YALE
INTERNATIONAL
CENTER
FOR FINANCE

Yale ICF Working Paper No. 04-26
July 2004

DYNAMIC BETA, TIME-VARYING RISK PREMIUM AND MOMENTUM

Hong Zhang
Yale School of Management

Dynamic Beta, Time-Varying Risk Premium, and Momentum

Hong Zhang *
Yale School of Management
P.O. Box 208200
New Haven, CT 06520-8200

This Draft: July 2004

*The author thanks Matthew Spiegel, Will Goetzmann, Zhiwu Chen, Geert Rouwenhorst, John Griffin, Jeff Wurgler, Harry Mamaysky, Arturo Bris, and seminar participants at Yale School of Management, INSEAD, University of British Columbia, University of Alberta, Fordham University, HKUST, and Hong Kong University for helpful comments and suggestions. All errors are my own. Email: bill.zhang@yale.edu, phone (203) 436-0648, fax (203) 436-0630.

Abstract

This article proposes a rational model to demonstrate that firm-specific risks can be priced in the equilibrium and can generate asset pricing anomalies such as momentum. In general, business risks at both the market level and firm level can affect a firm's investment decisions, and a firm usually has certain ability to forecast firm-level risks, such as demand changes or technology innovations. When a firm dynamically adjusts its business according to forecasted firm-level risks, investors face a beta risk (which proxies for firm-level risks) in addition to the market risk. These two risks jointly create a nonlinear risk premium, which simultaneously explains momentum and the Fama and French (1993) three-factor model. In other words, momentum profits and the size (value) premium in this model reflect reasonable rewards to compensate investors for the two risks. Empirically, the estimated risk premium contributes a large portion (in many cases a leading portion) of stock momentum profits and helps identify stocks likely to generate more momentum profits. Market-to-book, size, or two proxies of mispricing do not absorb the latter effect.

JEL Classification: G12

The phenomenon of momentum has perplexed financial economists, generating both excitement and controversy. On the one hand, since Jegadeesh and Titman (1993) first documented that a variety of strategies that buy past winners and sell past losers can produce significant profits, researchers have shown that the phenomenon is both robust for the "out-of-sample" U.S. data after 1993 (Jegadeesh and Titman (2001)), and pervasive in other parts of the world (Rouwenhorst (1998, 1999); Griffin, Ji, and Martin (2002)). On the other hand, according to traditional asset pricing theories, the momentum phenomenon should not exist. Recently, many behavioral, as well as rational, hypotheses have emerged to explain the source of momentum returns. But beyond the stylized empirical facts, the only consensus seems to be that traditional asset pricing models fail to explain momentum.[1] After ten years of research, momentum still remains one of the greatest challenges to rational asset pricing theories.

To investigate the potential economic source of momentum, this article explores the possibility that firm-level risks can affect a firm's investment decisions and, consequently, affect asset prices and lead to asset pricing anomalies. It contributes to the literature by demonstrating that a rational economy can have size, book-to-market, and several momentum-related phenomena because of the asset pricng impact of firm-level risks. Traditional asset pricing models focus on market-wide risks, assume a static variance-covariance matrix for asset returns, and, consequently, predict that a firm's risk exposure, or beta, is static and that the expected risk premium is linear in beta. If a firm's investment opportunity set changes over time, however, the firm's beta may be dynamic (see, for example, Berk, Green, and Naik (1999)), because a rational firm wants to explore good business opportunities whenever possible. Furthermore, since a firm is likely to have some information about or influence on its own business environment, anticipated fluctuations therein can easily alter the firm's investment policy and its systematic risk exposure. For example, a car manufacturer can survey its customers, anticipate increased demand and a higher profit margin for sports utility vehicles (SUVs) and start to expand the production line. It will then realize more earnings by selling more SUVs. Meanwhile, since SUV sales are more vulnerable to fluctuations in oil prices (an example of market-wide risk) than other models, expanding SUV production may introduce more oil risk to the firm's earnings.

[1] For example, Grundy and Martin (2001) show that residuals from a conditional version of the Capital Asset Pricing Model (CAPM) and the Fama and French (1993) three-factor model demonstrate pronounced momentum profits. Griffin, Ji, and Martin (2002) argue that macro and business cycle factors fail to explain momentum. Section 4 further surveys the literature.

In general, to determine which project it should invest and how large the investment scale should be, a firm needs to forecast the future cash flows generated by per dollar investment and risks that can affect future cash flows. The SUV example reminds us of one important property that traditional models ignore: all risks are not the same from a firm's perspective. Some risks, especially flucturations related to (though not restricted to) firm-level investment opportunities, such as demand changes or technology innovations, can be fairly forecasted by a firm's research branch. Others, such as market-wide risks or systematic risks, remain unpredictable or out of the control of a typical firm. Consequently, investment risks at the market level and the firm level can affect a firm's investment policy in different manners. A forecasted superior investment opportunity (or a forecasted positive shock of the firm-level risk) can generate more expected future earnings and induces a firm to invest more. Meanwhile, since a firm cannot predict or hedge out systematic risks, more investments are essentially a larger bet on systematic risks and will create more earnings variations. A rational firm needs to balance these benefits and costs in making investment decisions. But as the SUV example illustrates, even when a firm does not react dynamically to unpredictable systematic risks, its systematic risk exposure, or "beta," nevertheless varies when the firm adjusts its business in response to predictable risks. This observation suggests that static asset pricing models or models that focus only on unpredictable systematic risks may fail to fully describe the economic role of the production sector.

Ancitipating firms to ultimately distribute their earnings as dividends, investors now need to hedge against both market-wide and firm-level risks that affect future dividends. In practice, investors usually cannot forecast firm-level investment risks. But they can use beta dynamics to proxy for the latter. Consequently, investors' asset-allocation decisions depend on their inferences about beta, as well as those about the systematic risk.[2] This twofold uncertainty creates nonlinear dividend risks for investors. Generalizing Whittle (1990) and assuming that asset prices are determined such that investors are willing to hold all stocks issued by firms, this article demonstrates that the investors' stochastic control problem leads to two new components in the risk premium, in addition to the traditional CAPM element. These two components are related to the mean and variance of predicted beta dynamics, even when investors hold diversified portfolios. In other words, the model proposes that, when investors fully recognize the economic role of individual firms, asset prices should reflect the influence from corporate policies or factors affecting firms' decision-making process.

The new risk premium components rationalize momentum and the Fama and French (1993)

[2] In general beta dynamics reflect expected changes in the firm's business opportunity set and will be partly determined by firm investment policies.

model simultaneously. In the current model, momentum arises because firm-level business opportunities, as well as the corresponding corporate investments, are likely to persevere in a short period of time (SUV demands can last for a while; new technologies replace old ones gradually). Size (value) premium exists because investors discount stock prices for firms with bad firm-level investment opportunities. Both anomalies reflect reasonable rewards to compensate rational investors for nonlinear risks, a feature that linear asset pricing models fail to capture.[3] Using the terminology of the Taylor approximation, traditional asset pricing models consider first-order effects of the cash flows generated by the production sector of the economy. Although first-order terms can explain a large part of the time series and cross-sectional return variation, there can always exist higher-order effects or anomalies related to ignored corporate activities. In light of this theory, the article provides a coherent explanation for several well-known asset pricing anomalies, not just momentum. Furthermore, the equilibrium price system in this model already reflects all public information. Therefore, the existence of momentum or other anomalies does not necessarily mean a failure of the efficient market hypothesis. It is the linear pricing formula that fails in this model.

Besides momentum and the size (value) effect, a list of interesting phenomena can intuitively exist in this model, including industry momentum (Moskowitz and Grinblatt (1999)), Fama-French portfolio momentum (Lewellen (2002)), earnings momentum (Chan, Jegadeesh, and Lakonishok (1996)), positive relationships between market-to-book ratio and earnings (Fama and French (1995)), between market-to-book and investment (Xing (2002)), between market-to-book and momentum profitability (Daniel and Titman (1999); Sagi and Seasholes (2001)), and an inverted U-shape relationship between momentum and size (Hong, Lim, and Stein (2000)). The model also explains why momentum portfolios work better for a holding horizon between 3 and 12 months, why momentum profits are driven mostly by loser stocks (Hong, Lim, and Stein (2000)), and why momentum returns are more significant during the expanding period of a

[3]Priced firm-level risks create a conditional and cross-sectional dispersion for expected asset returns. This dispersion will last for a while, following the autocorrelation of firm investments, and lead to momentum. In a traditional asset pricing world, on the other hand, firm-level investment autocorrelations should be diversified away and will not lead to momentum. For the size (value) premium, investors do not like firms with bad investment opportunities and will push down the market price for these firms in order to guarantee a higher expected future return. As a result, smaller-size or low market-to-book firms are associated with higher future return mechanically. The link between firm-specific beta risks and firm characteristics explains the Daniel and Titman (1999) finding that the firm-specific component of characteristics can affect asset return. Upon these intuitions, the key difference between this model and the literature (Fama and French (1992, 1993), Conrad and Kaul (1998), Moskowitz and Grinblatt (1999), to name a few) is that asset pricing anomalies can exist because of priced firm-level risks, instead of market-level risks.

business cycle (Chordia and Shivakumar (2002)) or following a positive market return (Looper, Gutierrez and Hameed (2004)). This empirical evidence demonstrates the explanatory power of this model.

Since there may always be missing risk factors and behavioral effects, perhaps the most interesting empirical question is what fraction of momentum this simple model explains. The model can be estimated via an extended Kalman filter. For the New York (NYSE) and American Stock Exchange (AMEX) stocks, the risk-adjusted momentum profits based on this model usually drop to less than 40% of the raw return momentum profits during the period from 1965 to 1999 for a set of weighted relative strength strategies (WRSS). To further reveal the economic source of momentum, I decompose the WRSS momentum profits into four components. The decomposition clearly shows that a leading portion of the profitability is related to firm-level risks. The risk premium due to firm-level risks usually contributes more than half of the total momentum profits, whereas the traditional CAPM component explains a much smaller part (around 10% or less). The estimates for the S&P 500 index stocks attribute a similar fraction of stock momentum profits to risks. The remaining WRSS momentum is economically small, usually less than 4% a year before costs. The risk-adjusted momentum for winner-minus-loser (WML) strategies is more profitable, indicating that empirically the model less successfully explains the profits from strategies based upon the most extreme firms. Still the risk contribution to momentum is significant (up to 40%).

To further demonstrate that the model provides real explanatory power to momentum so that the decomposition is not an artifact of overfitting the data, I focus on a unique prediction of the model that firms with more dynamic behavior are likely to generate more momentum profits. In-sample tests confirm the hypothesis (sorting stocks into ex-post risk-premium-dynamics quartiles can create quartile-momentum-return differences up to 7% a year for WRSS strategies and 10% a year for WML strategies) and show that this effect is not absorbed by market-to-book, size, or two measures of mispricing. Out-of-sample firm-dynamic quartiles can still create significant momentum spreads between stock quartiles up to 5% a year for WRSS strategies and 7% a year for WML strategies. These tests also provide direct evidence that momentum can be associated with estimated firm-level risks. Overall, the empirical results strongly suggest that this model captures important and perhaps fundamental features of the asset return process.

In the rational-momentum literature, Berk, Green, and Naik (1999) first model the time-varying betas of a firm with the concept of a growth option, showing that their model can lead to momentum. More recently, Johnson (2002) shows that stochastic dividend growth rates lead to autocorrelation for firm returns. Sagi and Seasholes (2001) argue that momentum may be tied to the dynamics of firm-specific factors, especially through the measure of convexity. Although this article shares with these studies the intuition that firm behavior should be correlated with economic fundamentals and that characteristics or nonlinearity may be important in explaining momentum profits, it differs in several important respects.

First, since no model is likely to explain all momentum profits, the ultimate success of a model, rational or behavioral, should be judged by what scope of empirical phenomena and what fraction of momentum profits the model explains. The model proposed here is coherent for several CAPM anomalies, including momentum, and can explain a number of documented momentum phenomena in the literature. Furthermore, this study not only adopts a general equilibrium setup and offers a closed-form solution for the return process, but also provides an empirical method to directly test the model and estimate the percentage of momentum related to risks. This model also has the advantage of being easily tested by standard asset return data and expanded to bring more corporate fundamentals into the asset pricing framework. The empirical tests conducted here directly confirm the explanatory power of the model. As for the economic intuition, unlike other studies, the result here is largely driven by a firm's willingness and ability to achieve economies of scale. Finally, and perhaps most interestingly, this article demonstrates that, based on some general assumptions regarding the decision-making process of a firm, firm-specific risks can be priced and can lead to asset pricing anomalies. It therefore proposes a new economic insight.

The remaining paper proceeds as follows. Section 1 explains the model and solves for the equilibrium asset prices. Section 2 reveals the source of momentum profits. Section 3 empirically investigates the source of momentum profits. Section 4 surveys related literature. Section 5 concludes. The Appendix contains all proofs and some simulation results.

1 The Model

1.1 The Firms

There are I firms in the economy. A typical firm (firm i) lives forever and has one public share of stock outstanding. It has one unit of a fixed-scale project that will pay out a constant cash flow D_0^i at each period.[4] It can also invest in a risky project at the beginning of each period. At the end of the same period, the risky project will produce some cash flow, which is affected by different risks. As mentioned, a firm is likely to have different perceptions upon different investment risks. For example, the firm-level business environment (which can include

[4]A firm can, for example, hold some cash. This assumption avoids negative prices in the equilibrium.

firm-specific information such as customer demands and technology innovations) affects the profitability of any risky projects. But in reality a successful firm usually has some ability to forecast and even exploit possible changes in its firm-level business environment. A firm can, for instance, survey the market (or conduct research), predict the preferences of customers (or develop a promising new technology), change its business lines accordingly, and sell more favorable products in the future. Fluctuations in the firm-level business environment therefore illustrate an example of risks foreseeable from a firm's point of view. To emphasize the firm's role in shaping its own firm-level business environment, firm-level risks will also be referred to as firm-level investment opportunities. Meanwhile, as pointed out by traditional asset pricing models, there are also market-wide or systematic risks that are unpredictable or totally beyond the firm's control. These risks also affect future cash flows. Hence, without loss of generality, assume the cash flow generated by one unit of any risky project will be affected by two risks: a business environment risk, F_t^i , that is perfectly predictable to the firm (but not to investors and econometricians−therefore it is called a *latent* risk), and a market-wide systematic risk, X_t,which is totally unpredictable to the firm. Assume that $X_t \sim i.i.d.\ N(0, \sigma_x^2)$ and that the realization of X_t is observed by the firm and all investors. Analytically, one unit of a risky project costs k^i to invest and will produce a cash flow Y_t^i at the end of the period:

$$Y_t^i = Y_0^i + X_t + l^i F_t^i, \tag{1}$$

where Y_0^i is the unconditional expected payoff and l^i is the project's exposure to the foreseeable risk. Equation (1) can also be regarded as a first-order Taylor approximation of real production function, so all cash flow risk loadings are constants.[5] The exposure to the market risk has been normalized to one. Since it is forecasted by the firm, the latent risk is approximated by an AR(1) process.

$$F_t^i = \Gamma^i F_{t-1}^i + \varepsilon_{F,t}^i. \tag{2}$$

The AR(1) assumption seems to be a natural approximation of the ebb and flow of real-world business environments. For example, if consumers currently prefer SUVs to other car models, substantial earnings from SUV sales may last for a while, but not forever. For another example, a firm adopting a new technology gains a short-term competitive advantage. But competition makes it unlikely for the firm to stay at the innovation frontier forever. Original technology

[5]This decomposition simplifies potential risks and the decision-making process of a real firm, and omits labor costs and forecasting errors. But the intuition is rich enough for the purpose of this paper. Predicted risks can also be easily expanded to contain market-wide components (the business cycle, for example), industry components, or other predictable systematic risks. For simplicity, this paper focuses on the asset pricing impact of predicted firm-level risks.

repeatedly ages and costs the firm money to upgrade.[6] For simplicity, assume $0 < \Gamma^i < 1$, $F_0^i = 0$, and $\varepsilon_{F,t}^i$ follows an i.i.d. normal distribution $N(0, \sigma_{F,i}^2)$. To maintain tractability, assume that $k^i = k$, $Y_0^i = Y_0$, $\Gamma^i = \Gamma$, and $\sigma_{F,i}^2 = \sigma_F^2$ are the same for every firm. In other words, all latent factors are independent, and all the firms are i.i.d. at time zero (ex-ante). Assume that σ_x, Γ, and σ_F are publicly known to the firm and investors. Finally, in the economy there is also a risk-free good (bond), which can be sold long or short elastically at will.[7] The risk-free rate equals r.

The remaining part of this subsection examines the decision making process of a typical firm. For simplicity, the index i will be suppressed when there is no confusion. At the beginning of period $t+1$, a firm forecasts perfectly the next period's business environment F_{t+1} (or $\epsilon_{F,t+1}$) and determines m_{t+1}, the units of risky project it will invest in at $t+1$. It then borrows money from the bond market to finance the risky project. At the end of the $t+1$ period, it will pay back the borrowings and pass all the remaining cash flow to the investors as dividends. At the beginning of the next period $t+2$, the firm repeats this behavior. The firm maximizes the risk- or cost-adjusted earnings period by period, because, for example, management compensation is linked to earnings. For simplicity, assume both the firm and the investor have negative exponential utility functions (later I will show that a risk-neutral firm with a quadratic cost function leads to similar results).[8] The firm has a risk aversion parameter γ_F and solves the following problem:

$$\underset{m_{t+1}}{Max}\ E_t[-e^{-\gamma_F Z_{t+1}}] \tag{3}$$
$$s.t.\ Z_{t+1} = D_0 + (Y_0 + X_{t+1} + lF_{t+1})m_{t+1} - (1+r)km_{t+1},$$

where Z_{t+1} is the net cash flow by the end of the period $t+1$. The second equation simply states

[6] According to traditional financial theories, idiosyncratic latent risks should not have any asset pricing impact and will not automatically create momentum in a well-diversified economy. Later sections will explain why momentum exists in a well-diversified economy.

[7] The assumption of perfect elasticity avoids adding a hedging demand component to investors' endowments, allowing one to focus on the asset pricing impact due to the firm's dynamic investment behavior.

[8] This builds up a similarity between the production and the investor sectors of the economy. The risk aversion prevents the firm, which has a linear cost function, from investing infinitely in a favorable risky project. Meanwhile, because a firm will pay out all its net earnings as dividends to investors, maximizing either net earnings or asset prices will provide a firm the incentive to invest more into better investment opportunities (which is enough to generate momentum). Both approaches make economic sense because both earnings and share prices are commonly used measures for management performance. The approach adopted here is mathematically more appealing. In terms of the investment q theory, maximizing earnings enables the firm to invest according to the "marginal q," while maximizing share price is equivalent to investing according to the "average q." Both investment policies can generate momentum. However, the two q values may not match each other exactly. Interested readers may refer to Blanchard, Rhee, and Summers (1993), for instance, for why the two q values may differ.

that the net cash flow equals the sum of the cash flows from the risk-free and risky projects, minus the investment cost. All information is conditioned at the end of time period t. Assume there is no investment irreversibility, financial constraint, or any other friction. The optimal investment decision of the firm and its corresponding dividend policy are given by the following lemma:

Lemma 1 *(Optimal Investment Policy of the Firm)*

1. The optimal investment policy of the firm is given by:

$$m_{t+1} = \frac{Y_0 - k(1+r) + lF_{t+1}}{\gamma_F \sigma_x^2}, \tag{4}$$

where m_{t+1} has a static component $(Y_0 - k(1+r))/\gamma\sigma_x^2$ and a dynamic component $lF_{t+1}/\gamma\sigma_x^2$.

2. The optimal investment policy m_{t+1} follows an AR(1) process

$$m_{t+1} = m_0 + \Gamma m_t + \varepsilon_{m,t+1}, \tag{5}$$

where $m_0 = (Y_0 - k(1+r))(1-\Gamma)/\gamma_F \sigma_x^2$ and $\varepsilon_{m,t+1}$ follows the distribution $N(0, \sigma_m^2 = l^2\sigma_F^2/(\gamma_F^2\sigma_x^4))$.

3. The net earning and the dividend payoff of the firm will be

$$Y_{t+1} = D_{t+1} = D_0 + \gamma_F \sigma_x^2 m_{t+1}^2 + m_{t+1}X_{t+1} \tag{6}$$

4. If the firm is risk neutral, and if the total investment cost function $c(m_{t+1})$ is increasing and quadratic in m_{t+1} (i.e., $c(m) = Am + Bm^2$, where $A, B > 0$), then the firm's optimal investment will follow an AR(1) process, and the dividend policy will be

$$Y_{t+1} = D_{t+1} = D_0 + Bm_{t+1}^2 + m_{t+1}X_{t+1}. \tag{7}$$

Assume $Y_0 - k(1+r) + lF_{t+1} > 0$ so that firm investments are always positive. The lemma shows that the optimal scale of the risky project is jointly determined by the firm-level risk and the market risk. Additional investments help a firm fully utilize its expected investment opportunity in response to good news at the firm level, but will also result in extra earnings variations because of increased exposure to unpredictable systematic risk. Equation (4) states that when the expected firm-level business environment is more favorable, when the whole market is less risky, or when the firm is less risk averse, the firm will invest more in the risky project. Hence, a firm's optimal investment policy, or $m(t)$, correlates with anticipated future investment opportunities and also follows an AR(1) process. The autocorrelation of firm investments is well documented in the macroeconomics literature.

More interesting is the dividend policy. In the model, the dividend is simply the net earnings of the firm. Lemma 1 says that, although the firm does not purposefully respond to the market risk, the dividends nevertheless have a dynamic exposure to this market risk: since SUV sales have a greater exposure to oil price shocks, whenever the car maker expands SUV productions, its earnings will become more vulnerable to a sudden oil crisis. This exposure, or $m(t)$, can be thought of as a *cash flow beta*, since it measures the covariance between a firm's cash flow (earnings) and the systematic risk. Later sections will show that the expectation of the cash flow beta is related to the conditional CAPM beta. Because the cash flow beta is related to the latent risk, the $m(t)X(t)$ term combines two linear risks together and becomes itself nonlinear. The $m(t)^2$ term is also nonlinear and directly reflects the benefit from the firm's dynamic investment policies. Because a better investment opportunity leads to both more expected net cash flows per dollar investment and a larger investment scale, the cash flow (or earnings) sensitivity to investment increases in the investment scale itself. Needless to say, if a firm has no forecasting power, then $m(t)$ becomes constant and this term drops.

The model's main focus is on the economy with a risk-averse firm that has a linear cost function. But the last part of the lemma shows that the same functional dividend policy can come from a risk-neutral firm with a quadratic cost function. It can be proven that a second-order Taylor approximation of any increasing cost function is equivalent to a risk-aversion parameter of the model. In general, for any monotonically increasing cost function or monotonically increasing and concave utility function, if we allow the firm to achieve economies of scale, then $m(t)$ will be correlated with the latent factor (business risks) and the dividend policy will be nonlinear in the total risks.

1.2 Undiversified and Well-Diversified Investors

A typical investor lives two periods. Investor t enters the economy at time period t and receives a labor income or an endowment W_t for that period. At the beginning of period $t + 1$, the investor allocates her endowment between the risk-free asset and the stocks of the firm that she buys from the older generation. At the end of the period t+1 she collects the dividends, sells her firm to the new generation, and consumes her terminal wealth of W_{t+1}.[9] The investor has a negative exponential utility function and solves the following problem subject to some budget constraints:

$$\underset{n^i}{Max}\ E_t[-e^{-\gamma_I W_{t+1}}], \tag{8}$$

where γ_I measures risk aversion and n^i is the number of firm i shares she wants to buy.

To fully explore the asset pricing impact of firm decision making process under different economic consitions, let us consider the following two important cases. In the first case, asset

[9] Each generation holds the stock for one period. Hence, the model is essentially a repeated one-period model.

prices are determined by investors holding undiversified portfolios.[10] This economy is historically important because mutual funds became popular only in the late twentieth century. From another point of view, the concept of the "firm" in our model can also be applied to industries, or groups of real firms that are affected by similar technologies, or even stock indexes. For these macro "firms," investors might find it difficult to diversify their portfolios. For tractability, we assume that in this economy there are I investors in each generation, and that investor i can invest only in firm i and cannot trade with another investor. Mathematically, she solves out equation (8), subject to $W_{t+1}^i = (1+r)(W_t^i - nP_t^i) + n(D_{t+1}^i + P_{t+1}^i)$. In the equilibrium investors of each generation hold all firms.

On the other hand, for individual firms, institutional investors are well diversified. And they could be marginal investors who shape the asset prices. Hence, we also examine a second economy where a well-diversified representative investor of each generation holds all the firms in the equilibrium. I will call the first economy an undiversified economy; the second, a well-diversified economy. In the well-diversified economy, the representative investor's budget constraint will be determined by

$$W_{t+1} = (1+r)W_t + \sum_{i=1}^{I} n^i (D_{t+1}^i + P_{t+1}^i - (1+r)P_t^i). \tag{9}$$

Before she invests, the investor observes D_t^i and X_t. The investor perfectly understands the dividend policy of the firm as given by equation (6), but she cannot forecast the latent factor as a firm does. Nor will the firm reveal m_{t+1}^i to the economy until the $t+1$ return has been realized. Then, to the investor, m_{t+1}^i can be viewed as a random variable with a normal distribution $N(b_{t+1}^i,\ \sigma_\beta^2)$, where $b_{t+1}^i = E_t[m_{t+1}^i] = b_0 + \Gamma m_t^i$, $b_0 = m_0$ (same for all firms), and $\sigma_\beta^2 = \sigma_m^2$. Since the investor knows neither the beta uncertainty m_{t+1}^i nor the market risk X_{t+1} precisely, she must hedge against both. As a result, the investors' optimization problem contains nonlinear risks. Later sections will demonstrate that the nonlinear risk structure will lead to nonlinear risk premium, size (value) effects, and residual momentum in the equilibrium.

The time convention here explicitly follows the information lag existing in the real stock market. Since investors usually do not observe all useful business information collected by a firm, they cannot replicate the firm's investment decisions and need to hedge uncertainties related to these decisions. It is true that public firms issue quarterly, semiannual, and annual reports, so that investors are better informed around announcement times. But, in between, investors typically lack specific information to pin down the behavior of a firm. Furthermore,

[10] For literature on investors with undiversified portfolio holdings, one can refer to Goyal and Santa-Clara (2003) and the citations there. For the problem whether institutional investors set the stock prices or not, interested readers may refer to Bell and Jenkinson (2002).

accounting reports are backward looking. Even with accurate accounting information investors still need to make inferences about a firm's behavior in the future. In practice, analyst forecasts provide one comprehensive example of such inference.

Finally, a formal definition of the equilibrium is as follows:

Definition 1 *(Competitive Equilibrium) A competitive equilibrium is summarized by the optimal investment policy of all firms,* m^i_{t+1}*,* $i = 1, 2, , I$*, and the optimal investment decision of investors,* n^i*, and the firm's stock prices, such that*

1. (Optimality) A firm selects its optimal investment policy according to equation (4). Investors solves the optimization problem (8) with proper wealth constraint, and

2. (Market Clearing) The investor's optimal asset allocation satisfies $n^i = 1$*.*

1.3 Asset Prices in A One-Firm Economy

Before we solve the cross-sectional equilibrium asset prices in the two economies, let us first examine a one-firm and one-investor economy in order to get some basic ideas about the asset return processes. In this economy, the asset price of the firm is given in the following proposition.

Proposition 1 *(Equilibrium Asset Prices) In equilibrium, the firm's price will be*

$$P_t = A + Bb_{t+1} + Cb^2_{t+1}, \tag{10}$$

where $b_{t+1} = E_t[m_{t+1}]$ *is the investor's expectation of the firm's risk exposure in the next period, and A, B, and C are constant coefficients that can be solved recursively from (25) in Appendix A.*[11]

Based on the equilibrium prices, the next proposition describes the equilibrium return process. Given the overlapping time structure, I follow Lewellen and Shanken (2002) to study the

[11] When the dividend is $D_0 + m_{t+1}X_{t+1}$, when the firm can forecast only $\hat{F}_{t+1} = \Gamma F_t$, or when the firm reveals m_{t+1} at the beginning of the period $t+1$, the price function still contains all three terms. Especially, Appendix A proves that if m_{t+1} is revealed immediately, $b_{t+1} = E_t[m_{t+2}]$ as based on available information, and the equilibrium price will be a quadratic function of b_{t+1}. Hence, eliminating information asymmetry will not change the functional form of the equilibrium price. Economically, even when investors perfectly know the firm's behavior in the current period (so that earnings in the current period contain only the market risk), earnings in future periods will still be jointly determined by unknown market-wide and firm-level business risks. And future earnings, as well as both risk components affecting future earnings, will affect investors' current-period optimization problem through anticipated future prices at which investors can sell their stocks. In other words, investors will try to forecast the firm's future investment policy and the corresponding future earnings. But even the firm does not know exactly its future policy. Therefore, investors still face a beta uncertainty, which reflects anticipated fluctuations in the firm-level investment opportunity set in the future.

dollar return process instead of the return ratio process. Asset return is defined as the net capital gain from borrowing from the bond market to hold one share of the stock for one period. The capital gain is for a zero-cost investment portfolio and resembles the excess return in traditional asset pricing models.

Proposition 2 *(Equilibrium Asset Return) In equilibrium, the asset return can be written as*

$$r_{t+1} = D_{t+1} + P_{t+1} - (1+r)P_t \tag{11}$$

$$= m_{t+1}X_{t+1} + H_1 b_{t+1}^2 + H_2 b_{t+1} + H_3 b_{t+1}\varepsilon_{\beta,t+1} + H_4 \varepsilon_{\beta,t+1}^2 + H_5 \varepsilon_{\beta,t+1} + H_6, \tag{12}$$

where X_{t+1} is the market risk, and $\varepsilon_{\beta,t+1} = m_{t+1} - b_{t+1}$ is the prediction error for the firm's risk exposure. H_1 through H_6 are constants specified in Appendix A.

Since the analytical expression for C (and therefore B and A) is complex, it is not listed here. It suffices to know that under certain technical conditions A, B, and C are all constants. Coefficients H_1 through H_6 are functions of A, B, and C. The price contains a constant term, a linear beta term Bb_{t+1}, and a quadratic beta term Cb_{t+1}^2. Asset returns contain the market risk term $m_{t+1}X_{t+1}$, as well as a square beta term $H_1 b_{t+1}^2$ and a linear beta term $H_2 b_{t+1}$. When the firm behavior is not dynamic (σ_m^2 is small), the price simply becomes D/r, and the risk premium reverts to $\gamma_I b^2 \sigma_x^2$.

In this economy, equilibrium prices reflect all available information, and therefore the market is semi-efficient. To understand the implication of the two propositions, I will focus on one important approximation. Omitting high-order variance terms containing $\sigma_x^2\sigma_\beta^2$, one can show that

$$A \approx \frac{D}{r} + \frac{1+r+\Gamma}{1+r-\Gamma}Cb_0^2; B \approx \frac{2\Gamma C b_0}{1+r-\Gamma}; C \approx \frac{-(\gamma_I - \gamma_F)\sigma_x^2}{1+r-\Gamma^2}. \tag{13}$$

With these constants, Proposition 1 states that, for example, if investors are more risk averse than the firm, the constant C is negative and the price will be discounted. In this case the firm does not hedge enough in the opinion of the investors. Then, investors take actions to hedge more themselves by discounting the asset price. Not surprisingly, the price discount is proportional to the variance of the risk, and is quadratic in investors' expectation of the risk loading, because $b_{t+1}^2\sigma_x^2$ approximately measures the variance of the $m_{t+1}X_{t+1}$ part of the dividends. Furthermore, the equilibrium price should be discounted to the level where the corresponding risk premium matches investors' risk aversion. Proposition 2 confirms this intuition. The quadratic risk premium component in equation (12) is largely determined by the risk aversion of the investor, regardless of the firm's behavior: it is straightforward to show that $H_1 \approx \gamma_I \sigma_x^2$ and that the risk premium is $\gamma_I \sigma_x^2 b_{t+1}^2$.[12] When investors are more risk averse, or when the market is more risky, the risk premium required by the investor to hold the stock is greater.

Investors also directly ask for compensation of the beta risk, as reflected in the $H_4 \varepsilon_{\beta,t+1}^2$ term. Under the current approximation, $H_4 \approx (1+r)C + \gamma_I \sigma_x^2$. Interestingly, it will increase in both σ_x^2 and γ_I: hazards due to the unknown risk exposure become more troublesome when the market is more risky. Needless to say, more risk-averse investors require higher compensation.

1.4 The Cross-Sectional Dynamic CAPM Model

In accordance with the knowledge of the single-firm economy, we are ready to examine cross-sectional asset returns. In either the undiversified economy or the well-diversified economy, the i.i.d. assumption (all firms are i.i.d. ex-ante or at time zero) gurantees that the H_1 through H_6 parameters are the same for all firms. Hence, the market portfolio return is

$$r_{t+1}^M = \Sigma_i^I \, w_{i,t+1} r_{t+1}^i = \overline{m}_{t+1} X_{t+1} + H_1 \overline{b}_{t+1}^2 + H_2 \overline{b}_{t+1} + H_4 \overline{\varepsilon}_{\beta,t+1}^2 + H_6, \tag{14}$$

where the upper bar means the cross-sectional mean. Idiosyncratic risks have been diversified away. Normalize $\overline{m}_t$ to be one, and we have the following proposition:

Proposition 3 *(Dynamic CAPM) In both the undiversified economy and the well-diversified economy, the expected asset returns satisfy the following equation:*

$$E_t[r_{t+1}^i] = b_{i,t+1} E_t[r_{t+1}^M] + H_1(b_{i,t+1}^2 - \overline{b}_{t+1}^2) + H_2(b_{i,t+1} - \overline{b}_{t+1}), \tag{15}$$

where $E_t[.]$ is the expectation measure conditioned on end-of-period-t information, r_{t+1}^M is the return on the market portfolio, $b_{i,t+1} = E_t[m_{i,t+1}]$ is the expectation of i^{th} firm's risk exposure, and $\overline{b}_{t+1}^2$ and $\overline{b}_{t+1}$ are the cross-sectional mean of $b_{i,t+1}^2$ and $b_{i,t+1}$, respectively. The parameters H_1 and H_2 for the undiversified economy are provided by Proposition 1. The parameters for the well-diversified economy are specified in Appendix A.

This proposition builds a counterpart to the CAPM model in the economy. To make a distinction, call (15) a dynamic CAPM model, while $E_t[r_{t+1}^i] = b_{i,t+1} E_t[r_{t+1}^M]$ is a conditional CAPM. The parameter $b_{i,t+1}$ serves the role of the conditional CAPM beta, because it measures the covariance between the expected asset and market returns. Importantly, the conditional CAPM does not hold exactly, since the investors want extra compensation for nonlinear risks.

[12] The quadratic beta risk premium is a natural request by mean variance investors. For example, if a firm's return can be written as $r(t) = \alpha + \beta x(t)$, where α is the expected risk premium and $x(t)$ is an i.i.d. normal random risk with mean zero, then a positive or negative loading of the risk should carry the same degree of risk from the investor's point of view.

Given their exponential utility function, the extra compensation is related to the mean and variance of the beta risk investors face. One interesting implication of this property is that volatility can be endogenously related to expected asset return. A detailed study about volatility, however, is beyond the scope of this paper.

Proposition (3) essentially illustrates a rational framework where firm-specific risks matter in a well-diversified economy. Consistant with the traditional wisdom, linear and firm-specific investment risks will be diversified away. However, when a firm has the ability to identify superior investment opportunities and make investment decisions accordingly, expected total net cash flows or net earnings have an increasing sensitivity to better investment opportunities in a frictionless economy. Roughly speaking, this happens because a better investment opportunity generates a higher expected net cash flow per dollar investment, which induces a firm to invest more. Hence, a better investment opportunity leads to disproportionately more expected net cash flows in the future, and firms with bad investment opportunities will not be able to replicate the cash flow structure generated by a superior firm. Assuming firms ultimately distribute their net cash flows to investors, investors then ask for compensation to hold firms that cannot identify good investment opportunities. In this sense, the increasing cash flow sensitivity to investment (which is related to beta risks) will be priced. Firm-level investment risks, which are embodied in beta dynamics, now have nonlinear impact to earnings and will not be diversified away.[13]

Because traditional asset pricing models do not incorporate any firm-specific risk premium, these risk premium components become residuals based on traditional models. For example, the last two components of (15) now become the conditional CAPM model residual.[14] Since there are two sets of risks in the economy, it is reasonable that the CAPM cannot describe the return process. The surprising result is that, even when latent factors are added to the independent variable list, the linear APT model of Ross (1976) still fails to explain the whole return process. The risk premium component $H_1(b_{i,t+1}^2 - \overline{b}_{t+1}^2)$ becomes the APT model residual. Therefore,

[13]The above intuition implies the parameter $H_1 < 0$. In terms of the diversification benefit, cash flow components that are linear in firm-specific investment risks will sum up to zero with zero variation and, accordingly, have no impact to asset price. But the increasing cash flow sensitivity to investment (a result of the firm's investment activity) will generate cash flow components that are quadratic in firm-specific risks. These quadratic components will sum up to the cross-section variance of cash flow risks with nonzero variance and, consequently, have nontrivial impacts on the portfolio return. The prediction that firm-specific risks can be priced in a well-diversified economy is novel in the literature and adds to the research about "idiosyncratic" risks (Campbell, Lettau, Malkiel, and Xu (2001); Goyal and Santa-Clara (2003); Malkiel and Xu (2001); and others).

[14]Yet both the time series and cross-sectional means of the two components are zero. For example, $E_0[b_t^2] = \frac{\Gamma^2}{1-\Gamma^2}\sigma_\beta^2$ is the same for all firms. Consequently, if an econometrician is isolated from the economy and examines all the available data, he or she concludes that the time series regression does not reject the unconditional CAPM model based on Jensen's Alpha measure.

firms effectively create a set of new and nonlinear state variables in utilizing their reseach ability of forecasting investment risks. One can understand the difference between this model and the APT model as follows. In the APT framework, exposures to one systematic risk factor will earn one corresponding risk premium component. But in this model, since a firm can dynamically respond to some risk factors and adjust its investments accordingly, in general, not only the risk factors themselves, but also the interactions among risk factors are priced. These interactions allow the earnings risk exposure to depend on risks and create nonlinear asset return components that are difficult to explain by linear models.[15]

The failure of the conditional CAPM and the inadequacy of a linear APT model indicate that we might need to pay more attention to firm-specific information in order to describe the return process. This naturally leads to a favoring of the three-factor model proposed by Fama and French (1993). In the present economy a firm forms a zero-cost portfolio to finance the risky project and passes all the profits to the investor, so the book equity value will not change at the end of each period after the dividend is paid out. The market capital differs from the market-to-book equity (MB) by only a constant. Assuming the book value equals the value of the fixed-scale project in terms of risk-free asset, we have $MB_t^i = P_t^i - D_0/r = A - D_0/r + Bb_{i,t+1} + Cb_{i,t+1}^2$. Therefore, we can use either the size or the MB ratio to represent the firm's characteristics. These characteristics capture both the linear and nonlinear effect of the firm behavior. The return process with the firm characteristics can be rewritten as[16]

$$E_t[r_{t+1}^i] = b_{i,t+1}E_t[r_{t+1}^M] + H_1/C(MB_t^i - \overline{MB}_t) + (H_2 - H_1B/C)(b_{i,t+1} - \overline{b}_{t+1}). \tag{16}$$

Importantly, the first two components in (16) can be viewed as a degenerate version of the Fama and French (1993) model, because they demonstrate that besides the CAPM market risk, firm characteristics are also priced. Firm characteristics are priced because they are related to the firm-level business environment risks forecasted by a firm. In the undiversified economy, the previous approximation intuitively leads to $H_1 > 0$ and $C < 0$. Therefore, a firm's size (MB) at period t is negatively correlated with the expected return at both period t and period $t+1$, because investor discount market price and ask for high return for high risks. This is consistent

[15] Similar intuition can be found in conditional models such as Ferson and Harvey (1991). The difference between this model and the conditional beta literature is the economic intuition that firm decisions will be more easily affected by firm-specific risks than by systematic factors.

[16] Of course, we can write the formula in another way, letting the MB or size absorb the linear term and leaving an additional quadratic term on the right-hand side of the equation. But (16) has the advantage that we can sign H_1 and C. The degenerated size-MB effect follows from the simplified assumption of the model, which ignores important firm evolution features, such as startup period and distress risk. More detailed examinations of the two effects are summarized in Fama and French (1992, 1993).

with the empirical finding of the cross-sectional size (value) premium. Berk (1995) proposes a similar intuition. In the well-diversified economy, investors view firm-level business risks as valuable investment opportunities and a promise for future earnings. Investors then ask for return compensation for holding firms that cannot generate superior earnings by discounting the current market prices for such firms. As a result, $H_1 < 0$ and $C > 0$ and the size (MB) effect also exists in this economy. In both economies, equation (16) provides the economic foundation for the Fama-French model and explicitly demonstrates why characteristics are risk related.

The model thus far assumes that all predictable risks are firm-specific. As a result, firm characteristics are also firm-specific. More generally, business risks predicted by firms may contain both a systematic component (which may contain, for example, macro factors related to business cycles) and a firm-specific one. In this case, firm-specific component of characteristics still captures the asset pricing impact of forecasted firm-specific business risks, but a mimicking portfolio of firm characteristics will be able to capture that of forecasted market-wide risks. And both components of firm characteristics can be correlated with the return process. This suggests that Fama and French (1993) and Daniel and Titman (1997) reveal complementary, not necessarily contradicting, sources of return explanatory power due to firm characteristics.

But, even though firm characteristics such as book-to-market and size have a certain ability to capture the nonlinear risk premium and add explanatory power to asset returns, equation (16) suggests that they cannot perfectly address the risk premium. The last term in the equation becomes the Fama and French model residual and is crucial in explaining the three-factor-adjusted momentum in the next section. Economically, the Fama-French model residual and the APT model residual exist only when firms dynamically respond to forecasted firm-level business risks. If no firm is going to change its investment scale, then firm characteristics (or a proper APT model) will sufficiently absorb all effects of firm-level risks.

It is of course implausible if the dynamic CAPM model, aiming to explain momentum, builds its explanatory power on a relationship between firm fundamentals and returns that is inconsistent with the literature. The existence of the size (value) effect in the well-diversified economy therefore provides some support that this model inherits the intuitions from traditional asset pricing models. Other fundamentals in this economy include earnings and investments. In the context of this article, persistent high earnings result from valuable business opportunities (large m_t^2). Therefore, consistent with Fama and French (1995), a high MB ratio (size), which reveals similar information in this model ($C > 0$), signals high earnings. Thus this model properly captures the relationships between firm characteristics, earnings, and returns. Meanwhile, a valuable business opportunity induces more investments; therefore, a high investment ratio often is associated with a high MB and a low consequent return ($H_1 < 0$). Xing (2002) offers explicit evidence that supports these predictions.

2 The Existence of Momentum

This section applies the dynamic CAPM model to explain several momentum-related phenomena documented in the literature and will propose a unique model prediction to be tested. To help explain momentum profits, I briefly discuss the two most commonly studied momentum portfolio strategies: the weighted relative strength strategy (WRSS) and the winner-minus-loser strategy (WML). For the WRSS strategy, the investment weight (w_t) for any asset in period t is determined by the realized return in period $t-1$ (or r^i_{t-1}):

$$w^i_t = \frac{1}{I}(r^i_{t-1} - \overline{r}^i_{t-1}), \tag{17}$$

whereas a WML strategy sorts stocks into ten r^i_{t-1} deciles and longs (shorts) an equal-weighted portfolio of the winner (loser) deciles for the current period t. When the return process follows $r_t = \alpha_t + b_t r^M_t + \epsilon_t$, one can decompose the WRSS momentum profits as

$$\pi(R) = \sum_{i=1}^{I} \frac{1}{I}(r^i_{t-1} - \overline{r}_{t-1}) r^i_t = \pi(1) + \pi(2) + \pi(3) + \pi(4), \tag{18}$$

where $\pi(1) = cov\,(\alpha_t, \alpha_{t-1})$, $\pi(2) = cov\,(\beta_t, \beta_{t-1})\, r^M_t r^M_{t-1} + cov\,(\alpha_t, \beta_{t-1})\, r^M_{t-1} + cov\,(\beta_t, \alpha_{t-1})\, r^M_t$, $\pi(3) = cov\,(\varepsilon_t, \beta_{t-1})\, r^M_{t-1} + cov\,(\beta_t, \varepsilon_{t-1})\, r^M_t + cov\,(\alpha_t, \varepsilon_{t-1}) + cov\,(\varepsilon_t, \alpha_{t-1})$, and $\pi(4) = cov\,(\varepsilon_t, \varepsilon_{t-1})$, and $cov\,(\alpha_t, \alpha_{t-1}) \equiv \sum_{i=1}^{I} \frac{1}{I}(\alpha^i_{t-1} - \overline{\alpha}_{t-1})\alpha^i_t$ stands for the cross-section autocovariance. In the current model, $\alpha_t = H_1(b^2_{i,t+1} - \overline{b}^2_{t+1}) + H_2(b_{i,t+1} - \overline{b}_{t+1})$. The first term, $\pi(1)$, is the momentum contribution of the dynamic risk premium predicted by the model. The second term contains contributions related to a dynamic loading of the CAPM market factor. The third term represents the interaction between unknown factors and the market or latent risk factors, and the last term denotes the contribution due to missing factors. Economically, the first two momentum components can be clearly interpreted as compensations for firm-level risks and the market risk, respectively, while the last two components can be thought of as the risk-adjusted momentum return. This section mainly focuses on the first two components to theoretically explain documented empirical findings regarding momentum. Later sections will empirically estimate all the four components of momentum profits.

Momentum. Even when an unconditional asset pricing model is not rejected by Jensen's Alpha measure, the econometrician can still observe momentum in a semi-efficient market. First, the $\pi(2)$, or $cov\,(b_t, b_{t-1})\, r^M_t r^M_{t-1}$, component of momentum profits in (18) can explain a small fraction of momentum. Far more important, however, is the $\pi(1) = cov\,(\alpha_t, \alpha_{t-1})$ term measuring the momentum contribution of the time-varying risk premium not captured by the conditional CAPM model. It is straightforward to show that both the linear and quadratic terms of b_t are serially autocorrelated, following the autocorrelation of firm investments. In addition, there is a lead-lag effect in the economy, because the risk premium is determined by the beta's relative position $(b^2_{i,t+1} - \overline{b}^2_{t+1})$ in the economy. For example, in a two-firm economy, a high return in the current period for one firm in general "predicts" the low return for the other firm in the next period. Both autocorrelation and the cross-sectional lead-lag effect lead to momentum.

Since neither the CAPM, nor APT, nor the Fama-French model explain the total risk premium, residuals of the three models (now including the risk premium not captured by the models) demonstrate momentum. This prediction is supported by Grundy and Martin's (2001) finding that Fama-French model residuals exhibit significant momentum profits and Griffin, Ji, and Martin's (2002) evidence that APT models with macro and business cycle factors fail to explain momentum. Section 3 shows that the risk-adjusted momentum, according to the dynamic CAPM model, will be much less profitable.

Well-known properties of momentum. Consistent with this model, momentum profits are most favorable for a holding period between 3 and 12 months, because earnings uncertainties in the next quarter or year are likely to have the most significant asset pricing impact based on the current financial reporting scheme. Firms can change their investment or risk exposure during this period without alerting investors. Even when firm investments have a longer horizon, investors will care only about projected investment uncertainty contained in such a period, anticipating the uncertainty to be resolved by the next quarterly or annual report. Hence, momentum will also be most significant in the next 3 to 12 months. The dissipation of momentum at longer horizons is mathematically captured by the AR(1) property of business risks. To calibrate the model, let us focus on the quadratic APT anomaly term $H_1(b^2_t - \overline{b}^2_t)$. Suppose in the current period there is a Δ shock for asset i for this APT anomaly term. It is straightforward to show that $E_t(b^2_{t+k} - \overline{b}^2_{t+k}) = \Gamma^{2k}(b^2_t - \overline{b}^2_t)$. The expected excess return will decay to a fraction f in $lnf/(2ln\Gamma)$ periods. For example, if Γ is 0.75 and each period contains six months, it takes about four periods or two years for an excess return to decay to 10% of the original level. This dissipation rate seems to be reasonable.[17]

As for the economic magnitude of momentum profits, note that, when there is little time variation in beta, the quadratic risk premium for the investors to hold the whole market is

[17] The current model explains well momentum and the dissipation of momentum, but does not gurantee a long-run reversal. George and Hwang (2004) provides evidence that short-term momentum and long-term reversals are largely separate phenomena.

$H_1\bar{b}_t^2 \approx \gamma_I \sigma_x^2 \bar{b}_t^2$ in equation (14). Since $\bar{b} = 1$ on average, $\gamma_I \sigma_x^2$ approximately has the magnitude of the well-known equity premium (7% per year). Momentum returns for the quadratic term equal $\gamma_I \sigma_x^2 (b_{t+1}^2(winner_t) - b_{t+1}^2(loser_t))$. If the current period winners have an average beta of 1.5, and the losers have one of 0.5, then $b_t^2(winner) - b_t^2(loser)$ is approximately 2. For the next period, the winner-loser spread will decay to $2(.75)^2$ of the equity premium. This leads to about 8% per year for momentum profits, which is in proper order.[18]

Hong, Lim, and Stein (2000) illustrate that momentum profits are driven mostly by loser stocks rather than by winner stocks. This asymmetry is difficult to explain by traditional linear models, as noticed by these authors, but exactly implies a nonlinear risk premium as proposed by this article. To demonstrate the intuition, let us assume that $b_{i,t}$ has a normal cross-section distribution of $N(\bar{b}_{i,t}, \sigma_b^2)$. When expected market return is positive, winner stocks on average have a higher $b_{i,t}$ value. It is true that the $H_2(b_{i,t+1} - \bar{b}_{t+1})$ component in equation (15) is symmetric for winners and losers, but momentum profits also contain the contribution from the quadratic term: $\pi(R) \propto \sum_{i=1}^{I} \frac{1}{I}(b_{i,t}^2 - \bar{b}_{i,t}^2)(b_{i,t+1}^2 - \bar{b}_{i,t+1}^2) \approx \sum_{i=1}^{\Gamma^2} \frac{\Gamma^2}{I}(b_{i,t}^2 - \bar{b}_{i,t}^2)^2$. This quadratic contribution is asymmetric because b_i^2 has an asysmetric Chi-square distribution, and it is easy to check that $\sum_{loser}(b_{i,t}^2 - \bar{b}_{i,t}^2)^2 > \sum_{winner}(b_{i,t}^2 - \bar{b}_{i,t}^2)^2$. As a numerical example, suppose the economy contains three stocks with their betas to be 0 (loser), 1, and 2 (winner). Then, $(b_{i,t}^2 - \bar{b}_{i,t}^2)^2$ are 25/9 and 1/9 for the loser and the winner stock, respectively, where the loser obviously generates more momentum return. Intuitively, the cash flow generated by a firm has an increasing sensitivity to investment. The increasing sensitivity leads to a nonlinear sensitivity of expected asset returns to firm investment, and consequently creates a difference between winner's momentum and loser's momentum.

Even though this paper focuses on the cross-sectional properties of momentum, the model also provides a powerful tool with which to examine the time-series property of momentum. For example, the model can easily explain why momentum is more significant during business cycle expansions than during the contractions (Chordia and Shivakumar (2002)). Economically, when firms increase their investments, there will be more cross-sectional dispersion of the nonlinear risk premium component. This is because the momentum component $\sum_{i=1}^{I} \frac{1}{I}(b_{i,t}^2 - \bar{b}_{i,t}^2)^2$ can be

[18] It is well known that the exponential utility function usually cannot generate a significant equity premium (Mehra and Prescott (1985)). This paper therefore does not address the equity premium puzzle or calibrate the model directly from the γ_I and σ_x^2 parameters. Rather, the calibration here demonstrates that momentum returns are reasonable according to the observed magnitude of the U.S. stock market equity premium. The dynamic CAPM model implies that the magnitude of momentum and that of the equity premium can be endogenously related. Appendix A shows that H_1 in the diversified economy usually has the magnitude of $\gamma_F \sigma_x^2$, so momentum profits there can have a similar magnitude.

regarded as the cross-section variance of b_i^2 and is proportional to $\overline{b}_{i,t}^2$. In the expandion period, the aggregate investment of the economy (or $\overline{b}_{i,t}$) grows. As a result, both $\overline{b}_{i,t}^2$ and momentum profits will increase. Similarly, expanding investment may also help explain why momentum profits are more significant following positive market return (Cooper, Gutierrez, and Hameed (2004)).

Industry and Fama-French portfolio momentum. The existence of industry momentum (Moskowitz and Grinblatt (1999)) has an especially clean interpretation in this model, because industry by its definition may be viewed as a macro firm containing all smaller firms affected by a common technology or investment opportunity in the undiversified economy. From another point of view, predicted business environment risks can contain both industry-level and firm-level components. For example, SUV demand shocks may affect all car manufacturers, but each firm can develop its own model of SUV. Naturally both industry momentum and firm-level momentum can exist in the economy.

Interestingly, the model here implies that the Fama-French portfolios themselves will exhibit momentum, because average expected returns of each sorted portfolio approximately follow (16) and still have Fama-French model adjusted momentum. Lewellen (2001) provides empirical evidence that size and book-to-market portfolios exhibit momentum. This momentum can be regarded as a portfolio version of the Fama-French model adjusted momentum.

Earnings momentum. Equations (15) and (16) not only predict the existence of momentum, but also imply a link between momentum and corporate behavior in the well-diversified economy. As mentioned earlier, investors interpret high earnings as a result of good business opportunities. To the extent that return is also driven by a firm's forecasted risks, which are unlikely to change dramatically over time, the relationship between earnings and returns will persist for a while. Consequently, the model predicts the existence of earnings momentum. The empirical evidence of earnings momentum, such as that documented in Chan, Jegadeesh, and Lakonishok (1996), provides a hypothetical test for this model.

First, the market-adjusted unexpected earnings (as a proxy for earnings news) are $\gamma_F \sigma_x^2 (m_t^2 - m_{t-1}^2)$ in this model. The earnings momentum strategy therefore longs (shorts) stocks with high (low) $m_t^2 - m_{t-1}^2$ values. Since firms are ex-ante identical, those with positive earnings shocks are more likely firms with low previous earnings (mathematically, $E_t[m_{t+1}^2 - m_t^2] \approx (\Gamma^2 - 1)m_t^2 + E_t[\epsilon_{m,t+1}^2]$ decreases in m_t^2). Equation (15) implies that expected return $E_t[r_{t+2}]$ also decreases in m_t^2 in the well-diversified economy. As a result, large positive earnings shocks are associated with large positive future returns. Sorting on unexpected earnings will then introduce a spread on the expected return and generate significant momentum profits. Second, the return spread of the earnings momentum depends on the difference $m_{t+1}^2 - m_t^2$, whereas that of the price momentum depends on m_{t+1}^2. The two strategies are therefore different and cannot subsume each other. Roughly speaking, m_{t+1}^2 better correlates with future returns than the $m_{t+1}^2 - m_t^2$ term. Thus, this model implies that price momentum is more profitable and persistent than earnings momentum.[19] Finally, the earnings momentum is not absorbed by size or book-to-market. All these detailed predictions are consistent with Chan, Jegadeesh, and Lakonishok (1996). This article therefore offers an alternative explanation (or one that can be compared with their underreaction hypothesis).

Characteristics-based momentum. Equation (16) predicts that the two seemingly unrelated anomalies, firm characteristics and momentum, may economically share the same foundation. This interpretation provides new insight into the documented relationship between the two anomalies. Daniel and Titman (2001) and Sagi and Seasholes (2001) find that high MB (growth) firms tend to generate high momentum returns. This model generates precisely this effect. Equation (16) implies that $E_t[\pi(R)] \propto \frac{1}{I}\sum_{i\in I}[G_1(MB_t^i - \overline{MB})^2 + G_2 MB_t^i]$, where G_1 and G_2 are two positive parameters specified in Appendix A. Since $G_2 > 0$, momentum portfolios based on high MB firms are likely to generate more profits. Furthermore, the first component $G_1(MB_t^i - \overline{MB})^2$ suggests a possible (but not necessary) U-shaped relationship between momentum and MB.[20] Economically, high-MB firms are more likely firms with temperate good investment opportunities. These firms will probably invest more aggressively, anticipate more future earnings, invoke more CAPM deviations, and consequently generate higher momentum profits.

The relationship between size and momentum is more complex. Under the current assumptions, this model predicts that momentum profits should increase in firm size. But, though book-to-market equity generally measures the investment opportunity set of a firm, size has some other effects. For example, large firms usually have more analyst coverage and lead small firms in terms of information (Hong, Lim, and Stein (2001), Hou (2002)), so in reality it may be more difficult for investors to precisely predict small-firm behavior. If we relax the current

[19] A simple derivation might help in this case. Let us focus on the b_t^2 part of the return. We have $r_{t+2} \sim b_{t+2}^2 \sim m_{t+2}^2$. Expected earnings momentum profits are proportional to $\gamma_F \sigma_x^2 cov(m_{t+2}^2, m_{t+1}^2 - m_t^2)$, which equals $\gamma_F \sigma_x^2 cov(m_{t+2}^2, m_{t+1}^2) - \gamma_F \sigma_x^2 cov(m_{t+2}^2, m_t^2)$. Note the first term reflects the profits for price momentum. The second term roughly equals Γ^2 of the first term. Therefore, the earnings momentum profits are approximately $1 - \Gamma^2$ of (and smaller than) the price momentum profits.

[20] Appendix A provides more details. Sagi and Seasholes (2001) sort firms into four MB quartiles (their Table 3). The momentum profits from MB quartile 1 (low) and 4 (high) are higher than those generated from MB quartiles 1+2 and 3+4, respectively, suggesting that momentum in quartiles 2 and 3 is less profitable than in quartiles 1 and 4.

model assumption to allow investors' prediction error of a firm's behavior to depend on firm size (i.e., small firms have a larger prediction error variance), then investors will not only expect a positive return for real investments, but also require some additional return for small firms because of this enlarged prediction error (see Appendix A for more details). In this case, the extra premium associated with a larger level of information asymmetry in one period and expected positive return of real investments in the next period can effectively produce a positive return autocorrelation (other things equal). As a result, small firms can generate relative large momentum profits. Note that this information premium can lead to a negative relationship between momentum and size without resorting to investor under- or overreaction. Combining the positive and negative relationships between momentum and size can produce an inverted U-shape relationship between the two, as reported in Hong, Lim, and Stein (2000). Earlier studies usually focus on and explain the second part of the inverted U-shape. The omitted part, however, also seems to be a natural output of this model.

New prediction. Although all the empirical evidence listed above directly supports the model, the remaining section aims to achieve more by proposing unique predictions untouched by the literature. In the economy proposed here, the unconditional CAPM holds exactly. But conditionally, because of dynamic firm behavior, the risk premium α_t can deviate from zero at times and consequently generate momentum. In this case, based on firms with more dynamic investment policy (more dynamic α_t and more deviations from the conditional CAPM), momentum portfolios could be more profitable. With the standard deviation of α_t as a proxy for the asset pricing impact of firm dynamics, the model predicts that momentum portfolios based on a group of higher $\sigma_{\alpha(t)}$ stocks will be more profitable. Mathematically, $E_{t-1}[\pi(R)] \propto E_{t-1}[\sum_{i\in I} \alpha_{t-1}^i \alpha_t^i] \propto \sum_{i\in I}(\alpha_{t-1}^i)^2$. The last approximation follows because the risk premium is positively autocorrelated according to this model (the correlation itself is omitted because it is the same for all firms). Furthermore, the relationship between momentum profitability and $\sigma_{\alpha(t)}$ should not be subsumed by other characteristics, such as size and market-to-book. Section 3 empirically confirms these predictions. The measure of firm dynamics, albeit unique to this model, reflects an econometrician's estimation of the asset pricing impact of a firm's investment policies. With this economic explanation, we may regard it as a special firm characteristic.

3 Empirical Estimation of Momentum Profits

Empirically, one of the most interesting questions is what fraction of momentum profit can be attributed to risks by this simple model. This section mainly explores the in-sample explanatory power of the model but will also provide one out-of-sample test for the model prediction.

3.1 A Time Series Model and the Kalman Filter

The traditional rolling regression method may not be directly applicable to tests and estimations of a dynamic model, because it largely ignores short-term dynamics. To bypass problems related to rolling regressions, this section proposes an alternative computational technique in order to more precisely estimate firm dynamics. The structure of the model, especially the assumption of an AR(1) latent risk, invokes the possibility of using the Kalman filter to estimate the risk premium. Decomposing the CAPM beta into a static part and a dynamic part, $b_t = \beta_0 + \beta_t$, where β_t follows the AR(1) process of the latent factor (its unconditional mean is normalized to be zero), we can write the asset return process as (suppressing the index i for the i^{th} firm)

$$\begin{aligned} r_t &= \alpha_t + (\beta_0 + \beta_t) r_t^M + \epsilon_t, \qquad (19) \\ \alpha_t &= \alpha_0 + H_1 \beta_t^2 + H_2 \beta_t; \; \beta_t = \Gamma \beta_{t-1} + \varepsilon_{\beta,t}, \end{aligned}$$

where α_t is the conditional risk premium, $\epsilon_t \sim N(0, \sigma_\epsilon^2)$, and $\varepsilon_{\beta,t} \sim N(0, \sigma_\beta^2)$. The residual $\epsilon(t)$ reflects the risk-adjusted return implied by this model. Note that, under the model assumptions, the market is efficient and asset return in (19) should have already included all useful information.

Since equation (19) is nonlinear in the latent factor, this paper adopts the extended Kalman filtering technology.[21] The estimation of the risk premium takes two steps. First, the Kalman filter observes a firm's return and the market excess return and then estimates the Kalman parameters (as reported in Table 1, Panel B). Next, these parameters, together with the whole return process, give out a smoothed estimation of the α_t and β_t process (the best estimation based on all available data). A minimum of 60 data points is required for the Kalman estimation. Because of this requirement and the finite sample period, no cross-sectional or unconditional constraints will be imposed on the estimation. If the Kalman model is sufficient in describing the asset return process, we expect to see no significant risk-adjusted momentum profits.

3.2 Data Description and Momentum Portfolios

Monthly returns seem to be a natural choice for our estimation, not only because of the literature's conventions, but also because this horizon is likely to correlate with measurable changes in firm fundamentals. Though quarterly returns would be another good candidate, a minimum requirement of 60 quarters might be too long. Monthly stock return data and market portfolio returns come from the Center for Research in Security Prices (CRSP). To eliminate possible

[21] The basic idea (Harvey (1989)) is to set the quadratic term of the latent factor as ${\beta_t}^2 = (\beta_{t|t-1} + \beta_t - \beta_{t|t-1})^2$, where $\beta_{t|t-1}$ is the Kalman estimation of β_t given all $t-1$ information, and then omit the higher orders of the error term $\beta_t - \beta_{t|t-1}$. Details of the Kalman filter can be found in Harvey (1989) and Hamilton (1994).

survivorship bias and liquidity constraints, I also look at a special group of stocks: the S&P 500 index stocks. A list of the S&P 500 index firms at the end of the year 1999 is extracted from COMPUSTAT. The additions and deletion of the index firms (year 1976 to 1999) is downloaded from Jeff Wurgler's Web page.[22] I then replicate the S&P 500 index from January 1976 to December 1999. For earlier months, I use the S&P 500 firm list as it appeared in January 1976. All variables used to construct book equity come from the COMPUSTAT database.[23]

To be included in the sample, an asset should have at least 60 months of return data from January 1962 to December 2000 and converged Kalman estimations for all available returns during the same period. S&P firms include all members of the S&P 500 index during the period from 1975 to 1999. Panel A of Table 1 lists the mean excess return and Sharpe ratio of the pooled NYSE-AMEX and S&P stocks. This panel also reports the R^2 statistics of the OLS CAPM and Fama-French three-factor model regressions, and the effective R^2, as well as the revised R^2_D, of the Kalman estimation.[24] The panel shows that the CAPM explains a relatively small part of the return variations. The Kalman model generally improves the explanatory power by 20%. The Kalman R^2 is even larger than that of the Fama and French three-factor model, indicating that the Kalman model is able to capture more return variation than traditional models. Panel B reports the mean and cross-sectional standard deviation of the Kalman parameters for the pooled stocks. The Kalman estimations of α_t and β_t time series, though not listed here, will be used for future tests. [Insert Table 1 here.]

[22]The author thanks Jeff Wurgler for making the data available.

[23]Book equity is defined as the stockholders equity, plus balance sheet deferred taxes and investment tax credit, plus postretirement benefit liabilities, minus the book value of preferred stock (if available). Since I will use only the mean of book-to-market value for available periods, the COMPUSTAT fiscal year problem will not be important in this case. I simply match the year t reported book equity with the December market cap of the same year to get the BM ratio. The market capital in June of each year will be used as size.

[24]For the Kalman filter the effective R^2 measure, which reflects the goodness of fit, is defined as $R^2 = 1 - SSE/(\widetilde{r}'\widetilde{r})$, where SSE is the sum of residual squares and $\widetilde{r}$ is the demeaned excess return process. This definition is consistent with the OLS regression. To eliminate possible bias caused by time series with a trend, Harvey (1989, p. 268) proposes a revised measure $R^2_D = 1 - SSE/\sum_{t=2}^{T}(\Delta r_t - \overline{\Delta} r_t)^2$, where $\Delta r_t = r_t - r_{t-1}$ is the return increment and $\overline{\Delta} r_t$ is the mean of the first differences. I first estimate the Kalman model for available returns from Jan. 1962 to Dec. 2000 and then perturb estimated parameters for 20 times to get the global maximum log likelihood function. Momentum tests will be conducted from 1965 to 1999. The parameter estimating period is slightly longer than the testing period, in order to allow formation periods before 1965 and holding periods after 1999. The estimation from 1965 to 1999 gives almost identical results. As a robustness check, I also pick up the global maxima according to the two R^2 measures and the main conclusions of this study hold. An estimation is regarded as converged if $|\alpha_t| < 2, |\beta_t| < 2$ and σ_β will not hit the upper bound of 10 (more than 100 firms are discarded). If a firm has two valid return periods separated by missing data, then it will be counted twice in the pool, one for each period.

To form a momentum portfolio at the beginning of month t, I rank realized stock returns in a formation (or ranking) period of F months from month $t-F$ to $t-1$. Once the momentum portfolio is formed following either the WRSS or WML strategy, it will be held for H months (from month t to $t+H-1$). Sometimes a gap of G months (usually one, if not zero) is inserted between the ranking and the formation period. A complete momentum strategy can be described by PortfolioStrategy:F:G:H. When there is no confusion, I abbreviate WRSS:1:0:1 as WRSS101. Finally, the momentum return is normalized to be the profit realized per dollar long.[25]

The literature usually reports monthly momentum return based on overlapping portfolios for past winners (losers).[26] For the decomposition equation (18), however, this approach might be confusing when F and H are both greater than one, because more than one portfolio formation period is used to calculate the momentum return in each month. To make the interpretation straightforward, I instead focus on the decomposition of holding-period returns of nonoverlapping momentum portfolios.[27] Specifically, I construct only one momentum portfolio at the beginning of month t (based on realized returns of previous F months), hold it for H months, and then report the monthly return and its components as generated by the portfolio during the entire H months. This approach is economically appealing: just replace r_{t-1} with the formation-period return, replace r_t with the holding-period return, and (18) cleanly decomposes total momentum profits in terms of autocovariance among ranking and holding-period risk premium, market factor, and errors. The advantage, however, comes at some cost: the holding period for the month t momentum portfolio and month $t+1$ portfolio will overlap for $H-1$ months. These overlapping holding periods can induce an autocorrelation to the holding-period returns and bias the t statistics. To remedy the problem, whenever I report the returns for nonoverlapping momentum portfolios, the t statistics will be Newey-West adjusted by a lag of $H-1$ months (Newey and West (1987)). For results unrelated to the decomposition, I adopt the traditional overlapping portfolio method. Finally, following Grundy and Martin (1999), I calculate the formation period returns as the mean monthly returns during the ranking periods. Consistently, holding-period returns for nonoverlapping momentum portfolios are computed as the mean monthly returns of the entire H-month holding period. This method has the advantage of diversifying away the

[25]For the WML case, the momentum investor invests one dollar in the winner portfolio and shorts one dollar in the loser portfolio. For the WRSS case, all positive weights sum up to one for any momentum portfolio. Therefore, the weights can be directly interpreted as the dollar amount invested into individual stocks.

[26]Details are discussed by Jegadeesh and Titman (1993).

[27]As a robustness check, I look at the decomposition of both overlapping and nonoverlapping portfolio momentum returns. All results will be similar, with use of either method. Over a long run, the magnitude of momentum profitability should be the same, whether we view momentum profits as coming from monthly returns of overlapping portfolios, or from H-month returns of nonoverlapping portfolios.

idiosyncratic errors of the CAPM model, as discussed by Grundy and Martin (1999).

3.3 Decomposition of Momentum Profits: 1965-1999

Before we look at the real data, a simulation experiment helps explain how well the technology can capture the dynamics of momentum profits. All details are provided in Appendix B. The simulations demonstrate that the Kalman technology can estimate the time-varying betas and risk premium very well. For the $\pi(2)$ component of momentum profits, the Kalman model usually gives out a smoothed though unbiased estimation, whereas the $\pi(1)$ estimation is somehow downward biased under the null model. This is because a smoothed (though unbiased) estimation of the risk premium leads to a downward-biased estimation of the autocovariance and momentum profits.[28] Nevertheless, Panel A of Figure 1 illustrates that the $\pi(1)$ estimations (shadowed area) convincingly track the time series fluctuations of the true value (solid line). Hence, we can rely on the Kalman filter to capture the time series property, as well as a majority magnitude, of momentum profits.

In Table 2, all NYSE-AMEX listed firms with at least 60 months of returns and converged Kalman estimations are used. In forming momentum portfolios, I use only firms with a valid formation and holding period return (survivorship biased).[29] The formation periods are set to be 6 or 12 months, while the holding periods are 3, 6, and 12 months when $G = 0$, and 1, 3, 6, and 12 months when $G = 1$. The goal here is not to restate the well-documented relationship between holding horizon and momentum profitability, though the result confirms the relationship. The different lengths of F and H are designed primarily to have a wide spread over profitable momentum strategies. Panel A uses the WRSS strategy, whereas Panel B focuses on the WML strategy. This table thus covers 28 momentum strategies. Different momentum portfolios for the raw asset returns are formed from January 1965 to December 1999. The magnitudes of momentum returns are, as expected, very close, using overlapping or nonoverlapping portfolio methods. Therefore, just the nonoverlapping portfolios monthly returns ($\pi(R)$) are reported. To determine the significance level, I report the time series t statistics (T_{JT}) for overlapping portfolio momentum returns and the Newey-West $H-1$ lags-adjusted t statistics (T_{NW}) for nonoverlapping portfolio returns. The Newey-West t statistics appear to be slightly larger than the time series t statistics, but the difference is unlikely to drastically change the significance level of momentum profits. Panels A and B illustrate significant momentum profitability for

[28]The downward bias works only against the current model. Other filtering technologies may have similar or even more serious problems. For example, unconditional OLS regressions smooth the risk premium to a straight line and therefore miss most momentum returns.

[29]I will come back to the survivorship bias shortly.

a wide range of momentum strategies. When a lag of one month is inserted, the momentum strategies often become more profitable. These results are in general consistent with Jegadeesh and Titman (1993), where they use a lag of one week. [Insert Table 2 here.]

Next, the Kalman model is used to examine momentum profits. Recall that, if the model is sufficient, risk-adjusted momentum should not be profitable. In accordance with this intuition, I calculate two types of risk-adjusted momentum. For the first type, the winner-loser position (or the portfolio weight) is determined by the realized return, while the holding-period return, $\pi'(\epsilon)$, is computed from the Kalman ϵ_t residuals. Economically, $\pi'(\epsilon)$ is the risk-adjusted momentum return based on the model. As a sensitivity check, I also compute a second risk-adjusted momentum return $\pi''(\epsilon)$, where the Kalman ϵ_t residuals determine both investment weights and holding-period returns (replacing the return process in (18) with the ϵ_t residuals). Panels A and B report the risk-adjusted momentum in percentage of the original return momentum: $\eta' = \pi'(\epsilon)/\pi(R) \times 100\%$, and $\eta'' = \pi''(\epsilon)/\pi(R) \times 100\%$. Note that $1 - \eta'$ can also be interpreted as the fraction of $\pi(R)$ that is risk related according to the model. The Newey-West t statistics are omitted, while the superscripts 1 and 2 indicate a significance level of 1% and 5% for a two-tailed t test based on the Newey-West t measure, respectively.

The result is striking. In Panel A, for $\pi'(\epsilon)$, more than half of the strategies do not end up with significant risk-adjusted momentum returns. In other words, the dynamic CAPM model totally explains these momentum-strategy profits. For the remaining significant risk-adjusted momentum profits, the magnitude usually drops to around 30% of the original magnitude for a set of WRSS strategies. Equivalently, dynamic risk premium explains more than half of the momentum profitability. For $\pi''(\epsilon)$ we get similar results, though more momentum strategies remain profitable. Panel B, however, suggests that the model is less successful in explaining the return process of the most extreme stocks. For example, compared to the counterpart of Panel A, the range of η' suggests that the top and bottom 10% stocks generate more risk-adjusted momentum profits. Nevertheless, more than 40% of momentum profits are still risk related. If we look at the economic magnitude of the risk-adjusted momentum η', then it is usually less than 4% a year for WRSS strategies and 6% a year for WML strategies before transaction costs.

To further explain the source of momentum profits, I decompose the WRSS momentum returns according to equation (18). Table 3 provides an intuitive way to look at the relative contribution of the four components: $\eta_i = |\pi(i)|/(|\pi(1)| + |\pi(2)| + |\pi(3)| + |\pi(4)|) \times 100\%$. The η_i will be interpreted as the percentage or relative contribution of $\pi(i)$ to momentum return. In Panel A, relative contributions are calculated for nonoverlapping momentum portfolios. To understand these numbers, let us focus on the WRSS603 strategy. Panel A says that about

57% of momentum profits are due to risk premium, 11% are due to dynamic loading of the market factor, 16.5% are due to the interaction between unknown factors and the risk premium or market factor, and 16% are due to reasons purely outside this model. These numbers seem to be typical for all strategies. Therefore, according to the dynamic CAPM model, most of the explanatory power comes from the risk premium associated with firm-level risks ($\pi(1)$ is usually significant and contributes 50% to 70% of momentum returns). A better estimation of beta adds only marginal explanatory power (the relative magnitude of $\pi(2)$ is usually less than 10%). The error-related parts, $\pi(3)$ and $\pi(4)$, are relatively small (together 20% to 40% of momentum returns). In a robustness check, Panel B decomposes the returns for momentum strategies based on overlapping portfolios. The relative strength could be off around 10%, but this panel supports the conclusion that a leading part of momentum profits may be related to risks, especially to firm-level risks. Finally, different formation and holding-period horizons have some, but not prevailing, impacts on the relative contributions. [Insert Table 3 here.]

In applying the Kalman model in Table 2, I require a minimum of 60 months of valid returns for a stock to be included in the momentum portfolio. Since both the raw return and Kalman risk-adjusted momentum strategy are based on a common set of data, this requirement is unlikely to seriously undermine the decomposition results. Nevertheless, Table 4 explicitly examines the momentum of the S&P 500 index firms, firms least likely to be affected by this requirement, as well as the survivorship biases.

To be included into an SP momentum portfolio, a firm must be a member of the S&P 500 index in the month right before the holding period. During other months of the ranking or holding periods, the firm may or may not be in the index. The momentum portfolios are formed from January 1965 to December 1999. The Kalman ϵ_t residuals are then used to calculate SP risk-adjusted momentum profits for nonoverlapping momentum portfolios, based on formation-period realized returns.[30] Since this momentum is not well studied in the literature, I extend the holding period up to 2 years. In general, the SP momentum portfolios generate lower (but still significant) profits compared to their NYSE-AMEX counterparts. Although the SP momentum portfolios can earn significant profits up to a holding horizon of 2 years, the Kalman estimation shows that the risk-adjusted momentum return for a long holding period is very low, and may even be negative. Consistent with Table 2, the Kalman model sharply decreases the significance level of risk-adjusted momentum and usually explains more than 50% of the raw return momentum profits for WRSS strategies and more than 30% for WML strategies. The

[30] To save space I will not report the t statistic and the decomposition based on overlapping momentum portfolios, and will not report the type 2 risk-adjusted momentum η'': all these results are consistent with the conclusion.

relative contribution of the four components are plotted in Figure 3 for WRSS strategies with $F = 6$ and 12, and $G = 1$. The graph directly reflects that the fraction of momentum due to risk is enormous (around half) and does not diminish for long holding horizons. For a variety of other momentum strategies, a similar fraction of the SP momentum profits originates from the dynamic risk premium. [Insert Table 4 and Figure 3 here.]

The study of the S&P firms encompasses another advantage: these are the most important and liquid firms in the economy. It is therefore unlikely that the explanatory power of this model is totally subject to liquidity constraints. Though not reported, a study including NASDAQ stocks confirms that the previous decomposition results remain largely the same for stocks in different exchanges.

3.4 Empirical Test of the Model Prediction: 1965-1999

The theoretical model predicts that momentum portfolios should be more profitable when based on more dynamic firms, as measured by larger standard deviations of estimated $\alpha(t)$ time series. This section aims to explicitly test this hypothesis. Table 5 sorts firms into four quartiles based on the ex-post estimated $\sigma_{\alpha(t)}$ (Group 0, 1, 2, and 3 [highest $\sigma_{\alpha(t)}$]). It turns out that Group 0 firms have a virtually zero $\sigma_{\alpha(t)}$. We regard this zero $\sigma_{\alpha(t)}$ as an indication that the Kalman estimation fails to properly capture the firm dynamics and reverts to a static OLS model. For this reason Group 0 is discarded in the current table. Excluding these firms will not change the previous decomposition conclusions. [Insert Table 5 here.]

Panels A1 and A2 report the monthly momentum return for WRSS and WML overlapping portfolios. Nonoverlapping portfolios will generate similar results for all remaining tables. It is apparent that for all listed momentum strategies, profitability monotonically increases from Group 1 to 3. Furthermore, compared to Table 2, Group 1 generates fewer momentum profits than the pool of all firms, whereas Group 3 generates more. The return difference between the third and the first group is huge. Panel B calculates the mean and the t statistics of the differences. For example, the WRSS613 strategy generates a profit difference of more than 7% a year (comparable to the momentum profitability based on all firms). The WML613 profit difference is even higher (more than 10% a year). In general, the difference is more than half of the Group 3 profits and is both statistically and economically significant. These numbers indicate that both the theoretical model and the empirical tool are quite powerful in examinations of momentum profits.[31]

[31] If the model is completely misspecified, or if overfitting of data mainly drives the estimation of firm dynamics, then we should not observe significant difference here.

The dynamic CAPM model also predicts that firm characteristics such as size and MB are related to momentum. For example, there exists an increasing relationship between momentum and MB ratio, which should not subsume the relationship between σ_α and momentum. To test this hypothesis, Table 6 independently sorts firms into four $\sigma_{\alpha(t)}$ quartiles and three MB groups. Since the $\sigma_{\alpha(t)}$ measure is ex-post, the MB value is calculated as the mean of the inverse of available book-to-market values. All firms with negative mean MB ratios have been discarded. While the ex-post measure might induce an underestimation of the momentum sensitivity to MB, this approach eliminates possible measurement error in the MB ratio and is reasonable for an in-sample test. Momentum portfolios are formed for each of the 12 double-sorted pools. Each pool contains 400 to 600 firms. Table 6 focuses on the 606 and 616 strategies. Panel A1 and A2 report the time series mean and t statistics for WRSS606 and WRSS616 momentum profits based on overlapping portfolios. Panel B reports same result for WML strategies. Momentum profitability is increasing in $\sigma_{\alpha(t)}$ from Group 1 to 3 within each MB group. For 2 of the 3 MB groups, $\sigma_{\alpha(t)}$ will generate enough spread that the difference between Group 3 and Group 1 momentum returns is significant. Within each $\sigma_{\alpha(t)}$ group (including Group 0), high-MB firms are able to generate significant momentum profits. The low-MB firms usually generate more momentum profits than median-MB firms but fewer profits than high-MB firms. As mentioned before, this somehow U-shaped relationship is consistent with the dynamic CAPM model.[32] The relationship between the momentum profits, firm dynamics, and MB is more intuitively illustrated by Panel A and B of Figure 4. Clearly, the MB measure does not subsume the relationship between momentum and firm dynamics. [Insert Table 6 and Figure 4 here.]

Table 7 independently sorts firms into four $\sigma_{\alpha(t)}$ quartiles and three size groups. The mean of June market capital during the whole period is used as the ex-post size measure. Anticipated by this model and consistent with Hong, Lim, and Stein (2000), the relationship between momentum profits and size follows an inverse U-shape within $\sigma_{\alpha(t)}$ groups, while the monotonic increasing relationship between momentum return and $\sigma_{\alpha(t)}$ remains within each size group. The relationship is intuitively illustrated in Panels C and D of Figure 4. [Insert Table 7 here.]

The dynamic CAPM model suggests that more deviations from the conditional CAPM model lead to more momentum profits. But such deviations could also come from mispricing or missing factors. Table 8 explores whether the relationship between momentum and $\sigma_{\alpha(t)}$ will be absorbed by two mispricing measures. The first proxy for mismeasurement is $1 - R^2_{CAPM}$. The intuition is that, if $R^2_{CAPM} = 1$, then there should be no momentum. Presumably, if momentum is due to

[32]Both the magnitude of momentum profits and the relationship between momentum and MB indicates that Group 0 firms are average firms. Therefore, the Kalman technique does not fail for a particular group of firms.

mispricing, we expect to see more momentum profitability from low R^2_{CAPM} firms. The second proxy for missing factors is the standard deviation of the Kalman $\epsilon-$ residual. In Panels A and B, firms are independently sorted into four $\sigma_{\alpha(t)}$ quartiles and three groups based on $1 - R^2_{CAPM}$ and $\sigma_{\epsilon(t)}$, respectively. In Panel A, the momentum profitability does not decrease according to the explanatory power of the unconditional CAPM model. On the contrary, firms better described by the unconditional CAPM are able to generate more momentum profits. This result is surprising but does no harm to the Kalman model. In Panel B, however, both $\sigma_{\alpha(t)}$ and the mispricing measure correlate with momentum profitability. The current table, together with the results in Table 3, suggest that the empirical model properly identifies important sources of momentum. At the same time, there could be some momentum-related features that are outside the model and deserve future research. [Insert Table 8 here.]

Although this study mainly examines the in-sample explanatory power of the dynamic CAPM model, Table 9 provides an out-of-sample version of Table 5 to complete the analysis. Instead of estimating the Kalman model once for all available return data during the test period, this table estimates the Kalman model year by year for each stock, dividing stocks into different quartiles according to ex-ante $\sigma_{\alpha(t)}$ in each month. In order to get the ex-ante $\sigma_{\alpha(t)}$ measure in year t, I apply the Kalman model to all available return data after January 1962 and before January of year t (if there are more than 60 data points). Then, the standard deviation of the estimated risk premium ($\sigma_{\alpha(t)}$, based on prior-t information) is assigned to each month in year t as the ex-ante measure for firm dynamics. Next, in each month, firms are sorted into four quartiles according to the ex-ante $\sigma_{\alpha(t)}$ measure (Groups 0 to 3 [more dynamic firms]), and momentum strategies are applied to stocks within each of the four quartiles. Since more than one ranking period is involved for the overlapping-portfolio strategies in each month, stocks are sorted according to available $\sigma_{\alpha(t)}$ values at the end of the first (and the earliest) ranking period. Panel A reports the mean momentum return generated by Group 1 to 3 (Group 0 stocks are discarded as discussed). The return pattern is very close to that reported in Table 5: momentum profitability monotonically increases in firm dynamics. The WML:12:1:3 strategy, for example, generates a monthly momentum return of 1.5% based on most dynamic stocks and 0.87% for less dynamic stocks. Panel B further studies the momentum return difference between Group 3 and Group 1. Compared to Table 5, the return spreads here appear to be less significant. Nevertheless, for most combinations of ranking and holding-period lengths, the spreads are significant at the 5% level. The economic magnitudes of these spreads, ranging from 3% to 7% a year, are hardly negligible. Therefore, the theoretical model and empirical method provided here not only can explain a large part of momentum profitability in-sample, but also predict the group of stocks that are likely to generate more momentum returns out-of-sample. Tables 9 through 5 also provide direct evidence that, econometricians can properly anticipate some firms to generate more momentum profits based on a proper estimation of firm-specific risks. In other words, part of momentum profits are compensation for firm-level risks.

Overall, the evidence provided by this section indicates that a large fraction of momentum profits can be associated with firm-level risks in ways that the dynamic CAPM model predicts. Even a conservative interpretation of these tables confirms that the model captures several important properties of momentum. Allowing the traditional CAPM model to incorporate firms' decision-making process, we may move closer to the economic foundation of momentum profitability.

4 Related Literature

Well-established behavioral hypotheses that explain momentum returns include underreaction (Hong and Stein (1999)), representativeness and conservatism (Barberis, Shleifer, and Vishny (1998)), and overreaction and self-attribution (Daniel, Hirshleifer, and Subrahmanyam (1998)). More recently, Grinblatt and Bin (2002) utilize the existence of a disposition effect to explain momentum. There have also been continual efforts to understand momentum without resorting to bounded rationality, but some rational voices are not without dispute. Conrad and Kaul (1998), for example, attribute momentum profits to the cross-sectional dispersion of unconditional means of stock returns, but Jegadeesh and Titman (2002) illustrate that Conrad and Kaul's conclusion is premature because of the small sample biases of their tests and bootstrap experiments. Consistent with the APT framework, Moskowitz and Grinblatt (1999) conjecture that unpriced industry factors may lead to momentum, yet Grundy and Martin (2001) imply that stock and industry momentum components are different. In his recent examination of auto- and cross-serial correlation for size and book-to-market equity portfolios, Lewellen (2002) suggests that stocks covary very strongly and that it is the excess covariance, not underreaction to firm-specific news, that explains portfolio momentum. Chen and Hong (2002), however, demonstrate that Lewellen's empirical results do not necessarily refute the underreaction hypothesis. Chordia and Shivakumar (2002) investigate whether macro variables known to predict the business cycle can explain momentum. They show that momentum profitability is no longer significant when these macro variables are controlled. Their conclusion, however, is not supported by Griffin, Ji, and Martin (2002), who use data from both international and U.S. stock markets. Recently, Ang, Chen, and Xing (2002) demonstrate that momentum may be linked to a systematic factor of downside correlation. Ahn, Conrad, and Dittmar (2003) use a stochastic discount factor to explain half of momentum strategy profits. Their research adds to the rational explanations of momentum but differs from the traditional structural asset pricing approaches that this article adopts.

Besides the momentum literature, this work is also related to recent efforts attempting to explain size and book-to-market anomalies within a rational framework. Gomes, Kogan, and Zhang (2002), for example, extend the Berk, Green, and Naik (1999) intuition to a general equilibrium model. They show that firm characteristics such as size and book-to-market ratio will be correlated to the conditional CAPM betas and help explain the expected stock returns. Brennan, Wang, and Xia (2002) take inflation risk and the Sharpe ratio into consideration and conclude that the Fama-French mimicking portfolios carry information about latent investment opportunities. Berk (1995) provides a rational framework for understanding the size effect. Xing (2002), Zhang (2002), and Adrian and Franzoni (2002), among others, attempt to rationally explain the value premium. This article, however, starts from the intuition of economies of scale and uniquely proposes that the size, value, and momentum anomalies may all stem from the asset pricing impact of firm-level risks.

This model is essentially conditional. Empirically conditional models are arguably more successful than their unconditional counterparts (see, for example, Harvey (1989), Ferson and Harvey (1991), and a recent work of Lettau and Ludvigson (2001)). This paper differs from the conditional CAPM literature by pointing out that firm-specific information might be an important determinant of firm returns (Daniel and Titman (1997)). The empirical model proposed here does not rely on predetermined instruments (one purpose of this study is to show that the information set of an unconditional CAPM suffices to explain a large part of momentum profits), though it is quite easy to incorporate them into the two components of risk. In a similar way this paper differs from existing nonlinear CAPM researches, such as Harvey and Siddique (2000). The study of firm-specific components of risks is not new in the literature (Campbell, Lettau, Malkiel, and Xu (2001); Malkiel and Xu (2001); and Goyal and Santa-Clara (2003)). The novel feature of this study is that firm-specific risk is priced in the equilibrium even when investors hold well-diversified portfolios.

5 Conclusion

This article proposes a simple model in which a firm achieves economies of scale by dynamically adjusting its business according to anticipated firm-level business risks. Investors view the dynamic behavior of a firm as a beta risk, in addition to the market risk. Firms' dynamic investment policy ultimately leads to a dynamic beta and a nonlinear risk premium in the equilibrium.[33] This risk premium is not captured by traditional linear asset pricing models and will bring many interesting phenomena into the economy. Momentum, for example, can be explained mostly by the part of the risk premium ignored by traditional linear asset pricing models. Other momentum phenomena consistent with this story include industry momentum, Fama-French portfolio momentum, and earnings momentum.

This article reconciles several CAPM anomalies by pointing out that different anomalies capture different aspects of the asset pricing impact of firms' investment policies. Firm characteristics can explain time series and cross-sectional return variation because they are correlated with predicted investment opportunities. From this point of view, the two seemingly irrelevant phenomena, momentum and firm characteristics, may indeed be deeply related. Empirical findings about the relationship between the two support this interpretation.

On this evidence, it may appear premature to equal the existence of momentum and other asset pricing anomalies to irrationality and market inefficiency, though achievements from the behavioral literature are both necessary and promising. Momentum, for example, is likely to contain both rational and irrational components. To the extent that models usually provide a simplified description of the real world, this article suggests that CAPM and a possible theory for momentum might belong to different levels of approximations for real corporate business activities. This work therefore provides incentives to study the asset pricing impact of more complex corporate decision-making processes, in order to understand the power and limitation of traditional asset pricing approaches.

[33] The existence of such conditional and nonlinear risk premium itself is interesting. For example, the mutual fund literature has a long tradition of separating the so-called selection ability from the timing ability, whereas linear asset pricing theories deny the existence of such abilities.

Appendix A

Proof. (Lemma 1) The first result follows directly from a constrained optimization. The second and the third results follow the first one trivially. When the cost function is quadratic, the first-order condition gives out that $m_{t+1} = (C_0 - A + lF_{t+1}) / 2B$. So m_{t+1} still follows AR(1) process. The dividend policy follows trivially. ■

The proof of Proposition 1 depends on the following lemma, which extends Whittle (1990) Lemma 6.1.2 to more general stochastic control problems with separable quadratic risks.

Lemma 2 *(Control Problem with Separable Quadratic Risks) Let* $\mathbf{Y}$ *follow a multivariate normal distribution,* $\mathbf{Y} \sim N(\mu, \mathbf{\Omega})$*, and* $\mathbf{z}$ *be a vector control variable. Let* $Q(\mathbf{Y}, \mathbf{z})$ *be a quadratic function of* $\mathbf{Y}$ *and* $g(\mathbf{z})$ *be an arbitrary function of* $\mathbf{z}$*. If the function* $f(\mathbf{Y}, \mathbf{z}) = g(\mathbf{z})Q(\mathbf{Y}, \mathbf{z}) + 1/2 (\mathbf{Y} - \mu)' \Omega^{-1} (\mathbf{Y} - \mu)$ *is positive definite, then*

$$\max_{\mathbf{z}} E_{\mathbf{Y}} \left[\exp(-g(\mathbf{z})Q(\mathbf{Y}, \mathbf{z})) \right] \propto \exp\left(- \left[\min_{\mathbf{z}} \min_{\mathbf{Y}} \; f(\mathbf{Y}, \mathbf{z}) + \frac{1}{2} \ln(|H(\mathbf{z})|) \right] \right), \tag{20}$$

where $H(\mathbf{z}) = \frac{\partial^2 f(\mathbf{Y}, \mathbf{z})}{\partial \mathbf{Y} \partial \mathbf{Y}'}$ *denotes the Hessian matrix of* $f(\mathbf{Y}, \mathbf{z})$ *with respect to* $\mathbf{Y}$.

Especially, when x*,* y *are two independent normal variables,* $x \sim N(\mu_x, \sigma_x^2)$ *and* $y \sim N(\mu_y, \sigma_y^2)$*,* $Q(x, y)$ *is a quadric function of* x *and* y *such that* $f(x, y, z) = zQ(x, y) + \frac{(x - \mu_x)^2}{2\sigma_x^2} + \frac{(y - \mu_y)^2}{2\sigma_y^2}$ *is positive definite, then*

$$\max_{z} E_y \left[\exp(-zQ(x, y) \right] \propto \exp\left(- \left[\min_{z} \min_{x,y} \; f(x, y, z) - \ln(G(z)) \right] \right), \tag{21}$$

where $G(z) = \left[\left(zQ_{xx} + 1/\sigma_x^2 \right) \left(zQ_{yy} + 1/\sigma_y^2 \right) - z^2 Q_{xy} Q_{yx} \right]^{-1/2}$.

Proof. (Lemma 2) It is straightforward to show that

$$\max_{\mathbf{z}} E_{\mathbf{Y}} \left[\exp(-g(\mathbf{z})Q(\mathbf{Y}, \mathbf{z},) \right] \propto \max_{\mathbf{z}} \int \exp\left(- f(\mathbf{Y}, \mathbf{z}) \right) d\mathbf{Y}.$$

Let $\mathbf{Y}^*$ satisfy the following first order condition:

$$\frac{\partial f(\mathbf{Y}^*, \mathbf{z})}{\partial \mathbf{Y}} = 0. \tag{22}$$

Since $f(\mathbf{Y}, \mathbf{z})$ is quadric in $\mathbf{Y}$, it can be Taylor expanded as follows:

$$f(\mathbf{Y}, \mathbf{z}) = f(\mathbf{Y}^*, \mathbf{z}) + \frac{1}{2}(\mathbf{Y} - \mathbf{Y}^*)' \frac{\partial^2 f(\mathbf{Y}, \mathbf{z})}{\partial \mathbf{Y} \partial \mathbf{Y}'} (\mathbf{Y} - \mathbf{Y}^*) = f(\mathbf{Y}^*, \mathbf{z}) + \frac{1}{2}(\mathbf{Y} - \mathbf{Y}^*)' H(\mathbf{z}) (\mathbf{Y} - \mathbf{Y}^*),$$

and the original maximization problem becomes

$$\begin{aligned} \max_{\mathbf{z}} \int \exp\left(- f(\mathbf{Y}, \mathbf{z}) \right) d\mathbf{Y} &= \max_{\mathbf{z}} \; \exp\left(-f(\mathbf{Y}^*, \mathbf{z}) \right) \int \exp(-\frac{1}{2}(\mathbf{Y} - \mathbf{Y}^*)' H(\mathbf{z}) (\mathbf{Y} - \mathbf{Y}^*)) d\mathbf{Y} \\ &\propto \max_{\mathbf{z}} \exp\left(-f(\mathbf{Y}^*, \mathbf{z}) \right) |H(\mathbf{z})|^{-1/2}. \end{aligned}$$

The last step follows since $H(\mathbf{z})$ and $\mathbf{Y}^*$ can be treated as constants when one is performing the integration. This result and the first-order condition (22) lead to (20). The second part of the lemma is a trivial application of the first one. ■

Proof. (Proposition 1) The budget constraint can be written as

$$W_{t+1} = (1+r)W_t + n(D_0 + m_{t+1}^2\gamma_F\sigma_x^2 + m_{t+1}X_{t+1} + P_{t+1} - (1+r)P_t). \tag{23}$$

In the equilibrium, $n^* = 1$. Now, conjecture the price under the dynamic system to be $P_t = A + Bb_{t+1} + Cb_{t+1}^2$. Note that $b_{t+1} = E_t[m_{t+1}]$ reflects available information on the firm dynamics, while $b_{t+2} = b_0 + \Gamma m_{t+1}$ and $P_{t+1} = A + Bb_0 + Cb_0^2 + (B\Gamma + 2C\Gamma b_0)m_{t+1} + C\Gamma^2 m_{t+1}^2$ contain random variable m_{t+1}. Using the terminology of Lemma 2, we have (ignoring W_t, the nonstochastic part of the wealth)

$$\begin{aligned} f(n, X_{t+1}, m_{t+1}) = \gamma_I n(D_0 + \gamma_F\sigma_x^2 m_{t+1}^2 + m_{t+1}X_{t+1} + A + B(b_0 + \Gamma m_{t+1}) \\ + C(b_0 + \Gamma m_{t+1})^2 - (1+r)(A + Bb_{t+1} + Cb_{t+1}^2)) + X_{t+1}^2/2\sigma_x^2 + (m_{t+1} - b_{t+1})^2/2\sigma_\beta^2. \end{aligned}$$

An investor's original control problem is solved by

$$Exm_{n,X,m} \quad f(n, X_{t+1}, m_{t+1}) \ - \ln(G(n)), \tag{24}$$

where $Exm_{n,X,m}$ means to take first-order derivatives of the three variables. The three FOCs are

$$\begin{aligned} (1+r)(A + Bb_{t+1} + Cb_{t+1}^2) &= D_0 + \gamma_F\sigma_x^2 m_{t+1}^2 + m_{t+1}X_{t+1} - (\ln G(n^*))'/\gamma \\ &\quad + A + B(b_0 + \Gamma m_{t+1}) + C\Gamma(b_0 + \Gamma m_{t+1})^2; \\ X_{t+1} = -n\gamma_I\sigma_x^2 m_{t+1};\ \ m_{t+1} &= \frac{b_{t+1} - \gamma_I\sigma_\beta^2(B\Gamma + 2C\Gamma b_0)}{1 - n^2\gamma_I\sigma_x^2\sigma_\beta^2(\gamma_I - 2\gamma_F) + 2\Gamma^2\gamma_I\sigma_\beta^2 C}. \end{aligned}$$

Plugging the last two FOCs into the first one and matching the coefficient for b_{t+1}, b_{t+1}^2 and a constant, we get three equations that A, B, and C should satisfy

$$\begin{aligned} C &= \frac{(\gamma_F - \gamma_I)\sigma_x^2 + C\Gamma^2}{Z^2(1+r)}; B = \frac{2C\Gamma b_0/Z + 4\gamma_I\sigma_\beta^2 C^2\Gamma(1+r)}{1 + r - \Gamma/Z + 2\gamma_I\sigma_\beta^2(1+r)C}; \\ A &= (D_0 + Bb_0 + Cb_0^2 + \gamma_I\sigma_\beta^2[(1+r)C\gamma_I\sigma_\beta^2 - 1/Z](B\Gamma + 2C\Gamma b_0)^2)/r - (\ln G(n^*))'/\gamma r, \end{aligned} \tag{25}$$

where $Z = 1 - n^2\gamma_I\sigma_x^2\sigma_\beta^2(\gamma_I - 2\gamma_F) + 2\Gamma^2\gamma_I\sigma_\beta^2 C$ and $b_0 = m_0$. These equations will solve out the parameters C, B, and A recursively. Since $[\ln G(n^*)]_n'$ is a constant in this case, it enters the constant part of the price and will not have any impact on the main results. To save space, I omit the analytical expression for this term. It is worthwhile, however, to point out that, when $\sigma_x^2 \ll 1$ or $\sigma_\beta^2 \ll 1$, $G(n)$ will almost be a constant, since the integration is not sensitive to n. In this case we can totally ignore the $[\ln G(n^*)]_n'$ term. Finally, if investors know m_{t+1} at the beginning of period $t+1$, then denoting $b_{t+1} = E_t[m_{t+2}]$ and conjecturing $P_t = A + Bb_{t+1} + Cb_{t+1}^2$, one can easily calculate A, B, and C by repeating the above steps. So the functional form of the price function will not be affected by the information policy of the firm. ■

Proof. (Proposition 2) Notice that $P_{t+1} = A + Bb_0 + Cb_0^2 + (B\Gamma + 2C\Gamma b_0)m_{t+1} + C\Gamma^2 m_{t+1}^2$, let $\varepsilon_{\beta,t+1} = m_{t+1} - b_{t+1}$, and plug them into the return process, we get

$$\begin{aligned} r_{t+1} &= D_{t+1} + P_{t+1} - (1+r)P_t \\ &= D_0 + m_{t+1}X_{t+1} + m_{t+1}^2\gamma_F\sigma_x^2 + A + Bb_0 + Cb_0^2 + (B\Gamma + 2C\Gamma b_0)m_{t+1} + C\Gamma^2 m_{t+1}^2 \\ &\quad - (1+r)(A + Bb_{t+1} + Cb_{t+1}^2). \end{aligned}$$

Reorganizing the terms gives out

$$
\begin{aligned}
&H_1 = (1+r)C(Z^2-1) + \gamma_I \sigma_x^2; \;\; H_2 = (1+r)B(Z-1) + 2(1+r)\gamma_I \sigma_\beta^2 (B\Gamma + 2C\Gamma b_0)C; \\
&H_3 = 2[(1+r)CZ^2 + \gamma_I \sigma_x^2]; \qquad H_4 = (1+r)CZ^2 + \gamma_I \sigma_x^2; \\
&H_5 = B\Gamma + 2C\Gamma b_0; \quad H_6 = \gamma_I \sigma_\beta^2((1+r)C\gamma_I \sigma_\beta^2 - 1/Z)(B\Gamma + 2C\Gamma b_0)^2 - (\ln G(n^*))'/\gamma.
\end{aligned} \tag{26}
$$

Easy to check, when σ_β^2 is very small, $Z \approx 1$, $H_1 \approx \gamma_I \sigma_x^2$, and $H_6 \approx 0$. ■

Proof. (Proposition 3) For the undiversified economy, individual stock returns could be written as

$$
r_{t+1}^i = m_{t+1}^i r_{t+1}^M + H_1(b_{i,t+1}^2 - \overline{b}_{t+1}^2) + H_2(b_{i,t+1} - \overline{b}_{t+1}) + H_4(\varepsilon_{\beta_i,t+1}^2 - \overline{\varepsilon}_{\beta_i,t+1}^2) + H_3 b_{t+1}\varepsilon_{\beta_i,t+1} + H_5 \varepsilon_{\beta_i,t+1}.
$$

The proposition follows when one takes the expectation of the above equation. The $\varepsilon_{\beta_i,t}^2$ terms drops because $E[\varepsilon_{\beta_i,t+1}^2|t] = \sigma_{\beta_i}^2$ and will be the same for all firms.

For the well-diversified economy, it suffices to show that asset returns can be describe by equation (12) (omit index i for each firm) and that H_1 and H_2 are nonzero parameters. Conjecturing the pricing formula to be $P_t^i = A^i + B^i b_{i,t+1} + C^i b_{i,t+1}^2$, and following Lemma 2, the investor now solves the following problem:

$$
\begin{aligned}
Exm_{n^i, X_{t+1}, m_{i,t+1}} \quad & \gamma_I \{(1+r)W_t + \sum_{i=1}^I n^i [D_0^i + \gamma_F \sigma_x^2 m_{i,t+1}^2 + m_{i,t+1} X_{t+1} \\
&+ A^i + B^i(b_0^i + \Gamma m_{i,t+1}) + C(b_0^i + \Gamma m_{i,t+1})^2 - (1+r)(A^i + B^i b_{i,t+1} + C^i b_{i,t+1}^2)]\} \\
&+ \frac{X_{t+1}^2}{2\sigma_x^2} + \sum_{i=1}^I \frac{(m_{i,t+1} - b_{i,t+1})^2}{2\sigma_\beta^2} - \ln(G(n_1, ...n_I)).
\end{aligned} \tag{27}
$$

Since the portfolio is well diversified, assume $\sum_{i=1}^I n^i m_{i,t+1} = Ib_0$. Now the FOCs are ($n^{i*} = 1$):

$$
\begin{aligned}
\sum_{i=1}^I (1+r)(A^i + B^i b_{i,t+1} + C^i b_{i,t+1}^2) = \sum_{i=1}^I [&D_0^i + \gamma_F \sigma_x^2 m_{i,t+1}^2 + m_{i,t+1} X_{t+1} + A^i \\
&+ B^i(b_0^i + \Gamma m_{i,t+1}) + C(b_0^i + \Gamma m_{i,t+1})^2] - (\ln G(n^*))_i'/\gamma;
\end{aligned} \tag{28}
$$

$$
X_{t+1} = -\gamma_I \sigma_x^2 \sum_{i=1}^I m_{i,t+1} = -\gamma_I \sigma_x^2 I b_0; \tag{29}
$$

$$
m_{i,t+1} = \frac{b_{i,t+1} - \gamma_I \sigma_\beta^2 (B^i \Gamma + 2C^i \Gamma b_0)}{1 + \gamma_I \sigma_\beta^2 (2C^i \Gamma^2 + 2\gamma_F \sigma_x^2) - \gamma_I \sigma_x^2)}. \tag{30}
$$

In deriving the envelop condition for a single stock, the assumption is $X_{t+1} = -\gamma_I \sigma_x^2 \sum_{i=1}^I m_{i,t+1} \approx -\gamma_I \sigma_x^2 (m_{i,t+1} + (I-1)b_0) \approx -\gamma_I \sigma_x^2 (m_{i,t+1} + Ib_0)$, since all other stocks could be viewed as given. Now plugging the last two FOCs into the first one, strictly assuming $X_{t+1} = -\gamma_I \sigma_x^2 I b_0$ (the investor holds all the stocks), matching terms for the i^{th} asset for the constant term, the first- and second-order terms of $b_{i,t}$, and denoting $Z = 1 + \gamma_I \sigma_\beta^2 (2C^i \Gamma^2 + 2\gamma_F \sigma_x^2 - \gamma_I \sigma_x^2)$ and omitting the index i, we get

$$
\begin{aligned}
Z^2 C = \frac{\gamma_F \sigma_x^2 + C\Gamma^2}{1+r}; B = \frac{2C\Gamma b_0(\Gamma/Z - 2\gamma_I \sigma_\beta^2 (1+r)C) - \gamma_I \sigma_x^2 I b_0/Z}{1 + r - \Gamma/Z + 2\gamma_I \sigma_\beta^2 (1+r)C}; \\
rA = D_0 + Bb_0 + Cb_0^2 + \gamma_I \sigma_\beta^2 [(1+r)C\gamma_I \sigma_\beta^2 - 1/Z](B\Gamma + 2C\Gamma b_0)^2 \\
+ \gamma_I \sigma_x^2 I b_0 (B\Gamma + 2C\Gamma b_0) - (\ln G(n^*))_i'/\gamma.
\end{aligned} \tag{31}
$$

As seen before, the parameters can be solved in a sequence of C, B and A. Again I omit the analytical expression for $\ln G(n^*)$. It suffices to know that $(\ln G(n^*))'_i$ is a constant for each i.

As a result, return processes are described by equation (12). Furthermore, $H_1 = (\gamma_F \sigma_x^2 + C\Gamma^2 - (1+r)C) = (1+r)C(Z^2-1) = \gamma_F \sigma_x^2 (Z^2-1)/(Z^2 - \Gamma^2/(1+r))$ and $H_2 = B\Gamma + 2C\Gamma b_0 - (1+r)B = (1+r)B(Z-1) + 2\gamma_I \sigma_\beta^2 (B\Gamma + 2C\Gamma b_0)CZ + \gamma_I \sigma_x^2 I b_0$ and both are nonzero. Therefore there exists Fama-French residual momentum effect. When $1 > Z^2 > \Gamma^2/(1+r)$, $C > 0$ and $H_1 < 0$. When $Z^2 > 1 > \Gamma^2/(1+r)$ or $Z^2 < \Gamma^2/(1+r) < 1$, both C and H_1 are positive. The magnitude of H_1 in the diversified economy ($\sim \gamma_F \sigma_x^2$) can have the same order as that in the undiversified economy. Next, if the firm reveals m_{t+1} immediately at the beginning of the period $t+1$, then one can denote $b_{i,t+1} = E_t[m_{i,t+2}]$ and conjecture $P_t^i = A' + B' b_{i,t+1} + C' b_{i,t+1}^2$ as the equilibrium price. The parameters are generally nonzero. For example, one can show that $C' = \frac{\gamma_F \sigma_x^2}{\Gamma^2(1+r-\Gamma^2/(1+2C\Gamma^2\gamma_I\sigma_\beta^2)^2)} > 0$ and $H_1' = \frac{\gamma_F \sigma_x^2}{\Gamma^2}(1 - \frac{1+r-\Gamma^2}{1+r-\Gamma^2/(1+2C\Gamma^2\gamma_I\sigma_\beta^2)^2}) > 0$, where H_1' is the counterpart of H_1 parameter when there is no information lag. As a result, the functional form of the price function will not be affected by the information policy of the firm.

■

Proof. (Characteristic-based momentum)

1. The relationship between momentum and MB:

Expected momentum profits can be written as $E_t[\pi(R)] = \frac{1}{I}\sum_{i\in I}(r_t^i - \overline{r}_t)E_t[r_{t+1}^i]$. For simplicity, denoting $\overline{b}_t = \overline{b}$, $\overline{MB}_t = \overline{MB}$, and $A' = A - D_0/r$, we can Taylor expand MB around $\overline{b}$ as $MB_t^i = A' + Bb_{i,t} + Cb_{i,t}^2 = A' + B\overline{b} + C\overline{b}^2 + (2C\overline{b}+B)(b_{i,t}-\overline{b}) + C(b_{i,t}-\overline{b})^2$. From equation (16), one can easily get that $E_t[\pi(R)] \propto \frac{1}{I}\sum_{i\in I}[(H_1/C)^2(MB_t^i - \overline{MB})E_t[MB_{t+1}^i - \overline{MB}] + (H_2 - H_1 B/C)^2(b_{i,t}-\overline{b})E_t[b_{i,t+1}-\overline{b}]]$, and that $E_t[\pi(R)] \propto \frac{1}{I}\sum_{i\in I}[G_1(MB_t^i - \overline{MB})^2 + G_2 MB_t^i]$, where $G_1 = (H_1/C)^2 corr(MB_t^i, MB_{t+1}^i) > 0$ (because MB are positively correlated), $G_2 = (H_2 - H_1 B/C)^2\Gamma/C > 0$, and I omit the interaction between MB and $b_{i,t}$ since the model cannot sign this term. Since $G_2 > 0$, momentum profitability is in general increasing in MB. Furthermore, depending on the parameter values, there could be a second U-shaped relationship between momentum and MB: when MB is very low, momentum profits could increase because of the $G_1(MB_t^i - \overline{MB})^2$ component. In other words, expected momentum profits can be viewed as a convex quadratic function of MB. Since $G_1 > 0$, $\frac{\partial E_t[\pi(R)]}{\partial MB_t^i}|_{\overline{MB}} > 0$, and we refine $MB_t^i > 0$, momentum can either monotonically increase in MB (when $\overline{MB} - G_2/G_1 < 0$) or first decrease and then increase in MB (U-shape, when $\overline{MB} - G_2/G_1 > 0$). But even in the latter case momentum will increase in MB for most MB values.

2. The relationship between momentum and size:

The model will first predict an increasing relationship between momentum and size, because size and MB have similar roles in the economy. However, if size has other effects, the relationship will also be modified. For example, if small firms on average have more information asymmetry, then to investors the beta uncertainty will have higher variance (it is more difficult to predict the firm behavior). In this case, size negatively correlates with $\varepsilon_{\beta_i,t}^2$. Intuitively investors will ask for more return for more beta uncertainty (the H_4 component in [27] can be viewed as an information premium; easy to see $H_4 > 0$). Hence from equation (27) there will be an additional component of momentum profitability due to information premium: $E_t[\pi(R)] \propto \frac{1}{I}\sum_{i\in I}(\varepsilon_{\beta_i,t}^2 - \overline{\varepsilon}_{\beta_i,t}^2)E_t[r_{t+1}^i]$. Since $\varepsilon_{\beta_i,t}^2 - \overline{\varepsilon}_{\beta_i,t}^2$ is largely independent of $E_t[r_{t+1}^i]$, larger $\varepsilon^2_{\beta_i,t}$ (small firms) can generate more momentum profits, so long as $E_t[r^i_{t+1}] > 0$. Intuitively, for small firms, a positive information premium in one period and a positive overall return in the next period introduce a positive return autocorrelation, which contributes to momentum (the size premium and information premium in the next period return only exaggerate the autocorrelation). The combination of the two relationships can produce an overall inverse-U relationship between momentum and size, as observed by Hong, Lim, and Stein (2001).

■

Appendix B: Simulation Results

To better illustrate how well the Kalman filter estimates momentum profits, I conduct two stages of simulations. In Figure 1, a hypothetical economy is simulated for 100 times. This economy contains 30 firms, whose returns are determined by equations (19). More specifically, X_t is taken to be the realized monthly market excess return from January 1965 to December 1999. For i^{th} firm, $\alpha^i_0 = 0$, $\beta^i_0 = 1$, $\Gamma^i = 0.8$, $H^i_1 = 0.5$, $H^i_2 = 0.8$, $\epsilon^i_t \sim N(0, 0.03^2)$, and $\varepsilon^I_{\beta,t} \sim N(0, 0.1^2)$. These numbers are just illustrative. One can obtain similar simulation results for a wide range of parameter values. For the current stage of simulation, a random realization of β^i_t and α^i_t, according to the second and third equation of (19), is used for all simulations. Therefore, the latent factor of each stock, as well as the $\pi(1)$ and $\pi(2)$ components of WRSS101 momentum profits, are fixed for all 100 simulations. The cross-sectional mean of the β^i_t and α^i_t are plotted by the solid lines in Panels C and D in Figure 1, while the time series of $\pi(1)$ and $\pi(2)$ are plotted by the solid lines in Panels A and B in Figure 1. For each simulation, ϵ^i_t is randomly generated; therefore, the return process will differ according to the noise term.

For each simulation, the Kalman filter observes the market return and the simulated return processes of each stock, and estimates the β^i_t and α^i_t time series. Panels C and D plot the 10% and 90% confidence levels of estimated cross-sectional mean alphas and betas as the shadow regions, illustrating that under the current conditions the Kalman filter generally gives out very close and unbiased estimates of dynamic alphas and betas for individual firms. For individual firms similar results hold. The Kalman filter might smooth the time series of the latent factor and induce a downward bias in the estimated autocovariance and momentum profits. Panels A and B first calculate $\pi(1)$ and $\pi(2)$ from estimated alphas and betas for each simulation, and then plot the 10% and 90% distribution of the two estimated components. Panel A shows that in general $\pi(1)$ is underestimated under the null, while Panel B illustrates that, when the market factor itself does not bear momentum, the estimation of $\pi(2)$ is smoothed but unbiased. However, $\pi(1)$ estimations convincingly track the time series fluctuations of the true value. Therefore, we can rely on the Kalman filter to capture the time series property, as well as a majority magnitude, of momentum profits.

Figure 2 generalizes the previous experiment by randomizing the β^i_t and α^i_t time series and bootstrapping the market return (without replacement) for each simulation. The cross-sectional means of true alphas and betas, $\overline{\alpha}_t$ and $\overline{\beta}_t$, are different across simulations. Panels C and D report the cross-simulation mean (solid line) of $\overline{\alpha}_t$ and $\overline{\beta}_t$, the cross-simulation mean (dashed line) and the standard deviation (dotted line) of estimation error of $\overline{\alpha}_t$ and $\overline{\beta}_t$ ($\overline{\alpha}_t(Kal) - \overline{\alpha}_t(True)$ and $\overline{\beta}_t(Kal) - \overline{\beta}_t(Kal)$). It remains plausible that the Kalman estimations for latent factors are unbiased. In Panel A1, the solid lines plot the 10%

and 90% levels of true momentum profit distribution (across simulations) at any given time. The shadow area plots the same region for the Kalman estimations. Panel A2 plots the cross-simulation mean of $\pi(1)$ and the cross-simulation mean and the standard deviation of the estimation error $(\pi_{Kal}(1) - \pi(1))$. The gap between the two distributions and the negative value of the mean estimation error again indicate that smoothed estimations of latent factors will lead to a small downward bias in the estimation of $\pi(1)$, while the cross-simulation mean and standard deviation of estimation errors indicate that the Kalman model still tracks the time variations of $\pi(1)$ well. Panels B1 and B2 plot the same distributions for $\pi(2)$. The second component of momentum profits is generally estimated without significant bias. This stage of simulation illustrates that the first-stage simulation result is quite robust for different realizations of risks.

References

Adrian, Tobias and Francesco Franzoni, 2002, Learning About Beta: An Explanation of the Value Premium, Working Paper, MIT.

Ahn, Dong-Hyun, Jennifer Conrad and Robert F. Dittmar, 2003, Risk Adjustment and Trading Strategies, Review of Financial Studies 16, 459-485.

Ang, A., J. Chen and Y. Xing, Downside Correlation and Expected Stock Returns, Working paper, Columbia University.

Barberis, Nicholas, Andrei Shleifer, and Robert Vishny, 1998, A Model of Investor Sentiment, Journal of Financial Economics 49, 307-43.

Bell, L., and T. Jenkinson, 2002, New Evidence of the Impact of Dividend Taxation and on the Identity of the Marginal Investor, Journal of Finance 57, 1321-1345.

Blanchard, O., C. Rhee and L. Summers, 1993, The Stock Market, Profit, and Investment, The Quarterly Journal of Economics 108-1, 115-136.

Berk, J., 1995, A Critique of Size Related Anomalies, Review of Financial Studies 8, 275-286.

Berk, Jonathan B., Richard C. Green, and Vasant Naik, 1999, Optimal Investment, Growth Options, and Security Returns, Journal of Finance 54, 1553-1608.

Brennan, Michael J., Ashley W.Wang, and Yihong Xia, 2002, A Simple Model of Intertemporal Capital Asset Pricing and its Implications for the Fama-French Three-factor Model, Working paper.

Campbell, John Y., Martin Lettau, Burton G. Malkiel, and Yexiao Xu, 2001, Have Individual Stocks Become More Volatile? An Empirical Exploration of Idiosyncratic Risk, Journal of Finance 56, 1-43.

Chan, Louis K. C., Narasimhan Jegadeesh, and Josef Lakonishok, 1996, Momentum Strategies, Journal of Finance 51, 1681-1713.

Chen, Joseph, and Harrison Hong, 2002, Discussion of 'Momentum and Autocorrelation in Stock Returns', Review of Financial Studies 15, 565-573.

Conrad, Jennifer, and Gautam Kaul, 1998, An Anatomy of Trading Strategies, Review of Financial Studies 11, 489-519.

Chordia, Tarun, and Lakshmanan Shivakumar, 2002, Momentum, Business Cycle, and Time-varying Expected returns, Journal of Finance 57, 985-1020.

Cooper, J., Roberto C. Gutierrez JR, and Allaudeen Hameed, 2004, Market States and Momentum, Journal of Finance, forthcoming.

Daniel, Kent, David Hirshleifer, and Avanidhar Subrahmanyam, 1998, Investor Psychology and Security Market Under- and Overreactions, Journal of Finance 53, 1839-86.

Daniel, K. and S. Titman, 1997, Evidence on the Characteristics of Cross-Sectional Variation in Stock Returns, Journal of Finance 52, 1-33.

Daniel, K. and S., Titman, 1999, Market Efficiency in an Irrational World, Financial Analysts' Journal 55-6, 28-40.

Fama, Eugene F., and K. R. French, 1992, The Cross-Section of Expected Stock Returns, Journal of Finance 47, 427-65.

Fama, Eugene F., and K. R. French, 1993, Common Risk Factors in the Returns on Stock and Bonds, Journal of Financial Economics 33, 3-56.

Fama, Eugene F., and K. R. French, 1995, Size and Book-to-Market Factors in Earnings and Returns, Journal of Finance 50, 131-155.

Fama, Eugene F., and K. R. French, 1996, Multifactor Explanations of Asset Pricing Anomalies, Journal of Finance 51, 55-84.

Fama, Eugene F., and James Macbeth, 1973, Risk, Return and Equilibrium: Empirical Tests, Journal of Political Economy 81, 607-36.

Ferson,Wayne E., and Campbell R. Harvey, 1991, The Variation of Economic Risk Premiums, Journal of Political Economy 99, 385-415.

Geroge, Thomas J., and Chuan-Yang Hwang, 2004, The 52-week High and Momentum Investing, Journal of Finance, forthcoming.

Gomes, J. F., L. Kogan, and L. Zhang, 2001, Equilibrium Cross-Section of Returns, Journal of Political Economy, forthcoming.

Goyal, Amit and Pedro Santa-Clara, 2003, Idiosyncratic Risk Matters! Journal of Finance 58, 975-1008.

Grinblatt, Mark, and Bing Han, 2002, Disposition Effect and Momentum Profit, Working paper.

Griffin, John, Susan Ji, and Spencer Martin, 2003, Momentum Investing and Business Cycle Risk: Evidence from Pole to Pole, Journal of Finance 58-6, 2515-2547.

Grundy, B., and J. S. Martin, 2001, Understanding the Nature of the Risks and the Source of the Rewards to Momentum Investing, Review of Financial Studies 14, 29-78.

Hamilton, J., 1994, Time Series Analysis (Princeton University Press, Princeton, New Jersey).

Harvey, A., 1989, Forecasting, Structural Time Series Models and the Kalman Filter (Cambridge University Press, Cambridge, U.K.).

Harvey, Campbell R. and Akhtar Siddique, 2000, Conditional Skewness in Asset Pricing Tests, Journal of Finance 55, 1263-1295.

Hong, Harrison, Terence Lim, and Jeremy C. Stein, 2000, Bad News Travels Slowly: Size, Analyst Coverage, and the Profitability of Momentum Strategies, Journal of Finance 55, 265-95.

Hong, Harrison, and Jeremy C. Stein, 1999, A united theory of underreaction, momentum trading and overreaction in asset markets, Journal of Finance 54, 2143-84.

Hou, Kewei, 1999, Industry Information Diffusion and the Lead-Lag Effect in Stock Returns, working paper, Ohio State University.

Jegadeesh, Narasimhan, and Sheridan Titman, 1993, Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency, Journal of Finance 48, 65-91.

Jegadeesh, Narasimhan, and Sheridan Titman, 2001, Profitability of Momentum Strategies: an Evaluation of Alternative Explanations, Journal of Finance 56, 00-01.

Jegadeesh, Narasimhan, and Sheridan Titman, 2002, Cross-Sectional and Time-Series Determinants of Momentum Returns, The Review of Financial Studies 15, 143-157.

Johnson, Timothy C., 2002, Rational momentum effects, Journal of Finance 57, 585-608.

Korajczyk, Robert, and Ronnie Sadka, 2004, Are Momentum Profits Robust to Trading Costs? Journal of Finance, forthcoming.

Lettau, Martin, and Sydney Ludvigson, 2001, Resurrecting the (C)CAPM: A CrossSectional Test When Risk Premia Are Time-Varying , Journal of Political Economy 109(6), 1238-1287.

Lewellen, Jonathan, 2002, Momentum and Autocorrelation in Stock Returns, The Review of Financial Studies 15, 533-563.

Lewellen, Jonathan, and Jay Shanken, 2002, Learning, Asset-Pricing Tests, and Market Efficiency, Journal of Finance 57, 1113-1146.

Malkiel, Burton G. and Yexiao Xu, 2002, Idiosyncratic Risk and Security Returns, Working Paper.

Mehra, Rajnish, and Edward Prescott, 1985, The equity premium: A puzzle, Journal of Monetary Economics 15, 145-161.

Merton, R. C., 1973, An Intertemporal Capital Asset Pricing Model, Econometrica 41, 867-88.

Moskowitz, Tobias J., and Mark Grinblatt, 1999, Do Industries Explain Momentum? Journal of Finance 54, 1249-90.

Newey, Whitney K., and Kenneth D. West, 1987, A Simple Positive Semi-definite, Heteroscedasticity and Autocorrelation Consistent Covariance Matrix, Econometrica 55, 703-08.

Ross, Stephen, 1976, The Arbitrage Theory of Capital Asset Pricing, Journal of Economic Theory 13(December): 341-360.

Rouwenhorst, K. Geert, 1998, International Momentum Strategies, Journal of Finance 53, 267-84.

Rouwenhorst, K. Geert, 1999, Local Return Factors and Turnover in Emerging Stock Markets, Journal of Finance 54, 1439-64.

Sagi, Jacob S. and Mark S. Seasholes, 2001, Firm-Level Momentum: Theory and Evidence, NBER working paper.

Whittle, Peter, 1990, Risk-Sensitive Optimal Control (John Wiley & Sons, New York).

Xing, Y., 2002, Firm Investment and Expect Equity Returns, Columbia University, working paper.

Zhang, L. 2002, The Value Premium, Rochester University, working paper.

Table 1: **Summary Statistics.** Panel A reports the cross-sectional mean and standard deviation of several summary statistics for monthly returns of pooled NYEX-AMEX firms (ALL) and S&P index firms (SP). To be included in the sample, an asset must have at least 60 months of return data from January 1962 to December 2000. If a firm has two valid return periods separated by missing data, then it will be counted as two firms in the pool, one for each period. Each S&P firm must once be a member of the S&P 500 index during the period from 1975 to 1999. Column 1 (Number) reports the number of valid firms in the two pools that have converged Kalman estimation ($|\alpha_t| < 2, |\beta_t| < 2$ and σ_β will not hit the upper bound of 10. More than 100 firms are discarded.). Columns 2 and 3 list the mean excess return and the Sharpe ratio for available firm return data during the period from January 1962 to December 2000. The R^2_{CAPM} and R^2_{FF3} columns report the OLS R^2 measure when the CAPM model and the three-factor Fama and French (1993) model are applied to these firms, based on available return data during the period. The last two columns list the effective R^2 and the Harvey revised R^2_D measure when the Kalman model is applied to these firms. Panel B lists the cross-sectional means and standard deviations of the estimated Kalman parameters.

A. Summary Statistics

		Number	Mean	Sharpe	R^2_{CAPM}	R^2_{FF3}	R^2_{KAL}	$R^2_{D,KAL}$
ALL	mean	6363	0.0072	0.0669	0.1670	0.2339	0.3819	0.5852
	std		0.0113	0.0822	0.1102	0.1204	0.2840	0.0673
SP	mean	952	0.0110	0.1035	0.2146	0.2930	0.3805	0.6103
	std		0.0100	0.0657	0.1104	0.0875	0.2380	0.0617

B. Kalman Estimates

		α_0	β_0	Γ	σ^2_ϵ	σ^2_β	H_1	$atan(H_2)$
ALL	mean	-0.0005	1.0152	0.1979	0.0109	0.7562	0.0621	-0.0059
	std	0.2910	0.5258	0.3130	0.0124	1.3863	1.5180	0.5165
SP	mean	0.0049	1.0365	0.2118	0.0069	0.4846	0.0653	-0.0372
	std	0.0091	0.3883	0.3494	0.0059	0.9618	1.0381	0.6914

Table 2: **Momentum Profits of NYSE-AMEX Stocks.** This table examines stock momentum profitability over the horizon from 1965 to 1999. All NYSE and AMEX listed firms with at least 60 months of CRSP return data and converged Kalman estimations are used. Only stocks with valid formation- and holding-period returns are included in any momentum portfolio. Each panel reports monthly momentum portfolio return ($\pi(R)$), its corresponding time series t statistics (when overlapping portfolios are used), its corresponding Newey-West $H-1$ lags-adjusted t statistics (when nonoverlapping portfolios are used), and risk-adjusted momentum profits ($\pi'(\epsilon)$ and $\pi''(\epsilon)$) in percentage of original return momentum profits ($\eta' = \pi'(\epsilon)/\pi(R) \times 100\%$ and $\eta'' = \pi''(\epsilon)/\pi(R) \times 100\%$). For calculating $\pi'(\epsilon)$ and $\pi''(\epsilon)$, momentum portfolio weights are determined by the ranking-period raw returns and Kalman residuals, respectively, while holding-period returns are calculated from Kalman residuals. Panels A and B are based on WRSS and WML (long [short] top [bottom] 10% stocks) strategies, respectively. F, H, and G denote the length (in months) of the formation period, the holding period, and the gap between the two periods.

		G = 0			G = 1			
	F	H = 3	H=6	H=12	H=1	H = 3	H=6	H=12
A. WRSS strategies with G = 0								
$\pi(R)$	6	0.0039[2]	0.0056[1]	0.0053[1]	0.0063[1]	0.0068[1]	0.0077[1]	0.0052[1]
$\pi(R)$	12	0.0084[1]	0.0074[1]	0.0042[2]	0.0103[1]	0.0090[1]	0.0072[1]	0.0035[2]
T_{JT}	6	2.02	3.28	3.77	3.26	3.73	4.77	3.78
T_{JT}	12	4.15	3.90	2.51	5.07	4.60	3.91	2.10
T_{NW}	6	2.29	3.89	4.87	3.25	4.52	5.73	4.84
T_{NW}	12	5.11	5.29	3.37	5.04	5.80	5.32	2.69
η'	6	1.90	24.46	21.78	26.35	34.09[2]	37.99[1]	20.13
η'	12	26.96[2]	21.87	−14.57	35.16[1]	31.96[1]	20.77	−32.22
η''	6	10.12	29.69[1]	29.41[1]	29.04	37.21[1]	40.01[1]	29.06[1]
η''	12	29.53[1]	28.48[1]	4.39	34.64[1]	35.17[1]	28.53[1]	−7.27
B. WML with G = 0								
$\pi(R)$	6	0.0062[2]	0.0083[1]	0.0075[1]	0.0100[1]	0.0103[1]	0.0110[1]	0.0072[1]
$\pi(R)$	12	0.0114[1]	0.0103[1]	0.0059[1]	0.0142[1]	0.0123[1]	0.0101[1]	0.0049[2]
T_{JT}	6	2.53	3.83	4.22	4.11	4.46	5.45	4.22
T_{JT}	12	4.50	4.32	2.81	5.63	5.02	4.39	2.38
T_{NW}	6	2.96	4.61	5.37	4.14	5.57	6.62	5.39
T_{NW}	12	5.63	5.97	4.04	5.60	6.43	6.10	3.36
η'	6	19.80	48.66[1]	40.07[1]	54.37[2]	60.08[1]	62.89[1]	37.86[2]
η'	12	40.43[1]	37.85[2]	−9.83	50.73[1]	47.65[1]	36.60[2]	−30.97
η''	6	39.31	72.28[1]	63.08[1]	72.06[1]	82.80[1]	84.35[1]	62.32[1]
η''	12	61.23[1]	59.61[1]	13.11	67.12[1]	71.28[1]	57.68[1]	−5.32

1 and 2: Significant at the 1% and 5% levels, two-tailed t test.

Table 3: **Decomposition of NYSE-AMEX Stock Momentum Profits.** This table decomposes WRSS momentum strategy returns for NYSE-AMEX stocks over the period from 1965 to 1999 into four components: $\pi(R) = \pi(1) + \pi(2) + \pi(3) + \pi(4)$, where $\pi(1) = cov\,(\alpha_t, \alpha_{t-1})$, $\pi(2) = cov\,(\beta_t, \beta_{t-1})\, r_t^M r_{t-1}^M + cov\,(\alpha_t, \beta_{t-1})\, r_{t-1}^M + cov\,(\beta_t, \alpha_{t-1})\, r_t^M$, $\pi(3) = cov\,(\varepsilon_t, \beta_{t-1})\, r_{t-1}^M + cov\,(\beta_t, \varepsilon_{t-1})\, r_t^M + cov\,(\alpha_t, \varepsilon_{t-1}) + cov\,(\varepsilon_t, \alpha_{t-1})$, and $\pi(4) = cov\,(\varepsilon_t, \varepsilon_{t-1})$ for the dynamic CAPM model of $r_t = \alpha_t + \beta_t r_t^M + \epsilon_t$. All components will be scaled by a same constant so that the portfolio will have one dollar in the long position. Panels A and B report the percentage contribution of each component, $\eta_i = |\pi(i)|/(|\pi(1)| + |\pi(2)| + |\pi(3)| + |\pi(4)|) \times 100\%$, for returns from nonoverlapping momentum portfolios and overlapping momentum portfolios, respectively.

		G = 0			G = 1			
	F	H = 3	H=6	H=12	H=1	H = 3	H=6	H=12
A. WRSS strategies, Nonoverlapping Portfolios								
η_1	6	56.90[1]	70.86[1]	70.10[1]	71.53[1]	64.83[1]	62.47[1]	69.37[1]
η_1	12	72.77[1]	71.35[1]	57.08[1]	67.50[1]	66.92[1]	70.23[1]	53.96[1]
η_2	6	10.66[2]	11.03	7.27	12.30	10.74	7.48	7.30
η_2	12	7.50	7.44	6.87[2]	6.29	6.80	7.25	7.19[2]
η_3	6	16.47[1]	16.99[1]	20.21[1]	13.53[2]	11.03[2]	10.01[2]	20.46[1]
η_3	12	15.81[1]	18.59[1]	21.28[1]	11.10[2]	12.96[1]	18.38[1]	21.31[1]
η_4	6	15.98	1.12	2.41	2.64	13.40	20.05[1]	2.88
η_4	12	3.93	2.62	14.77[2]	15.11[2]	13.33[2]	4.14	17.54[1]
B. WRSS strategies, Overlapping Portfolios								
η_1	6	68.47[1]	61.73[1]	65.28[1]	60.64[1]	55.87[1]	56.16[1]	64.67[1]
η_1	12	64.52[1]	65.28[1]	73.99[1]	61.13[1]	60.66[1]	64.67[1]	68.44[1]
η_2	6	16.65	12.78	8.32	13.67	12.74	9.36	8.02
η_2	12	8.53	8.32	9.70	7.11	7.95	8.02	9.71
η_3	6	14.39	7.66	8.28	7.25	5.36	1.92	8.55
η_3	12	6.47	8.28	15.73[2]	2.93	5.23	8.55	16.09[2]
η_4	6	0.48	17.83	18.12	18.44	26.04[1]	32.56[1]	18.76[2]
η_4	12	20.48[2]	18.12	0.58	28.83[1]	26.16[1]	18.76[2]	5.75

1 and 2: Significant at the 1% and 5% levels, two-tailed t test.

Table 4: **Momentum Profits of S&P 500 Stocks.** This table examines momentum strategy returns based on S&P500 stocks. In each month, stocks that have been members of the S&P 500 index in the previous month are used to form nonoverlapping momentum portfolios. Each panel reports the mean and Newey-West t statistics of the monthly momentum profits ($\pi(R)$), as well as risk-adjusted momentum profits ($\pi'(\epsilon)$) in percentage of original return momentum profits ($\eta' = \pi'(\epsilon)/\pi(R) \times 100\%$). For calculating $\pi'(\epsilon)$, momentum portfolios weights are determined by the ranking-period raw returns, whereas holding-period returns are calculated from Kalman residuals. Panels A and B are based on WRSS and WML (long [short] top [bottom] 10% stocks) strategies with no lag between the ranking and holding periods. Panels C and D report the results when $G = 1$.

	F	H=1	H = 3	H=6	H=12	H = 18	H=24
A. WRSS strategies with G = 0							
$\pi(R)$	6	0.0006	0.0032	0.0049^1	0.0054^1	0.0025^2	0.0017^2
$\pi(R)$	12	0.0067^1	0.0083^1	0.0078^1	0.0050^1	0.0024^2	0.0014
T_{NW}	6	0.31	1.86	3.07	4.20	2.45	2.12
T_{NW}	12	3.19	4.36	4.24	3.39	2.04	1.33
η'	6	–	18.08	43.26^2	49.97^1	19.49	−3.18
η'	12	38.86	50.69^1	49.86^1	35.64	−5.88	−65.30
B. WML with G = 0							
$\pi(R)$	6	0.0014	0.0045^2	0.0065^1	0.0072^1	0.0035^1	0.0023^2
$\pi(R)$	12	0.0083^1	0.0105^1	0.0099^1	0.0066^1	0.0031^2	0.0016
T_{NW}	6	0.57	2.04	3.32	4.56	2.83	2.52
T_{NW}	12	3.14	4.39	4.29	3.59	2.26	1.37
η'	6	–	31.53	58.89^2	67.74^1	33.10	4.30
η'	12	45.12	63.16^1	63.59^1	47.71^2	−8.91	−95.03
C. WRSS strategies with G = 1							
$\pi(R)$	6	0.0038^2	0.0049^1	0.0065^1	0.0052^1	0.0025^1	0.0015^2
$\pi(R)$	12	0.0091^1	0.0087^1	0.0076^1	0.0043^1	0.0019	0.0011
T_{NW}	6	2.06	2.87	4.24	4.14	2.58	1.98
T_{NW}	12	4.49	4.54	4.25	3.03	1.69	1.04
η'	6	27.78	44.10	53.77^1	49.48^1	20.44	−9.47
η'	12	54.12^1	53.70^1	50.09^1	29.49	−25.67	−101.14
D. WML strategies with G = 1							
$\pi(R)$	6	0.0058^2	0.0068^1	0.0085^1	0.0069^1	0.0035^1	0.0021^2
$\pi(R)$	12	0.0116^1	0.0110^1	0.0098^1	0.0057^1	0.0025	0.0012
T_{NW}	6	2.42	3.11	4.51	4.42	2.97	2.31
T_{NW}	12	4.57	4.62	4.34	3.23	1.89	1.01
η'	6	50.63	64.97^2	71.84^1	66.84^1	35.67	−3.05
η'	12	67.88^1	68.71^1	65.05^1	39.59	−34.00	−142.44

1 and 2: Significant at the 1% and 5% levels, two-tailed t test.

Table 5: **Momentum Profits for More Dynamic versus Less Dynamic Stocks.** This table examines momentum profits generated by ex-post sorted NYSE-AMEX stocks from 1965 to 1999. Firms are sorted into four quartiles according to the standard deviation of estimated Kalman alphas, $\sigma_{\alpha(t)}$. Firms in Group 0 to 3 have $\sigma_{\alpha(t)}$ between 0 and 8.2×10^{-11}, between 8.2×10^{-11} and 0.0067, between 0.0067 and 0.0269, and above 0.0269, respectively. Each group has 1,595 firms. Group 0 is discarded since the Kalman model fails to properly capture the dynamics of firms in that group. Momentum strategies are then applied to firms within each of the three remaining groups. Panels A1 and A2 report the (overlapping-portfolio) WRSS and WML momentum strategy profits generated by firms in Groups 1 to 3 ($\pi(R1)$ to $\pi(R3)$). To save space, t statistics are not reported. Panel B reports the time series mean and t statistics for differences between Group 3 and Group 1 momentum profits ($\pi(R3) - \pi(R1)$).

		G = 0			G = 1		
	F	H = 3	H=6	H=12	H = 3	H=6	H=12
A1. WRSS strategies							
$\pi(R1)$	6	0.0005	0.0026[2]	0.0033[1]	0.0033[2]	0.0046[1]	0.0034[1]
$\pi(R2)$	6	0.0034[2]	0.0051[1]	0.0050[1]	0.0061[1]	0.0070[1]	0.0049[1]
$\pi(R3)$	6	0.0062[1]	0.0077[1]	0.0065[1]	0.0093[1]	0.0097[1]	0.0062[1]
$\pi(R1)$	12	0.0048[1]	0.0046[1]	0.0027[2]	0.0059[1]	0.0048[1]	0.0024
$\pi(R2)$	12	0.0076[1]	0.0070[1]	0.0043[1]	0.0082[1]	0.0070[1]	0.0037[1]
$\pi(R3)$	12	0.0105[1]	0.0089[1]	0.0050[1]	0.0109[1]	0.0085[1]	0.0040[2]
A2. WML strategies							
$\pi(R1)$	6	0.0005	0.0037[2]	0.0046[1]	0.0051[1]	0.0064[1]	0.0050[1]
$\pi(R2)$	6	0.0048[2]	0.0072[1]	0.0070[1]	0.0085[1]	0.0098[1]	0.0069[1]
$\pi(R3)$	6	0.0091[1]	0.0112[1]	0.0090[1]	0.0140[1]	0.0139[1]	0.0086[1]
$\pi(R1)$	12	0.0060[1]	0.0061[1]	0.0037[2]	0.0077[1]	0.0067[1]	0.0034[2]
$\pi(R2)$	12	0.0100[1]	0.0095[1]	0.0059[1]	0.0110[1]	0.0096[1]	0.0051[1]
$\pi(R3)$	12	0.0137[1]	0.0125[1]	0.0072[1]	0.0150[1]	0.0120[1]	0.0059[2]
B1. Difference between 3 and 1: WRSS							
$\pi(R3) - \pi(R1)$	6	0.0058[1]	0.0051[1]	0.0033[1]	0.0060[1]	0.0052[1]	0.0028[1]
$\pi(R3) - \pi(R1)$	12	0.0057[1]	0.0043[1]	0.0023	0.0050[1]	0.0037[1]	0.0016
$T_{\pi(R3)-\pi(R1)}$	6	4.10	4.06	3.07	4.37	4.21	2.65
$T_{\pi(R3)-\pi(R1)}$	12	3.74	3.02	1.71	3.35	2.57	1.22
B2. Difference between 3 and 1: WML							
$\pi(R3) - \pi(R1)$	6	0.0086[1]	0.0075[1]	0.0044[1]	0.0089[1]	0.0075[1]	0.0037[1]
$\pi(R3) - \pi(R1)$	12	0.0077[1]	0.0064[1]	0.0035	0.0073[1]	0.0053[2]	0.0025
$T_{\pi(R3)-\pi(R1)}$	6	4.43	4.34	3.05	4.67	4.44	2.59
$T_{\pi(R3)-\pi(R1)}$	12	3.47	3.07	1.80	3.30	2.55	1.30

1 and 2: Significant at the 1% and 5% levels, two-tailed t test.

Table 6: **Momentum Profits (Dynamic versus Market-to-Book).** This table examines momentum profits generated by ex-post double-sorted NYSE-AMEX stocks from 1965 to 1999. Firms are sorted independently into four $\sigma_{\alpha(t)}$ quartiles (Groups 0 to 3 [most dynamic firms]) and three market-to-book groups (MB1 to MB3 [high-MB firms]), where MB value is calculated as the mean of the inverse of available book-to-market values (all firms with negative means have been discarded). Each group contains from 400 to 600 firms. Momentum strategies are then applied to firms within each of the 12 double-sorted groups. Panels A1 and A2 report the time series mean and t statistics for WRSS606 and WRSS616 momentum profits (based on overlapping portfolios). Panel B reports WML strategy profits.

	Strategy:6:0:6			Strategy:6:1:6		
	MB1	MB2	MB3	MB1	MB2	MB3
A1. WRSS strategies						
$\pi(R0)$	0.0054^{2}	0.0042^{2}	0.0083^{1}	0.0087^{1}	0.0061^{1}	0.0101^{1}
$\pi(R1)$	0.0016	0.0013	0.0048^{1}	0.0039^{2}	0.0033^{2}	0.0062^{1}
$\pi(R2)$	0.0069^{1}	0.0041^{2}	0.0059^{1}	0.0091^{1}	0.0059^{1}	0.0075^{1}
$\pi(R3)$	0.0064^{1}	0.0071^{1}	0.0086^{1}	0.0087^{1}	0.0090^{1}	0.0104^{1}
$\pi(R3) - \pi(R1)$	0.0049^{2}	0.0059^{1}	0.0038	0.0048^{2}	0.0057^{1}	0.0042^{2}
A2. Corresponding t statistics						
$T_{\pi(R0)}$	2.04	2.04	3.63	3.46	3.07	4.46
$T_{\pi(R1)}$	0.77	0.75	2.83	1.98	2.07	3.81
$T_{\pi(R2)}$	3.39	2.39	3.00	4.80	3.51	3.97
$T_{\pi(R3)}$	2.66	3.21	3.67	3.81	4.28	4.51
$T_{\pi(R3)-\pi(R1)}$	2.22	3.12	1.95	2.20	3.13	2.16
B1. WML strategies						
$\pi(R0)$	0.0078^{2}	0.0058^{2}	0.0111^{1}	0.0129^{1}	0.0084^{1}	0.0139^{1}
$\pi(R1)$	0.0027	0.0023	0.0064^{1}	0.0057^{2}	0.0050^{2}	0.0088^{1}
$\pi(R2)$	0.0091^{1}	0.0062^{1}	0.0079^{1}	0.0126^{1}	0.0085^{1}	0.0101^{1}
$\pi(R3)$	0.0108^{1}	0.0096^{1}	0.0126^{1}	0.0133^{1}	0.0122^{1}	0.0146^{1}
$\pi(R3) - \pi(R1)$	0.0080^{1}	0.0073^{1}	0.0062^{2}	0.0077^{2}	0.0072^{1}	0.0058^{2}
B2. Corresponding t statistics						
$T_{\pi(R0)}$	2.20	2.12	3.52	3.77	3.22	4.45
$T_{\pi(R1)}$	1.06	1.14	3.07	2.25	2.55	4.34
$T_{\pi(R2)}$	3.36	2.76	2.92	4.94	3.86	3.92
$T_{\pi(R3)}$	3.20	3.17	3.98	4.09	4.29	4.70
$T_{\pi(R3)-\pi(R1)}$	2.66	2.80	2.24	2.52	2.91	2.11

1 and 2: Significant at the 1% and 5% levels, two-tailed t test.

Table 7: **Momentum Profits (Dynamic versus Size).** This table examines momentum profits generated by ex-post double-sorted NYSE-AMEX stocks from 1965 to 1999. Firms are sorted independently into four $\sigma_{\alpha(t)}$ quartiles (Groups 0 to 3 [more dynamic firms]) and three size groups (Size1 to Size3 [large-sized firms]) according to the average market capital of a firm during the period. Each group contains 400 to 600 firms. Momentum strategies are then applied to firms within each of the 12 double-sorted groups. Panel A1 and A2 report the time series mean and t statistics for WRSS606 and WRSS616 momentum profits (based on overlapping portfolios). Panel B reports WML strategy profits.

	Strategy:6:0:6			Strategy:6:1:6		
	Size1	Size2	Size3	Size1	Size2	Size3
A1. WRSS strategies						
$\pi(R0)$	0.0010	0.0105[1]	0.0092[1]	0.0047[2]	0.0116[1]	0.0103[1]
$\pi(R1)$	0.0016	0.0052[1]	0.0020	0.0040	0.0063[1]	0.0039[1]
$\pi(R2)$	0.0030	0.0057[1]	0.0069[1]	0.0055[1]	0.0075[1]	0.0085[1]
$\pi(R3)$	0.0028	0.0121[1]	0.0109[1]	0.0054[2]	0.0131[1]	0.0125[1]
$\pi(R3) - \pi(R1)$	0.0012	0.0069[1]	0.0089[1]	0.0013	0.0069[1]	0.0087[1]
A2. Corresponding t statistics						
$T_{\pi(R0)}$	0.40	4.81	4.26	2.05	5.33	4.89
$T_{\pi(R1)}$	0.63	2.89	1.36	1.68	3.68	2.75
$T_{\pi(R2)}$	1.35	3.26	3.82	2.57	4.50	4.85
$T_{\pi(R3)}$	1.14	5.44	4.61	2.30	6.14	5.45
$T_{\pi(R3)-\pi(R1)}$	0.43	3.64	4.84	0.47	3.72	4.69
B1. WML strategies						
$\pi(R0)$	0.0010	0.0161[1]	0.0130[1]	0.0067[2]	0.0179[1]	0.0143[1]
$\pi(R1)$	0.0025	0.0078[1]	0.0029	0.0067[2]	0.0095[1]	0.0054[1]
$\pi(R2)$	0.0044	0.0080[1]	0.0091[1]	0.0079[1]	0.0106[1]	0.0110[1]
$\pi(R3)$	0.0046	0.0180[1]	0.0146[1]	0.0078[2]	0.0196[1]	0.0168[1]
$\pi(R3) - \pi(R1)$	0.0021	0.0102[1]	0.0117[1]	0.0012	0.0101[1]	0.0114[1]
B2. Corresponding t statistics						
$T_{\pi(R0)}$	0.28	5.46	4.61	1.99	6.22	5.14
$T_{\pi(R1)}$	0.72	3.42	1.59	1.98	4.37	3.01
$T_{\pi(R2)}$	1.45	3.39	3.83	2.67	4.70	4.75
$T_{\pi(R3)}$	1.36	5.83	4.58	2.37	6.58	5.39
$T_{\pi(R3)-\pi(R1)}$	0.50	3.82	4.43	0.29	3.79	4.30

1 and 2: Significant at the 1% and 5% levels, two-tailed t test.

Table 8: **Momentum Profits (Dynamic versus Mispricing).** This table examines momentum profits generated by ex-post double-sorted NYSE-AMEX stocks studied from 1965 to 1999. In Panel A, firms are sorted independently into four $\sigma_{\alpha(t)}$ quartiles (Groups 0 to 3 [more dynamic firms]) and three mispricing groups (E1 to E3 [more mispricing]) according to $1 - R^2_{CAPM}$ (R^2_{CAPM} is the ex-post R^2 measure from the CAPM regression, based on all available firm returns from 1965 to 1999). Momentum strategies are then applied to firms within each of the 12 double-sorted groups. Panel A1 and A2 report the time series mean and t statistics for WRSS606 and WRSS616 momentum profits (based on overlapping portfolios). In Panel B firms are sorted independently into four $\sigma_{\alpha(t)}$ quartiles and three mispricing groups according to $\sigma(\epsilon_{Kal})$ (Group E1 to E3 [highest value]). The $E3 - E1$ column reports the profit difference between column $E3$ and column $E1$.

	Strategy:6:0:6				Strategy:6:1:6			
	E1	E2	E3	$E3-E1$	E1	E2	E3	$E3-E1$
A1. WRSS strategies, $\sigma(\alpha_t)$ vs. $1-R^2_{CAPM}$								
$\pi(R0)$	0.0099[1]	0.0079[1]	0.0033	−0.0066[2]	0.0105[1]	0.0098[1]	0.0061[2]	−0.0045
$\pi(R1)$	0.0030[2]	0.0026	0.0013	−0.0018	0.0049[1]	0.0043[2]	0.0039	−0.0010
$\pi(R2)$	0.0060[1]	0.0057[1]	0.0032	−0.0028	0.0076[1]	0.0077[1]	0.0054[2]	−0.0023
$\pi(R3)$	0.0108[1]	0.0066[1]	0.0046	−0.0062[2]	0.0127[1]	0.0086[1]	0.0072[1]	−0.0055[2]
$\pi(R3)-\pi(R1)$	0.0078[1]	0.0040[2]	0.0033	–	0.0078[1]	0.0042[2]	0.0032	–
A2. Corresponding t statistics								
$T_{\pi(R0)}$	4.95	3.53	1.22	−2.44	5.44	4.60	2.35	-1.70
$T_{\pi(R1)}$	2.14	1.39	0.48	−0.74	3.59	2.38	1.52	-0.42
$T_{\pi(R2)}$	3.48	3.06	1.35	−1.44	4.65	4.31	2.30	-1.19
$T_{\pi(R3)}$	4.88	2.70	1.88	−2.53	5.95	3.73	2.99	-2.27
$T_{\pi(R3)-\pi(R1)}$	4.52	2.10	1.17	–	4.62	2.25	1.11	–
B1. WRSS strategies, $\sigma(\alpha_t)$ vs. $\sigma(\epsilon_{Kal})$								
$\pi(R0)$	−0.0004	0.0041[1]	0.0072[1]	0.0075[1]	0.0013	0.0054[1]	0.0096[1]	0.0083[1]
$\pi(R1)$	−0.0008	0.0023	0.0065[1]	0.0073[1]	0.0011	0.0042[1]	0.0085[1]	0.0073[1]
$\pi(R2)$	0.0030[2]	0.0059[1]	0.0064[1]	0.0035	0.0048[1]	0.0077[1]	0.0084[1]	0.0036
$\pi(R3)$	0.0053[1]	0.0074[1]	0.0106[1]	0.0053	0.0078[1]	0.0092[1]	0.0116[1]	0.0038
$\pi(R3)-\pi(R1)$	0.0061[1]	0.0051[1]	0.0040	–	0.0067[1]	0.0049[1]	0.0031	–
B2. Corresponding t statistics								
$T_{\pi(R0)}$	-0.26	2.78	3.45	3.53	0.92	3.72	4.90	3.96
$T_{\pi(R1)}$	-0.66	1.68	2.76	3.50	0.99	3.19	3.78	3.65
$T_{\pi(R2)}$	2.25	3.53	2.68	1.75	3.83	4.82	3.64	1.83
$T_{\pi(R3)}$	2.64	3.49	3.29	1.92	4.03	4.56	3.68	1.37
$T_{\pi(R3)-\pi(R1)}$	3.51	2.98	1.35	–	3.87	2.95	1.04	–

1 and 2: Significant at the 1% and 5% levels, two-tailed t test.

Table 9: **Out-of-Sample Momentum Profits for More Dynamic versus Less Dynamic Stocks.** This table examines momentum profits generated by ex-ante sorted NYSE-AMEX stocks from 1965 to 1999. Instead of estimating the Kalman model once for all available return data during the test period, this table estimates the Kalman model year by year for each stock, and divides stocks into different quartiles according to ex-ante $\sigma_{\alpha(t)}$ in each month. In order to attain the ex-ante $\sigma_{\alpha(t)}$ measure in year t, the Kalman model is applied to all available return data after January 1962 and before January of year t (if there are more than 60 data points). Then, the standard deviation of the estimated risk premium ($\sigma_{\alpha(t)}$, based on prior-t information) is assigned to each month in year t as the ex-ante measure for firm dynamics. Next, in each month, firms are sorted into four quartiles according to the ex-ante $\sigma_{\alpha(t)}$ measure (Groups 0 to 3 [more dynamic firms]) , and momentum strategies are applied to stocks within each of the four quartiles. More specifically, since more than one ranking period is involved for the overlapping-portfolio strategies in each month, stocks are sorted according to available $\sigma_{\alpha(t)}$ values at the end of the first (earliest) ranking period. Panels A1 and A2 report the (overlapping-portfolio) WRSS and WML momentum strategy profits generated by firms in Groups 1 to 3 ($\pi(R1)$ to $\pi(R3)$). To save space, t statistics are not reported. Panel B reports the time series mean and t statistics for differences between Group 3 (more dynamic firms) and Group 1 (less dynamic firms) momentum profits ($\pi(R3) - \pi(R1)$).

	F	G = 0 H = 3	H=6	H=12	G = 1 H = 3	H=6	H=12
A1. WRSS strategies							
$\pi(R1)$	6	0.0020	0.0037[2]	0.0035[2]	0.0045[2]	0.0058[1]	0.0040[1]
$\pi(R2)$	6	0.0036[2]	0.0049[1]	0.0057[1]	0.0055[1]	0.0066[1]	0.0053[1]
$\pi(R3)$	6	0.0040	0.0062[1]	0.0060[1]	0.0079[1]	0.0087[1]	0.0060[1]
$\pi(R1)$	12	0.0063[1]	0.0061[1]	0.0026	0.0068[1]	0.0061[1]	0.0024
$\pi(R2)$	12	0.0082[1]	0.0074[1]	0.0051[1]	0.0086[1]	0.0071[1]	0.0044[2]
$\pi(R3)$	12	0.0095[1]	0.0087[1]	0.0055[1]	0.0108[1]	0.0090[1]	0.0051[1]
A2. WML strategies							
$\pi(R1)$	6	0.0039	0.0057[2]	0.0052[1]	0.0074[1]	0.0082[1]	0.0057[1]
$\pi(R2)$	6	0.0050[2]	0.0071[1]	0.0081[1]	0.0080[1]	0.0097[1]	0.0076[1]
$\pi(R3)$	6	0.0067[2]	0.0096[1]	0.0085[1]	0.0123[1]	0.0128[1]	0.0084[1]
$\pi(R1)$	12	0.0077[1]	0.0078[1]	0.0037	0.0087[1]	0.0081[1]	0.0035
$\pi(R2)$	12	0.0111[1]	0.0103[1]	0.0072[1]	0.0122[1]	0.0105[1]	0.0061[1]
$\pi(R3)$	12	0.0133[1]	0.0127[1]	0.0085[1]	0.0149[1]	0.0128[1]	0.0078[1]
B1. Difference between 3 and 1: WRSS							
$\pi(R3) - \pi(R1)$	6	0.0021	0.0025	0.0025[2]	0.0034[2]	0.0029[2]	0.0020[2]
$\pi(R3) - \pi(R1)$	12	0.0032[2]	0.0027	0.0029[2]	0.0039[1]	0.0028[2]	0.0027[2]
$T_{\pi(R3)-\pi(R1)}$	6	1.44	1.85	2.41	2.36	2.25	2.00
$T_{\pi(R3)-\pi(R1)}$	12	2.18	1.83	2.43	2.65	2.03	2.23
B2. Difference between 3 and 1: WML							
$\pi(R3) - \pi(R1)$	6	0.0028	0.0039[2]	0.0034[2]	0.0049[2]	0.0046[1]	0.0028[2]
$\pi(R3) - \pi(R1)$	12	0.0056[1]	0.0049[2]	0.0048[1]	0.0062[1]	0.0048[2]	0.0043[2]
$T_{\pi(R3)-\pi(R1)}$	6	1.43	2.19	2.43	2.47	2.66	2.01
$T_{\pi(R3)-\pi(R1)}$	12	2.71	2.45	2.87	2.98	2.49	2.57

1 and 2: Significant at the 1% and 5% levels, two-tailed t test.

Figure 1: **Simulation of Momentum Profits with Fixed Factors.** Below, a hypothetical economy is simulated for 100 times. The economy contains 30 firms, whose returns are determined by: $r_t = \alpha_t + (\beta_0 + \beta_t)r_t^M + \epsilon_t; \alpha_t = \alpha_0 + H_1\beta_t^2 + H_2\beta_t; \beta_t = \Gamma\beta_{t-1} + \varepsilon_{\beta,t}$, where r_t^M is the realized monthly market excess return from January 1965 to December 1999. For i^{th} firm, $\alpha_0^i = 0$, $\beta_0^i = 1$, $\gamma^i = 0.8$, $H_1^i = 0.5$, $H_2^i = 0.8$, $\epsilon_t^i\ N(0, 0.03^2)$, and $\varepsilon_{\beta,t}^i\ N(0, 0.1^2)$. For each simulation, α_t^i and β_t^i are fixed, while ϵ_t^i are different to generate different return processes. In Panel A and B the solid lines plot the true value of momentum components $\pi(1)$ and $\pi(2)$. The shadow region represents the 10% to 90% distribution of the estimated $\pi(1)$ and $\pi(2)$. In Panel C and D the solid lines plot the true value of cross-sectional mean of α_t^i and β_t^i, $\overline{\alpha}_t$ and $\overline{\beta}_t$, while the shadow region represents the 10% to 90% distribution of the estimated $\overline{\alpha}_t$ and $\overline{\beta}_t$.

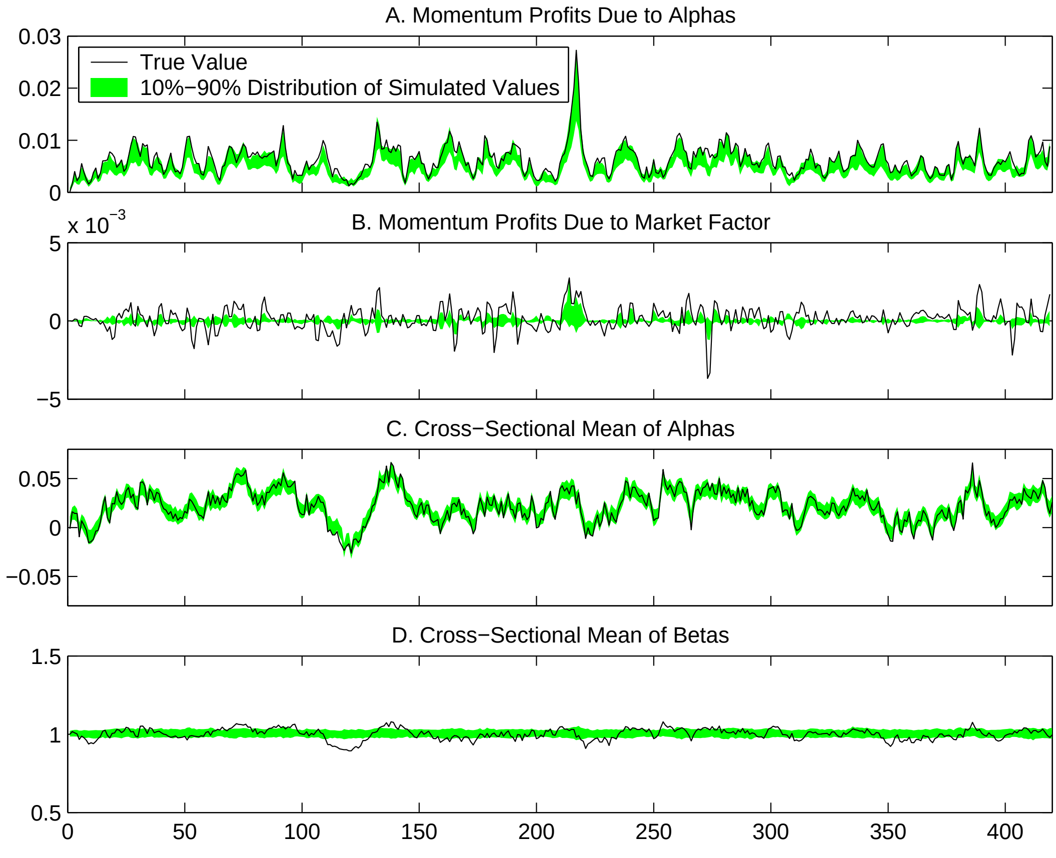

Figure 2: **Simulation of Momentum Profits with Random Factors.** Below, a hypothetical economy is simulated for 200 times. The economy contains 30 firms, whose returns are determined by $r_t = \alpha_t + (\beta_0 + \beta_t) r_t^M + \epsilon_t; \alpha_t = \alpha_0 + H_1 \beta_t^2 + H_2 \beta_t; \beta_t = \Gamma \beta_{t-1} + \varepsilon_{\beta,t}$, where for i^{th} firm, $\alpha_0^i = 0$, $\beta_0^i = 1$, $\gamma^i = 0.8$, $H_1^i = 0.5$, $H_2^i = 0.8$, $\epsilon_t^i \; N(0, 0.03^2)$, and $\varepsilon_{\beta,t}^i \; N(0, 0.1^2)$. For each simulation, new α_t^i, β_t^i time series, as well as ϵ_t^i, will be generated, while r_t^M, the market excess return from 1965 to 1999, is bootstrapped without replacement. In Panel A1 and B1 the solid lines plot the 10% and 90% true value of cross-simulation momentum components $\pi(1)$ and $\pi(2)$. The shadow region represents the 10% to 90% distribution of the estimated $\pi(1)$ and $\pi(2)$. Panel A2 plots the cross-simulation mean of $\pi(1)$, the cross-simulation mean and standard deviation of estimation error $(\pi(1, Kal) - \pi(1))$. Panel B2 plots the cross-simulation mean estimation error, $\pi(2, Kal) - \pi(2)$. Panels C and D report the cross-simulation mean (solid line) of cross-sectional mean of α_t^i and β_t^i for each month ($\overline{\alpha}_t$ and $\overline{\beta}_t$), the cross-simulation mean (dashed line) and standard deviation (dotted line) of estimation error of $\overline{\alpha}_t$ and $\overline{\beta}_t$, or $\overline{\alpha}_t(Kal) - \overline{\alpha}_t(True)$ and $\overline{\beta}_t(Kal) - \overline{\beta}_t(True)$.

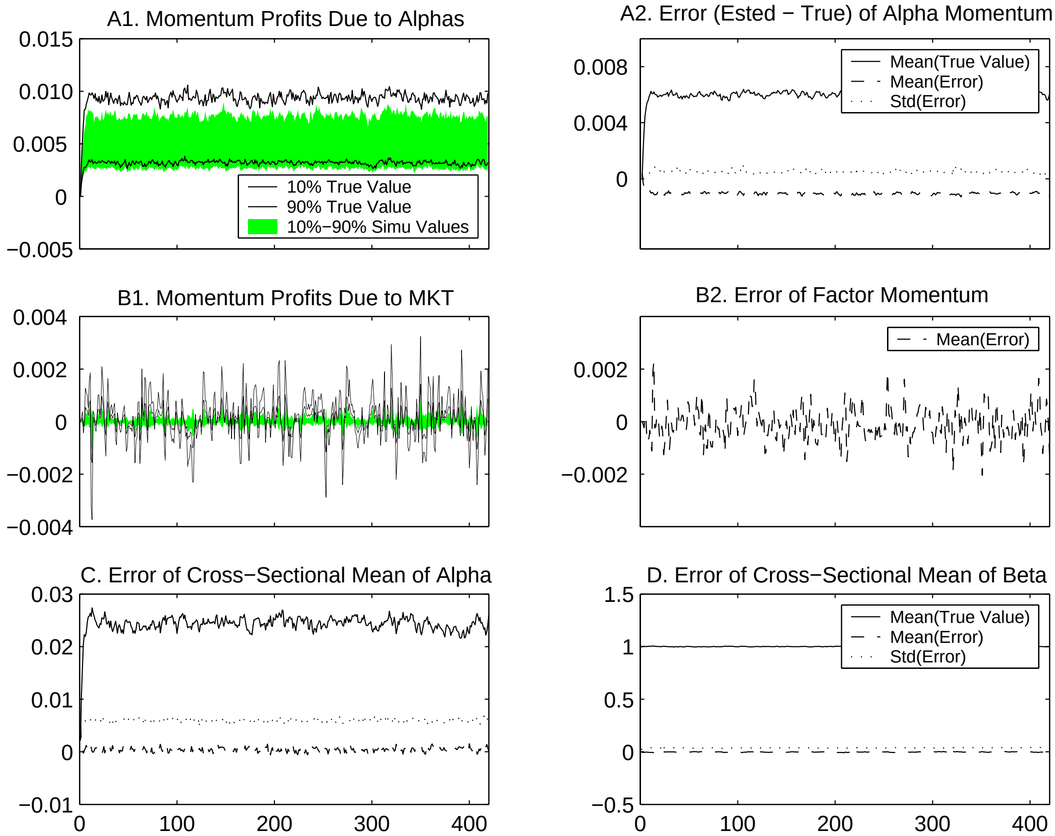

Figure 3: **Components of S&P 500 Momentum Profits.** This figure decomposes a set of nonoverlapping WRSS S&P 500-stock momentum strategy profits from 1965 to 1999 into four components: $\pi(r) = \pi(1) + \pi(2) + \pi(3) + \pi(4)$. The percentage contribution of each component is calculated as $\eta_i = |\pi(i)|/(|\pi(1)| + |\pi(2)| + |\pi(3)| + |\pi(4)|) \times 100\%$. In Panels A and B, the formation horizons are 6 and 12 months, respectively, while $G = 1$ month. The two panels plot the four components with holding horizons from 1 to 24 months.

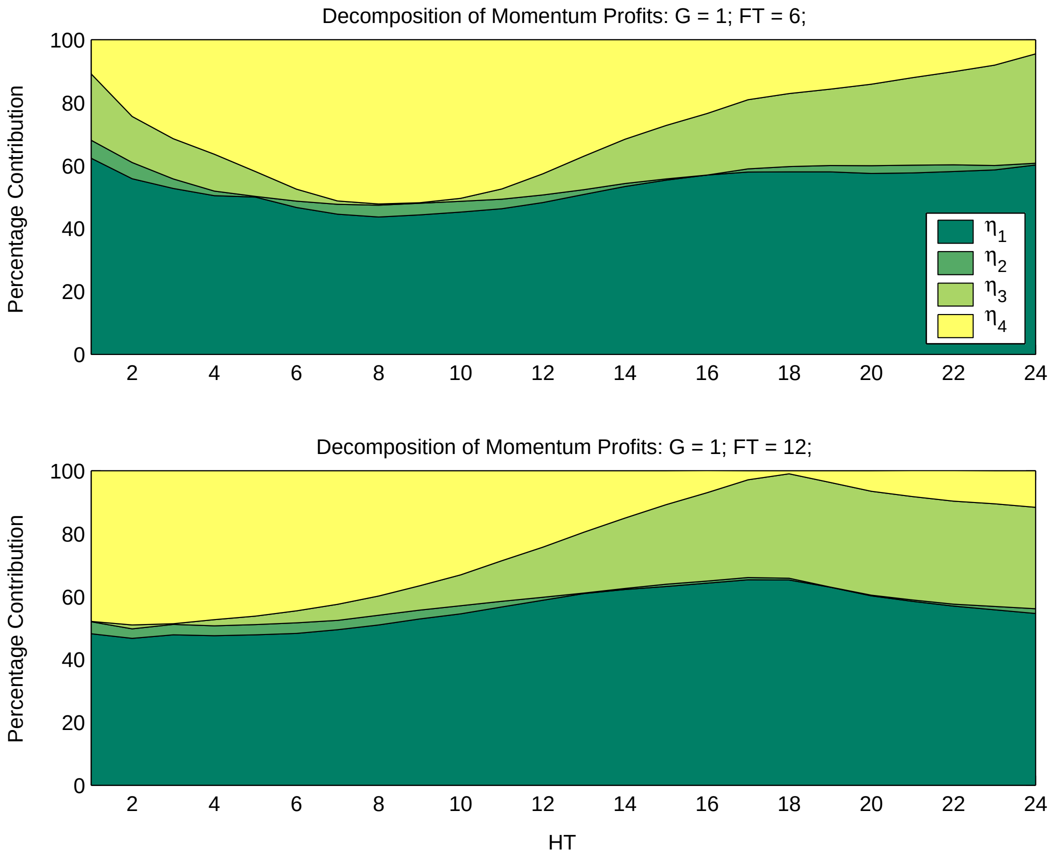

Figure 4: **Sorted Momentum Profits.** Panels A and B independently sort firms into four $\sigma_{\alpha(t)}$ quartiles (Groups 0 to 3 [more dynamic firms]) and three MB groups (MB1 to MB3 [high-MB firms]), where MB value is calculated as the mean of the inverse of available book-to-market values (all firms with negative means have been discarded). Firms within each of the 12 double-sorted groups are used to form momentum portfolios. Panels A and B plot the WML606 and WML616 momentum strategy profits, respectively. Details are in Table 6). Panels C and D independently sort firms into four $\sigma_{\alpha(t)}$ quartiles and three size groups (Size1 to Size3 [large firms]). Momentum profits for WML606 and WML616 strategies are plotted. Details are in Table 7.