Volatility of TFP	0.03	Volatility of GDP (%)	1.61	1.70
ρ_a	Persistence of TFP	0.92	Persistence of Solow Residual	0.92	0.87
ϕ	Adjustment Cost	19	Volatility of Cons. Growth (%)	1.32	1.38
Firm Level Data					
ρ_b	Persist of idio prod	0.65	Persistence of Firm ROE	0.40	0.48
σ_b	Vol of idio prod	0.80	Vol Firm Stock Return (%)	0.35	0.34
Chosen Outside of the Model					
γ	Utility Curvature	2.00	For ease of comparison with Campbell-Cochrane (1999)		

Table 4: Unconditional Aggregate Asset Price Moments

All figures are annual. Data moments are taken from Beeler and Campbell (2009) and correspond to 1947-2008. The model columns show means and percentiles from many simulations of the same length as the empirical sample.

	US Data	Model			
		mean	5 %	50 %	95 %
Calibrated Moments					
$\mathbb{E}(r_f)$ (%)	0.89	0.96	-3.98	0.34	8.45
$AC1(r_f)$	0.84	0.83	0.57	0.86	0.99
$\mathbb{E}(R_m - R_f)/\sigma(R_m)$	0.44	0.46	0.26	0.46	0.69
Untargeted Moments					
$\mathbb{E}(r_m - r_f)$ (%)	6.36	7.42	4.95	7.44	9.91
$\sigma(r_m - r_f)$ (%)	16.52	18.20	12.25	17.78	25.19
$AC1(r_m - r_f)$	0.08	-0.06	-0.28	-0.06	0.14
$\sigma(r_f)$ (%)	1.82	3.87	0.55	2.25	12.20
$\mathbb{E}(p_m - d_m)$	3.36	2.57	2.00	2.57	3.04
$\sigma(p_m - d_m)$	0.45	0.43	0.24	0.41	0.70
$AC1(p_m - d_m)$	0.87	0.88	0.72	0.90	0.95

Table 5: Predicting Dividend Growth and Excess Returns with the Price-Dividend Ratio

All figures are annual. Data moments are taken from Beeler and Campbell (2009) and correspond to 1947-2008. The model columns show means and percentiles from many simulations of the same length as the empirical sample.

Panel A: Predicting Dividend Growth						
$\sum_{j=1}^L \Delta d_{m,t+j} = \alpha + \beta(p_{m,t} - d_{m,t}) + \epsilon_{t+L}$						
	L	US Data	mean	5 %	50 %	95 %
$\hat{\beta}$	1	0.003	0.009	-0.009	0.008	0.033
	3	0.012	0.013	-0.015	0.013	0.048
	5	0.044	0.016	-0.036	0.017	0.057
t-stat	1	0.112	0.492	-0.616	0.550	1.399
	3	0.193	0.669	-0.828	0.754	2.055
	5	0.482	0.678	-1.608	0.857	2.408
R^2	1	0.000	0.010	0.000	0.006	0.032
	3	0.001	0.017	0.000	0.012	0.049
	5	0.011	0.026	0.000	0.015	0.080

Panel A: Predicting Excess Returns						
$\sum_{j=1}^L (r_{m,t+j} - r_{f,t+j}) = \alpha + \beta(p_{m,t} - d_{m,t}) + \epsilon_{t+L}$						
	L	US Data	mean	5 %	50 %	95 %
$\hat{\beta}$	1	-0.12	-0.12	-0.26	-0.11	-0.01
	3	-0.27	-0.23	-0.47	-0.22	-0.03
	5	-0.42	-0.33	-0.64	-0.30	-0.02
t-stat	1	-2.63	-1.91	-3.30	-1.99	-0.24
	3	-3.19	-2.50	-4.44	-2.54	-0.40
	5	-3.37	-3.13	-6.72	-2.98	-0.31
R^2	1	0.09	0.07	0.00	0.07	0.16
	3	0.19	0.14	0.01	0.13	0.28
	5	0.26	0.19	0.01	0.19	0.39

Table 6: Business Cycle Moments

All figures are annual. Data moments correspond to 1947-2011. The model columns show means and percentiles from many simulations of the same length as the empirical sample.

	US Data	Model			
		mean	5 %	50 %	95 %
Calibrated Moments					
$\sigma(y_{hp})$ (%)	1.50	1.70	1.20	1.66	2.20
$\sigma(\Delta c)$ (%)	1.32	1.38	0.91	1.35	1.95
Untargeted Moments					
$\sigma(c_{hp})/\sigma(y_{hp})$	0.49	0.47	0.39	0.47	0.56
$\sigma(i_{hp})/\sigma(y_{hp})$	2.68	3.45	3.04	3.44	3.90
$\rho(y_{hp}, c_{hp})$	0.84	0.99	0.98	1.00	1.00
$\rho(y_{hp}, i_{hp})$	0.58	1.00	0.99	1.00	1.00
AC1(Δc)	0.52	0.04	-0.20	0.04	0.28
$\mathbb{E}(\text{Adj Cost}/Y)$ (%)		1.01	0.63	0.96	1.64
$\mathbb{E}(\text{Adj Cost}/I)$ (%)		5.92	2.89	5.48	10.61

Table 7: Regressions of Future Returns on Book-to-Market

All figures are annual. Data moments correspond to 1947-2012. The model columns show means and percentiles from many simulations of the same length as the empirical sample.

Dependent Var: $R_{i,t+1}$					
	US Data	Model			
		mean	5 %	50 %	95 %
intercept	18.62	18.34	13.33	18.03	24.14
t-stat	5.74	3.88	2.99	3.88	4.80
$\log(B/M)_{i,t}$	5.70	5.84	3.58	5.84	7.86
t-stat	4.88	3.34	2.17	3.30	4.57

Table 8: Summary Statistics from 10 Book-to-Market Sorted Portfolios

All figures are annual. Returns are value-weighted. Data moments correspond to 1947-2012. The model columns show means and percentiles from many simulations of the same length as the empirical sample.

port	$\mathbb{E}(R_{\text{port}})$		$\sigma(R_{\text{port}})$	
	US Data	Model	US Data	Model
Lo	7.7	9.6	20.9	19.9
2	8.0	10.3	17.1	21.4
3	8.1	10.6	16.7	22.5
4	8.6	10.9	17.6	22.9
5	9.5	11.1	18.3	23.4
6	9.7	11.3	17.6	23.9
7	9.8	11.5	19.3	24.5
8	11.6	11.7	21.1	24.9
9	11.9	12.0	20.4	25.6
Hi	13.3	12.3	25.7	26.7

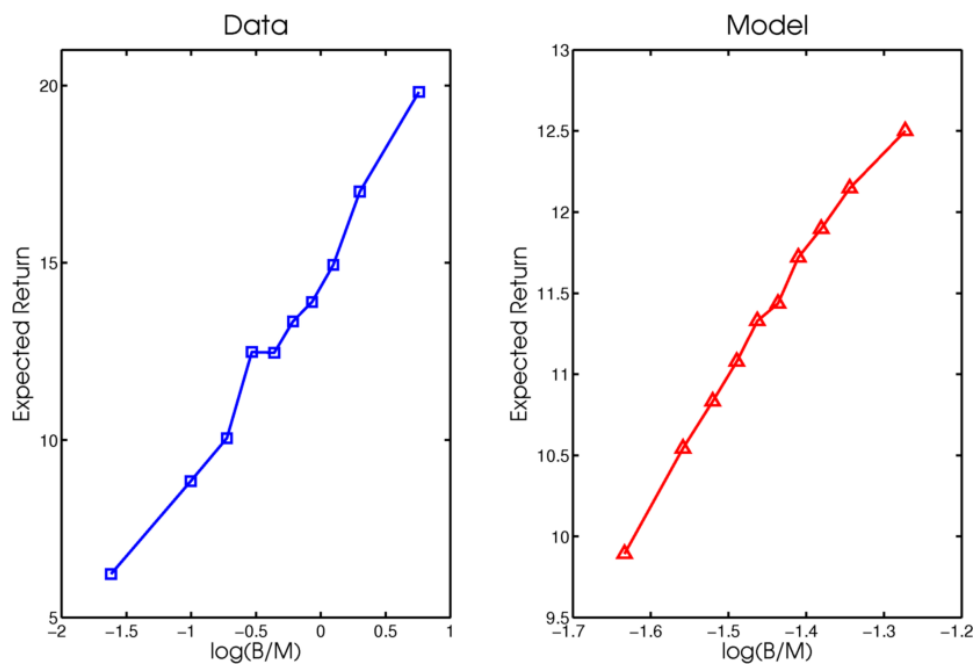


Figure 1: B/M Decile Returns as a Function of B/M. Figures are annual. Returns are equal-weighted.

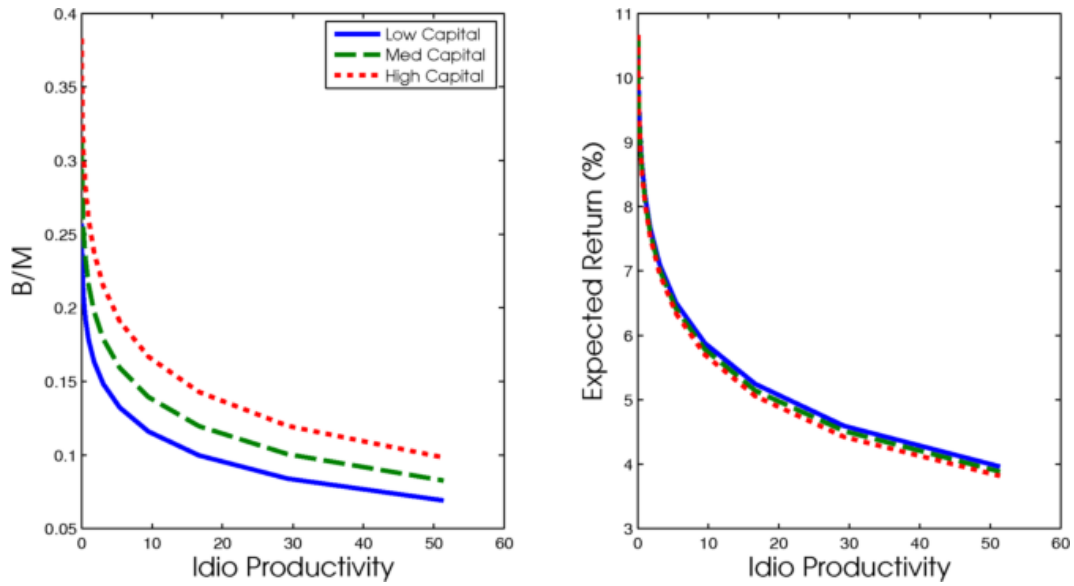


Figure 2: B/M and Expected Returns as a Function of Firm States. Figures are annual and computed from the model solution.

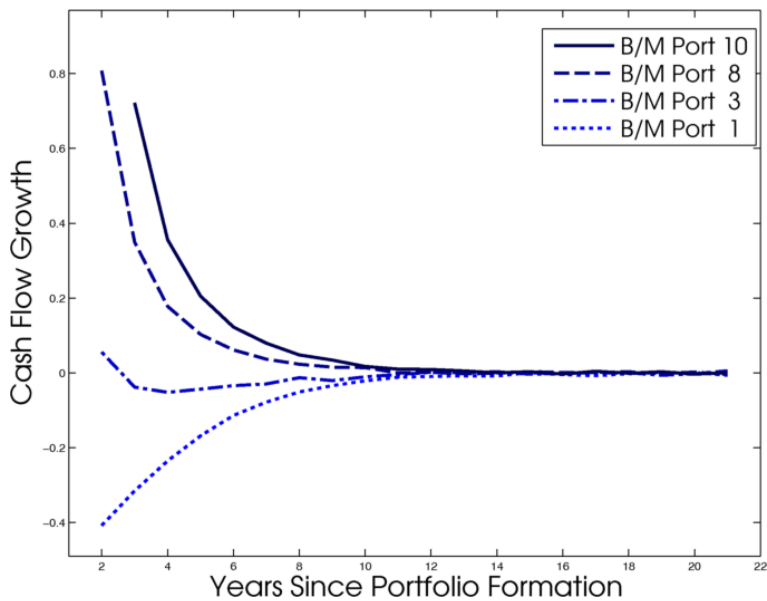


Figure 3: Cash-Flow Growth of Book-to-Market Sorted Portfolios. Figures are annual and computed from model simulations.

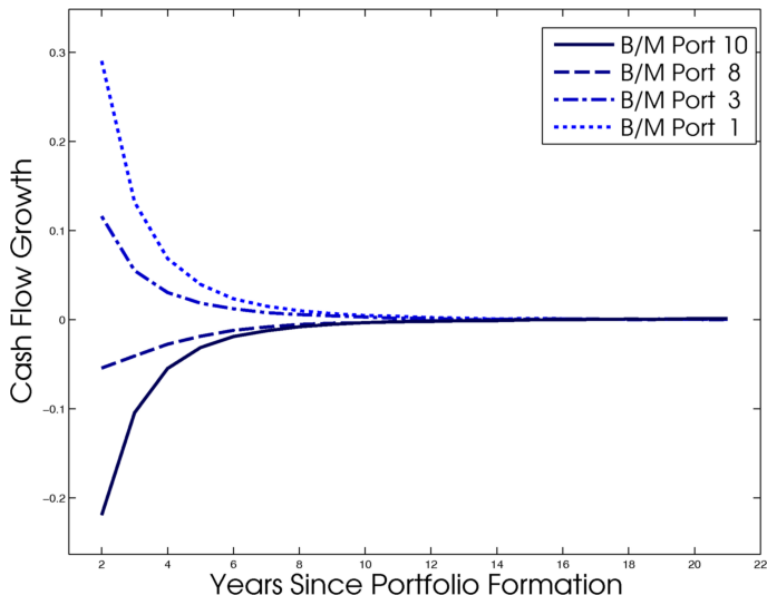


Figure 4: Cash-Flow Growth of Book-to-Market Sorted Portfolios: No Habit. Figures are annual and computed from model simulations.

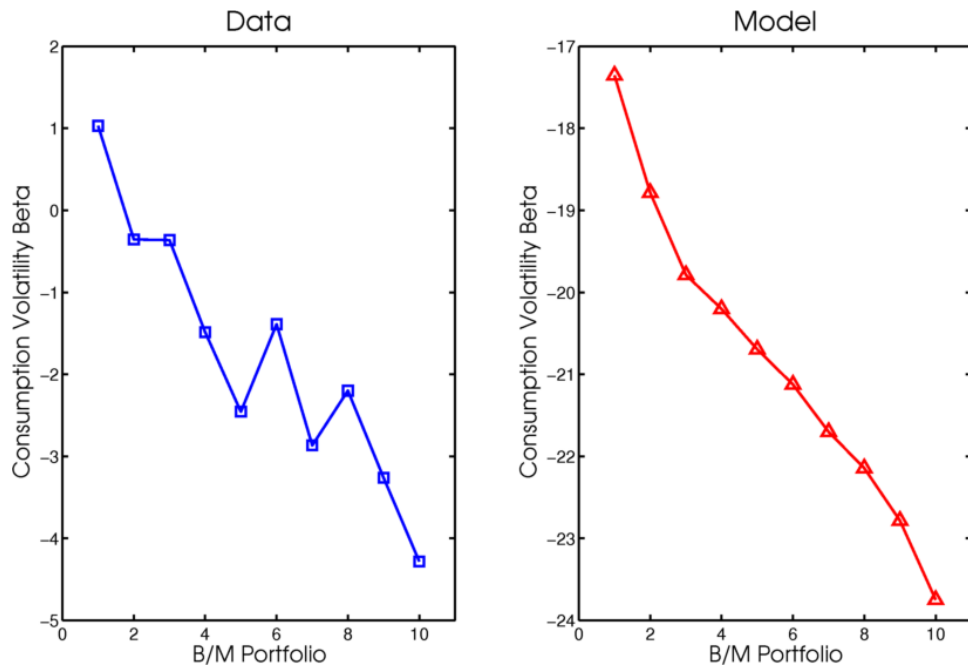


Figure 5: Consumption Volatility Betas of Book-to-Market Sorted Portfolios. Betas are from regressions of returns on changes in consumption volatility. Changes in consumption volatility are normalized by their standard deviation.

Table 9: Regressions of Consumption Volatility Betas on Book-to-Market

Figures are annualized. Consumption volatility in the data is Boguth and Kuehn (2013)'s estimate. Consumption volatility in the model is the true consumption volatility. Standard errors are Newey-West with 12 lags. Regressions are firm-level Fama-Macbeth using weighted least squares where the weights are the inverse of the squared standard error of the consumption volatility beta estimate. Betas are constructed by regressing excess returns on changes in consumption volatility for 40 quarters into the future.

Dependent Var: Consumption Vol Beta					
	US Data	Model			
		mean	5%	50%	95%
$\log(B/M)_{i,t}$	-1.02	-2.68	-4.20	-2.64	-1.24
t-stat	-2.57	-3.83	-6.08	-3.67	-1.56

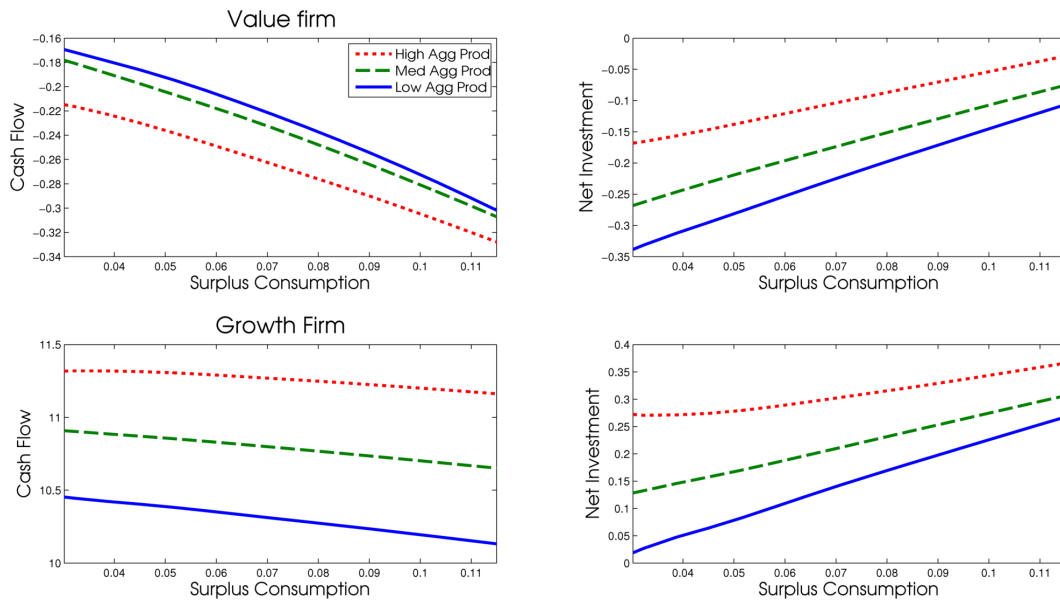


Figure 6: The Cyclicity of Value and Growth Cash Flows. Value firm plots are calculated using the median capital and productivity of firms in the 10th decile of B/M-sorted portfolios. Growth firm plots are from the first decile. Net investment is investment net of depreciation. Aggregate capital is fixed at its mean.

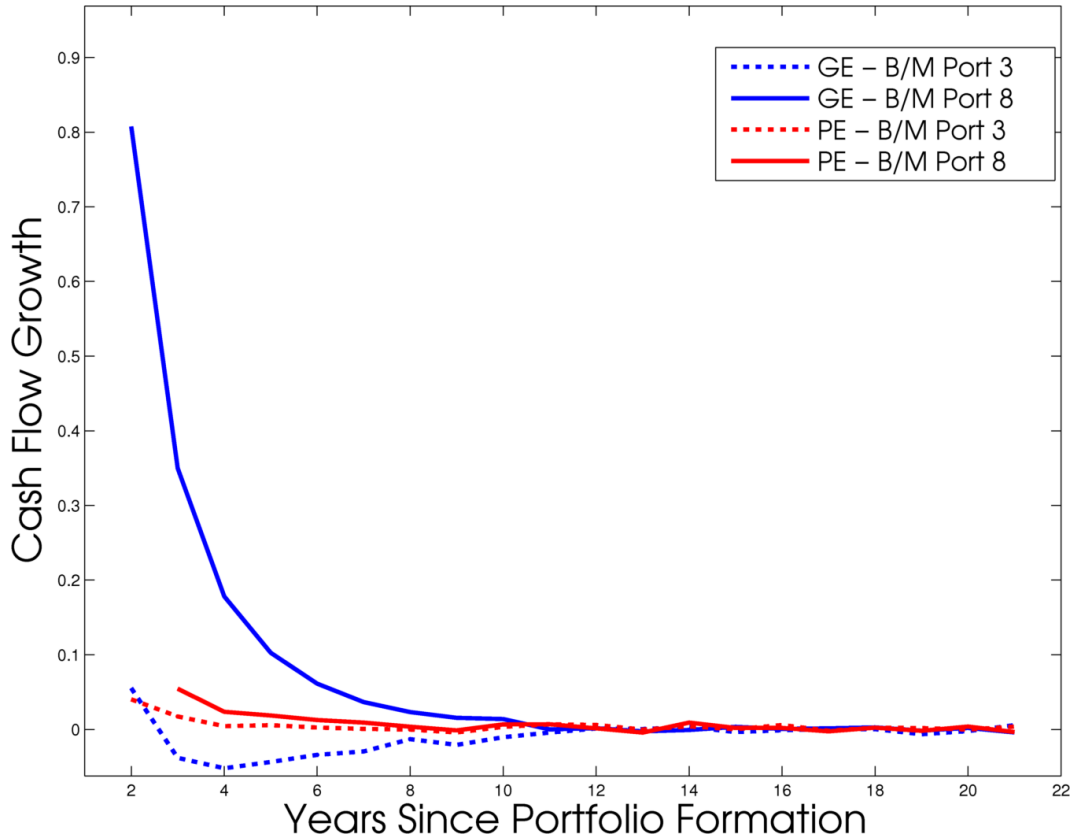


Figure 7: Partial Equilibrium: Cash-Flow Growth of Book-to-Market Sorted Portfolios. ‘GE’ uses calibrated parameter values from Table 3 and equilibrium laws of motion. ‘PE’ uses equilibrium aggregate laws of motion, but firm-level decision rules consistent with adjustment costs which are 1/20th of the value from Table 3.

Table 10: Partial Equilibrium Experiment: Regressions of Future Returns on Book-to-Market

All figures are annual. Data moments correspond to 1947-2012. The model columns show means and percentiles from many simulations of the same length as the empirical sample.

	Dependent Var: $R_{i,t+1}$		
	US data	GE model	PE model
intercept	18.62	18.34	10.82
t-stat	5.74	3.88	3.60
$\log(B/M)_{i,t}$	5.70	5.84	0.53
t-stat	4.88	3.34	0.83