

Conditional Risk and Performance Evaluation: Volatility Timing, Overconditioning, and New Estimates of Momentum Alphas

Oliver Boguth, Murray Carlson, Adlai Fisher, and Mikhail Simutin

The University of British Columbia*

October 27, 2009

Abstract

Unconditional alpha estimates are biased when conditional beta covaries with the market risk premium (“market-timing”) or volatility (“volatility-timing”). Whereas prior literature focuses on market-timing, we demonstrate that volatility-timing has a plausible impact 2 to 10 times larger. Moreover, we identify a novel and potentially substantial bias (“overconditioning”) that can occur any time an empiricist estimates conditional risk using information unavailable to investors – for example proxying with contemporaneous realized beta when asset returns are nonlinear. To correct market- and volatility-timing biases without overconditioning, we show that incorporating realized betas into instrumental variables estimators is effective. Empirically, instrumentation reduces momentum alphas by 20-40% relative to unconditional, while overconditioned alphas are up to 2.5 times larger. Volatility-timing inflates unconditional momentum performance because the formation-period market return (i) positively predicts holding-period beta (Grundy and Martin, 2001) and (ii) negatively predicts holding-period market volatility (French, Schwert, and Stambaugh, 1987), inducing negative covariation between conditional momentum beta and market volatility.

*Sauder School of Business, University of British Columbia, 2053 Main Mall, Vancouver, BC, V6T 1Z2. We thank for helpful comments Jonathan Berk, Ekkehart Boehmer, Long Chen, Zhi Da, Engelbert Dockner, Wayne Ferson, Lorenzo Garlappi, Ron Giammarino, Ro Gutierrez, Cam Harvey, Ravi Jagannathan, Marcin Kacperczyk, Jonathan Lewellen, Stefan Nagel, Jacob Sagi, Rob Stambaugh, Steve Thorley, Keith Vorkink, Kevin Wang, Moto Yogo, and seminar participants at Brigham Young University, HEC Lausanne, New York University, Rice University, Texas A&M, UBC, UCLA, the University of Minnesota, the University of Vienna, the 2007 Fall Meeting of the NBER Asset Pricing Program, the 2007 Meetings of the Northern Finance Association, and the 2009 Western Finance Association Meetings. Support for this project from the Social Sciences and Humanities Research Council of Canada and the UBC Bureau of Asset Management is gratefully acknowledged.

Conditional Risk and Performance Evaluation: Volatility Timing, Overconditioning, and New Estimates of Momentum Alphas

Abstract

Unconditional alpha estimates are biased when conditional beta covaries with the market risk premium (“market-timing”) or volatility (“volatility-timing”). Whereas prior literature focuses on market-timing, we demonstrate that volatility-timing has a plausible impact 2 to 10 times larger. Moreover, we identify a novel and potentially substantial bias (“overconditioning”) that can occur any time an empiricist estimates conditional risk using information unavailable to investors – for example proxying with contemporaneous realized beta when asset returns are nonlinear. To correct market- and volatility-timing biases without overconditioning, we show that incorporating realized betas into instrumental variables estimators is effective. Empirically, instrumentation reduces momentum alphas by 20-40% relative to unconditional, while overconditioned alphas are up to 2.5 times larger. Volatility-timing inflates unconditional momentum performance because the formation-period market return (i) positively predicts holding-period beta (Grundy and Martin, 2001) and (ii) negatively predicts holding-period market volatility (French, Schwert, and Stambaugh, 1987), inducing negative covariation between conditional momentum beta and market volatility.

1. Introduction

Under the conditional CAPM, risk equals the conditional exposure to market returns given the information available to investors. As is well known, time-variation in risk can impact unconditional estimates of investment strategy performance (Jensen, 1968; Dybvig and Ross, 1985) and asset pricing tests (Jagannathan and Wang, 1996). We call the conditioning problems studied in prior literature *underconditioning* because – with the notable exception of Jagannathan and Korajczyk (1986) discussed below – the empiricist is assumed to work with a subset of investor information, as in the canonical study of Hansen and Richard (1987). To empirically address underconditioning, Shanken (1990) and others allow estimated loadings to depend on lagged data observable to investors, such as the dividend yield.¹

Empiricists are not restricted, however, to using lagged data known to be in the investor information set. For example, an alternative approach directly estimates conditional factor loadings using “realized betas” estimated from short-window regressions simultaneous with or subsequent to the returns to be risk-adjusted (e.g., Chan, 1988; Grundy and Martin, 2001, “GM”; Lewellen and Nagel, 2006, “LN”). Using a realized beta to risk-adjust has a natural appeal, and may appear to be fully justified by the theoretical linear relationship between conditional risk and expected return.² With daily and higher-frequency data increasingly available, we anticipate growing use of realized betas for performance measurement and asset pricing tests.

In this paper, we identify a novel source of alpha bias that may occur any time an empiricist uses a conditional risk proxy not entirely contained in the investor information set. This potential empirical problem is the complement of underconditioning, and we call it *overconditioning*. While the concept is general and we discuss other examples, we focus on the overconditioning bias generated by using contemporaneous realized beta as a proxy for conditional beta. Note that any empirical realized beta estimate cannot be fully anticipated by investors.³ The estimation error, or “noise”, can be substantial in short windows and can impact alpha even under the optimistic assumption

¹Other studies using lagged instruments for alpha estimation include Avramov and Chordia (2006), Bollerslev, Engle, and Wooldridge (1988), Campbell (1987), Cochrane (1996), Duffee (2005), Ferson and Harvey (1991, 1993, 1999), Ferson, Kandel, and Stambaugh (1987), Ferson and Schadt (1996), Harvey (1989), Lettau and Ludvigson (2001*b*), Petkova and Zhang (2005), Santos and Veronesi (2006), Wang (2003), and the textbook of Cochrane (2001).

²GM proxy for month τ momentum betas using loadings estimated in the holding period τ to $\tau + 5$, and explain “the relevant risk to an investor... is the strategy’s factor exposure during the investment window.” (p. 43) LN state, “Our methodology... does not require any conditioning information. As long as betas are relatively stable within a month or quarter, simple CAPM regressions estimated over a short window – using no conditioning variables – provide direct estimates of assets’ conditional alphas.” (p. 291) Ang, Chen, and Xing (2006) similarly explain, “The CAPM predicts an increasing relationship between realized average returns and realized factor loadings... More generally, a multifactor model implies that we should observe patterns between average returns and sensitivities to different sources of risk over the same time period used to compute the average returns and the factor sensitivities.” (p. 1201)

³In ideal settings more restrictive than needed for a conditional CAPM, local quadratic variations and covariations are observable (e.g., Foster and Nelson, 1996), but microstructure effects remain important empirically.

that it has mean zero. We show that the overconditioning bias generated when using a realized beta is tied to nonlinearity in the relation between asset and factor returns.⁴ Intuitively, if an asset payoff is concave (convex) in market returns, the noise in realized beta negatively (positively) correlates with market return surprises, biasing alpha.

Theoretically, payoff nonlinearities can and should occur for many reasons. In managed portfolios, financial option holdings induce convex or concave payoffs. More pertinent to our study, stock returns can be decomposed into real and financial options (Brennan and Schwartz, 1985; McDonald and Siegel, 1985, 1986; Black and Scholes, 1973), which produce nonlinearities. Behavioral biases might also create nonlinearities, for example if past returns or other characteristics cause a stock to respond differently to positive versus negative systematic news, due to a disposition effect or biased self-attribution (Grinblatt and Han, 2005; Daniel, Hirshleifer, and Subrahmanyam, 1998). In our paper, it is immaterial why nonlinearities occur, but any reasonable theoretical prior should strongly favor the existence of nonlinearities in many stock returns and style portfolios.⁵

Empirically, abundant evidence of nonlinearities is provided by Ang and Chen (2002), Ang, Chen, and Xing (2006, “ACX”), Hong, Tu, and Zhou (2007), and other authors. These studies show that many individual stocks and style portfolios covary differently with negative and positive market surprises.⁶ For example, ACX sort days within a year according to whether the market return is below or above the average, and calculate down- and up-betas for each group. For the highest quintile of stocks, this beta asymmetry is larger than one. The point of our paper is not to explain these asymmetries, or, following ACX, to determine whether large down betas lead to higher returns. Instead, we seek to understand the implications of nonlinearities for performance measurement under the conditional CAPM, where there is no risk premium for beta asymmetry, yet nonlinearity determines the alpha bias from overconditioning with realized beta.

To be clear, many pricing models imply a linear correspondence between expected asset and factor returns, but the realized return relation may generally be nonlinear. For example, under quadratic preferences the CAPM holds for arbitrary return specifications. Similarly, while early APT formulations assume a strict factor structure, extensions are compatible with nonlinearities for an arbitrary number of assets provided these average out in random large portfolios.⁷ Payoff

⁴A *payoff nonlinearity* occurs when, conditional on the *contemporaneous* factor return, the relation between the expected return on an asset and the realized factor return is nonlinear. In a single-factor setting if one projects an asset return onto the factor and the residuals are correlated with any function of the factor, then payoffs are nonlinear.

⁵In randomly formed portfolios the convexities in some stocks will tend to cancel the concavities in others with increasing aggregation. Nonlinearities may remain strong however in any portfolio formed on firm characteristics related to real or financial options or behaviorally motivated nonlinearities. Jagannathan and Korajczyk (1986) discuss theoretical causes of nonlinearities in stock returns, emphasizing operating and financial leverage.

⁶A closely related measure of nonlinearity is coskewness – the covariation of a return with the squared market innovation. Harvey and Siddique (2000) and others provide empirical evidence of nonlinearity from this perspective.

⁷See, e.g., Chamberlain and Rothschild (1983) generalizing Ross (1976), and Grinblatt and Titman (1985).

nonlinearities should therefore not be ruled out when calculating alpha for general factor models.

If contemporaneous realized betas produce biased alphas due to overconditioning, then how should an empiricist measure conditional risk? One possibility is to use a lagged beta estimate, a common approach following Fama and MacBeth (1973). Previous authors point out a problem, however, with using lagged beta directly. For example, Chan (1988) shows that the market betas of winners decline on average from the formation to the holding period, while loser betas increase, which he explains through leverage changes. Similarly, GM observe predictable changes in the size loadings of winners and losers during and after formation. Using a historical beta as a risk proxy clearly biases alpha if the holding-period beta differs predictably under investor information.

We propose a simple solution to this problem, which to our knowledge has not been previously suggested or applied in performance evaluation, by using lagged loadings as instruments rather than direct proxies for the conditional loading. The instrumental variables (IV) approach allows combination of beta estimates from multiple prior windows – which may be useful if stocks have short- and long-run components in risk (e.g., Ghysels, Santa Clara, and Valkanov, 2005) – with traditional instruments such as the dividend yield, and other risk predictors such as the formation-period return. The IV method solves the problems noted by Chan (1988) and GM because it adjusts the conditional beta estimate for predictable changes from the formation to the holding period. The remedy combines the traditional method of using lagged instruments (Shanken, 1990) with the more recent literature emphasizing realized betas (LN), and solves both the problems of overconditioning and of predictable changes in beta from formation to holding periods.

We use this IV approach and related methods to demonstrate a substantial *overconditioning* problem in momentum portfolios, consistent with prior evidence of dramatic nonlinearities in portfolios sorted on past returns (De Bondt and Thaler, 1987; Chan, 1988; ACX; Hong, Tu, and Zhou, 2007). We also uncover a significant and previously undocumented *underconditioning* problem in unconditional estimates of the momentum strategy alpha.

Prior literature (e.g., Jagannathan and Wang, 1996) shows that unconditional alphas are biased when conditional beta covaries with the market risk premium (“market-timing”) or market volatility (“volatility-timing”). Previous studies focus almost exclusively on market-timing to evaluate the importance of underconditioning. For example, LN bound the market-timing bias in style portfolios given plausible ranges of variation in beta and the market risk premium. GM give specific evidence that market-timing is not significant for momentum. We agree with these prior findings.

We show that volatility timing has a plausible impact on alpha 2 to 10 times larger than market timing, using a model-free analytical approximation and parameter estimates from Brandt and Kang (2004). The substantial bound on the volatility-timing bias is easily confirmed non-parametrically,

following from the well-known fact that market volatility is highly variable (e.g., Schwert, 1989). Thus, whereas prior literature focuses on the underconditioning bias caused by market-timing, the volatility-timing bias is likely to be more important in practice.

A novel and significant empirical finding in our study is that the momentum strategy possesses remarkably strong volatility-timing, which inflates its unconditional alpha. The cause of volatility-timing in momentum is simple and robust, deriving from two well-documented and widely-accepted regularities: i) GM prove theoretically and empirically that the formation period market return significantly predicts the holding period momentum beta due to selection – when formation-period market returns are high, winners tend to have high betas; ii) high formation-period market returns also predict low holding-period market volatility, a consequence of more general “predictive asymmetry” in volatility (e.g., French, Schwert, and Stambaugh, 1987; Schwert, 1989; Campbell and Hentschel, 1992; Glosten, Jagannathan, and Runkle, 1994). Combining these facts, the holding-period beta and market volatility negatively covary, inflating the unconditional alpha.

We document these effects and their magnitudes. Instrumenting with lagged realized betas corrects the volatility-timing bias, significantly reducing momentum alphas by 20-40% relative to unconditional.⁸ By contrast, overconditioned alphas are up to 2.5 times larger, due to nonlinearities. Thus, both underconditioning and overconditioning lead to incorrect inference about the strategy’s conditional CAPM alpha. Instrumenting as executed in a variety of ways in our study corrects for the volatility-timing and nonlinearities inherent in momentum returns.

Some caveats about interpretation of our study are in order. Our purpose is not to defend the conditional CAPM or argue that correct use of conditioning information should generally improve its fit. Rather, we seek to generate accurate alpha estimates under the model, which is the canonical formulation of time-varying risk in finance. We take a major step in disentangling, both theoretically and empirically, the difference between conditional beta (which may predictably move with the market risk-premium and volatility) and nonlinearities (e.g., differences in “up-market” and “down-market” betas where “up” and “down” realizations are unpredictable). Separating conditional beta from nonlinearity is both important and difficult (Ferson and Schadt, 1996), and our methods should therefore be of interest to any empiricist using dynamic asset pricing models.

Consistent with LN, our momentum alpha estimates are still significantly positive. We nonetheless view the performance reduction, 20 basis points per month relative to unconditional and up to 90 basis points relative to overconditioned, as both statistically and economically meaningful. Ko-

⁸Existing theories of rational momentum effects (Berk, Green, and Naik, 1999; Johnson, 2002; Sagi and Seasholes, 2007) are unrelated to market- or volatility-timing. In these theories, risk varies over time for individual stocks, but on average winners load more heavily on risk than losers, and higher average risk is the source of momentum profits. Given access to the correct factor in these models, unconditional risk adjustment would explain momentum profits.

rajczyk and Sadka (2004) suggest that high trading costs substantially offset abnormal momentum profits. Overconditioned alphas of almost 1.5% per month would make such arguments seem much less relevant, but by substantially reducing alpha, our IV methods enhance the relative importance of considering trading costs. More broadly, the conditional CAPM alpha has an important interpretation as the average profit of a zero-cost conditionally market-neutral position in a strategy. Whether a strategy has an alpha of 60 or 80 or 140 basis points per month before costs is meaningful to an investor making an asset allocation decision. We therefore focus on testing the significance of the difference in alpha estimates, rather than simply testing whether alpha equals zero.

In Section 2, we distinguish overconditioning from underconditioning in a general theoretical framework. Overconditioning previously generated controversy regarding long-run reversals (De Bondt and Thaler, 1987; Chan, 1988), and can play an important role in market-timing (e.g., Henriksson and Merton, 1981; Jagannathan and Korajczyk, 1986; Ferson and Schadt, 1996), and methodologies using kernel estimates of beta (Li and Yang, 2008; Ang and Kristensen, 2009).

Section 3 shows that the plausible magnitudes of the volatility-timing and overconditioning biases exceed the market-timing bias, which has been the focus of prior literature. We calibrate a conditional CAPM matching the dynamics of the market premium and volatility, and asset return nonlinearities, to previous studies. Using simulation, we assess proxy and IV methods of estimating alpha, including techniques that incorporate contemporaneous, lagged, and filtered betas.

Section 4 contains our empirical analysis of momentum. All IV approaches confirm our primary result that appropriate conditioning reveals the volatility-timing bias in momentum alphas, whereas overconditioning with realized beta greatly overstates momentum performance. The IV methods we use include incorporating lagged betas, filtering, using a two-sided kernel, and using contemporaneous realized beta while instead instrumenting for the market risk premium. The IV alphas are all similar, significantly lower than either the unconditional or overconditioned estimates.

Our empirical analysis focuses on a style portfolio where turnover occurs monthly, formation rules are mechanistic, and as a consequence daily holdings are known. Additional and thornier issues may arise in managed portfolios where the unobservability of interim holdings is important (e.g., Goetzmann, Ingersoll, Spiegel, and Welch, 2007).⁹ Appendix A contains all proofs. Appendix B contains other details, and shows robustness under extension to conditional 3-factor performance.

⁹In managed portfolios, alpha can be biased when trading occurs at a higher frequency than risk is measured, and the pricing model holds over the same higher frequency (e.g., Leland, 1999; Goetzmann, Ingersoll, Spiegel, and Welch, 2007). The manipulation of alpha in this manner combines underconditioning (the empiricist lacks information about risk) with misspecification of the pricing model horizon. We follow the vast majority of the empirical asset pricing literature by using a monthly horizon for alpha measurement, while acknowledging that the relevant horizon(s) of marginal investor(s) remains an open question. A monthly measurement interval also matches well with the monthly turnover in momentum and other style portfolios, an important driver of their beta dynamics.

2. Conditioning Biases in Performance Measurement

For $t = 1, 2, \dots$, let conditional expected excess returns on asset i be

$$\mathbb{E}(R_{it} | \mathcal{F}_{t-1}) = \alpha_{it}^{t-1} + \beta_{it}^{t-1} \mathbb{E}(R_{Mt} | \mathcal{F}_{t-1}), \quad (2.1)$$

where $\{\mathcal{F}_t\}_{t=1}^\infty$ is a filtration, R_{Mt} is the excess market return, α_{it}^{t-1} is the conditional intercept, and $\beta_{it}^{t-1} \equiv \text{Cov}(R_{it}, R_{Mt} | \mathcal{F}_{t-1}) / \text{Var}(R_{Mt} | \mathcal{F}_{t-1})$ is the conditional beta. If $\alpha_{it}^{t-1} = 0$ and $\mathcal{F}_{t-1}$ represents investor information, the conditional CAPM is satisfied.

The bias from estimating unconditional alpha when a conditional model holds is well understood. Let $\sigma_M^2 \equiv \text{Var}(R_{Mt})$, and consider the market model $\bar{R}_i = \alpha_i^{UC} + \beta_i^{UC} \bar{R}_M$, where α_i^{UC} is the intercept and $\beta_i^{UC} \equiv \text{Cov}(R_{it}, R_{Mt}) / \sigma_M^2$. Grant (1977) shows by taking expectations of (2.1) that $\bar{R}_i = \bar{\alpha}_i + \text{Cov}(\beta_{it}^{t-1}, R_{Mt}) + \bar{\beta}_i \bar{R}_M$, where $\bar{\alpha}_i \equiv \mathbb{E}(\alpha_{it}^{t-1})$ is the mean conditional alpha, $\bar{\beta}_i \equiv \mathbb{E}(\beta_{it}^{t-1})$ is the mean conditional beta, $\bar{R}_i \equiv \mathbb{E}(R_{it})$, and $\bar{R}_M \equiv \mathbb{E}(R_{Mt})$. The bias is thus

$$\alpha_i^{UC} - \bar{\alpha}_i = \text{Cov}(\beta_{it}^{t-1}, R_{Mt}) - (\beta_i^{UC} - \bar{\beta}_i) \bar{R}_M, \quad (2.2)$$

where the first term captures the direct effect of market timing and the second reflects that the unconditional beta is generally a biased measure of average risk.

We use the following result, which is known from prior literature (Grant, 1977; Jagannathan and Wang, 1996; LN), but has not been fully exploited.¹⁰

Proposition 1. *If the conditional alpha and market premium are uncorrelated, the beta bias is $\beta_i^{UC} - \bar{\beta}_i = -(\bar{R}_M / \sigma_M^2) \text{Cov}(\beta_{it}^{t-1}, R_{Mt}) + \text{Cov}(\beta_{it}^{t-1}, R_{Mt}^2) / \sigma_M^2$. The alpha bias can be expressed:*

$$\alpha_i^{UC} - \bar{\alpha}_i = (1 + \bar{R}_M^2 / \sigma_M^2) \text{Cov}(\beta_{it}^{t-1}, R_{Mt}) - (\bar{R}_M / \sigma_M^2) \text{Cov}(\beta_{it}^{t-1}, R_{Mt}^2) \quad (2.3)$$

as a sum of market-timing and volatility-timing components.

We calibrate the potential magnitudes of both sources of alpha bias in Section 3 and find that volatility-timing can plausibly impact alpha by 2 to 10 times as much as market-timing, even though market-timing has received considerably more attention in the literature. Empirically, we

¹⁰ Jagannathan and Wang (1996) show in equation (A12) of their Appendix that covariation between conditional beta and market-volatility impacts unconditional alpha, but since this channel would create a complication to the main point of their empirical analysis, they assume it equal to zero in equation (A14). Lewellen and Nagel similarly recognize volatility-timing in their equation (2), but in the final paragraph of their Section II determine not to evaluate the importance of this channel in their calibration. Grant (1977) includes volatility-timing in his equation (12.1), but assumes it away by continuing his analysis with equation (12.2) under the assumption of homoskedasticity.

show in Section 4 that volatility-timing significantly inflates the unconditional momentum alpha, due to negative covariation between conditional beta and market volatility.

2.1. General Conditioning Information

We generalize the unconditional bias identified by previous authors. Assume an empirical estimate $\hat{\beta}_{it}$ of the conditional beta β_{it}^{t-1} , where the estimate may or may not belong to investor information $\mathcal{F}_{t-1}$. For example, if the empiricist evaluates the data at $T > 0$, then $\hat{\beta}_{it}$ must be measurable with respect to $\mathcal{G}_T \subseteq \mathcal{F}_T$, but in general $\mathcal{G}_T \not\subseteq \mathcal{F}_{t-1}$. In particular, the realized beta from a window containing t may be calculated by an empiricist at $T > t$, but is not available to investors at $t - 1$.

To capture this idea, decompose the beta estimate and the market return into parts that are predictable to investors at time $t - 1$ and residuals:

$$\hat{\beta}_{it} = \hat{\beta}_{it}^{t-1} + \varepsilon_{\beta t}, \quad R_{Mt} = \bar{R}_{Mt} + \varepsilon_{Mt}, \quad (2.4)$$

where $\hat{\beta}_{it}^{t-1} \equiv \mathbb{E}(\hat{\beta}_{it}|\mathcal{F}_{t-1})$ and $\bar{R}_{Mt} \equiv \mathbb{E}(R_{Mt}|\mathcal{F}_{t-1})$. Note that if $\hat{\beta}_{it}$ is measurable with respect to the investor information $\mathcal{F}_{t-1}$, then by definition the residual $\varepsilon_{\beta t}$ is identically zero.

Denote the conditional and unconditional alpha estimates $\hat{\alpha}_{it} = R_{it} - \hat{\beta}_{it}R_{Mt}$ and $\bar{\alpha}_i \equiv \mathbb{E}(\hat{\alpha}_{it})$.

Proposition 2. *The bias in $\hat{\alpha}_{it}$ under investor information is $\mathbb{E}(\hat{\alpha}_{it} - \alpha_{it}^{t-1}|\mathcal{F}_{t-1}) = \Delta_{\alpha t}^{UC} + \Delta_{\alpha t}^{OC}$, where $\Delta_{\alpha t}^{UC} \equiv (\beta_{it}^{t-1} - \hat{\beta}_{it}^{t-1})R_{Mt}^{t-1}$ and $\Delta_{\alpha t}^{OC} \equiv -Cov(\varepsilon_{\beta t}, \varepsilon_{Mt}|\mathcal{F}_{t-1})$. Taking expectations gives the bias in mean alpha: $\bar{\alpha}_i - \bar{\alpha}_i = \bar{\Delta}_{\alpha}^{UC} + \bar{\Delta}_{\alpha}^{OC}$, where*

$$\bar{\Delta}_{\alpha}^{UC} \equiv \mathbb{E}(\beta_{it}^{t-1} - \hat{\beta}_{it}^{t-1})\bar{R}_M + Cov(\beta_{it}^{t-1} - \hat{\beta}_{it}^{t-1}, R_{Mt}^{t-1}) \quad (2.5)$$

$$\bar{\Delta}_{\alpha}^{OC} \equiv -Cov(\varepsilon_{\beta t}, \varepsilon_{Mt}). \quad (2.6)$$

We call the first term the *underconditioning* bias, because the difference $\beta_{it}^{t-1} - \hat{\beta}_{it}^{t-1}$ is predictable under $\mathcal{F}_{t-1}$. By contrast, the second term reflects *overconditioning* because it can be nonzero only when $\hat{\beta}_{it}$ depends on information not available to investors.

Proposition 2 nests a number of special cases. When $\hat{\beta}_{it}$ is the unconditional beta, then the overconditioning bias is zero and the underconditioning bias is the formula (2.2) known from prior literature. More broadly, if $\hat{\beta}_{it}$ is measurable with respect to investor information then the overconditioning bias $\bar{\Delta}_{\alpha}^{OC}$ is zero and the underconditioning bias $\bar{\Delta}_{\alpha}^{UC}$ generalizes (2.2). Conversely, if the empirical beta is an unbiased estimate of the true investor beta, i.e., $\mathbb{E}(\hat{\beta}_{it}|\mathcal{F}_{t-1}) = \beta_{it}^{t-1}$, the underconditioning bias $\bar{\Delta}_{\alpha}^{UC}$ is zero but an alpha bias may still be present due to overconditioning.

A simple example illustrates how beta asymmetry and overconditioning interact to produce a

bias. Assume a static CAPM: $\bar{R}_i = \beta_i \bar{R}_M$ where $\beta_i \equiv \text{Cov}(R_i, R_M) / \sigma_M^2$. Let $S \in \{G, B\}$ be a variable that is not available to investors at time zero, but is observable ex post. For example, G might be the event that excess market returns are greater than $\bar{R}_M$, and B its complement. For $s \in \{G, B\}$, define the overconditioned betas $\beta_i^s \equiv \text{Cov}(R_i, R_M | S = s) / \text{Var}(R_M | S = s)$, and denote overconditioned abnormal returns by $\alpha_i^s \equiv \mathbb{E}(R_i | S = s) - \beta_i^s \mathbb{E}(R_M | S = s)$. We show:

Proposition 3. *The overconditioned alphas satisfy $\mathbb{E}(\alpha_i^S) = [\beta_i - \mathbb{E}(\beta_i^S)] \bar{R}_M - \text{Cov}[\beta_i^S, \mathbb{E}(R_M | S)]$. If S contains information about the market return and $\beta_i^G \neq \beta_i^B$, the mean overconditioned alpha is generally biased.*

Figure 1 illustrates this proposition in a four-state setting. The CAPM holds with a slope of one, shown by the solid line going through the origin. In the example, $\beta_i = \mathbb{E}(\beta_i^S)$ so no beta bias is present. The conditional regressions correspond to the two dashed lines with slopes $\beta_i^B > \beta_i^G$. The concave payoffs imply $\text{Cov}[\beta_i^S, \mathbb{E}(R_M | S)] < 0$, and following Proposition 3 the mean overconditioned alpha is positive, consistent with the intercepts $\alpha_i^B, \alpha_i^G > 0$. Under convex payoffs $\beta_i^B < \beta_i^G$, the mean overconditioned alpha would be negative. In general, using a beta estimate conditioned on information not available to investors causes an alpha bias that depends on the degree of concavity or convexity in payoffs.

2.2. Types of Overconditioning

Aided by the previous example, we discuss methodologies that may produce overconditioning biases:

Contemporaneous realized beta: Suppose an empiricist has a sample of N independent draws from the Figure 1 example, and calculates a realized beta to use in estimating alpha. For a sample with a large (small) proportion of G draws, the realized market return will tend to be high (low) and the realized beta will tend to be low (high) relative to their ex ante expectations, inducing covariation between the beta error and the unexpected market return as in (2.6).¹¹

In addition to LN and GM, an earlier example of using contemporaneous realized beta to measure abnormal returns is Chan (1988), who focuses on long-run reversals. He runs market model regressions using 36-month windows of monthly returns, and interprets the average intercept as the average conditional alpha, a procedure identical to LN except that Chan uses longer windows of monthly rather than daily returns. Chan's results contradict De Bondt and Thaler (1985) by making the reversal alpha insignificant. Thus, Chan reduces the alpha of a loser minus winner portfolio by overconditioning with contemporaneous realized beta – consistent with what we find for momentum

¹¹In this setting, as N grows the overconditioning bias falls. A large sample can however worsen the underconditioning problem when beta is time-varying, a tradeoff we quantify in Section 3.

where overconditioning increases alpha. Of course, our paper provides a very different interpretation of Chan’s result as the consequence of payoff nonlinearities and overconditioning.

Subsequent to our work, Li and Yang (2008) and Ang and Kristensen (2009) extend the LN approach. Rather than calculating a day t alpha as the intercept from an OLS market-model regression in a fixed window, Li and Yang allow window size to vary and use weighted least squares where weights decline with distance from t . Ang and Kristensen use a related weighted least squares approach. The econometric results in Ang and Kristensen assume a strict factor structure, which eliminates the possibility of payoff nonlinearities and is narrower than required for a conditional factor model to hold.¹² Empirically, overconditioning remains an important consideration in both approaches. Following our results, the potential overconditioning bias is larger in small windows, and expanding the window size increases the relative importance of underconditioning.

Up-market and Down-market Betas: In the setting of Proposition 3, if G is the event that the realized market return exceeds its expectation, then β_i^G and β_i^B have the interpretation of up- and down-market betas. As shown in Proposition 3, it would be incorrect to interpret the intercepts α_i^G or α_i^B as CAPM (or conditional CAPM) performance measures.

De Bondt and Thaler (1987, “DT”) intuitively suggest that using up- and down-betas for performance measurement is a mistake, and draw a link with using contemporaneous realized betas, in responding to the criticism of Chan (1988). DT run up- and down-beta regressions on reversal portfolios and find alphas similar to Chan’s estimates. They comment, “*Only when the betas are allowed to vary with the level of the market is the alpha of the arbitrage portfolio no longer positive... These time-varying ‘split’ betas are questionable measures of risk... It seems odd to say that a portfolio with a beta of 1.602 in up markets and .591 in down markets is riskier than one with up and down betas of .854 and 1.439.*” The intuition of DT is correct. The realized betas estimated by Chan and the up- and down-betas that DT estimate are not conditional CAPM betas, as interpreted by Chan, but rather capture nonlinearities or slopes in different parts of the return distribution. The portfolio with betas of 1.602 and .591 in the DT study consists of reversal portfolio losers, whose convex payoffs lead to a downward bias in overconditioned alphas, consistent with Proposition 3. By contrast, the reversal winners have concave payoffs and overconditioning inflates their alphas.

¹² Ang and Kristensen also develop their theoretical results under the null of a constant market premium, which reduces the importance of measuring conditional beta correctly since an underconditioning bias is not present. They note that interacting factors with predictors of risk premia can justify an unconditional model with managed portfolios as factors, but do not operationalize what managed portfolios should be used. The approach thus relies on having adequate instruments for risk premia. By contrast, our study focuses on finding good instruments for conditional beta while remaining agnostic about predictors for the market premium. Our alpha estimates are nonetheless empirically robust to using contemporaneous realized beta and instead instrumenting for market returns with standard market premium predictors.

In the momentum setting, Rouwenhorst (1998, p. 278) uses an up- and down-beta regression and reports that alphas, which he interprets as measures of abnormal performance, increase substantially, opposite to the effect in reversal strategies. These results are all consistent with the overconditioning biases we document.

Henriksson-Merton and Treynor-Mazuy Regressions: To decompose managed fund performance into timing ability and selection components, Henriksson and Merton (“HM”, 1981) suggest regressing portfolio returns on the market return and a nonlinear function of the market return, such as an option payoff. The coefficient on the option payoff is interpreted as timing ability and the regression intercept identifies selection skill.¹³ Jagannathan and Korajczyk (1986) point out that when payoffs are in fact nonlinear, for example due to option holdings, the HM regressions lead to incorrect inference about both timing and selection. Specifically if the manager has no timing or selection advantage and payoffs are convex, HM regressions incorrectly identify positive timing ability and negative selection. Under concave payoffs one incorrectly infers negative timing and positive selection. Overconditioning in an HM regression occurs because, as with contemporaneous realized beta or up- and down-betas, one uses on the right-hand-side a regressor that conditions on the realized level of market returns.

Ferson and Schadt (1996) address “conditional market-timing”. They allow conditional betas to vary with instruments, and incorporate a call payoff as in HM, but acknowledge that their alphas must be cautiously interpreted because of the Jagannathan and Korajczyk critique.

Two-sided Kernel Estimates: A different kind of overconditioning bias may occur when one uses future realized betas as proxies or instruments for conditional beta. Specifically, in the presence of real or financial options, contemporaneous market news may be correlated with innovations in true conditional beta, either through changes in financial leverage or changes in asset risk induced by option exercise (e.g., Hamada, 1972; Carlson, Fisher, and Giammarino, 2004). The importance of such channels has not yet been evaluated for performance evaluation methods that use future realized betas as instruments for conditional beta. We can say, however, that if a difference in alpha is obtained when using one-sided (backward only) vs. two-sided beta estimates, it will be a thorny issue to determine whether the apparent change in beta was anticipated by investors at the beginning of the return-measurement interval, and certainly identifying instruments for the beta change (e.g., leverage, growth options) would be an important empirical issue.

Fortunately, this type of difference between one-sided and two-sided beta estimates is not present in our empirical work on momentum. We obtain the same alphas using past or future realized

¹³See also Treynor and Mazuy (1966), who suggest using a different nonlinear function of market returns – the squared market – for a similar purpose.

betas as instruments (we also show robustness to using contemporaneous returns with appropriate filtering). Robustness to using future betas as instruments confirms that the overconditioning bias in momentum portfolios derives from payoff nonlinearities related to the unexpected market return, rather than to permanent changes in beta that might or might not be anticipated by investors.

Solutions: Examination of (2.6) shows that the overconditioning bias can be eliminated by instrumenting for either i) the contemporaneous realized beta, or ii) the realized market return. Consistent with prior literature (e.g., Andersen, Bollerslev, Diebold, and Wu, 2006; Ghysels and Jacquier, 2006), realized beta is highly predictable from prior window beta estimates and other variables, hence we focus on instrumenting for realized beta. However, our empirical results are robust to using contemporaneous realized beta and instead instrumenting for market returns with standard risk premium predictors. These solutions are new to the performance evaluation literature.

Summary: Our theoretical analysis provides a general framework within which to understand both underconditioning and overconditioning biases. Unconditionally evaluating a conditional pricing model (e.g., Jagannathan and Wang, 1996) is a special case of using too little information, or underconditioning. By contrast, the problems observed by Jagannathan and Korajczyk (1986) and De Bondt and Thaler (1987) and the bias caused by using contemporaneous realized beta that we identify, are all special cases of the general concept of overconditioning. Our framework distinguishes overconditioning – a new concept – from underconditioning, which is well understood in the literature but generalized by our analysis.

3. Magnitudes of the Conditioning Biases and Empirical Methods

We analytically calculate model-free bounds on the market- and volatility-timing biases, and use a straightforward calibration with empirical estimates from prior literature to evaluate the importance of overconditioning. Using simulation, we assess proxy and IV methods for estimating conditional alpha and beta.

A back-of-the-envelope calculation shows that volatility-timing can impact alpha much more than market timing. We rewrite the alpha bias (2.3) as¹⁴ $\alpha_i^{UC} - \bar{\alpha}_i = \alpha_M^{UC} + \alpha_\sigma^{UC}$, where

$$\alpha_M^{UC} = (1 + \bar{R}_M^2/\sigma_M^2) Cov(\beta_{it}^{t-1}, \bar{R}_{Mt}) - (\bar{R}_M/\sigma_M^2) Cov(\beta_{it}^{t-1}, \bar{R}_{Mt}^2) \quad (3.1)$$

$$\alpha_\sigma^{UC} = -(\bar{R}_M/\sigma_M^2) Cov(\beta_{it}^{t-1}, \sigma_{Mt}^2). \quad (3.2)$$

¹⁴We use the facts that $Cov(\beta_{it}^{t-1}, R_{Mt}) = Cov(\beta_{it}^{t-1}, \bar{R}_{Mt})$ and $Cov(\beta_{it}^{t-1}, R_{Mt}^2) = Cov(\beta_{it}^{t-1}, \sigma_{Mt}^2) + Cov(\beta_{it}^{t-1}, \bar{R}_{Mt}^2)$. The second equality follows from $Cov(\beta_{it}^{t-1}, R_{Mt}^2) = E((\beta_{it}^{t-1} - \bar{\beta}_i) R_{Mt}^2)$, which using the law of iterated expectations is equal to $E((\beta_{it}^{t-1} - \bar{\beta}_i) E(R_{Mt}^2 | \mathcal{F}_{t-1}))$, simplifying to $E((\beta_{it}^{t-1} - \bar{\beta}_i) (\sigma_{Mt}^2 + \bar{R}_{Mt}^2))$.

All covariances related to the market premium are contained in α_M^{UC} , while the covariance of beta with market volatility drives α_σ^{UC} .

LN quantify the potential magnitude of α_M^{UC} . At a monthly horizon, $\bar{R}_M^2/\sigma_M^2 \approx (0.005/0.05)^2 = 10^{-2}$. The ratio $\bar{R}_M/\sigma_M^2 \approx 0.005/(0.05^2) = 2$ and the potential magnitude of the second term of (3.1) is clearly small relative to the first since $Var(\bar{R}_{Mt}^2) \ll Var(\bar{R}_{Mt})$.¹⁵ A close approximation of the magnitude of the market-timing bias and an upper bound is thus

$$\begin{aligned} |\alpha_M^{UC}| &\approx |\bar{R}_M Corr(\beta_{it}^{t-1}, \bar{R}_{Mt}) std(\beta_{it}^{t-1}) std(\bar{R}_{Mt}/\bar{R}_M)| \\ &\leq \bar{R}_M \times 1 \times std(\beta_{it}^{t-1}) std(\bar{R}_{Mt}/\bar{R}_M) \equiv \alpha_M^{\max}. \end{aligned}$$

With a standard deviation of conditional beta of 0.5, assuming the standard deviation of $\bar{R}_{Mt}$ is not larger than $\bar{R}_M$ bounds the market-timing bias at half the market risk premium.

We show the volatility-timing bias has a larger potential magnitude. Note that

$$\begin{aligned} |\alpha_\sigma^{UC}| &= |\bar{R}_M Corr(\beta_{it}^{t-1}, \sigma_{Mt}^2) std(\beta_{it}^{t-1}) std(\sigma_{Mt}^2/\sigma_M^2)| \\ &\leq \bar{R}_M \times 1 \times std(\beta_{it}^{t-1}) std(\sigma_{Mt}^2/\sigma_M^2) \equiv \alpha_\sigma^{\max}. \end{aligned} \tag{3.3}$$

The relative size of the two bounds is

$$\frac{\alpha_\sigma^{\max}}{\alpha_M^{\max}} = \frac{std(\sigma_{Mt}^2/\sigma_M^2)}{std(\bar{R}_{Mt}/\bar{R}_M)}. \tag{3.4}$$

Market volatility is widely known to be highly variable (e.g., Schwert, 1989), so we may intuitively expect this ratio to be larger than one. Direct evidence of the quantities in (3.4) is available from Brandt and Kang (2004), who estimate a latent-variable specification of joint variations in the market risk premium and volatility. Using their results, the numerator is approximately 1.2 with a two standard deviation range of about ± 0.20 , while the denominator is approximately 0.3 with a two standard deviation range of about ± 0.15 . We infer that the plausible magnitude of the volatility-timing bias is about 2 to 10 times larger than the market-timing bias.

We non-parametrically confirm the magnitude of the numerator of (3.4). The standard deviation of VIX-squared, since its inception and adjusted to monthly variance, is approximately 0.00337, roughly equal to the mean of (monthly adjusted) VIX-squared over the same period, and about 1.5 times the variance of the CRSP value-weighted index from 1925-present. Whereas the upper bound for the market-timing bias suggested by LN was based on very aggressive assumptions about

¹⁵The squared conditional market returns is of the order .000025, and the range of plausible values is narrow relative to the range of the conditional market return since $|\bar{R}_{Mt}| \ll 1$.

variability of the market premium, an estimate of $1 \leq \text{std}(\sigma_{Mt}^2/\sigma_M^2) \leq 1.5$ is entirely reasonable for bounding volatility-timing. Thus, if we are to find evidence of an underconditioning bias in any asset return, it will more likely be due to volatility timing than market timing. Following (3.3), the impact of volatility timing on alpha can reach 0.5 to 0.75 times the market risk premium, which is certainly enough to be meaningful to investors.

3.1. A Model of Market-timing, Volatility-timing, and Nonlinearities

We provide a simple model of underconditioning and overconditioning biases using a realistic calibration of market returns from prior literature, and evaluate different empirical methods of estimating alpha. The key components of such a model are the dynamics of the market risk premium and volatility, and the specification of beta dynamics and nonlinearities in individual asset returns.

Following Brandt and Kang (2004), we choose the conditional mean and variance of the market return to be a bivariate lognormal process.¹⁶ This ensures a positive market premium consistent with theoretical priors, as in Bekaert and Harvey (1995) and De Santis and Gerard (1997), and also captures the approximate lognormality of volatility documented by Andersen, Bollerslev, Diebold, and Ebens (2001). These choices have many precedents in prior literature.¹⁷

Let $t = 0, 1, \dots, T$, index days, and let $\tau(t)$ map days to “months”. Market returns follow $R_{Mt} = \bar{R}_{Mt} + \varepsilon_{Mt}$, where ε_{Mt} is normally distributed with mean zero and variance σ_{Mt}^2 . We assume that the conditional mean and variance are determined by state variables observable to investors at the end of the prior month. That is,

$$\bar{R}_{Mt} \equiv \mathbb{E}(R_{Mt}|\mathcal{F}_{t-1}) \equiv \bar{R}_M \exp[\lambda_M X_{\tau(t)-1} - \lambda_M^2/2] \quad (3.5)$$

$$\sigma_{Mt}^2 \equiv \text{Var}(R_{Mt}|\mathcal{F}_{t-1}) \equiv \bar{\sigma}_M^2 \exp[\lambda_\sigma Y_{\tau(t)-1} - \lambda_\sigma^2/2], \quad (3.6)$$

where $\bar{\sigma}_M^2$ is the average conditional variance and the state variables follow $AR(1)$ processes:

$$X_\tau = \varphi_x X_{\tau-1} + \sigma_x \varepsilon_{x\tau}, \quad Y_\tau = \varphi_y Y_{\tau-1} + \sigma_y \varepsilon_{y\tau}, \quad (3.7)$$

where $-1 < \varphi_x, \varphi_y < 1$, the innovations $\varepsilon_{x\tau}$ and $\varepsilon_{y\tau}$ are standard normals with correlation ρ_ε , and

¹⁶Many specifications of market return dynamics are available in the literature. For example, the market risk premium can be modeled as a linear, Markov-switching, or log-linear process, and similar choices can be made for market volatility. In prior versions we have used both linear and Markov-switching specifications for the market premium and obtained results similar to those reported here. Following Proposition 1, the market- and volatility-timing biases depend on the variability of the market risk premium and volatility, not on a particular choice of functional form.

¹⁷Latent time-variation in the conditional mean is modelled by Lamoureux and Zhou (1996) and others. Stochastic volatility in market returns is estimated by numerous authors (e.g., Wiggins, 1987; Jacquier, Polson, and Rossi, 1994; Kim, Shephard, and Chib, 1998).

the normalizations $\sigma_x = (1 - \varphi_x^2)^{1/2}$ and $\sigma_y = (1 - \varphi_y^2)^{1/2}$ ensure unit unconditional variance.

The stock return R_{it} satisfies $\bar{R}_{it} \equiv \mathbb{E}(R_{it}|\mathcal{F}_{t-1}) = \alpha_{it}^{t-1} + \beta_{it}^{t-1} \bar{R}_{Mt}$, where α_{it}^{t-1} is the conditional intercept and the conditional beta may vary with either of the state variables:

$$\beta_{it}^{t-1} \equiv \text{Cov}(R_{it}, R_{Mt}|\mathcal{F}_{t-1}) / \sigma_{Mt}^2 = \bar{\beta} + b_x X_{\tau(t)-1} + b_y Y_{\tau(t)-1}. \quad (3.8)$$

Note that β_{it}^{t-1} is known to investors at the end of the month prior to t .

3.1.1. Underconditioning Biases

We decompose the alpha bias caused by unconditional evaluation:

Proposition 4. *Under the dynamics (3.5)-(3.8) and the conditional CAPM ($\alpha_{it}^{t-1} = 0$), the unconditional alpha satisfies $\alpha^{UC} = \alpha_{M,direct}^{UC} + \alpha_{loading}^{UC}$, where the direct market-timing alpha is*

$$\alpha_{M,direct}^{UC} = \text{Cov}(\beta_{it}^{t-1}, R_{Mt}) = \lambda_M \bar{R}_M (b_x + b_y \text{Cov}(X, Y)),$$

and the loading-mismeasurement alpha is $\alpha_{loading}^{UC} = -\bar{R}_M (\beta^{UC} - \bar{\beta})$. Expressions for the unconditional beta bias and $\text{Cov}(X, Y)$ are given in the Appendix.

To better understand these biases, we compare with an almost equivalent decomposition:

Proposition 5. *The unconditional alpha bias can be decomposed $\alpha^{UC} = \alpha_M^{UC} + \alpha_\sigma^{UC}$, where*

$$\alpha_M^{UC} = \left(1 - \frac{k_\mu}{\sigma_M^2}\right) \text{Cov}(\beta_{it}^{t-1}, \bar{R}_{Mt}) = \lambda_M \bar{R}_M \left(1 - \frac{k_\mu}{\sigma_M^2}\right) (b_x + b_y \text{Cov}(X, Y)), \quad (3.9)$$

$$\alpha_\sigma^{UC} = -(\bar{R}_M / \sigma_M^2) \text{Cov}(\beta_{it}^{t-1}, \sigma_{Mt}^2) = -\lambda_\sigma \bar{R}_M (\bar{\sigma}_M^2 / \sigma_M^2) (b_y + b_x \text{Cov}(X, Y)), \quad (3.10)$$

are respectively the total market-timing and volatility-timing biases and k_μ is given in the Appendix.

The total market-timing alpha α_M^{UC} is proportional to the direct market-timing alpha $\alpha_{M,direct}^{UC}$, with their difference due to the impact of market timing on the beta bias. The ratio of total to direct market-timing alphas is $\alpha_M^{UC} / \alpha_{M,direct}^{UC} \approx 1$ when alpha is measured at a monthly frequency.¹⁸ Thus, if the loading-measurement alpha has any practical consequence, it must be due to volatility timing as captured by α_σ^{UC} . The tight link between loading mismeasurement and volatility timing is explained in more detail in Appendix B. Because the total and direct market-timing alphas are empirically very close, we do not distinguish between the two in subsequent discussion.

¹⁸This follows from the calculation $\alpha_M^{UC} / \alpha_{M,direct}^{UC} = 1 - k_\mu / \sigma_M^2$. Under a reasonable calibration, discussed below, the ratio $k_\mu / \sigma_M^2 = \bar{R}_M^2 (2e^{\lambda_M^2} - 1) / \sigma_M^2 \approx .006^2 (2e^{0.09} - 1) / .0025 \approx .02$.

From Proposition 5 and consistent with the previous model-free calibration, the market-timing alpha is proportional to the standard deviation λ_M of the log market risk premium, and the volatility-timing alpha is proportional to the standard deviation λ_σ of log variance. A comparison of the magnitudes of these effects is therefore straightforward, as shown in Table 1 where we assume $Cov(X, Y) = 0$ for simplicity.

The market-timing bias in Panel A satisfies $\alpha_M^{UC} \approx \lambda_M \bar{R}_M b_x$. We consider values of λ_M ranging from 0.10 to 0.50. At the upper end $\lambda_M = 0.50$, which is slightly more than two standard deviations above the point estimate in Brandt and Kang (2004), the 95% confidence interval for the market risk premium goes from 3% to 20% annually.¹⁹ The results are consistent with Table 1 in LN, showing that for reasonable magnitudes of variation in the market risk premium and beta, the underconditioning bias is moderate, at the upper end about 0.2% per month.²⁰

Panel B shows the volatility-timing bias. We can calibrate the magnitude of λ_σ in several ways. Using the estimates provided in Brandt and Kang gives an estimate of λ_σ of about 1.2. (See footnote 19). This implies that the 95% confidence interval for volatility differs by a factor of approximately 10 from its upper to lower ends, consistent with observed values of the VIX, which has ranged from a maximum of 90 to a minimum of 10 over the five year period ending December 31, 2008.²¹ Panel B considers values of λ_σ from 0.8 to 1.6 to reflect a range of more or less conservative possibilities. The magnitudes of the volatility-timing alphas in Panel B are considerably larger than the market-timing alphas in Panel A. For example, when $\lambda_\sigma = 1.2$ and $b_y = 0.6$, the volatility-timing alpha of 0.46% per month is more than twice as large as any of the plausible calibrations in Panel A.

Table 1 therefore confirms our model-free calibration, and verifies an important new result. While prior authors have been aware of the theoretical possibility that a volatility-timing bias can arise from failing to account for conditional beta dynamics, no previous study has emphasized the quantitative significance of this channel.²²

One can obviously combine the biases in Panels A and B. For any individual investment strategy, these effects may reenforce or counteract one another depending on the signs of b_x and b_y and covariance in the state-variable innovations. The cumulative effects of market- and volatility-timing

¹⁹The model we use in calibration is a special case of the model estimated in Brandt and Kang, Table 3, Model A. From their estimates, the point estimate of the variability of the market risk premium is about $\lambda_M \approx 0.3$, with a two standard deviation range of about ± 0.15 . Similarly, $\lambda_\sigma \approx 1.2$ with a two standard deviation range of about ± 0.20 .

²⁰For example, when $\lambda_M = 0.5$ and $b_z = 0.6$, implying a 95% confidence interval for beta from -0.2 to 2.2 , the estimated underconditioning bias is 0.19% per month.

²¹The upper end of the 95% confidence interval for conditional volatility is given by $\bar{\sigma}_M \exp [1.96\lambda_\sigma/2 - \lambda_\sigma^2/4] \approx 2.26\bar{\sigma}_M$ and the lower end is $\bar{\sigma}_M \exp [-1.96\lambda_\sigma/2 - \lambda_\sigma^2/4] \approx 0.22\bar{\sigma}_M$.

²²Previous research has shown the importance of volatility-timing in other contexts. For example, Fleming, Kirby, and Ostdiek (2001, 2003) demonstrate significant value to volatility-timing in a portfolio choice context, and Busse (1999) shows that mutual fund managers engage in volatility timing. Carlson, Fisher, and Giammarino (2009) show that corporate seasoned equity offerings time low points of firm and market volatility.

can clearly be empirically meaningful.

3.1.2. Overconditioning Bias

To complete our specification of the returns R_{it} we permit nonlinearity:

$$R_{it} = \alpha_{it}^{t-1} + \beta_{it}^{t-1} R_{Mt} - \Delta_\beta \sigma_{Mt} \{|\varepsilon_{Mt}| - \mathbb{E}(|\varepsilon_{Mt}|)\} + \varepsilon_{it}, \quad (3.11)$$

where $\varepsilon_{it} \sim \mathcal{N}(0, \sigma_{it}^2)$ is independent of other variables and the parameter Δ_β has critical importance. When $\Delta_\beta = 0$, returns are linear. If $\Delta_\beta > 0$ the return is concave in the market return and if $\Delta_\beta < 0$ the relationship is convex, permitting a wide range of nonlinearities. Denote the down-market beta $\beta_{it}^- \equiv \text{Cov}(R_{it}, R_{Mt} | R_{Mt} < \bar{R}_{Mt}) / \text{Var}(R_{Mt} | R_{Mt} < \bar{R}_{Mt}) = \beta_{it}^{t-1} + \Delta_\beta$ and the analogous up-market beta $\beta_{it}^+ = \beta_{it}^{t-1} - \Delta_\beta$. We show:

Proposition 6. *The return on asset i satisfies*

$$R_{it} = \begin{cases} a_{it}^- + \beta_{it}^- R_{Mt} + \varepsilon_{it} & \text{if } R_{Mt} \leq \bar{R}_{Mt} \\ a_{it}^+ + \beta_{it}^+ R_{Mt} + \varepsilon_{it} & \text{if } R_{Mt} \geq \bar{R}_{Mt} \end{cases},$$

where $a_{it}^- = \alpha_{it}^{t-1} + \Delta_\beta [\sqrt{2/\pi} \sigma_{Mt} - \bar{R}_{Mt}]$ and $a_{it}^+ = \alpha_{it}^{t-1} + \Delta_\beta [\sqrt{2/\pi} \sigma_{Mt} + \bar{R}_{Mt}]$.

Thus, if an empiricist uses the up- and down- betas to risk adjust, the conditional intercept differs from the investor alpha α_{it}^{t-1} . The overconditioned intercepts a_{it}^- and a_{it}^+ are proportional to beta asymmetry, determined by the parameter Δ_β .

Similar biases occur if one overconditions by using contemporaneous realized beta. To see this, we calibrate the model. We normalize $\bar{\beta} = 1$, assume $\bar{R}_M = 0.0003$ (about 7.5% annually), $\bar{\sigma}_M = 0.01$ per day (about 16% per year), and $\sigma_{it} = \sigma_{Mt}$ (implying total volatility of 23% annually). This idiosyncratic volatility is small for an individual stock, but represents well the volatility of many portfolios. We consider a range of beta asymmetries $\Delta_\beta \in \{0, 0.2, 0.5, 1.0, -0.2\}$ consistent with the empirical results of ACX.²³ For long-short portfolios beta asymmetries may plausibly be larger.

For each model specification, we simulate 10^8 months of $n = 21$ daily returns and calculate unconditional alphas and betas from both monthly and daily returns. Following LN, we partition the data into windows of N months (nN days), for $N = 1, 3, 6$. In each window, indexed by θ , we run a market model regression $R_{it} = a_{i\theta} + \beta_{i\theta}^{CP} R_{Mt} + \varepsilon_{it}$, where the estimated loading

²³ ACX sort days within a year by whether the excess market return is below or above its within-year average, and run market model regressions on the subsets of “down” and “up” days. They sort companies by the difference between the down and up betas and find considerable dispersion. For the highest quintile, this beta asymmetry is almost one, roughly consistent with the specification $\Delta_\beta = \pm 0.5$.

$\beta_{i\theta}^{CP}$ is the contemporaneous realized beta. We estimate alpha using the buy-and-hold abnormal return within the window, rescaled to monthly units: $\alpha_{i\theta}^{CP} \equiv (R_{i\theta} - \beta_{i\theta}^{CP} R_{M\theta}) / N$.²⁴ We first set $\lambda_M = \lambda_\sigma = b_x = b_y = 0$. Returns are then iid, and the unconditional CAPM holds, which allows us to isolate overconditioning.

For different values of Δ_β , Table 2 shows the unconditional (UC) alphas, and the contemporaneously risk-adjusted alphas $\alpha_{i\theta}^{CP}$. When $\Delta_\beta = 0.5$, the overconditioned alpha is biased by 0.42 for monthly windows, 0.13 for quarterly windows, and 0.07 for semiannual windows, where we henceforth report all alphas in percent per month. Overconditioning thus has a large impact in small windows, and as N grows the conditional regressions converge to the unconditional case. Consistent with Proposition 4, the magnitude of the overconditioning bias when using contemporaneous realized beta risk-adjustment is linear in the beta asymmetry Δ_β .

3.1.3. The Overconditioning vs. Underconditioning Tradeoff

In the presence of both nonlinearities and time-variation in conditional beta, overconditioning and underconditioning can simultaneously be important. Using contemporaneous beta, as window size N increases the underconditioning problem (diluting information about conditional beta) becomes more important relative to overconditioning (measurement error due to nonlinearities). Alternatively, if state variables are persistent, using a lagged beta eliminates the possibility of overconditioning while capturing most of the useful information about conditional beta. We define alphas from lagged portfolio betas by $\alpha_{i\theta}^{LP} \equiv \frac{1}{N} [R_{i\theta} - \beta_{i\theta}^{LP} R_{M\theta}]$, where $\beta_{i\theta}^{LP} \equiv \beta_{i,\theta-1}^{CP}$.

Table 3 shows alphas from unconditional (UC), contemporaneous portfolio (CP), and lagged portfolio (LP) risk-adjustment. All specifications set $\varphi_x = \varphi_y = 0.9$, $\lambda_M = 0.3$, $\lambda_\sigma = 1.2$, equal to their approximate point estimates from Brandt and Kang (2004) and other literature.²⁵ The specifications thus capture realistic variability of the conditional mean and volatility of the market and persistence of the state variables. Across specifications, we vary b_x , b_y , and Δ_β , since assets generally differ in market-timing, volatility-timing, and payoff nonlinearities.

Cases 1-3 consider respectively an asset with only nonlinearities but no conditional beta dynam-

²⁴LN instead multiply $a_{i\theta}$ by the number of days in the month n and call this the conditional alpha. To distinguish the two approaches, we call $\alpha_{i\theta}^{CP}$ a “buy-and-hold” alpha, and $na_{i\theta}$ a “rescaled daily” alpha. Longstaff (1989) shows that when the CAPM holds for a given observation interval, it need not be satisfied at other horizons. In our model, the CAPM holds exactly at a daily frequency, but we find no practical distinction between buy-and-hold or rescaled daily results in the context of the model, and hence report only one set of results. In empirical data where microstructure effects are important, the distinction can be significant as we later discuss.

²⁵For all specifications considered in Brandt and Kang (2004), the persistence parameters for both the conditional mean and the conditional volatility are estimated to be within one standard deviation of 0.9, measured at a monthly frequency. Substantial related literature establishes that market returns are predictable primarily at low frequencies (e.g., Lettau and Ludvigson, 2001a; Cochrane, 2001), and that monthly stock market volatility is highly persistent (e.g., Schwert, 1989).

ics ($\Delta\beta = 0.5, b_x = b_y = 0$), only market-timing ($b_x = 0.5, \Delta\beta = b_y = 0$), and only volatility-timing ($b_y = 0.5, \Delta\beta = b_x = 0$). The results show that using lagged realized beta to risk adjust offers a useful compromise. The CP approach does well when there are no nonlinearities (cases 2 and 3), but suffers badly when nonlinearities are present. The UC alpha is unbiased when there are no conditional beta dynamics even in the presence of nonlinearities (case 1). The LP alpha does well in all three cases, eliminating the possibility of overconditioning, while also capturing most of the useful information about contemporaneous realized beta because the state variables are persistent.

Specifications 5-7 show the bias in each method when there is time-variation in conditional beta as well as payoff nonlinearities. The results confirm that both unconditional and overconditioned alphas can have substantial biases for the same asset, the UC due to underconditioning and the CP due to overconditioning. The lagged portfolio alpha robustly has low alphas in all cases, because it eliminates the possibility of overconditioning while capturing most of the useful information about conditional beta. In the final two specifications (8-9) we consider the possibility that the state variables X and Y are positively or negatively correlated. The correlation of the market risk premium and volatility can either enhance or diminish the underconditioning bias, depending on whether the signs of b_x and b_y are the same or opposite.

3.2. Instrumental Variables and Filtering

Using lagged beta as a risk proxy has the disadvantage that risk may predictably change between the prior and contemporaneous window, as noted by Chan (1988) and GM. To correct this problem, following Shanken (1990) and others, we specify the conditional return regression:

$$R_{i\tau} = \alpha_i^{IV1} + \beta_i \begin{bmatrix} 1 & Z_{\tau-1} \end{bmatrix}' R_{M\tau} + e_{i\tau}, \quad (3.12)$$

where β_i is a $1 \times (k+1)$ parameter vector, $Z_{\tau-1}$ is a $1 \times k$ instruments vector, τ indexes months, and α_i^{IV1} is the alpha.²⁶ We denote the conditional beta estimate $\beta_{i\tau}^{IV1} \equiv \beta_i [1 \ Z_{\tau-1}]'$. Common instruments $Z_{\tau-1}$ are risk premium predictors such as the dividend yield (DY), term spread (TS), T-bill rate (TB), and default spread (DS). Since we have shown that volatility-timing has a potentially larger impact than market-timing for performance measurement, instruments are also needed for joint movements in conditional beta and market volatility. We recommend in particular that lagged betas should be useful instruments in performance analysis.

Instrumenting with lagged betas in (3.12) differs importantly from proxying as in the last section.

²⁶The specification is identical to equation (4) in Ferson and Schadt (1996). Shanken permits the intercept to also be linear in the instruments, and notes that the coefficients are zero under the null. Ferson and Harvey (1999) use the coefficient restrictions in a time-varying intercept specification to reject the conditional Fama-French model.

Proxying is equivalent to setting $Z_{\tau-1} = \beta_{\tau-1}^{CP}$ with $\beta_i = [0 \ 1]$. By contrast, instrumenting allows conditional beta to be a linear combination of beta estimates, possibly from multiple prior windows, as well as other potential predictors in $Z_{\tau-1}$. Proxying suffers when beta changes predictably, as noted by Chan (1988) and GM, whereas the IV method corrects this problem.

We also use a closely related two-step IV method, restricting (3.12).²⁷ Consider the first-stage predictive regression $\beta_{i\tau}^{CP} = \gamma_{i0} + \gamma_{i1}Z'_{\tau-1} + \varepsilon_{i\tau}$, as in Ghysels and Jacquier (2006). The second-stage return regression specifies conditional beta to be linear in the fitted first-stage CP beta:

$$R_{i\tau} = \alpha_i^{IV2} + \left(\phi_{i0} + \phi_{i1}\hat{\beta}_{i\tau}^{CP} \right) R_{M\tau} + u_{i\tau}, \quad (3.13)$$

where the conditional beta estimate is $\beta_{i\tau}^{IV2} \equiv \phi_{i0} + \phi_{i1}\hat{\beta}_{i\tau}^{CP}$. If the fitted CP beta is an unbiased predictor of conditional beta then $\phi_{i0} + \phi_{i1} = 1$, and if it is efficient then additionally $\phi_{i0} = 0$. Estimating ϕ_{i0} and ϕ_{i1} is sensible when CP beta and conditional beta are correlated, but rescaling or translating offer potential for improvement. One important reason why rescaling can help is that the realized beta estimated from daily data tends to underestimate the appropriate loading at longer horizons such as one month (e.g., Dimson, 1979). Mitigating the microstructure bias in high-frequency realized betas is thus an additional benefit of using an IV approach.²⁸

We assess the IV alphas and two filtering approaches, using as instruments betas lagged one, two, and three months ($\beta_{\tau}^{LP1}, \beta_{\tau-1}^{LP1}, \beta_{\tau-2}^{LP1}$) and betas calculated from larger three- and six-month windows ($\beta_{\tau}^{LP3}, \beta_{\tau}^{LP6}$). For comparison, we also use the contemporaneous beta β_{τ}^{CP} . To provide benchmarks, the Appendix applies the extended Kalman filter to obtain optimally filtered beta estimates, as well as an approximate linear filter. From the exact filter, denote $\beta_{\tau|\tau-1}^{FX}$ the predicted beta using information up to the end of month $\tau - 1$, and $\beta_{\tau|\tau}^{FX}$ the estimate at the end of month τ . Similarly, $\beta_{\tau|\tau-1}^{FL}$ and $\beta_{\tau|\tau}^{FL}$ denote betas from the linear filter. We use a model specification that includes market- and volatility-timing and nonlinearities ($b_x = b_y = \Delta\beta = 0.5$), simulate data, and regress the known conditional beta on different instrument sets in order to assess accuracy from a beta estimation perspective. We then carry out the one-step and two-step IV procedures for each instrument set, including the filtered betas, to evaluate effectiveness in measuring alpha.

²⁷Whereas IV1 allows conditional beta to be any linear function of the instruments, IV2 requires the instruments to first be projected on the CP beta. As a consequence, under the null that the IV2 model is correct, the one-step coefficients are the product of the IV2 first stage parameters and the second stage coefficient.

²⁸In a model where there are no microstructure effects, enforcing restrictions on ϕ_{i0} and ϕ_{i1} in the second stage regression may be beneficial. The first-stage regression provides an unbiased forecast $\hat{\beta}_{i\tau}^{CP}$ of the daily realized beta, guaranteeing that the loading-mismeasurement bias is minimized. The second-stage regression introduces the possibility that average estimated beta can differ from the true conditional beta, and hence trades off the benefit of accounting for potential impacts of microstructure biases against the cost of any loading bias caused by volatility timing. We leave a quantitative investigation of the practical magnitude of this tradeoff for future research.

Table 4 gives results. In all cases, the IV1 and IV2 alphas are indistinguishable so we report only one set of IV alphas. With no instruments (regression 1), we reproduce the unconditional alpha bias from Table 3. In (2), the CP beta is a noisy but unbiased estimator of the true conditional beta (weights on the constant and β_{τ}^{CP} are 0.122 and 0.878 respectively), and produces a high R^2 of 0.878. Nonetheless, the alpha bias of 0.359 is substantial due to overconditioning. By contrast, the one month lagged beta is noisier (3), with a lower weight in the conditional beta regression (0.789) and a lower R^2 (0.710), but a substantially smaller conditional alpha (0.138).

Regressions (4-9) show that by using various combinations of lagged betas from different window sizes, one can marginally increase the overall R^2 of the conditional beta regression, and produce an alpha of 0.128.²⁹ All of the filtered beta estimates (regressions 10-13) are unbiased and efficient (coefficients in the beta regression of approximately 0 and 1). The linear filter with information up to $\tau - 1$ produces almost the same R^2 and alpha estimate as the regressions that use combinations of lagged betas, which is expected since the predictions of the linear filter are primarily driven by lagged realized betas. The $\tau - 1$ predictions of the exact filter give additional small improvements in R^2 and alpha. Incorporating information through the end of τ into the exact filter gives the best R^2 as it should, and an alpha very close to zero. We note that without direct access to the state variables used by investors, even exact filtering does not produce an R^2 of 1 or an alpha of 0 – there is no universal solution to the Hansen and Richard critique, even with exact filtering.

Linear filtering with information including τ increases the beta regression R^2 almost to the level of exact filtering, but achieves a similar alpha estimate to using lagged data only. This occurs because the linear filter is approximate, and while it increases useful information its filtering of the overconditioning problem is not complete, and a tradeoff between overconditioning and underconditioning remains.³⁰ In empirical settings, the exact structure of the underlying data-generating process is not known with certainty, making exact filtering considerably more complex. The results discussed in this subsection show, however, that the less model-specific linear filtering or IV procedures can substantially improve alpha estimates by capturing useful information without introducing a large overconditioning bias, providing an effective and practical empirical approach.

4. The Conditional Performance of Momentum Strategies

We introduce the momentum data including returns and betas, and then discuss some simple methodological adaptations that allow us to move from the single-asset simulation environment

²⁹For additional discussion of prediction using regressors derived from different window sizes, see, e.g., Ghysels, Santa-Clara, and Valkanov (2005), and in the specific context of betas Ghysels and Jacquier (2006).

³⁰In untabulated results, we verify that depending on parameters, incorporating contemporaneous information into the linear filter can either slightly improve or slightly worsen alpha estimates relative to using only lagged information.

of Section 3 to the empirical analysis of high-turnover momentum portfolios. We then present the empirical results, including conditional performance measures and decompositions of the alpha differences into market-timing, volatility-timing, and overconditioning components. We focus in the main text on conditional CAPM performance, and show in Appendix B that the methods are easily extended to conditional 3-factor model performance, giving robust results.

We consider three momentum strategies (Jegadeesh and Titman, 1993), denoted 6- d - h , with common 6 month formation periods over which stocks are sorted into winners and losers. We assume d months delay between the end of the formation period and the initial investment, after which stocks are held without rebalancing for h months. Details are given in Appendix B. The specific portfolios are 6-0-6, 6-1-1, and 6-1-6, which aids comparison with LN (6-0-6) and GM (6-1-1). We calculate the returns for winners (W) and losers (L) from January 1930 to December 2005, and their difference (WL). Our results are robust to variations of the portfolio rules.³¹

Table 5 reports mean returns and betas. Panel A shows means in excess of the T-bill rate for each portfolio at horizons of one day and one, three, and six months.³² The excess winner minus loser returns are large and positive across strategies and horizons. For example, the 6-0-6 average profit is $1.30 - 0.76 = 0.54$ percent at the one month horizon with similar rescaled quarterly and semi-annual averages. The 6-1-1 and 6-1-6 profits are respectively smaller and larger. The unconditional betas in Panel B are loadings from standard market model regressions using daily and monthly returns. We also report Dimson (1979) “sum” betas from daily data using the lag structure suggested by LN,³³ which helps to mitigate the effects of asynchronous trading. Dimson adjustment has a stronger impact on losers than winners, consistent with the lower liquidity of losers. The unconditional winner minus loser loadings are then negative for both daily and monthly horizons. The average contemporaneous portfolio (CP) betas reported in Panel B are calculated by forming non-overlapping windows of $N \in \{1, 3, 6\}$ months. Within each window we estimate a Dimson regression from daily momentum returns, and the contemporaneous beta is the sum beta. The winner minus loser portfolio has a lower average CP than UC beta, consistent with a UC loading bias generated by negative volatility-timing.

Panel C shows correlations between the formation-period market return and holding-period be-

³¹We confirm robustness to (i) including only NYSE and AMEX stocks, (ii) imposing a minimum price screen of \$1, (iii) restricting the sample period to January 1964-December 2005, and (iv) combinations of the above.

³²The daily mean is multiplied by the average number of trading days in one month. The approximate number of days in a month is 24.5 for months prior to 1952, and 21 thereafter, due to the end of Saturday trading. The overall average is approximately 22 days in a month. For a horizon of N months, we scale each mean return by $1/N$.

³³We run the regression $R_{it} = a_i + \beta_{i0}R_{M,t} + \beta_{i1}R_{M,t-1} + \beta_{i2}\sum_{k=2}^4 R_{M,t-k}/3 + \varepsilon_{it}$, where R_{it} and $R_{M,t}$ are respectively excess returns on portfolio i and the value-weighted CRSP index. The Dimson “sum” beta is $\beta_{i0} + \beta_{i1} + \beta_{i2}$. Our results are robust to using (i) no Dimson leads or lags, (ii) one lead of market returns in addition to lags, and (iii) alternatively using the Fowler and Rorke (1983) adjustment for asynchronous trading.

tas, market returns, and squared market returns. RU6 is the prior 6 month market return, which relates to the formation period market returns. Consistent with GM, RU6 strongly predicts winner and loser betas with positive and negative signs respectively. RU6 also negatively predicts future squared market returns, confirming a well-known regularity (e.g., French, Schwert, and Stambaugh, 1987). Combining these two phenomena generates negative covariation between holding-period momentum betas and market volatility. Following (3.2), such volatility-timing in momentum should cause an upward bias in the unconditional momentum alpha. Our conditional performance evaluation methods allow us to quantify the magnitude of this effect.

Panel D of Table 5 shows asymmetric betas. For the winner minus loser portfolios, the down-market betas are uniformly larger than the up-market betas, consistent with De Bondt and Thaler (1987) and Hong, Tu, and Zhou (2007). Following Section 3, such beta asymmetries will lead to substantial biases in alphas calculated from uninstrumented contemporaneous realized betas.

4.1. Empirical Implementation: Portfolios vs. Individual Stocks

To apply our empirical methods to high-turnover momentum portfolios, we adapt the procedures discussed previously. We distinguish between two types of empirical beta estimates for portfolios, the lagged portfolio (LP) and lagged component (LC) betas. The lagged portfolio beta is defined as in Section 3, treating the portfolio as a single asset by calculating the beta of the portfolio returns in a prior window: $\beta_{i\tau}^{LP} \equiv \beta_{i,\tau-1}^{CP}$. The LP beta does not account for changing portfolio weights, and if the portfolio beta changes due to changing composition (i.e., turnover of high vs. low beta stocks), the LP beta will not accurately reflect portfolio risk in the holding period.

Our lagged component betas, by contrast, are calculated as portfolio-weighted averages of beta estimates from the individual stocks that will be in the portfolio in the next period. The LC beta thus uses an important piece of investor information – the portfolio weights of individual stocks at the beginning of the investment period. To summarize the LC procedure, at the end of calendar month $\tau - 1$ we estimate betas of the individual stocks (components) that will belong to a portfolio in holding month τ , and sum over all components the product of (1) the estimated component beta and (2) the beginning of month τ component portfolio weight. Obviously, many different ways of estimating component betas are possible, for example by using different window sizes or daily vs. monthly data. Once LC betas are calculated, a single estimate may proxy for conditional beta, or alternatively one or more betas can be used as instruments in the IV approach, with weights optimally allocated to the instruments most informative about portfolio risk in the holding period.

In this study, we define $\beta_{i\tau}^{LC6}$ and $\beta_{i\tau}^{LC36}$ as the LC betas where component betas are calculated respectively from (i) six months of daily returns with loadings estimated as Dimson sum betas using

the LN lag structure; and (ii) thirty-six month windows of monthly returns. We have calculated betas many different ways and found no meaningful difference in alphas or R^2 when using any more than two LC betas calculated from different window sizes. We also consider as instruments standard predictors of the market premium: dividend yield (DY), term spread (TS), one-month T-bill rate (TB), and default spread (DS).³⁴ Finally, we include as potential instruments 6- and 36-month prior window market returns (RU6 and RU36), motivated by the observation in Table 5 that prior window market runup predicts both future beta and market volatility, consistent with prior literature (e.g., GM; French, Schwert, and Stambaugh, 1987).

Implementing filtering as in Section 3 must also account for the high turnover of momentum portfolios. It is beyond the scope of the present study to empirically implement exact nonlinear filtering for momentum.³⁵ We do however adapt an approximate linear filter, which has the potential to capture useful information in the contemporaneous realized beta that is not included in lagged instruments, while filtering out the part of contemporaneous beta that is correlated with the surprise in contemporaneous market returns. Following our Monte Carlo study, the simple-to-implement lagged beta, IV, and linear filtering methods should correct underconditioning biases related to market- or volatility-timing, while avoiding the overconditioning bias.

4.2. New Estimates of Momentum Alphas

We calculate momentum alphas from 1) proxy methods, using contemporaneous and lagged betas; 2) IV methods, using lagged betas and other instruments; 3) methods that incorporate contemporaneous and future information while instrumenting, including using contemporaneous realized beta while instead instrumenting for the market return, a two-sided kernel, and filtering. All methods that use instrumentation consistently reduce momentum alphas by 20-40% relative to unconditional, while overconditioned alphas are up to 2.5 times larger.

4.2.1. Proxy Methods

Table 6 shows CAPM alphas using empirical proxies for conditional beta. Column (i) reports unconditional alphas. Consistent with the negative market exposure of the winner minus loser

³⁴Dividend yield is computed following Fama and French (1988). Term spread is from Robert Shiller's web-site <http://www.econ.yale.edu/~shiller/data.htm>, measured at the end of the previous year. Default spread is the difference between BAA and AAA corporate bond yields, obtained from the Federal Reserve, <http://research.stlouisfed.org/fred2>. T-bill is the 30-day yield from CRSP.

³⁵Such an endeavor would require assumptions about the nonlinear data generating process for each individual stock entering and leaving the momentum portfolios, and would be much more computationally intensive than the single asset simulation study in Section 3. Moreover, a fully structural approach would require a complete specification of not only conditional beta dynamics for each stock, but also the dynamics of beta asymmetries.

portfolios, UC alphas increase relative to raw profits. At a one month horizon, the 6-0-6 strategy has an alpha of 0.57 for winners, -0.24 for losers, and a momentum alpha of 0.81.

Columns (ii) and (iii) use contemporaneous realized betas as proxies for conditional beta. In column (ii), alphas are calculated from rescaled daily average returns: i.e., $\alpha_{i\tau}^{CP, RD} = R_{i\tau}^{RD} - \beta_{\tau}^{CP} R_{M\tau}^{RD}$, where $R_{i\tau}^{RD}$ and $R_{M\tau}^{RD}$ are respectively the average daily return in window τ of asset i and the market. The alphas in column (iii) use buy-and-hold returns over the same interval: $\alpha_{i\tau}^{CP, BH} = R_{i\tau} - \beta_{\tau}^{CP} R_{M\tau}$. We view the RD alphas as unreliable because daily return averages can be highly impacted by microstructure effects, and in particular average daily returns of illiquid portfolios such as the loser side of the momentum strategy are biased downwards.³⁶ For example, the raw returns of the momentum strategy reported in Table 5 are substantially larger for daily returns than any other horizon. Rescaled daily-return alphas are the focus of recent studies using realized betas (e.g., LN; Li and Yang, 2008; Ang and Kristensen, 2009), and we report this measures simply for comparison purposes. The one-month horizon buy-and-hold alphas are easy to compare with the one-month unconditional alpha because both performance measures use identical returns $R_{i\tau}$, and differ only by their risk measures. For these reasons, we henceforth focus our discussion on the buy-and-hold alphas.

For the one-month horizon $N = 1$, the buy-and-hold CP alpha is substantially larger than unconditional (1.09 vs. 0.81 for 6-0-6). By contrast, the lagged portfolio alpha in column (v) is substantially smaller (0.47). If one believes that the state variables driving joint movements in conditional beta and market returns are persistent, this dramatic difference between CP and LP alphas suggests an overconditioning problem. At three- and six-month horizons, the difference between using a contemporaneous and lagged beta becomes gradually smaller. With $N = 3$ the contemporaneous portfolio alpha is 0.86 for winners minus losers while the lagged portfolio alpha is 0.50. For $N = 6$ the contemporaneous and lagged portfolio alphas are almost identical (0.61 vs. 0.58) and are both smaller than unconditional (0.81). These results can be explained by the tradeoffs between overconditioning and underconditioning (Sections 2 and 3) and the large beta asymmetry of momentum portfolios (Table 5). As window size increases ($N = 1, 3, 6$), the impact of overconditioning becomes smaller and the contemporaneous portfolio alphas fall dramatically (1.09, 0.86, 0.61). By contrast, larger windows worsen the underconditioning problem by blending months with different portfolio holdings and risk, and lagged portfolio alphas rise moderately with

³⁶In an earlier working paper (Boguth, Carlson, Fisher, and Simutin, 2007), we showed theoretically the impact of return autocorrelation on the ratio of rescaled-daily to buy-and-hold returns. RD returns are biased downward (upward) relative to BH returns when autocorrelations are positive (negative). Due to space constraints and to maintain focus, we have removed more detailed discussion of the RD bias from this paper and plan to treat this issue separately.

N (0.47, 0.50, 0.58). The pattern of these results is similar across all three momentum strategies, and consistent with the Monte Carlo experiment in Section 3.

Average lagged component (LC) alphas are reported in columns (vi) and (vii). For all strategies, the LC winner minus loser alphas are close to the one-month LP alphas. From Panel B, the 6-0-6 LC36 winner beta is smaller than the loser beta (1.30 vs. 1.37), which almost entirely explains the larger winner minus loser alpha for LC36 relative to LC6 (0.49 vs. 0.43). To mitigate concerns about microstructure biases impacting high-frequency betas, to make use of multiple beta estimates and other predictors, and to eliminate the possibility of predictable changes in risk from prior windows to the holding period, we next use lagged betas as instruments within the IV framework.

4.2.2. IV Alphas

Table 7 presents our IV risk-adjustment results for the 6-0-6 strategy. We fully report both stages of the two-step procedure (beta and return regressions), as well as the alpha and R^2 from the one-step method. In all cases, the IV1 and IV2 alphas are very close, and our discussion focuses on the two-step results. With no instruments (specification 1), the return-regression intercepts are the unconditional alphas previously reported in Table 6, and the net momentum alpha is 0.81.

In (2), the standard instruments help somewhat to predict beta ($R^2 = 2.85$ and 9.73 percent for winners and losers). The first-stage fitted beta is significant in the second-stage and efficiently predicts conditional beta for winners ($\phi_{i0}, \phi_{i1} = -0.12, 1.20$), while showing downward bias for losers since the sum of the two second-stage coefficients significantly exceeds one ($\phi_{i0} + \phi_{i1} = 0.51 + 0.79 > 1$). For winners and losers, the second-stage R^2 increases by approximately two percentage points relative to (1), and the alphas attenuate towards zero. Net alpha decreases from 0.81 unconditionally to 0.70 conditionally.

Our paper emphasizes that lagged-component betas are useful instruments for performance evaluation.³⁷ Relative to the standard instruments, the isolated LC6 and LC36 betas (3 and 4) substantially improve the first- and second-stage R^2 . For either instrument, the winner and loser alphas move toward zero and the net alpha falls to 0.65 and 0.64 respectively. Combining LC6 and LC36 (5), both are significant with LC6 more heavily weighted by about 2:1, and net alpha drops to 0.62. Adding the market runups RU6 and RU36 with the LC betas in (6) has a small impact on R^2 and slightly reduces the net alpha to 0.59. In untabulated results, the runups alone predict beta and reduce alpha, but their effect after controlling for LC betas is marginal. Combining all

³⁷Ghysels and Jacquier (2006) regress contemporaneous realized betas on lagged portfolio betas and other predictor variables, similar to our first stage beta regression. They do not consider lagged component betas, which may be justifiable given their focus on forecasting relatively low-turnover industry portfolio betas. Because of the emphasis of their paper, they also do not consider the second-stage performance regression.

instruments (7), the lagged component betas are stable and significant, the runups are driven out, and among the standard instruments only dividend yield for winners and T-bill rate for losers remain significant. Relative to (1), the momentum alpha falls by approximately 30% to 0.59.

We abbreviate the IV analysis for 6-1-1 and 6-1-6 strategies in Table 8. The results are qualitatively similar to those reported for 6-0-6, and the net alpha reductions from conditioning information are approximately 20% for 6-1-6 (0.93 vs. 1.14) and 40% for 6-1-1 (0.36 vs. 0.57).

To test the statistical significance of the alpha reductions from conditioning, we use the GMM procedure in Appendix B, which accounts for measurement error in the betas used as instruments. For each return regression in Tables 8 and 9, an asterisk beside the winner minus loser alpha denotes rejection of the null hypothesis that $\alpha^{IV} \geq \alpha^{UC}$ at the 5% level. With the standard instruments alone, the decrease in alpha is significant only for 6-1-1. In all specifications with at least one lagged-component beta as an instrument, the conditional alpha is significantly lower than the unconditional alpha for all strategies.

In untabulated results we added additional lagged betas to the regressions in Tables 7 and 8. For example, the LP6 beta, calculated from six months of daily lagged-portfolio returns, is significant independently, but is driven out when combined with LC6. In general, LC betas dominate LP betas, and the alphas reported in Tables 7 and 8 with LC6 and LC36 betas are robust to adding component or portfolio betas measured over other horizons.

The results of this section show that overconditioned alphas using uninstrumented CP betas are much larger than unconditional. By contrast, the IV procedure, which eliminates overconditioning while capturing predictable variation in conditional beta, decreases alpha by a statistically and economically significant 20-40% relative to unconditional.

4.3. Do Current or Future Returns Have Additional Information About Beta?

The IV alphas calculated so far use information known to be in the investor's information set, such as lagged betas and other instruments, to estimate conditional beta. One might wonder whether one can incorporate information from current or future returns into beta estimates and still get reliable alphas. We use three such methods: 1) using contemporaneous realized beta while instead instrumenting for market returns with predictors of the risk premium, 2) incorporating future realized betas into the IV alpha estimates; and 3) using linear filtering to extract additional information from contemporaneous returns while eliminating variations related to the market return innovation. We briefly discuss the methodological issues related to each approach and show that all three methods produce additional small declines in estimated momentum alphas.

4.3.1. Using Contemporaneous Beta and Instrumenting for Market Returns

As shown in Section 2, when one calculates alpha using contemporaneous realized beta as a proxy for conditional beta, i.e., $\alpha_{i\tau} \equiv R_{i\tau} - \beta_{i\tau}^{CP} R_{M\tau}$, an overconditioning bias occurs when payoffs are nonlinear because measurement error in $\beta_{i\tau}^{CP}$ covaries with the unexpected market return. We have thus far focused on solving this problem by using instruments for conditional beta that are known to investors ex ante, which eliminates any overconditioning bias.

If one wants to use the contemporaneous realized beta, an alternative that also removes the overconditioning bias is to instead instrument for the realized market return with predictors of the risk premium. That is, instrumenting for *either* the realized beta $\beta_{i\tau}^{CP}$ or $R_{M\tau}$ eliminates overconditioning. We correspondingly calculate the market-instrumented (MI) conditional alpha as $\alpha_{i\tau}^{MI} \equiv R_{i\tau} - \beta_{i\tau}^{CP} \hat{R}_{M\tau}$, where $\hat{R}_{M\tau}$ is the fitted value from a first-stage regression of market returns on instruments. In implementation, we use the full instrument set from Tables 7 and 8.

To place the MI alpha estimate in a broad context, Table 10 reports it alongside momentum performance estimates calculated previously in the paper. The key finding is that using contemporaneous realized beta and instead instrumenting for the market return gives almost the same alpha as instrumenting for conditional beta. In either case, instrumenting reduces the alpha by 20-40% below unconditional, and is much lower than the overconditioned alphas that instrument for neither market returns nor conditional beta.

4.3.2. Instrumenting with Future Realized Betas

To better understand the source of overconditioning in momentum portfolios, we consider using *future* realized betas as instruments for conditional beta. To do this, we define forward component (FC) betas analogously to lagged component betas. For each stock (component) i that will be in the portfolio in period τ , we estimate the component beta in a window starting immediately after the return interval τ . The FC beta is then the sum over i of the product of (i) the portfolio weight of stock i at the beginning of τ , and (ii) the component beta in the forward window.

We consider using as “instruments” in the standard IV regression (3.12) a combination of LC, CP, and FC betas, i.e., $\hat{\beta}_{i\tau} = w_{i0} + \sum_l w_{il} \beta_{i\tau}^{LCl} + w_{ic} \beta_{i\tau}^{CP} + \sum_f w_{if} \beta_{i\tau}^{FCf}$. Such an approach resembles the procedures of Li and Yang (2008) and Ang and Kristensen (2009) because when multiple betas are used the aggregate weights on returns may vary non-parametrically, and we expect returns closer to the date τ to receive heavier aggregate weights.

In this framework, two different types of overconditioning biases can occur, consistent with the discussion in Section 2. Using our previous notation, denote for $m \in \{LC, CP, FC\}$ the resid-

uals of the empirical beta estimates after projecting onto investor information, i.e., $\varepsilon_{\beta\tau}^m \equiv \beta_{\tau}^m - E(\beta_{\tau}^m | \mathcal{F}_{\tau-1})$. Obviously, LC betas cannot produce an overconditioning bias since they are in the investor information set and their projection residuals are zero: $\varepsilon_{\beta\tau}^{LC} = 0$, implying $Cov(\varepsilon_{\beta\tau}^{LC}, \varepsilon_{M\tau}) = 0$. However, neither CP nor FC betas are measurable under investor information, and either of their residuals $\varepsilon_{\beta\tau}^{CP}$ or $\varepsilon_{\beta\tau}^{FC}$ could covary with the market return surprise. We previously focused on the overconditioning bias from using a CP beta and showed that when the asset i return is nonlinear in the realized market return, which may occur under general factor models and is empirically likely for many assets, then $Cov(\varepsilon_{\beta\tau}^{CP}, \varepsilon_{M\tau}) \neq 0$ and the CP alpha is biased.

A different type of overconditioning bias can occur when using an FC beta. The identifying condition required for an FC beta to be a valid instrument, giving an unbiased alpha is $Cov(\varepsilon_{\beta\tau}^{FC}, \varepsilon_{M\tau}) = 0$. However, many models of asset price dynamics imply that the unexpected market return drives movements in future conditional beta, implying $Cov(\varepsilon_{\beta\tau}^{FC}, \varepsilon_{M\tau}) \neq 0$. For example, if the asset i is more levered than the market, then a surprise positive (negative) market return decreases (increases) conditional beta, implying $Cov(\varepsilon_{\beta\tau}^{FC}, \varepsilon_{M\tau}) < 0$ (Hamada, 1972). Similarly, if firm i possesses a growth option, then locally $Cov(\varepsilon_{\beta\tau}^{FC}, \varepsilon_{M\tau}) < 0$ (Carlson, Fisher, and Giammarino, 2004), and for a contraction option $Cov(\varepsilon_{\beta\tau}^{FC}, \varepsilon_{M\tau}) > 0$.

This discussion raises a challenging question for any study using a two-sided estimate of beta, even if only FC and not CP betas are included. Specifically, if an alpha estimate using FC betas differs significantly from an alpha using only lagged information, then how should this be interpreted? One possibility is that the FC beta contains information about a change in conditional beta not known in advance to investors and correlated with the market innovation $\varepsilon_{M\tau}$, in which case the one-sided kernel beta estimate is more appropriate. The other possibility is that the beta change was known in advance to investors, but the empiricist lacks sufficient instruments to capture it. Ang and Kristensen (2009) assume this problem away by specifying innovations in beta to be uncorrelated with market return surprises. To empirically address the issue is more difficult, requiring either identification of lagged instruments that predict the beta change, or else a structural approach where one can make inferences about predictability of beta changes, market return state variables, and any variables that drive payoff nonlinearities. The latter approach is beyond the frontier of current research due to our limited understanding of the joint dynamics of payoff nonlinearities with market returns and conditional betas, but can be investigated in future work.

Fortunately, this issue is not a problem in our empirical results. We calculate alphas using both LC and FC betas as instruments and report the results in Table 10, finding very little difference with IV alphas that use only LC betas as instruments. This result has an important implication: The overconditioning bias in momentum is due to a true nonlinearity, where the measurement error

in contemporaneous realized beta covaries with the surprise in market returns ($Cov(\varepsilon_{\beta\tau}^{CP}, \varepsilon_{M\tau}) \neq 0$). In order to believe that the uninstrumented CP alpha is correct, one would have to believe that the IV alphas using only lagged information are incorrect, *and also* that the IV alphas using *both* LC and FC betas are incorrect. We find such a scenario highly unlikely, and see no possible theoretical explanation for such a pattern.³⁸ By contrast, simple contemporaneous nonlinearities in momentum returns (De Bondt and Thaler, 1987; Hong, Tu, and Zhou, 2007), the predictability of momentum beta with formation-period market returns (GM), and the inverse predictability of market volatility with formation-period market returns (e.g., French, Schwert, and Stambaugh, 1987), combine to explain all of the momentum alphas we have calculated. In a world where all of these well-documented regularities are important, appropriate instrumentation gives the most accurate estimate of conditional performance.

4.3.3. Filtering

We can also extract marginal information about conditional beta β_{it}^{t-1} from the contemporaneous realized beta $\beta_{i\tau}^{CP}$, while removing variation linearly related to the surprise in market returns. To do so, we regress the forward-component beta $\beta_{i\tau}^{FC}$ on the standard IV beta that uses only lagged information, the contemporaneous realized beta, and the unexpected market return as inferred from a predictive regression of market returns on instruments:

$$\beta_{i\tau}^{FC} = \gamma_0 + \gamma_1 \beta_{i\tau}^{IV} + \gamma_2 \beta_{i\tau}^{CP} + \gamma_3 \varepsilon_{M\tau} + u_{i\tau}.$$

The fitted value $\hat{\beta}_{i\tau}^{FC}$ includes, in addition to the lagged information $\beta_{i\tau}^{IV}$, any useful information from $\beta_{i\tau}^{CP}$ while linearly removing measurement error in $\beta_{i\tau}^{CP}$ that covaries with the market surprise $\varepsilon_{M\tau}$. The fitted value $\hat{\beta}_{i\tau}^{FC}$ is a valid instrument (i.e., will not produce an overconditioning bias) in the standard IV regression (3.12) under the identifying assumption $Cov(\varepsilon_{\beta\tau}^{FC}, \varepsilon_{M\tau}) = 0$, which is consistent with the empirical results in the last section. Details of the linear filtering procedure, a thorough explanation of the identification requirements, and complete empirical results including an additional test of the identifying assumption are given in Appendix B.

Table 10 includes the final alpha estimates for the filtering procedure (FI) for all strategies. A strong message emerges. The filtering results are again very close to all other methods that

³⁸This belief would imply that, prior to date τ , investors have information unavailable to the empiricist that predicts movements in conditional beta and expected market returns i) in opposite directions at the beginning of window τ , and ii) that these movements will predictably reverse at the end of window τ . For example, investors would sometimes have to know that the market return would be unusually high over the next month while the conditional momentum beta is lower than forecast from a predictive regression, and that these patterns will predictably reverse at the end of the month. A much simpler and more plausible explanation for our empirical findings is that measurement error in realized beta is correlated with the unexpected market return due to payoff nonlinearities.

instrument for either the realized beta or the realized market return. The instrumented conditional alphas are consistently 20-40% below the unconditional alpha, and in many cases less than half the level of overconditioned estimates.

4.4. Decomposing the Alpha Biases

We now decompose the differences between alpha estimates into components due to market-timing, volatility-timing, and overconditioning. We take as a benchmark the one-stage instrumental variables alpha $\bar{\alpha}_i^{IV}$ calculated with the full set of instruments in Tables 7 and 8, and consider comparisons with: i) the unconditional alpha α_i^{UC} , and ii) the uninstrumented estimate $\bar{\alpha}_i^{CP}$.

We first break $\alpha_i^{UC} - \bar{\alpha}_i^{IV}$ into market- and volatility-timing components: $\alpha_i^{UC} - \bar{\alpha}_i^{IV} = \alpha_{iM}^{UC} + \alpha_{i\sigma}^{UC}$, where α_{iM}^{UC} and $\alpha_{i\sigma}^{UC}$ are defined as in (3.1) and (3.2) substituting $\beta_{i\tau}^{IV}$ for conditional beta. Nearly equivalently, we can separate according to uncentered first (M1) and second (M2) sample moments of the market: $\alpha_i^{UC} - \bar{\alpha}_i^{IV} = \alpha_{iM1}^{UC} + \alpha_{iM2}^{UC}$, where $\alpha_{iM1}^{UC} = (1 + \bar{R}_M^2/\sigma_M^2) Cov(\beta_{i\tau}^{IV}, R_{M\tau})$ and $\alpha_{iM2}^{UC} = -(\bar{R}_M/\sigma_M^2) Cov(\beta_{i\tau}^{IV}, R_{M\tau}^2)$ as in Proposition 1. The two decompositions are practically identical,³⁹ and we refer to both as separating into market- and volatility-timing components.

Table 10, Panel A shows the market- and volatility-timing components for each portfolio. Volatility timing has a larger impact than market timing in all cases. For example, in the 6-0-6 winner minus loser portfolio volatility-timing contributes 21 basis points per month, relative to a total alpha difference of 24 basis points per month. In all cases volatility-timing biases upward unconditional winner alphas and biases downward unconditional loser alphas. In other words, unconditional momentum alphas are inflated by volatility-timing.

The importance of volatility timing in momentum is consistent with prior evidence that the formation period market return (i) positively (negatively) predicts the winner (loser) holding-period beta (GM), and (ii) negatively predicts holding-period volatility (e.g., French, Schwert, and Stambaugh, 1987). Combining these regularities explains the negative (positive) volatility-timing for winners (losers) in Panel A, depicted in Figure 2.

We can also decompose the difference between the overconditioned alpha $\bar{\alpha}_i^{CP}$, formed from the contemporaneous realized beta, and the IV conditional alpha, following the logic of Section 2:

$$\begin{aligned}\bar{\alpha}_{i1}^{CP} - \bar{\alpha}_i^{IV} &= -Cov(\beta_{i\tau}^{CP-IV}, R_{M\tau}) - (\bar{\beta}_{i\tau}^{CP-IV}) \bar{R}_M \\ &= -Cov(\beta_{i\tau}^{CP-IV}, \varepsilon_{M\tau}) - Cov(\beta_{i\tau}^{CP-IV}, \bar{R}_{M\tau}) - (\bar{\beta}_{i\tau}^{CP-IV}) \bar{R}_M,\end{aligned}\quad (4.1)$$

³⁹The two decompositions differ only due to the allocation of a term involving covariation of beta with $\bar{R}_M^2$, i.e., $\alpha_{iM}^{UC} = \alpha_{iM1}^{UC} - (\bar{R}_M/\sigma_M^2) Cov(\beta_{i\tau}^{IV}, \bar{R}_{M\tau}^2)$. Empirically, $-(\bar{R}_M/\sigma_M^2) Cov(\beta_{i\tau}^{IV}, \bar{R}_{M\tau}^2) \approx 0$, as discussed in Section 2 and confirmed in Table 10.

where $\bar{R}_{M\tau} \equiv \mathbb{E}_{\tau-1} R_{M\tau}$, $\varepsilon_{M\tau} \equiv R_{M\tau} - \mathbb{E}_{\tau-1} R_{M\tau}$, and $\beta_{i\tau}^{CP-IV} \equiv \beta_{i\tau}^{CP} - \beta_{i\tau}^{IV}$. The first component captures the covariance of the noise in CP beta with the surprise in market returns, and reflects overconditioning. The second part measures covariation of the beta difference with the predictable part of market returns, and may represent useful performance-related information in the CP beta not captured by the IV beta. Our earlier results showed that linear filtering to extract additional information in the CP beta produced little change relative to the IV beta, and hence we expect the second term to be small. The final term is the difference in the average betas.

Panel B of Table 11 shows the decomposition (4.1) for all momentum portfolios. The covariance between unexpected market returns and the difference of CP and IV betas is large and negative for winners (-0.65), and almost zero for losers (-0.08), consistent with the beta asymmetries reported in Section 4. The main message of Panel B is reenforced in Figure 3, which shows that the overconditioning bias from using an uninstrumented contemporaneous realized beta is overwhelmingly explained by two facts: i) losers have similar loadings on negative and positive market news, and ii) winners are much more heavily exposed to negative versus positive market surprises.

5. Conclusion

We show that overconditioning – a new concept introduced in this paper – and volatility-timing can plausibly bias conditional CAPM alphas by several times more than market-timing. Empirically, negative volatility-timing substantially inflates the unconditional momentum alpha, and occurs because the formation-period market return positively predicts holding period beta, while also negatively predicting holding-period market volatility.

Attempting to incorporate information about time-varying risk by using a contemporaneous realized beta without instrumentation produces an overconditioning bias related to nonlinearity in the relation between asset and factor returns. In momentum, the overconditioning bias is large and inflates alpha because the strategy loads much more heavily on negative versus positive market news. We propose a variety of instrumental variables estimators, using lagged realized betas, two-sided kernels, the contemporaneous realized beta while instrumenting for the market return, and filtering. Our new estimates of momentum alpha correct the volatility-timing and overconditioning biases, and are significantly below both unconditional and overconditioned estimates.

We see several directions in which this literature should continue to develop. 1) More research can be done to understand the dynamics of return nonlinearities, in particular joint dynamics with conditional beta and the conditional mean and variance of market returns. Such advances would permit structural estimation of conditional factor model alphas, complementing the reduced form

approach of this paper. 2) We calculate alphas under the conditional CAPM, where there is no reward for payoff nonlinearities. If nonlinearities carry a premium (e.g., Kraus and Litzenberger, 1973; Bawa and Lindenberg, 1978) then our estimates of risk-adjusted momentum performance represent an upper bound – returns such as momentum that are concave in the realized market (equivalently negative coskewness) should command higher expected returns than predicted by the conditional CAPM, consistent with the results of Harvey and Siddique (2000). Future research can improve conditional performance evaluation for models where return nonlinearities earn a premium, for example by incorporating realized betas as instruments and attempting to better model the dynamics of return nonlinearities. 3) The concepts and methodologies developed in this paper in the context of a style portfolio are also applicable to managed portfolios, with additional challenges because of the incentives of the fund manager to manipulate alpha through unobservable dynamic trading strategies (Goetzmann, Ingersoll, Spiegel, and Welch, 2007). In particular, volatility-timing is easier to implement than market-timing, and has been documented in mutual funds (Busse, 1999). Given our result that volatility-timing has a larger potential impact on alpha than market-timing, the managed portfolio setting is a natural one in which to extend this research.

Appendix

A. Proofs of Propositions

The following lemma is used to simplify the proofs:

Lemma 1: For any three random variables X , Y , and Z ,

$$Cov(X, YZ) = Cov(XY, Z) - E(X)Cov(Y, Z) + E(Z)Cov(X, Y).$$

Proof of Lemma 1: $Cov(X, YZ) = E(XYZ) - E(X)E(YZ) = Cov(XY, Z) + E(XY)E(Z) - E(X)E(YZ)$, which leads to the result by applying the definition of covariance.

Proof of Proposition 1: To derive the beta bias, note that $\beta_i^{UC} = Cov(\beta_{it}^{t-1} R_{Mt}, R_{Mt}) / \sigma_M^2$. Applying Lemma 1, $\beta_i^{UC} = [\sigma_M^2 \bar{\beta}_i - \bar{R}_M Cov(\beta_{it}^{t-1}, R_{Mt}) + Cov(\beta_{it}^{t-1}, R_{Mt}^2)] / \sigma_M^2$, which reduces to the beta bias in the Proposition. The alpha bias follows.

Proof of Proposition 2: The conditional difference is

$$\begin{aligned} \mathbb{E}(\hat{\alpha}_{it} - \alpha_{it}^{t-1} | \mathcal{F}_{t-1}) &= \mathbb{E}\left(R_{it} - \hat{\beta}_{it} R_{Mt} - \mathbb{E}(R_{it} | \mathcal{F}_{t-1}) + \beta_{it}^{t-1} \bar{R}_{Mt} \middle| \mathcal{F}_{t-1}\right) \\ &= \beta_{it}^{t-1} \bar{R}_{Mt} - \mathbb{E}(\hat{\beta}_{it} | \mathcal{F}_{t-1}) \bar{R}_{Mt} - Cov(\hat{\beta}_{it}, R_{Mt} | \mathcal{F}_{t-1}). \end{aligned} \quad (A.1)$$

Rearranging and simplification gives the result. The unconditional result follows immediately from taking the unconditional expectation.

Proof of Proposition 3: Recall the definition $\alpha^s \equiv \mathbb{E}(R_i | S) - \beta^s \mathbb{E}(R_M | S)$. Taking expectations and using the CAPM yields:

$$\begin{aligned} \mathbb{E}(\alpha_i^S) &\equiv \bar{R}_i - \mathbb{E}(\beta_i^S) \bar{R}_M - Cov[\beta_i^S, \mathbb{E}(R_M | S)] \\ &= [\beta_i - \mathbb{E}(\beta_i^S)] \bar{R}_M - Cov[\beta_i^S, \mathbb{E}(R_M | S)]. \end{aligned}$$

If S contains information about R then without loss of generality assume $\mathbb{E}(R_M | G) > \mathbb{E}(R_M | B)$. Since $\beta_i^G \neq \beta_i^B$ then $Cov[\beta_i^S, \mathbb{E}(R_M | S)]$ will not be zero and $\mathbb{E}(\alpha_i^S)$ will generally be biased.

Proof of Proposition 4: Equation (2.2) provides a general form of the alpha decomposition:

$$\alpha^{UC} = Cov(\beta_{it}^{\tau-1}, R_{Mt}) - (\beta^{UC} - \bar{\beta}) \bar{R}_M = \alpha_{M,direct}^{UC} + \alpha_{loading}^{UC}.$$

The unconditional beta from Proposition 1 can be further decomposed:⁴⁰

$$\beta^{UC} = \bar{\beta} - (\bar{R}_M / \sigma_M^2) Cov(\beta_{it}^{\tau-1}, R_{Mt}) + Cov(\beta_{it}^{\tau-1}, \bar{R}_{Mt}^2) / \sigma_M^2 + Cov(\beta_{it}^{\tau-1}, \sigma_{Mt}^2) / \sigma_M^2. \quad (A.2)$$

In our model, all parts of these expressions can be analytically calculated using Stein's Lemma and repeated application of Lemma 1. Note that R_{Mt} can be expressed in terms of an *i.i.d.* standard

⁴⁰Equation (A.2) is equivalent to equation 2 in LN. Note that $Cov(\beta_{it}^{\tau-1}, R_{Mt}) = Cov(\beta_{it}^{\tau-1}, \bar{R}_{Mt})$ since ε_{Mt} is uncorrelated to $\beta_{it}^{\tau-1}$. Demeaning the market premium in the second term, $Cov(\beta_{it}^{\tau-1}, \bar{R}_{Mt}^2) = Cov(\beta_{it}^{\tau-1}, (\bar{R}_{Mt} - \bar{R}_M)^2) + 2\bar{R}_M Cov(\beta_{it}^{\tau-1}, \bar{R}_{Mt})$, immediately results in the LN relation: $\beta^{UC} = \bar{\beta} + (\bar{R}_M / \sigma_M^2) Cov(\beta_{it}^{\tau-1}, \bar{R}_{Mt}) + Cov(\beta_{it}^{\tau-1}, (\bar{R}_{Mt} - \bar{R}_M)^2) / \sigma_M^2 + Cov(\beta_{it}^{\tau-1}, \sigma_{Mt}^2) / \sigma_M^2$.

normal innovation η_t , i.e. $R_{Mt} = \bar{R}_{Mt} + \sigma_{Mt}\eta_t$.

$$\begin{aligned} Var(R_{Mt}) &= Var(\bar{R}_M e^{\lambda_M X_{\tau-1} - \frac{1}{2}\lambda_M^2} + \bar{\sigma}_M e^{\frac{1}{2}\lambda_\sigma Y_{\tau-1} - \frac{1}{4}\lambda_\sigma^2} \eta_t) \\ &= \bar{R}_M^2 (e^{\lambda_M^2} - 1) + \bar{\sigma}_M^2 \end{aligned} \quad (A.3)$$

$$\begin{aligned} Cov(\beta_{it}^{\tau-1}, R_{Mt}) &= Cov(\bar{\beta} + b_x X_{\tau-1} + b_y Y_{\tau-1}, \bar{R}_M e^{\lambda_M X_{\tau-1} - \frac{1}{2}\lambda_M^2} + \bar{\sigma}_M e^{\frac{1}{2}\lambda_\sigma Y_{\tau-1} - \frac{1}{4}\lambda_\sigma^2} \eta_t) \\ &= \lambda_M \bar{R}_M (b_x + b_y Cov(X, Y)) \end{aligned} \quad (A.4)$$

$$\begin{aligned} Cov(\beta_{it}^{\tau-1}, \bar{R}_{Mt}^2) &= Cov(\bar{\beta} + b_x X_{\tau-1} + b_y Y_{\tau-1}, \bar{R}_M^2 e^{\lambda_M (2X_{\tau-1} - \lambda_M)}) \\ &= 2\lambda_M e^{\lambda_M^2} \bar{R}_M^2 (b_x + b_y Cov(X_{\tau-1}, Y_{\tau-1})) \end{aligned} \quad (A.5)$$

$$\begin{aligned} Cov(\beta_{it}^{\tau-1}, \sigma_{Mt}^2) &= Cov(\bar{\beta} + b_x X_{\tau-1} + b_y Y_{\tau-1}, \bar{\sigma}_M^2 e^{\lambda_\sigma Y_{\tau-1} - \frac{1}{2}\lambda_\sigma^2}) \\ &= \lambda_\sigma \bar{\sigma}_M^2 (b_y + b_x Cov(X_{\tau-1}, Y_{\tau-1})) \end{aligned} \quad (A.6)$$

The beta bias is $\beta^{UC} - \bar{\beta} = \lambda_M (k_\mu / \sigma_M^2) (b_x + b_y Cov(X, Y)) + \lambda_\sigma (\bar{\sigma}_M^2 / \sigma_M^2) (b_y + b_x Cov(X, Y))$, where $k_\mu \equiv \bar{R}_M^2 (2e^{\lambda_M^2} - 1) > 0$ and $Cov(X, Y) = \rho_\varepsilon \sigma_x \sigma_y / (1 - \varphi_x \varphi_y)$.

Proof of Proposition 5: This decomposition isolates components that relate to covariation with $\bar{R}_{Mt}$ (market-timing) and σ_{Mt}^2 (volatility-timing). Substituting (A.3)–(A.6) into (3.1) and (3.2), noting that $Cov(\beta_{it}^{\tau-1}, \bar{R}_{Mt}^2) = 2\bar{R}_M e^{\lambda_M^2} Cov(\beta_{it}^{\tau-1}, R_{Mt})$, and simplifying gives (3.9) and (3.10).

Proof of Proposition 6: Conditioning on $R_{Mt} \geq \bar{R}_{Mt}$, the return on asset i is the random variable $\alpha_{it}^{\tau-1} + \beta_{it}^{\tau-1} R_{Mt} - \Delta_\beta \sigma_{Mt} \varepsilon_{Mt} + \Delta_\beta \sigma_{Mt} \sqrt{2/\pi} + \varepsilon_{it} = \alpha_{it}^{\tau-1} + \Delta_\beta (\sigma_{Mt} \sqrt{2/\pi} + \bar{R}_{Mt}) + (\beta_{it}^{\tau-1} - \Delta_\beta) R_{Mt} + \varepsilon_{it}$. The proof for $R_{Mt} < \bar{R}_{Mt}$ is analogous.

B. Details of Calculations and Empirical Procedures

B.1. The Overconditioning Example (Section 2.2, Figure 1)

We define four possible market payoffs $\bar{R}_M - kx/2$, where $k = -3, -1, 1, 3$. We solve for α_i^G and α_i^B and the returns of asset i in the four states. Conditional on each state, β_i^s and α_i^s are the conditional beta and intercept. The unconditional CAPM must hold $\bar{R}_i = \beta_i \bar{R}_M$. We impose $\beta_i = (\beta_i^B + \beta_i^G)/2$. Setting $Std(R_M) = .05$ and $\bar{R}_M = .01$, we get $R_M(\Omega) \equiv [-0.057, -0.012, 0.032, 0.077]$ and $R_i(\Omega) = [-0.068, -0.001, 0.044, 0.066]$.

B.2. The Link Between Loading Mismeasurement and Volatility Timing

To provide general intuition about why the loading measurement alpha is tightly linked to volatility timing, consider the following simple example. Suppose a constant market risk premium, which rules out market timing. At $\tau - 1$ investors observe one of two equally likely states $\{v, nv\} \in \mathcal{F}_{\tau-1}$ with market volatility strictly higher in the volatile than the non-volatile state $\sigma_v^2 > \sigma_{nv}^2$. Assume that in different market-volatility states, an asset's conditional betas $\{\beta_v, \beta_{nv}\}$ may differ. The unconditional beta then satisfies $\beta^{UC} = (\sigma_v^2 / \sigma_M^2) \beta_v / 2 + (\sigma_{nv}^2 / \sigma_M^2) \beta_{nv} / 2$. The unconditional beta exceeds average beta $\bar{\beta} = \beta_v / 2 + \beta_{nv} / 2$ if and only if $\beta_v > \beta_{nv}$, and is less than average beta if and only if $\beta_v < \beta_{nv}$. Thus, returns from high volatility periods are more influential in a simple OLS time-series regression. Positive volatility timing produces unconditional betas that overstate average portfolio risk and understate unconditional alphas, whereas negative volatility timing yields the opposite effect.⁴¹ Considering the high observed predictability of market volatility and the

⁴¹The volatility-timing example is easily generalized and relates to a much broader statistics literature that seeks to identify subsets of data with high influence. For example, a common measure of the influence that a given pair of

large alpha impacts of volatility timing demonstrated in Table 2, this general intuition shows that volatility timing should be taken into account whenever evaluating investment performance.

B.3. Exact Filtering and Linear Filtering (Section 3)

Exact Filtering: Let $\mathbb{X}_\tau \equiv [X_\tau, Y_\tau]$. We assume that $\mathbb{X}_0$ is drawn from the unconditional distribution of states, i.e., $[X_0, Y_0]$ are drawn as bivariate standard normals with correlation $\rho_u \equiv \rho_\varepsilon \sigma_x \sigma_y / (1 - \varphi_x \varphi_y)$. The empiricist observes R_{it} and R_{Mt} , $t > 0$ and updates conditional probabilities. Let $\mathcal{G}_t \equiv \{R_{is}, R_{Ms} | s \leq t\}$ and $G_\tau \equiv \{R_{is}, R_{Ms} | s \leq \tau(s)\}$. We filter using Bayes' rule, first using the state equation:

$$f(\mathbb{X}_1 | \mathcal{G}_0) = \int f(\mathbb{X}_1 | \mathbb{X}_0) f(\mathbb{X}_0 | \mathcal{G}_0) d\mathbb{X}_0 \quad (\text{B.1})$$

and then updating after observing new data:

$$\begin{aligned} f(\mathbb{X}_1 | \mathcal{G}_0, R_{M1}) &= f(R_{M1} | \mathbb{X}_1, \mathcal{G}_0) f(\mathbb{X}_1 | \mathcal{G}_0) / f(R_{M1} | \mathcal{G}_0) \\ &= f(R_{M1} | \mathbb{X}_1) f(\mathbb{X}_1 | \mathcal{G}_0) / \int f(R_{M1} | \mathbb{X}_1) f(\mathbb{X}_1 | \mathcal{G}_0) d\mathbb{X}_1, \end{aligned} \quad (\text{B.2})$$

$$\begin{aligned} f(\mathbb{X}_1 | \mathcal{G}_0, R_{M1}, R_{i1}) &= f(R_{i1} | \mathbb{X}_1, \mathcal{G}_0, R_{M1}) f(\mathbb{X}_1 | \mathcal{G}_0, R_{M1}) / f(R_{i1} | \mathcal{G}_0, R_{M1}) \\ &= f(R_{i1} | \mathbb{X}_1, R_{M1}) f(\mathbb{X}_1 | \mathcal{G}_0, R_{M1}) / \int f(R_{i1} | \mathbb{X}_1, R_{M1}) f(\mathbb{X}_1 | \mathcal{G}_0, R_{M1}) d\mathbb{X}_1. \end{aligned} \quad (\text{B.3})$$

For $t > 1$, the observation equations (B.2) and (B.3) are iterated within a month, while the state equation (B.1) is updated at the end of each month.

To implement the filtering equations we use recursive numerical integration as in Fridman and Harris (1998), building on Kitagawa (1987). We discretize the two-dimensional state-space $[X_\tau, Y_\tau]$ into 31 points in each dimension, and find little change in the accuracy of the filter or the associated alphas relative to coarser discretizations using as little as half as many integration points in each dimension. This confirms that the calculations are more than sufficient to approximate well the exact non-linear filtering equations.

Linear Filtering: We can also implement computationally simpler linear filtering using realized beta. Assume the following state equation and observation equation:

$$\beta_{\tau+1}^\tau = \rho \beta_\tau^{\tau-1} + (1 - \rho) \bar{\beta} + \varepsilon_{\tau+1} \quad (\text{B.4})$$

$$\beta_\tau^{CP} = \beta_\tau^{\tau-1} + a(R_{M\tau} - \bar{R}_{M\tau}) + e_t, \quad (\text{B.5})$$

where ε_t and e_t have mean zero, variances σ_ε^2 and σ_e^2 , and are independent. The state equation (B.4) applies exactly to the model in Section 3 if $\phi_z = \phi_y$ by setting $\rho = \phi_x = \phi_y$.⁴² The observation equation allows the noise in realized beta to be correlated with the market return surprise, as will occur when payoffs are nonlinear.

For given parameter vector $\{\rho, \bar{\beta}, a, \sigma_\varepsilon^2, \sigma_e^2\}$ we can recursively calculate the conditional probabilities of the unobserved state variable $\beta_\tau^{\tau-1}$. Let $E_{\tau|\tau^*}^\beta \equiv \mathbb{E}(\beta_\tau^{\tau-1} | G_{\tau^*})$ and $Var_{\tau|\tau^*}^\beta \equiv Var(\beta_\tau^{\tau-1} | G_{\tau^*})$.

independent and dependent variables have on slope coefficients is the statistic $DFBETAS_i$, (see, e.g., Belsley, Kuh, and Welch, 1980), which measures the degree to which the unconditional slope coefficient changes when datapoint i is dropped. This statistic is driven by the extent to which the dependent variable is an outlier and the size of the i^{th} diagonal of the $(X'X)^{-1}$ matrix associated with the independent variables.

⁴²When the state variables X and Y have different degrees of persistence, then additional lags of investor-conditioned beta can be added to the state equation (B.4).

We recursively update the system

$$\begin{aligned}
E_{\tau|\tau-1}^\beta &= \rho E_{\tau-1|\tau-1}^\beta + (1-\rho) \bar{\beta} \\
Var_{\tau|\tau-1}^\beta &= \rho^2 Var_{\tau-1|\tau-1}^\beta + \sigma_\varepsilon^2 \\
E_{\tau|\tau}^\beta &= \omega_\tau E_{\tau|\tau-1}^\beta + (1-\omega_\tau) (\beta_\tau^{CP} - a(R_{M\tau} - \bar{R}_{M\tau})) \\
\omega_\tau &= \sigma_e^2 / (Var_{\tau|\tau-1}^\beta + \sigma_e^2) \\
Var_{\tau|\tau}^\beta &= \omega_\tau^2 Var_{\tau|\tau-1}^\beta + (1-\omega_\tau)^2 \sigma_e^2,
\end{aligned}$$

which is initialized with $E_{1|0}^\beta = \bar{\beta}$, $Var_{1|0}^\beta = \sigma_\varepsilon^2 / (1-\rho^2)$.

In practice the parameter vector $\{\rho, \bar{\beta}, a, \sigma_\varepsilon^2, \sigma_e^2\}$ must be estimated which is easily implemented using GMM or simple least squares. Define the forecasting errors $u_\tau = \beta_\tau^{CP} - E_{\tau|\tau-1}^\beta$ whose variance is $Var(u_\tau) = Var_{\tau|\tau-1}^\beta + \sigma_e^2$. We choose the parameters $\{\rho, \bar{\beta}, a, \sigma_\varepsilon^2, \sigma_e^2\}$ to minimize the sum of squared forecasting errors u_τ^2 .⁴³

B.4. Momentum Portfolio Construction

At the beginning of calendar month τ , we sort stocks into deciles based on their return over the formation period $\tau-d-6$ to $\tau-d-1$. To be included in the sort, stocks must have (i) valid monthly returns on the CRSP database over the entire formation period, (ii) at least 12 additional valid monthly returns in the thirty months prior to formation, (iii) at least 15 non-missing daily returns in each month of the formation period. Immediately following the sort, the winner portfolio (W) makes a fixed \$1 investment with equal weights in the top decile stocks, and sells stocks that were added to the portfolio at the beginning of month $\tau-h$. The loser portfolio (L) is defined by similarly timed investments and liquidations in the bottom decile stocks. Momentum (WL) profits are the difference between W and L returns. The portfolios are seasoned by implementing the strategies with holding period h for $h-1$ months prior to the sample start date.

B.5. A GMM Test of the Difference in Alphas

We compare the alphas of a long-short position in portfolios $i = 1, 2$ under two different performance specifications $j = 1, 2$. Let $R_i = [1_T \quad \mathbf{X}_{ij}] [\alpha_{ij} \quad \beta_{ij}]' + \varepsilon_{ij}$, where 1_T , R_i , and ε_{ij} are column vectors of length T , α_{ij} are scalars, $\mathbf{X}_{ij}$ are T by $(k_{ij}-1)$ matrices, and β_{ij} are column vectors of size $k_{ij}-1$.

Define the moment conditions for asset i

$$g_i \equiv \mathbb{E} \begin{bmatrix} R_1 - \alpha_{11} - \mathbf{X}_{11}\beta_{11} \\ (R_1 - \alpha_{11} - \mathbf{X}_{11}\beta_{11})' \mathbf{X}_{11} \\ R_1 - \alpha_{12} - \mathbf{X}_{12}\beta_{12} \\ (R_1 - \alpha_{12} - \mathbf{X}_{12}\beta_{12})' \mathbf{X}_{12} \end{bmatrix},$$

the coefficient vector $b_i \equiv [\alpha_{11} \quad \beta_{11}' \quad \alpha_{12} \quad \beta_{12}']'$, and the matrix

$$d_i = \frac{\partial g_i}{\partial b} = \begin{bmatrix} \mathbf{D}_{11} & \mathbf{0}_{k_{11}, k_{12}} \\ \mathbf{0}_{k_{12}, k_{11}} & \mathbf{D}_{12} \end{bmatrix},$$

⁴³ An additional moment is needed to separately identify σ_ε^2 and σ_e^2 , but this has no impact on predicted or filtered betas in a long sample after conditional variance levels reach their steady state.

where $\mathbf{0}_{n1,n2}$ denotes a matrix of zeros of dimensions $n1$ by $n2$, and

$$\mathbf{D}_{ij} = - \begin{bmatrix} 1 & \mathbb{E}(\mathbf{X}_{ij}) \\ \mathbb{E}(\mathbf{X}'_{ij}) & \mathbb{E}(\mathbf{X}'_{ij}\mathbf{X}_{ij}) \end{bmatrix}$$

are symmetric squared matrices of size k_{ij} . Let $g = [g'_1 \ g'_2]'$, $b = [b'_1 \ b'_2]'$, and $d = \begin{bmatrix} d_1 & \mathbf{0} \\ \mathbf{0} & d_2 \end{bmatrix}$.

Using standard GMM results, $V \equiv \text{Var}(\hat{b}) = d^{-1}Sd^{-1}/T$, where $S = \sum_{k=-\infty}^{\infty} \mathbb{E}(u_t u'_{t-k})$ and $u_t = [\varepsilon_{11t} \ \varepsilon_{11t}\mathbf{X}_{11t} \ \varepsilon_{12t} \ \varepsilon_{12t}\mathbf{X}_{12t} \ \varepsilon_{21t} \ \varepsilon_{21t}\mathbf{X}_{21t} \ \varepsilon_{22t} \ \varepsilon_{22t}\mathbf{X}_{22t}]'$. We estimate $\hat{V} = d^{-1}\hat{S}d^{-1}/T$ following Newey and West (1987) with Bartlett kernel weights $\omega(k, m) = 1 - k/(m + 1)$. Let $\alpha_j \equiv \alpha_{1j} - \alpha_{2j}$. The test statistic $\hat{\alpha}_2 - \hat{\alpha}_1$ is asymptotically normally distributed with a mean of $\alpha_2 - \alpha_1$ and a variance of $c'Vc$, where $c \equiv [1 \ 0'_{k_{11}-1} \ -1 \ 0'_{k_{12}-1} \ -1 \ 0'_{k_{21}-1} \ 1 \ 0'_{k_{22}-1}]'$.

Applying this methodology to test the difference between conditional and unconditional momentum alphas, we set $R_1 = R_W$, $R_2 = R_L$, $\mathbf{X}_{11} = \mathbf{X}_{21} = R_M$, $\mathbf{X}_{12} = (R_M 1'_{k_{12}}) * [1_T \ Z_W]$, and $\mathbf{X}_{22} = (R_M 1'_{k_{22}}) * [1_T \ Z_L]$, where $*$ denotes element-by-element multiplication. To test the null hypothesis that the conditional alpha is greater than or equal to the unconditional alpha, we use a one-tailed test. We implement the Newey-West procedure with $m = 5$. Our results are unaffected by other choices of $m \leq 12$.

B.6. Filtering: Application to Momentum Strategies

Consider the predictive regression

$$\beta_{i\tau}^{FC} = \gamma_0 + \gamma_1 \beta_{i\tau}^{IV} + \gamma_2 \beta_{i\tau}^{CP} + \gamma_3 \varepsilon_{M\tau} + u_{i\tau} = \gamma_0 + \gamma_1 \beta_{i\tau}^{IV} + \gamma_2 (\beta_{i\tau}^{\tau-1} + \varepsilon_{\beta\tau}) + \gamma_3 \varepsilon_{M\tau} + u_{i\tau},$$

where $\varepsilon_{M\tau} \equiv R_{M\tau} - \mathbb{E}_{\tau-1}(R_{M\tau})$ and $\varepsilon_{\beta\tau}$ represents measurement error. Denoting the fitted values $\hat{\beta}_{i\tau}^{FC}$, we run the second-stage performance regression $R_{i\tau} = \alpha_i^{FI} + (\delta_0 + \delta_1 \hat{\beta}_{i\tau}^{FC}) R_{M\tau} + \nu_{i\tau}^{FI}$, where $\beta_{i\tau}^{FI} = \delta_0 + \delta_1 \hat{\beta}_{i\tau}^{FC}$ and α_i^{FI} are respectively the filtered conditional beta and mean alpha. The key to this procedure is the first-stage. The IV and CP betas should have useful information about the forward beta ($\gamma_1, \gamma_2 > 0$). However, the CP beta contains measurement error $\varepsilon_{\beta\tau}$ potentially correlated with the contemporaneous market innovation $\varepsilon_{M\tau}$. Under reasonable assumptions, discussed in detail below, including $\varepsilon_{M\tau}$ in the regression removes noise in the CP beta that is correlated with the market innovation and not useful for predicting the forward beta.

Table A.3. presents complete results. The IV beta used as an input in this table is obtained from the one-step instrumental variables procedure with all instruments (Tables 7 and 8). Panel A shows results for 6-0-6. Specification (1) imposes $\gamma_2 = \gamma_3 = 0$, and shows that the IV beta is an unbiased and efficient predictor of the forward beta ($\gamma_0 \approx 0, \gamma_1 \approx 1$), even though it uses no information from month τ . The second-stage alpha is by construction equal to the IV1 alpha of 0.57 (see Table 7 regression (7)). In (2), adding the CP beta slightly increases the R^2 , and the coefficients on both regressors are significant with weights favoring IV relative to CP by about 9:1. Using the CP beta introduces the possibility of overconditioning, and we cannot be sure to what extent the increase in the second-stage alpha (to 0.62) should be attributed to improved information about the conditional beta or overconditioning. In (3), the IV beta and market surprise $\varepsilon_{M\tau}$ are included but the CP beta is omitted ($\gamma_2 = 0$). The market surprise is not significant in the first-stage regression for either winners or losers. By contrast, in the full specification (4), the coefficient on $\varepsilon_{M\tau}$ is about zero for losers, but positive and significant for winners, revealing that when market returns are high, the forward beta is larger than suggested by the CP beta alone, and when market returns are low the opposite holds. Thus, measurement error in the CP beta negatively relates to the market return surprise, consistent with the strong beta asymmetry of the winner portfolio. Under the full

specification (4), the alpha falls to 0.48.

In Panels B and C we abbreviate the analysis for the 6-1-1 and 6-1-6 strategies by showing results only for the full specification. In both cases, filtering similarly produces a small drop in alpha relative to the base IV results from the previous section.

B.6.1. Identification Requirements

We establish conditions under which $\hat{\beta}_{i\tau}^{FC}$ is orthogonal to the market surprise $\varepsilon_{M\tau}$. Let

$$\beta_{i\tau}^{FC} \equiv E(\beta_{i\tau}^{FC} | \beta_{i\tau}^{\tau-1}) + \delta_{\beta\tau}^{FC} + \varepsilon_{\beta\tau}^{FC},$$

where $\delta_{\beta\tau}^{FC}$ is an innovation to conditional beta and $\varepsilon_{\beta\tau}^{FC}$ is measurement error (i.e., $\mathbb{E}(\delta_{\beta\tau}^{FC} | F_\tau) = \delta_{\beta\tau}^{FC}$ and $\mathbb{E}(\varepsilon_{\beta\tau}^{FC} | F_\tau) = 0$ under investor information F_τ). We show that $\text{corr}(\hat{\beta}_{i\tau}^{FC}, \varepsilon_{M\tau}) = 0$ if $\text{corr}(\delta_{\beta\tau}^{FC}, \varepsilon_{M\tau}) = 0$.

Consider the linear relation $\beta_{i\tau}^{FC} = \rho\beta_{i\tau}^{CP} + \hat{\gamma}\varepsilon_{M\tau} + u_{i\tau}^{FW}$, where ρ is any constant and $\hat{\gamma}$ is the OLS estimate. We allow that the measurement error in $\beta_{i\tau}^{CP}$ is correlated with $\varepsilon_{M\tau}$: $\beta_{i\tau}^{CP} = \beta_{i\tau}^{\tau-1} + a\varepsilon_{M\tau} + e_{i\tau}$. Using $\text{corr}(\delta_{\beta\tau}^{FC}, \varepsilon_{M\tau}) = 0$, we obtain $\text{var}(\varepsilon_{M\tau})\hat{\gamma} = \text{cov}(\beta_{i\tau}^{FC} - \rho\beta_{i\tau}^{CP}, \varepsilon_{M\tau}) = \text{cov}(-\rho\beta_{i\tau}^{CP}, \varepsilon_{M\tau}) = -\rho a(\text{var}(\varepsilon_{M\tau}))$. Thus, for any ρ , $\hat{\gamma}$ converges to $-\rho a$, which guarantees that $\hat{\beta}_{i\tau}^{FW}$ and in turn the filtered beta $\hat{\beta}_{i\tau}^{FI}$ are uncorrelated with the market surprise $\varepsilon_{M\tau}$. The procedure thus permits information in the CP beta to be incorporated into the performance measure while eliminating overconditioning.

The identifying assumption $\text{corr}(\delta_{\beta\tau}^{FC}, \varepsilon_{M\tau}) = 0$ is plausible for the FC beta. The empirical results obtained in Section 4.3.2 support this assumption. Further, regression (3) in Table A.3. indicates that the market surprise has little ability to independently predict the forward-component beta $\beta_{i\tau}^{FC}$ of momentum portfolios, consistent with the assumption.

B.7. The Conditional 3-Factor Model and Momentum Performance

To obtain conditional FF performance measures, in each non-overlapping window θ of length $N \in \{1, 3, 6\}$ months, we run a Fama-French daily regression for each factor $j \in \{MKT, HML, SMB\}$: $R_{it} = \alpha_{i\theta}^{CPRD}/n + \sum_j (\beta_{ij1\theta}F_{jt} + \beta_{ij2\theta}F_{j,t-1} + \beta_{ij3\theta} \sum_{k=2}^4 F_{j,t-k}/3) + \varepsilon_{it}$ using the same structure of Dimson adjustments as in our CAPM results.⁴⁴ Denoting $\beta_{ij\theta}^{CP} \equiv \beta_{ij1\theta} + \beta_{ij2\theta} + \beta_{ij3\theta}$ as the sum betas from this regression, we calculate buy-and-hold and rescaled daily alphas for contemporaneous and lagged betas analogously to the conditional CAPM.

We obtain partial time-series for the three factors and historical book equity values from Ken French's website. We create the pre-1963 daily factors following the procedure outlined by Fama and French (1993). Table A.1 shows for each momentum strategy the FF alphas obtained from the methods UC (column i), CP (ii-iii), LP (iv-v), and LC (vi-vii). Consistent with prior research (e.g., Fama and French, 1996), UC risk adjustment produces larger momentum alphas under the FF model than the CAPM. For 6-0-6 at a one month horizon, the FF winner alpha is lower than the CAPM alpha reported in Table 6 (0.44 vs. 0.57), the loser alpha is lower by a greater margin (-0.65 vs. -0.24), and the net WL alpha increases by 0.29 to 1.10. As in the conditional CAPM, overconditioning is a significant problem. For 6-0-6, the difference between CP and LP alphas exceeds 1.0 for one month windows, is about 0.6 for $N = 3$, and ranges from 0.05 to 0.36 for $N = 6$.

To calculate LC loadings and performance measures we again use either six months of lagged daily data (LC6) or 36 months of lagged monthly data (LC36) for each component in the W and

⁴⁴Eliminating the Dimson lags, or alternatively adding a lead to account for asynchronous trading delays in the relatively illiquid long side of HML and SMB, does not substantially alter our results.

L portfolios, using Dimson sum betas for LC6 as previously. The LC6 alphas (column vi) are moderately smaller than the LPBH alphas (0.71 vs. 0.79 for 6-0-6) and considerably smaller than UC (1.10). The LC36 method (vii) further reduces alphas.

To implement the IV method, we focus on one-step instrumentation. The conditional regression is $R_{i\tau} = \alpha_i^{IV1} + \sum_j \beta_{ij} F_{j\tau} ([1 \quad Z_{j,\tau-1}]') + \varepsilon_{i\tau}$, where $j \in \{MKT, HML, SMB\}$ are the Fama-French factors. Table A.2 presents results for 6-0-6. The unconditional regression (1) shows that loadings are uniformly larger for L than W (1.21 vs. 1.00 for MKT, 0.51 vs -0.06 for HML, and 1.52 vs. 0.88 for SMB). The standard instruments (2) appear especially useful for predicting HML loadings. The regression R^2 improve from 85.8 to 88.5 for winners, and from 82.3 to 83.4 for losers. The winner minus loser alpha falls to 0.97 from the unconditional 1.10. Instrumenting with the lagged component loadings (3), both LC6 and LC36 are always highly significant for W with roughly equal weightings for all factors. For L, the weightings are higher on LC36 than LC6, and the latter are insignificant for HML and SMB. Relative to (1) and (2), the R^2 improve considerably, increasing to 92.9 and 86.0 for W and L. The alphas further attenuate toward zero for winners and losers, and the winner minus loser alpha is 0.90. Combining the standard instruments and LC betas (4), the significance of the standard instruments generally moderates for SMB and HML, and is mixed for MKT. The lagged component coefficients are more stable. Relative to (3), the R^2 improves marginally for winners and is approximately constant for losers. The alphas for winners and losers attenuate slightly towards zero, and the winner minus loser alpha is 0.87.

Results for the 6-1-1 and 6-0-6 strategies are similar and are omitted for brevity. We conclude that proper use of conditioning information reduces three-factor momentum performance by a statistically significant 20% to 25%, while overconditioned estimates can overstate performance by more than 2.5 times.

References

- Andersen, Torben G., Tim Bollerslev, Francis X. Diebold, and Jin Wu, 2006, Realized Beta: Persistence and Predictability, *Advances in Econometrics*, edited by Thomas B. Fomby, Dek Terrell, vol. 20B, 1-39. Emerald Group Publishing.
- Andersen, Torben G., Tim Bollerslev, Francis X. Diebold, and Heiko Ebens, 2001, The distribution of realized stock return volatility, *Journal of Financial Economics* 61, 43-76.
- Ang, Andrew, and Joseph Chen, 2002, Asymmetric correlations of equity portfolios, *Journal of Financial Economics* 63, 443-494.
- Ang, Andrew, Joseph Chen, and Yuhang H. Xing, 2006, Downside risk, *Review of Financial Studies* 19, 1191-1239.
- Ang, Andrew, and Dennis Kristensen, 2009, Testing conditional factor models, Columbia University Working Paper.
- Avramov, Doron, and Tarun Chordia, 2006, Asset pricing models and financial market anomalies, *Review of Financial Studies* 19, 1001-1040.
- Bawa, Vijay S., and Eric B. Lindenberg, 1977, Capital market equilibrium in a mean-lower partial moment framework, *Journal of Financial Economics* 5, 189-200.
- Bekaert, Geert, and Campbell R. Harvey, 1995, Time-varying world market integration, *Journal of Finance* 50, 952-952.
- Belsley, David A., Edwin Kuh, and Roy E. Welsch, 1980, Regression diagnostics: Identifying influential data and sources of collinearity. New York, John Wiley.
- Berk, Jonathan, Richard Green, and Vasant Naik, 1999, Optimal investment, growth options, and security returns, *Journal of Finance* 54, 1553-1607.
- Black, Fisher and Myron Scholes, 1972, Valuation of option contracts and a test of market efficiency, *Journal of Finance* 27, 399-417.
- Boguth, Oliver, Murray Carlson, Adlai Fisher, and Mikhail Simutin, 2007, Conditional risk, overconditioning, and the performance of momentum strategies, Working Paper, University of British Columbia.
- Bollerslev, Tim, Robert F. Engle, and Jeffrey M. Wooldridge, 1988, A capital asset pricing model with time-varying covariances, *Journal of Political Economy* 96, 116-131.
- Brandt, Michael W., and Qiang Kang, 2004, On the relationship between the conditional mean and volatility of stock returns: A latent VAR approach, *Journal of Financial Economics* 72, 217-257.
- Brennan, Michael J., and Eduardo S. Schwartz, 1985, Evaluating natural-resource investments, *Journal of Business* 58, 135-157.
- Busse, Jeffrey A., 1999, Volatility timing in mutual funds: Evidence from daily returns, *Review of Financial Studies* 12, 1009-1041.
- Campbell, John Y., 1987, Stock returns and the term structure, *Journal of Financial Economics* 18, 373-399.
- Campbell, John and Ludger Hentschel, 1992, No news is good news - an asymmetric model of changing volatility in stock returns, *Journal of Financial Economics* 31, 281-318.

- Carlson, Murray, Adlai Fisher, and Ron Giammarino, 2004, Corporate investment and asset price dynamics: Implications for the cross-section of returns, *Journal of Finance* 59, 2577-2603.
- Carlson, Murray, Adlai Fisher, and Ron Giammarino, 2009, SEO risk dynamics, University of British Columbia Working Paper.
- Chamberlain, Gary, and Michael Rothschild, 1983, Arbitrage, factor structure, and mean-variance analysis on large asset markets, *Econometrica* 51, 1281-1304.
- Chan, K. C., 1988, On the contrarian investment strategy, *Journal of Business* 61, 147-163.
- Cochrane, John H., 1996, A cross-sectional test of an investment-based asset pricing model, *Journal of Political Economy* 104, 572-621.
- Cochrane, John H., 2001, *Asset Pricing*, Princeton University Press.
- Daniel, Kent, David Hirshleifer, and Avanidhar Subrahmanyam, 1998, Investor psychology and security market under- and overreactions, *Journal of Finance* 53, 1839-1885.
- De Bondt, Werner F.M., and Richard Thaler, 1985, Does the stock-market overreact?, *Journal of Finance* 40, 793-805.
- De Bondt, Werner F.M., and Richard Thaler, 1987, Further evidence on investor overreaction and stock-market seasonality, *Journal of Finance* 42, 557-581.
- De Santis, Giorgio, and Bruno Gerard, 1997, International asset pricing and portfolio diversification with time-varying risk, *Journal of Finance* 52, 1881-1912.
- Dimson, Elroy, 1979, Risk measurement when shares are subject to infrequent trading, *Journal of Financial Economics* 7, 197-226.
- Duffee, Gregory R., 2005, Time variation in the covariance between stock returns and consumption growth, *Journal of Finance* 60, 1673-1712.
- Dybvig, Philip H., and Stephen A. Ross, 1985, Differential information and performance measurement using a security market line, *Journal of Finance* 40, 383-399.
- Fama, Eugene F., and Kenneth R. French, 1988, Dividend yields and expected stock returns, *Journal of Financial Economics* 22, 3-25.
- Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3-56.
- Fama, Eugene F., and Kenneth R. French, 1996, Multifactor explanations of asset pricing anomalies, *Journal of Finance* 51, 55-84.
- Fama, Eugene F., and James D. MacBeth, 1973, Risk, return, and equilibrium: Empirical tests, *Journal of Political Economy* 81, 607-636.
- Ferson, Wayne E., and Campbell R. Harvey, 1991, The variation of economic risk premiums, *Journal of Political Economy* 99, 385-415.
- Ferson, Wayne E., and Campbell R. Harvey, 1993, The risk and predictability of international equity returns, *Review of Financial Studies* 6, 527-566.

Ferson, Wayne E., and Campbell R. Harvey, 1999, Conditioning variables and the cross section of stock returns, *Journal of Finance* 54, 1325-1360.

Ferson, Wayne E., Shmuel Kandel, and Robert F. Stambaugh, 1987, Tests of asset pricing with time-varying expected risk premiums and market betas, *Journal of Finance* 42, 201-220.

Ferson, Wayne E., and Rudi W. Schadt, 1996, Measuring fund strategy and performance in changing economic conditions, *Journal of Finance* 51, 425-461.

Fleming, Jeff, Chris Kirby, and Barbara Ostdiek, 2001, The economic value of volatility timing, *Journal of Finance* 56, 329-352.

Fleming, Jeff, Chris Kirby, and Barbara Ostdiek, 2003, The economic value of volatility timing using “realized” volatility, *Journal of Financial Economics* 67, 473-509.

Foster, Dean P., and Daniel B. Nelson, 1996, Continuous record asymptotics for rolling sample variance estimators, *Econometrica* 64, 139-174.

Fowler, David J., and C. Harvey Rorke, 1983, Risk measurement when shares are subject to infrequent trading: Comment, *Journal of Financial Economics* 12, 279-283.

French, Kenneth R., G. William Schwert, and Robert F. Stambaugh, 1987, Expected stock returns and volatility, *Journal of Financial Economics* 19, 3-29.

Fridman, Moshe, and Lawrence Harris, 1998, A maximum likelihood approach for non-Gaussian stochastic volatility models, *Journal of Business and Economic Statistics* 16, 284-91.

Ghysels, Eric, and Éric Jacquier, 2006, Market beta dynamics and portfolio efficiency, Working Paper, University of North Carolina at Chapel Hill.

Ghysels, Eric, Pedro Santa-Clara, and Rossen Valkanov, 2005, There is a risk-return trade-off after all, *Journal of Financial Economics* 76, 509-548.

Glosten, Lawrence R., Ravi Jagannathan, and David E. Runkle, 1993, On the relation between the expected value and the volatility of the nominal excess return on stocks, *Journal of Finance* 48, 1779-1801.

Goetzmann, William, Jonathan Ingersoll, Matthew Spiegel, and Ivo Welch, 2007, Portfolio performance manipulation and manipulation-proof performance measures, *Review of Financial Studies* 20, 1503-1546.

Grant, Dwight, 1977, Portfolio performance and the “cost” of timing decisions, *Journal of Finance* 32, 837-846.

Grinblatt, Mark, and Bing Han, 2005, Prospect theory, mental accounting, and momentum, *Journal of Financial Economics* 78, 311-339.

Grinblatt, Mark, and Sheridan Titman, 1985, Approximate factor structures: Interpretations and implications for empirical tests, *Journal of Finance* 40, 1367-1373.

Grundy, Bruce D., and J. Spencer Martin, 2001, Understanding the nature of the risks and the source of the rewards to momentum investing, *Review of Financial Studies* 14, 29-78.

Hamada, Robert S., 1972, Effect of firms capital structure on systematic risk of common stocks, *Journal of Finance* 27, 435-452.

Hansen, Lars Peter, and Scott F. Richard, 1987, The role of conditioning information in deducing testable restrictions implied by dynamic asset pricing models, *Econometrica* 55, 587-613.

- Harvey, Campbell R., 1989, Time-varying conditional covariances in tests of asset pricing models, *Journal of Financial Economics* 24, 289-317.
- Harvey, Campbell R., and Akhtar Siddique, 2000, Conditional skewness in asset pricing tests, *Journal of Finance* 55, 1263-1295.
- Henriksson, Roy, and Robert Merton, 1981, On market timing and investment performance II, statistical procedures for evaluating forecasting skills. *Journal of Business* 54, 513-533.
- Hong, Yongmiao, Jun Tu, and Guofu Zhou, 2007, Asymmetries in stock returns: Statistical tests and economic evaluation, *Review of Financial Studies* 20, 1547-1581
- Jacquier, Éric, Nicholas G. Polson, and Peter E. Rossi, 1994, Bayesian analysis of stochastic volatility models, *Journal of Business and Economic Statistics* 12, 371-389.
- Jagannathan, Ravi, and Robert Korajczyk, 1986, Assessing the market timing performance of managed portfolios, *Journal of Business* 59, 217-235.
- Jagannathan, Ravi, and Zhenyu Wang, 1996, The conditional CAPM and the cross-section of expected returns, *Journal of Finance* 51, 3-53.
- Jegadeesh, Narasimhan, and Sheridan Titman, 1993, Returns to buying winners and selling losers: Implications for stock-market efficiency, *Journal of Finance* 48, 65-91.
- Jensen, Michael C., 1968, Problems in selection of security portfolios: Performance of mutual funds in the period 1945-1964, *Journal of Finance* 23, 389-416.
- Johnson, Timothy, 2002, Rational momentum effects, *Journal of Finance* 57, 585-608.
- Kim, Sangjoon, Neil Shephard, and Siddhartha Chib, 1998, Stochastic volatility: Likelihood inference and comparison with ARCH models, *Review of Economic Studies* 65, 361-393.
- Kitagawa, Genshiro, 1987, Non-Gaussian state-space modeling of nonstationary time-series, *Journal of the American Statistical Association* 82, 1032-1063.
- Korajczyk, Robert A., and Ronnie Sadka, 2004, Are momentum profits robust to trading costs?, *Journal of Finance* 59, 1039-1082.
- Kraus, Alan, and Robert Litzenberger, 1976, Skewness preference and valuation of risky assets, *Journal of Finance* 31, 1085-1100.
- Lamoureux, Christopher G., and Guofu F. Zhou, 1996, Temporary components of stock returns: What do the data tell us?, *Review of Financial Studies* 9, 1033-1059.
- Leland, Hayne E., 1999, Beyond mean-variance: Performance measurement in a nonsymmetrical world, *Financial Analysts Journal* 55, 27-36.
- Lettau, Martin, and Sydney Ludvigson, 2001a, Consumption, aggregate wealth, and expected stock returns, *Journal of Finance* 56, 815-849.
- Lettau, Martin, and Sydney Ludvigson, 2001b, Resurrecting the (C)CAPM: A cross-sectional test when risk premia are time-varying, *Journal of Political Economy* 109, 1238-1287.
- Lewellen, Jonathan, and Stefan Nagel, 2006, The conditional CAPM does not explain asset-pricing anomalies, *Journal of Financial Economics* 82, 289-314.

- Li, Yan and Liyan Yang, 2009, Underconditioning and overconditioning: Testing the conditional CAPM and the conditional Fama-French three-factor model, Working Paper, Cornell University.
- Longstaff, Francis A., 1989, Temporal aggregation and the continuous-time capital asset pricing model, *Journal of Finance* 44, 871-887.
- McDonald, Robert, and Daniel Siegel, 1985, Investment and the valuation of firms when there is an option to shut down, *International Economic Review* 26, 331-349.
- McDonald, Robert, and Daniel Siegel, 1986, The value of waiting to invest, *Quarterly Journal of Economics* 101, 707-727.
- Newey, Whitney K., and Kenneth D. West, 1987, A simple, positive semidefinite, heteroskedasticity and autocorrelation consistent covariance-matrix, *Econometrica* 55, 703-708.
- Petkova, Ralitsa, and Lu Zhang, 2005, Is value riskier than growth?, *Journal of Financial Economics* 78, 187-202.
- Ross, Stephen A., 1976, The arbitrage theory of capital asset pricing, *Journal of Economic Theory* 13, 341-360.
- Rouwenhorst, K. Geert, 1998, International momentum strategies, *Journal of Finance* 53, 267-284.
- Sagi, Jacob, and Mark Seasholes, 2007, Firm specific attributes and the cross-section of momentum, *Journal of Financial Economics* 84, 389-434.
- Santos, Tano, and Pietro Veronesi, 2006, Labor income and predictable stock returns, *Review of Financial Studies* 19, 1-44.
- Schwert, G. William, 1989, Why does stock-market volatility change over time?, *Journal of Finance* 44, 1115-1153.
- Shanken, Jay, 1990, Intertemporal asset pricing: An empirical investigation, *Journal of Econometrics* 45, 99-120.
- Treynor, Jack L., and Kay K. Mazuy, 1966, Can mutual funds outguess market, *Harvard Business Review* 44, 131-136.
- Wang, Kevin Q., 2003, Asset pricing with conditioning information: A new test, *Journal of Finance* 58, 161-196.
- Wiggins, James B., 1987, Option values under stochastic volatility - theory and empirical estimates, *Journal of Financial Economics* 19, 351-372.

Table 1. THE UNCONDITIONAL ALPHA BIAS

A. Market-timing					
	b_x				
λ_M	0.2	0.4	0.6	0.8	1.0
0.1	0.01	0.03	0.04	0.05	0.06
0.2	0.03	0.05	0.08	0.10	0.13
0.3	0.04	0.08	0.11	0.15	0.19
0.4	0.05	0.10	0.15	0.20	0.25
0.5	0.06	0.13	0.19	0.25	0.31
B. Volatility-timing					
	b_y				
λ_σ	-0.2	-0.4	-0.6	-0.8	-1.0
0.8	0.10	0.20	0.30	0.40	0.50
1.0	0.13	0.25	0.38	0.50	0.63
1.2	0.15	0.30	0.45	0.60	0.76
1.4	0.18	0.35	0.53	0.71	0.88
1.6	0.20	0.40	0.60	0.81	1.01

Notes: This table reports unconditional alphas (in % monthly), calculated analytically as in Proposition 6, from the calibrated dynamic CAPM described in equations (3.5)-(3.8) and (3.11). The model parameters are $\bar{R}_M = 0.0003$, $\bar{\sigma}_M = 0.01$, and $\Delta_\beta = 0$. In Panel A, only information about the conditional mean of market returns is relevant, i.e., $\lambda_\sigma = 0$, while Panel B only considers information about the conditional market volatility, i.e., $\lambda_M = 0$.

Table 2. THE OVERCONDITIONING BIAS

Δ_β	Daily Betas		Alpha Bias (α^{RD} , %/m)				
	β^-	β^+	α_1^{UC}	$\bar{\alpha}_1^{CP}$	$\bar{\alpha}_3^{CP}$	$\bar{\alpha}_6^{CP}$	$\bar{\alpha}_{12}^{CP}$
0.00	1.00	1.00	0.00	-0.00	0.00	-0.00	-0.00
0.20	1.20	0.80	0.00	0.17	0.05	0.03	0.01
0.50	1.50	0.50	0.00	0.42	0.13	0.07	0.03
1.00	2.00	0.00	0.00	0.84	0.27	0.13	0.07
-0.20	0.80	1.20	0.00	-0.17	-0.05	-0.03	-0.01

Notes: This table shows the effects of changes in Δ_β in the unconditional version of the calibrated dynamic CAPM ($\lambda_M = \lambda_\sigma = b_x = b_y = 0$). The model parameters are $\bar{R}_M = 0.0003$, $\bar{\sigma}_M = 0.01$, $\sigma_i = \sigma_M$, and $\bar{\beta} = 1$. The reported statistics are obtained by simulating 10^8 months of $n = 21$ daily returns. The reported rescaled daily (RD) alphas are in percent per month. Risk adjustment is done either unconditionally (UC) or using the contemporaneous portfolio method (CP) using non-overlapping windows of length $N = 1, 3, 6$ or 12 months. The buy-and-hold (BH) alphas are virtually identical to the RD alphas and are not reported.

Table 3. THE OVERCONDITIONING VS. UNDERCONDITIONING TRADEOFF

	Market Parameters			Stock Parameters				N = 1	N = 3	N = 6
	λ_M	λ_σ	ρ_ϵ	b_x	b_y	Δ_β				
(1)	Overconditioning Only						UC	0.00		
	0.30	0.60	0.00	0.00	0.00	0.50	CP	0.42	0.14	0.07
							LP	0.00	0.00	0.00
(2)	Market-Timing Only						UC	0.09		
	0.30	0.60	0.00	0.50	0.00	0.00	CP	0.00	0.01	0.02
							LP	0.01	0.03	0.04
(3)	Volatility-Timing Only						UC	0.38		
	0.30	0.60	0.00	0.00	-0.50	0.00	CP	0.00	0.03	0.06
							LP	0.00	0.03	0.06
(4)	Market- and Volatility-Timing, No Nonlinearities						UC	0.47		
	0.30	0.60	0.00	0.50	-0.50	0.00	CP	0.00	0.04	0.08
							LP	0.01	0.06	0.10
(5)	Market- and Volatility-Timing, Nonlinearities						UC	0.47		
	0.30	0.60	0.00	0.50	-0.50	0.50	CP	0.35	0.16	0.14
							LP	0.01	0.06	0.10
(6)	Market- and Volatility-Timing, Nonlinearities, $\rho_\epsilon > 0$						UC	0.23		
	0.30	0.60	0.50	0.50	-0.50	0.50	CP	0.35	0.14	0.10
							LP	0.00	0.03	0.05
(7)	Market- and Volatility-Timing, Nonlinearities, $\rho_\epsilon < 0$						UC	0.70		
	0.30	0.60	-0.50	0.50	-0.50	0.50	CP	0.35	0.18	0.18
							LP	0.02	0.08	0.16

Notes: This table reports buy-and-hold alphas (in % monthly) from the calibrated dynamic CAPM with parameters $\bar{R}_M = 0.0003$, $\bar{\sigma}_M = 0.01$, $\sigma_i = \sigma_M$, and $\bar{\beta} = 1$. The persistence of the conditioning variables X and Y is $\varphi_x = \varphi_y = 0.9$. Estimates are obtained from 10^8 months of $n = 21$ daily returns. Alphas are rescaled to monthly equivalents for data windows of $N = 1, 3$, or 6 months, and the performance measures UC (unconditional), CP (contemporaneous risk adjustment), and LP (lagged risk adjustment).

Table 4. EVALUATION OF IV AND FILTERING METHODS

Conditional Beta Regression													
	b_0	β_τ^{CP}	β_τ^{LP1}	$\beta_{\tau-1}^{LP1}$	$\beta_{\tau-2}^{LP1}$	β_τ^{LP3}	β_τ^{LP6}	$\beta_{\tau \tau-1}^{FL}$	$\beta_{\tau \tau}^{FL}$	$\beta_{\tau \tau-1}^{FX}$	$\beta_{\tau \tau}^{FX}$	R^2	α^{IV}
(1)	0.998											0.000	0.469
(2)	0.122	0.878										0.878	0.359
(3)	0.210		0.789									0.710	0.138
(4)	0.162		0.608	0.229								0.733	0.128
(5)	0.151		0.592	0.187	0.070							0.735	0.127
(6)	0.202					0.838						0.680	0.147
(7)	0.167		0.496			0.354						0.733	0.128
(8)	0.268						0.809					0.568	0.190
(9)	0.165		0.499			0.326	0.027					0.733	0.128
(10)	0.000							0.999				0.738	0.128
(11)	0.002								0.998			0.913	0.128
(12)	0.001									0.999		0.758	0.105
(13)	0.000										1.000	0.938	-0.024

Notes: This table reports the results from the two-step instrumental variables conditioning method in the simulated data. The model parameters are $R_M = 0.0003$, $\bar{\sigma}_M = 0.01$, $\sigma_i = \sigma_M$, $\beta = 1$, $\lambda_m = 0.30$, $\lambda_\sigma = 0.60$, $\rho_\epsilon = 0$, $b_x = 0.50$, $b_y = -0.50$, and $\Delta_\beta = 0.50$. The persistence of the conditioning variables X and Y is $\varphi_x = \varphi_y = 0.9$. The first set of columns reports the intercept, coefficients, and R^2 statistics from regressions of true conditional beta on the instruments. The instruments considered include: contemporaneous beta (β_t^{CP}); monthly betas lagged one, two, and three months (β_τ^{LP1} , $\beta_{\tau-1}^{LP1}$, $\beta_{\tau-2}^{LP1}$); betas calculated from the prior three- and six-month windows (β_τ^{LP3} , β_τ^{LP6}); betas from the linear (β^{FL}) and exact (β^{FX}) filter using information only up to the end of month $\tau - 1$ or through the end of month τ . The final column shows the IV2 performance estimate (in % monthly) calculated as in equations (3.16) and (3.17). The IV1 estimates from equation (3.15) are indistinguishable and a separate column is not shown.

Table 5. MOMENTUM SUMMARY STATISTICS

	6-0-6			6-1-1			6-1-6				
	W	L	WL	W	L	WL	W	L	WL	R_M	R_M^2
A. Mean Returns											
daily	1.18	0.35	0.84	1.26	0.62	0.64	1.30	0.12	1.18	0.59	22.77
1 month	1.30	0.76	0.54	1.38	1.13	0.24	1.43	0.52	0.91	0.62	30.15
3 month	1.40	0.88	0.52	1.46	1.29	0.18	1.54	0.62	0.91	0.64	37.08
6 month	1.42	0.79	0.64	1.49	1.23	0.26	1.57	0.53	1.04	0.62	29.21
B. Unconditional and Average CP Betas											
β^{UC} , 1 month	1.17	1.61	-0.43	1.16	1.67	-0.52	1.20	1.57	-0.37	-	-
β^{UC} , daily	1.16	1.38	-0.22	1.15	1.41	-0.26	1.17	1.36	-0.19	-	-
β^{CP}	1.14	1.16	-0.01	1.14	1.19	-0.05	1.15	1.14	0.00	-	-
C. Correlation of CP Betas, R_M and R_M^2 with Formation Period Market Return											
Corr($\cdot$,RU6)	0.20	-0.15	0.30	0.26	-0.22	0.37	0.15	-0.10	0.21	0.01	-0.20
D. Asymmetric Betas											
β^-	1.51	1.31	0.20	1.52	1.29	0.23	1.50	1.31	0.19	-	-
β^+	0.89	1.19	-0.29	0.88	1.26	-0.38	0.93	1.15	-0.22	-	-
$\beta^- - \beta^+$	0.61	0.12	0.49	0.64	0.04	0.60	0.57	0.16	0.41	-	-

Notes: This table reports summary statistics for momentum portfolios and the market excess returns over the sample period from January 1930 to December 2005. Daily returns are computed as average daily returns scaled by average number of days in one month. 1-, 3-, and 6-month returns are computed by compounding monthly returns in overlapping windows of $N = 1, 3, 6$ months and dividing by N . Squared market returns are computed in a similar manner. Returns in Panel A are reported in percent, and squared market returns are in percent squared. All returns are in excess of the T-bill rate from Ken French's website. In Panel B, 1-month unconditional (UC) betas are the slope coefficients from market model regressions on monthly data. Daily unconditional betas are computed as the sum of the slope coefficients from regressing daily portfolio excess returns on market excess return, its lag, and the average of lags 2 through 4 of market excess return. β^{CP} are average loadings from market model regressions of daily returns in each calendar month, computed using the same lag structure as daily UC betas. Panel C shows correlations of 6-month ($\tau - 6$ to $\tau - 1$) market return with contemporaneous portfolio momentum betas, market return and squared market return in month τ . In Panel D, β^- and β^+ are sum of the slope coefficients from regressions of portfolio excess return on market excess return, its lag, and the average of lags 2 through 4 of market excess return. β^- and β^+ are calculated in every calendar year using daily data and then averaged.

Table 6. MOMENTUM CAPM ALPHAS: PROXY METHODS

		Conditional Models											
		Contemporaneous Portfolio (CP)						Lagged Portfolio (LP)					
		(ii)			(iii)			(iv)			(v)		
		W	L	WL	W	L	WL	W	L	WL	W	L	WL
		UC	L	WL	RD	BH	BH	RD	BH	BH	W	L	WL
		W	L	WL	W	L	WL	W	L	WL	W	L	WL
		A. Alphas											
6-0-6 Strategy													
1 month	0.57	-0.24	0.81	1.18	-0.25	1.43	1.24	0.15	1.09	0.77	0.54	0.07	0.47
3 month	0.56	-0.33	0.89	0.77	-0.39	1.16	0.86	0.00	0.86	0.80	0.59	0.09	0.50
6 month	0.55	-0.32	0.87	0.56	-0.34	0.90	0.65	0.04	0.61	0.88	0.64	0.06	0.58
6-1-1 Strategy													
1 month	0.66	0.09	0.57	1.29	-0.13	1.43	1.36	0.39	0.97	0.65	0.61	0.36	0.25
3 month	0.67	-0.05	0.72	0.87	-0.14	1.01	0.96	0.34	0.62	0.61	0.65	0.43	0.21
6 month	0.66	-0.01	0.67	0.58	-0.02	0.60	0.67	0.45	0.21	0.73	0.74	0.41	0.33
6-1-6 Strategy													
1 month	0.68	-0.46	1.14	1.26	-0.43	1.69	1.34	-0.06	1.39	1.08	0.66	-0.14	0.81
3 month	0.66	-0.54	1.20	0.90	-0.61	1.51	0.99	-0.24	1.23	1.13	0.70	-0.15	0.85
6 month	0.64	-0.53	1.17	0.71	-0.58	1.29	0.80	-0.22	1.02	1.21	0.76	-0.18	0.94
		B. LC Betas											
6-0-6 Strategy													
6-1-1 Strategy													
6-1-6 Strategy													
				1.23	1.22	0.01	1.30	1.37	-0.08				
				1.23	1.18	0.05	1.29	1.35	-0.06				
				1.23	1.23	0.00	1.30	1.37	-0.07				

Notes: This table reports the unconditional and conditional CAPM alphas, in percent per month, and lagged component (LC) betas of winners (W), losers (L), and winners minus losers (WL) of the three momentum strategies. N -month unconditional (UC) alphas are taken as intercepts from the regression $R_{i\theta} = N\alpha_i^{UC} + \beta_i R_{M,\theta} + \eta_{i\theta}$, where $R_{i\theta}$, $i \in \{W, L\}$, and $R_{M,\theta}$ are N -month portfolio and market excess returns. To perform contemporaneous portfolio (CP) and lagged portfolio (LP) risk adjustment, data is partitioned into non-overlapping windows of length $N \in \{1, 3, 6\}$ months indexed by θ , and the following regressions are run: $R_{i\theta} = \alpha_{i\theta}/n + \beta_{i\theta} R_{M,t-1} + (\beta_{i\theta}/3)\Sigma_{k=2}^4 R_{M,t-k} + \varepsilon_{it}$, where n is the average number of trading days in one month. CP rescaled daily (RD) alphas are averages over $\alpha_{i\theta}$ from these regressions. CP buy-and-hold (BH) alphas are averages over $(1/N)(R_{i\theta} - \beta_{i\theta}^{CP} R_{M,\theta})$, where $\beta_{i\theta}^{CP} = \beta_{i\theta} + \beta_{i10} + \beta_{i20}$. LP rescaled daily and buy-and-hold alphas are calculated by averaging over $(1/N)\Sigma_{t \in \theta}[R_{it} - \beta_{i\theta}^{LP} R_{M,t}]$ and $(1/N)[R_{i\theta} - \beta_{i\theta}^{LP} R_{M,\theta}]$, respectively, where $\beta_{i\theta}^{LP} = \beta_{i,\theta-1}^{CP}$. To estimate the month τ lagged component (LC) risk exposure, at the end of each calendar month $\tau - 1$, betas of the individual stocks (components) that will belong to a portfolio in month τ are computed. The component loading estimations use either daily returns from the beginning of $\tau - 6$ to the end of $\tau - 1$ with the Dimson lag structure; or monthly returns from $\tau - 36$ to $\tau - 1$. Portfolio loadings are calculated as the sum over components the product of (1) the component beta and (2) the beginning of month τ component portfolio weight. For lags $l = 6, 36$, the reported LC alphas are averages over $R_{i\tau} - \beta_{i\tau}^{LC} R_{M\tau}$. The sample period is from January 1930 to December 2005.

Table 7. MOMENTUM CAPM ALPHAS: IV METHOD, 6-0-6

		IV2 (Two-Step)										IV1					
		Stage 1										Stage 2					
		Beta Regression										Return Regression					
		γ_0	β^{LC6}	β^{LC36}	RU6	RU36	DY	TS	TB	DS	R^2	α^{IV2}	ϕ_0	ϕ_1	R^2	α^{IV1}	R^2
(1)	W	1.14										0.57		1.02	72.1	0.57	72.1
		[60]										[4.3]		[48]		[4.3]	
	L	1.16										-0.24		1.39	60.9	-0.24	60.9
		[50]										[-1.0]		[38]		[-1.0]	
	WL											0.81				0.81	
(2)	W	1.41					-4.58	1.13	-23.7	-1.89	2.85	0.53	-0.12	1.20	73.8	0.56	74.1
		[21]					[-3.4]	[0.7]	[-2.7]	[-0.6]		[4.1]	[-0.7]	[7.9]		[4.3]	
	L	1.39					-3.14	-7.93	-86.9	18.2	9.73	-0.17	0.51	0.79	62.7	-0.14	63.3
		[18]					[-2.0]	[-4.2]	[-8.4]	[5.3]		[-0.8]	[3.1]	[6.7]		[-0.6]	
	WL											0.70				0.69	
(3)	W	0.19	0.77								26.7	0.52	0.13	1.00	76.4	0.52	76.4
		[3.6]	[18]									[4.3]	[1.6]	[13]		[4.3]	
	L	0.31	0.70								23.0	-0.12	0.51	0.89	63.0	-0.12	63.0
		[5.5]	[17]									[-0.5]	[3.2]	[7.1]		[-0.5]	
	WL											0.65*				0.65*	
(4)	W	0.26		0.68							20.1	0.50	0.02	1.06	76.0	0.50	76.0
		[4.3]		[15]								[4.0]	[0.2]	[12]		[4.0]	
	L	0.06		0.80							17.9	-0.14	0.19	1.16	63.2	-0.14	63.2
		[0.7]		[14]								[-0.6]	[1.0]	[7.6]		[-0.6]	
	WL											0.64*				0.64*	
(5)	W	0.06	0.59	0.28							28.5	0.51	0.17	0.96	76.7	0.51	76.7
		[1.0]	[10]	[4.8]								[4.2]	[2.1]	[14]		[4.2]	
	L	0.10	0.53	0.30							24.1	-0.12	0.42	0.96	63.3	-0.12	63.4
		[1.3]	[8.6]	[3.7]								[-0.5]	[2.6]	[7.7]		[-0.5]	
	WL											0.62*				0.63*	
(6)	W	0.10	0.61	0.22	0.28	-0.01					28.7	0.50	0.21	0.94	77.0	0.48	77.1
		[1.5]	[11]	[3.3]	[2.1]	[-0.3]						[4.1]	[3.0]	[14]		[4.0]	
	L	0.22	0.50	0.28	-0.08	-0.13					24.4	-0.09	0.31	1.01	63.7	-0.10	64.1
		[2.4]	[7.9]	[3.5]	[-0.5]	[-2.3]						[-0.4]	[1.9]	[8.4]		[-0.5]	
	WL											0.59*				0.58*	
(7)	W	0.22	0.58	0.24	0.19	-0.04	-2.17	1.02	-1.17	-1.11	28.8	0.50	0.30	0.87	77.0	0.51	77.3
		[2.3]	[9.7]	[3.6]	[1.4]	[-0.8]	[-1.8]	[0.7]	[-0.2]	[-0.4]		[4.1]	[4.7]	[14]		[4.2]	
	L	0.41	0.40	0.32	-0.16	-0.08	-1.16	-2.68	-44.2	4.98	25.9	-0.09	0.30	0.98	64.1	-0.06	64.9
		[3.2]	[6.0]	[3.9]	[-1.0]	[-1.3]	[-0.8]	[-1.5]	[-4.4]	[1.4]		[-0.4]	[2.0]	[9.0]		[-0.3]	
	WL											0.59*				0.57*	

Notes: This table reports the results for the instrumental variables (IV) conditioning method under the 6-0-6 momentum strategy. The first set of columns gives estimates, t -statistics, and adjusted R^2 values from the first stage beta regression, $\beta_{i\tau}^{CP} = \gamma_{i0} + \gamma_{i1}Z_{\tau-1} + \varepsilon_{i\tau}$, where $i \in \{W, L\}$, τ indexes months, and instruments $Z_{\tau-1}$ include 6- and 36-month LC betas (β^{LC6} and β^{LC36}), 6- and 36-month market runup (RU6 and RU36), dividend yield (DY), term spread (TS), 30-day T-bill rate (TB), and default spread (DS). The second set of columns presents the results from the second stage return regression $R_{i\tau} = \alpha_i^{IV2} + (\phi_{i0} + \phi_{i1}\hat{\beta}_{i\tau}^{CP})R_{M\tau} + u_{i\tau}$. The third set of columns reports alphas and adjusted R^2 values from a single-step regression, $R_{i\tau} = \alpha_i^{IV1} + \beta_i[1 - Z_{\tau-1}]R_{M\tau} + \varepsilon_{i\tau}$. The performance measures α_i^{IV2} and α_i^{IV1} are in percent. Conditional winner minus loser IV alphas that are significantly smaller than UC alphas at the 5% level are marked with an asterisk. The sample period is from January 1930 to December 2005.

Table 8. MOMENTUM CAPM ALPHAS: IV METHOD, 6-1-1 AND 6-1-6

		IV2 (Two-Step)										IV1					
		Stage 1								Stage 2							
		Beta Regression								Return Regression							
		γ_0	β^{LC6}	β^{LC36}	RU6	RU36	DY	TS	TB	DS	R^2	α^{IV2}	ϕ_0	ϕ_1	R^2	α^{IV1}	R^2
A. 6-1-1 Strategy																	
(1)	W	1.14										0.66		1.01	68.9	0.66	68.9
		[57]										[4.6]		[45]		[4.6]	
	L	1.19										0.09		1.40	57.8	0.09	57.8
		[47]										[0.3]		[35]		[0.3]	
	WL											0.57				0.57	
(2)	W	1.43					-4.67	0.74	-27.9	-1.70	2.73	0.62	-0.11	1.18	70.6	0.64	70.7
		[20]					[-3.3]	[0.4]	[-3.0]	[-0.5]		[4.5]	[-0.6]	[7.2]		[4.6]	
	L	1.37					-2.32	-6.97	-83.3	18.5	7.90	0.15	0.52	0.81	59.6	0.20	60.3
		[16]					[-1.4]	[-3.3]	[-7.2]	[4.8]		[0.6]	[2.8]	[6.4]		[0.8]	
	WL											0.46*				0.44*	
(3)	W	0.07	0.59	0.27							31.4	0.61	0.25	0.90	74.0	0.60	73.9
		[1.2]	[11]	[4.8]								[4.7]	[3.5]	[13]		[4.7]	
	L	0.08	0.54	0.35							24.8	0.23	0.47	0.92	60.6	0.22	60.7
		[1.1]	[7.9]	[4.2]								[0.9]	[3.0]	[8.1]		[0.9]	
	WL											0.38*				0.39*	
(4)	W	0.22	0.59	0.24	0.17	-0.03	-2.37	0.91	1.14	-1.80	31.9	0.59	0.41	0.79	74.2	0.61	74.9
		[2.4]	[10]	[3.5]	[1.1]	[-0.6]	[-2.0]	[0.6]	[0.1]	[-0.6]		[4.6]	[6.9]	[14]		[4.8]	
	L	0.33	0.40	0.40	-0.08	-0.11	-0.51	-1.40	-45.1	5.28	26.9	0.23	0.41	0.88	61.3	0.26	61.9
		[2.4]	[5.5]	[4.6]	[-0.4]	[-1.6]	[-0.3]	[-0.7]	[-4.2]	[1.4]		[0.9]	[2.9]	[9.2]		[1.1]	
	WL											0.36*				0.35*	
B. 6-1-6 Strategy																	
(1)	W	1.15										0.68		1.05	73.0	0.68	73.0
		[62]										[5.1]		[50]		[5.1]	
	L	1.14										-0.46		1.37	60.9	-0.46	60.9
		[50]										[-2.0]		[38]		[-2.0]	
	WL											1.14				1.14	
(2)	W	1.41					-4.14	0.79	-24.2	-2.07	2.65	0.64	-0.21	1.30	74.7	0.68	75.3
		[22]					[-3.2]	[0.5]	[-2.8]	[-0.7]		[4.9]	[-1.2]	[7.9]		[5.3]	
	L	1.39					-3.34	-7.80	-86.8	18.0	9.80	-0.40	0.52	0.77	62.6	-0.37	63.1
		[18]					[-2.2]	[-4.2]	[-8.5]	[5.3]		[-1.8]	[3.1]	[6.5]		[-1.6]	
	WL											1.03				1.05	
(3)	W	0.07	0.58	0.28							29.7	0.63	0.18	0.96	77.7	0.62	77.7
		[1.3]	[10]	[4.9]								[5.2]	[2.4]	[14]		[5.1]	
	L	0.12	0.54	0.26							24.0	-0.34	0.38	0.99	63.4	-0.35	63.7
		[1.6]	[9.1]	[3.4]								[-1.5]	[2.4]	[8.0]		[-1.6]	
	WL											0.97*				0.97*	
(4)	W	0.22	0.58	0.25	0.19	-0.05	-1.82	0.66	-3.95	-1.24	29.9	0.61	0.30	0.88	78.0	0.64	78.4
		[2.4]	[9.7]	[3.9]	[1.5]	[-1.1]	[-1.6]	[0.5]	[-0.5]	[-0.5]		[5.1]	[4.5]	[14]		[5.3]	
	L	0.41	0.42	0.29	-0.08	-0.07	-1.20	-2.65	-42.2	5.06	25.6	-0.32	0.23	1.03	64.3	-0.30	65.1
		[3.3]	[6.4]	[3.7]	[-0.5]	[-1.2]	[-0.8]	[-1.5]	[-4.2]	[1.5]		[-1.5]	[1.6]	[9.3]		[-1.4]	
	WL											0.93*				0.94*	

Notes: This table reports the results for the instrumental variables (IV) conditioning method under the 6-1-1 and 6-1-6 momentum strategies. The first set of columns gives estimates, t -statistics, and adjusted R^2 values from the first stage beta regression, $\beta_{i\tau}^{CP} = \gamma_{i0} + \gamma_{i1}Z_{\tau-1} + \varepsilon_{i\tau}$, where $i \in \{W, L\}$, τ indexes months, and instruments $Z_{\tau-1}$ include 6- and 36-month LC betas (β^{LC6} and β^{LC36}), 6- and 36-month market runup (RU6 and RU36), dividend yield (DY), term spread (TS), 30-day T-bill rate (TB), and default spread (DS). The second set of columns presents the results from the second stage return regression $R_{i\tau} = \alpha_i^{IV2} + (\phi_{i0} + \phi_{i1}\hat{\beta}_{i\tau}^{CP})R_{M\tau} + u_{i\tau}$. The third set of columns reports alphas and adjusted R^2 values from a single-step regression, $R_{i\tau} = \alpha_i^{IV1} + \beta_i[1 - Z_{\tau-1}]R_{M\tau} + \varepsilon_{i\tau}$. The performance measures α_i^{IV2} and α_i^{IV1} are in percent. Conditional winner minus loser IV alphas that are significantly smaller than UC alphas at the 5% level are marked with an asterisk. The sample period is from January 1930 to December 2005.

Table 9. MOMENTUM CAPM ALPHAS: COMPARISON OF METHODS

Strategy	UC	Proxy Methods				Instrumented Methods			
		Overconditioned		Lagged		IV	MI	FC	FI
		CPRD	CPBH	LP	LC				
6-0-6	0.81	1.43	1.09	0.47	0.43	0.57	0.52	0.60	0.51
$p(\alpha \leq 0)$	0.000	0.000	0.000	0.013	0.016	0.002	0.013	0.001	0.005
$p(\alpha \leq \alpha^{UC})$		0.000	0.022	0.998	1.000	0.992	1.000	0.977	0.999
$p(\alpha \leq \alpha^{IV})$	0.008	0.000	0.000	0.869	0.982		0.682	0.157	0.968
6-1-1	0.57	1.43	0.97	0.25	0.15	0.35	0.31	0.33	0.22
$p(\alpha \leq 0)$	0.015	0.000	0.000	0.151	0.271	0.065	0.126	0.071	0.171
$p(\alpha \leq \alpha^{UC})$		0.000	0.008	0.984	0.999	0.972	0.998	0.974	0.999
$p(\alpha \leq \alpha^{IV})$	0.028	0.000	0.000	0.769	0.997		0.613	0.676	0.999
6-1-6	1.14	1.69	1.39	0.81	0.81	0.94	0.86	0.98	0.88
$p(\alpha \leq 0)$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
$p(\alpha \leq \alpha^{UC})$		0.000	0.031	0.998	0.999	0.980	1.000	0.941	0.995
$p(\alpha \leq \alpha^{IV})$	0.020	0.000	0.000	0.922	0.984		0.734	0.046	0.972

Notes: This table reports magnitude and significance of alphas from three momentum strategies. For each portfolio, the first row shows the following alphas (in percent monthly): UC is unconditional alpha; CPRD is rescaled daily contemporaneous portfolio alpha calculated using one-month windows; CPBH is buy-and-hold contemporaneous portfolio alpha calculated using one-month windows; LP is lagged portfolio buy-and-hold alpha calculated using one-month windows; LC is the lagged component (LC6) alpha; IV is alpha from a one-step instrumental variables approach; MI is market-instrumented alpha; FC is forecast component (FC6) alpha; and FI is alpha from the filtering approach. Alphas reported under instrumented methods use the full set of instruments: LC6, LC36, RU6, RU36, DY, TS, TB, and DS, and calculation of FC alpha uses as an additional instrument the forward component (β^{FC6}) beta. The bottom three columns for each strategy show p-values for the test that alpha in the column heading is less than or equal to either (1) zero, (2) the UC alpha, or (3) the IV alpha, respectively. The sample period is from January 1930 to December 2005.

Table 10. ALPHA BIAS DECOMPOSITION

A. Underconditioning											
Alpha Difference					Market Timing				Volatility Timing		
					M1 Timing			M2 Timing			
α_i^{UC}	-	$\bar{\alpha}_i^{IV}$	=	UC Bias	=	$\left(1 + \frac{\bar{R}_M^2}{\sigma_M^2}\right)Cov(\beta_{i\tau}^{IV}, R_{M\tau})$	+	$-\frac{\bar{R}_M}{\sigma_M^2}Cov(\beta_{i\tau}^{IV}, \bar{R}_{M\tau}^2)$	+	$-\frac{\bar{R}_M}{\sigma_M^2}Cov(\beta_{i\tau}^{IV}, \sigma_{M\tau}^2)$	
6-0-6 Strategy											
W	0.57	-	0.51	=	0.06	=	-0.01	+	0.00	+	0.07
L	-0.24	-	-0.06	=	-0.17	=	-0.04	+	0.00	+	-0.13
WL	0.81	-	0.57	=	0.23	=	0.03	+	0.00	+	0.20
6-1-1 Strategy											
W	0.66	-	0.61	=	0.04	=	-0.03	+	0.00	+	0.07
L	0.09	-	0.26	=	-0.17	=	-0.02	+	0.00	+	-0.15
WL	0.57	-	0.35	=	0.22	=	-0.01	+	0.00	+	0.22
6-1-6 Strategy											
W	0.68	-	0.64	=	0.04	=	-0.01	+	0.00	+	0.05
L	-0.46	-	-0.30	=	-0.16	=	-0.04	+	0.00	+	-0.12
WL	1.14	-	0.94	=	0.20	=	0.03	+	0.00	+	0.17
B. Overconditioning											
Alpha Difference					Overconditioning		Possible Underconditioning		Beta Difference		
					$\bar{\alpha}_i^{CP}$	-	$\bar{\alpha}_i^{IV}$	=	OC Bias	=	$-Cov(\beta_{i\tau}^{CP-IV}, \varepsilon_{M\tau})$
6-0-6 Strategy											
W	1.24	-	0.51	=	0.73	=	0.65	+	-0.01	-	-0.09
L	0.15	-	-0.06	=	0.22	=	0.08	+	-0.01	-	-0.14
WL	1.09	-	0.57	=	0.51	=	0.57	+	0.00	-	0.06
6-1-1 Strategy											
W	1.36	-	0.61	=	0.75	=	0.66	+	0.02	-	-0.08
L	0.39	-	0.26	=	0.13	=	-0.01	+	-0.01	-	-0.15
WL	0.97	-	0.35	=	0.62	=	0.66	+	0.03	-	0.07
6-1-6 Strategy											
W	1.34	-	0.64	=	0.70	=	0.63	+	-0.02	-	-0.09
L	-0.06	-	-0.30	=	0.24	=	0.10	+	0.00	-	-0.15
WL	1.39	-	0.94	=	0.46	=	0.53	+	-0.01	-	0.06

Notes: This table provides decompositions that demonstrate the magnitudes of the biases due to market-timing, volatility-timing, and overconditioning. Panel A provides two equivalent decompositions of the unconditional alpha bias $\alpha_i^{UC} - \bar{\alpha}_i^{IV}$ into either (i) the sum of market-timing bias $\alpha_M^{UC} = (1 + \bar{R}_M^2/\sigma_M^2)Cov(\beta_{i\tau}^{IV}, R_{M\tau}) - (\bar{R}_M/\sigma_M^2)Cov(\beta_{i\tau}^{IV}, \bar{R}_{M\tau}^2)$ and volatility-timing bias $\alpha_\sigma^{UC} = -(\bar{R}_M/\sigma_M^2)Cov(\beta_{i\tau}^{IV}, \sigma_{M\tau}^2)$, or (ii) the sum of biases due to covariance with with first and second uncentered moments of market return, $\alpha_{M1}^{UC} = (1 + \bar{R}_M^2/\sigma_M^2)Cov(\beta_{i\tau}^{IV}, R_{M\tau})$ and $\alpha_{M2}^{UC} = -(\bar{R}_M/\sigma_M^2)Cov(\beta_{i\tau}^{IV}, \bar{R}_{M\tau}^2) - (\bar{R}_M/\sigma_M^2)Cov(\beta_{i\tau}^{IV}, \sigma_{M\tau}^2) = -(\bar{R}_M/\sigma_M^2)Cov(\beta_{i\tau}^{IV}, R_{M\tau}^2)$. Panel B decomposes the difference between CP and IV alphas $\bar{\alpha}_i^{CP} - \bar{\alpha}_i^{IV}$ into the sum of three components: overconditioning, $-Cov(\beta_{i\tau}^{CP-IV}, \varepsilon_{M\tau})$, where $\beta_{i\tau}^{CP-IV} = \beta_{i\tau}^{CP} - \beta_{i\tau}^{IV}$ and $\varepsilon_{M\tau}$ is the unexpected market return; possible underconditioning, $-Cov(\beta_{i\tau}^{CP-IV}, R_{M\tau}^2)$; and the effect of the difference in average betas, $-\bar{\beta}_i^{CP-IV} \cdot \bar{R}_M$, where $\bar{\beta}_i^{CP-IV} = \bar{\beta}_i^{CP} - \bar{\beta}_i^{IV}$. CP betas are calculated in windows of one month. IV alphas and betas are from the IV1 approach that uses the full set of instruments: 6- and 36-month lagged component betas, 6- and 36-month market runup, dividend yield, term spread, T-bill rate, and default spread. The same variables are used as regressors to obtain unexpected market returns. The sample period is from January 1930 to December 2005.

Table A1. MOMENTUM 3-FACTOR MODEL ALPHAS: PROXY METHODS

Conditional Models																					
Contemporaneous Portfolio (CP)						Lagged Portfolio (LP)						Lagged Component (LC)									
(i)		(ii)		(iii)		(iv)		(v)		(vi)		(vii)									
UC		RD	L	BH	WL	W	RD	L	BH	W	6-month betas	W	36-month betas								
W	L	W	L	W	L	W	W	L	W	L	W	W	L								
A. Alphas																					
6-0-6 Strategy																					
1 month	0.44	-0.65	1.10	0.90	-1.17	2.07	0.88	-0.95	1.83	0.16	-0.89	1.06	0.28	-0.51	0.79	0.32	-0.39	0.71	0.02	-0.42	0.44
3 month	0.49	-0.65	1.15	0.46	-0.95	1.42	0.45	-0.74	1.20	0.17	-0.66	0.84	0.29	-0.25	0.54						
6 month	0.42	-0.63	1.05	0.25	-0.63	0.88	0.24	-0.44	0.69	0.20	-0.88	1.08	0.32	-0.46	0.77						
6-1-1 Strategy																					
1 month	0.51	-0.34	0.85	1.03	-1.10	2.13	1.06	-0.72	1.78	0.37	-0.78	1.15	0.49	-0.31	0.80	0.56	-0.04	0.60	0.25	-0.03	0.28
3 month	0.60	-0.44	1.03	0.56	-0.69	1.26	0.56	-0.40	0.96	0.28	-0.51	0.79	0.40	0.00	0.40						
6 month	0.57	-0.41	0.97	0.25	-0.26	0.52	0.24	0.01	0.23	0.26	-0.61	0.88	0.38	-0.10	0.48						
6-1-6 Strategy																					
1 month	0.54	-0.87	1.40	0.92	-1.30	2.22	0.91	-1.16	2.08	0.25	-1.06	1.31	0.37	-0.69	1.06	0.44	-0.64	1.07	0.11	-0.63	0.73
3 month	0.58	-0.83	1.41	0.58	-1.18	1.75	0.57	-0.99	1.56	0.27	-0.87	1.14	0.40	-0.48	0.87						
6 month	0.48	-0.79	1.26	0.40	-0.88	1.28	0.40	-0.71	1.11	0.32	-1.11	1.43	0.44	-0.71	1.15						

Notes: This table reports the unconditional and conditional Fama-French three-factor alphas, in percent per month, of winners (W), losers (L), and winners minus losers (WL) of the three momentum strategies. N -month unconditional (UC) alphas are average intercepts from the regression $R_{i\theta} = N\alpha_i^{UC} + \Sigma_j(\beta_{ij}F_{j\theta}) + \eta_{i\theta}$, where $R_{i\theta}, i \in \{W, L\}$ and $F_{j\theta}, j \in \{MKT, HML, SMB\}$ are N -month portfolio excess and factor returns, respectively. To perform contemporaneous portfolio (CP) and lagged portfolio (LP) risk adjustment, data is partitioned into non-overlapping windows of length $N \in \{1, 3, 6\}$ months indexed by θ , and the following regressions are run: $R_{it} = \alpha_{i\theta}/n + \Sigma_j[\beta_{ij\theta}F_{jt} + \beta_{ij1\theta}F_{j,t-1} + (\beta_{ij2\theta}/3)\Sigma_{k=2}^4 F_{j,t-k}] + \varepsilon_{it}$, where t refers to days and n is the average number of trading days in one month. CP rescaled daily (RD) alphas are the averages over $\alpha_{i\theta}$ from these regressions. CP buy-and-hold (BH) alphas are averages over $(1/N)\Sigma_j(R_{i\theta} - \beta_{ij\theta}^{CP}F_{j\theta})$, where $\beta_{ij\theta}^{CP} = \beta_{ij0\theta} + \beta_{ij1\theta} + \beta_{ij2\theta}$. Lagged portfolio rescaled daily and buy-and-hold alphas are calculated by averaging over $(1/N)\Sigma_{t \in \theta}\Sigma_j(R_{it} - \beta_{ij\theta}^{LP}F_{jt})$ and $(1/N)\Sigma_j(R_{i\theta} - \beta_{ij\theta}^{LP}F_{j\theta})$, respectively, where $\beta_{ij\theta}^{LP} = \beta_{ij\theta-1}^{CP}$. To estimate the month τ lagged component (LC) risk exposure, at the end of each calendar month $\tau - 1$, betas of the individual stocks (components) that will belong to a portfolio in month τ are computed. The component loading estimations use either daily returns from the beginning of $\tau - 6$ to the end of $\tau - 1$ with the Dimson lag structure; or monthly returns from $\tau - 36$ to $\tau - 1$. Portfolio loadings are calculated as the sum over components the product of (1) the component beta and (2) the beginning of month τ component portfolio weight. For lags $l = 6, 36$, the reported LC alphas are averages over $\alpha_{i\tau}^{LC} = R_{i\tau} - \Sigma_j(\beta_{ij\tau}^{LC}F_{j\tau})$. The sample period is from January 1930 until December 2005.

Table A2. MOMENTUM 3-FACTOR MODEL ALPHAS: IV, 6-0-6

		α^{IV1}	$MKT_{\tau \times}$				$HML_{\tau \times}$				$SMB_{\tau \times}$				R^2							
			1	$\beta_{MKT}^{LC\%}$	DY	TS	TB	DS	1	$\beta_{HML}^{LC\%}$	DY	TS	TB	DS		1	$\beta_{SMB}^{LC\%}$	DY	TS	TB	DS	
(1)	W	0.44	1.00					-0.06						0.88						85.8		
		[4.7]	[53]					[2.1]						[30]								
	L	-0.65	1.21					[12]						1.52						82.3		
		[-4.1]	[39]											[31]								
	WL	1.10																				
(2)	W	0.41	1.24		-5.65	1.96	-10.9	2.64	0.19		2.36	5.53	-36.1	-7.96	0.92		-4.78	5.18	74.4	-7.20	88.5	
		[4.7]	[23]		[-5.0]	[1.3]	[-1.2]	[1.1]	[2.3]		[1.6]	[2.2]	[-2.5]	[-2.3]	[9.7]		[-2.9]	[1.9]	[4.6]	[-2.5]		
	L	-0.55	1.37		-4.35	-4.00	-27.6	1.74	0.28		10.9	-2.48	-75.4	-6.77	1.20		4.54	-0.75	11.7	3.92	83.4	
		[-3.5]	[14]		[-2.1]	[-1.4]	[-1.7]	[0.4]	[1.9]		[4.2]	[-0.6]	[-2.9]	[-1.1]	[7.1]		[1.5]	[-0.2]	[0.4]	[0.7]		
	WL	0.97																				
(3)	W	0.38	0.10	0.53	0.38				-0.14	0.41	0.26				0.00	0.57	0.23				92.9	
		[5.7]	[1.5]	[8.1]	[6.0]				[-5.4]	[5.8]	[6.3]				[-0.1]	[8.5]	[3.7]					
	L	-0.52	-0.19	0.46	0.72				0.21	0.27	0.50				1.19	-0.03	0.28				86.0	
		[-3.7]	[-1.4]	[5.2]	[5.6]				[3.9]	[2.9]	[6.1]				[8.6]	[-0.3]	[2.1]					
	WL	0.90*																				
(4)	W	0.35	0.36	0.42	0.37	-2.79	-15.20	-1.38	0.02	0.37	0.32	-0.21	-7.89	-1.75	-0.01	0.34	0.38	1.67	-8.49	38.24	6.65	93.8
		[5.4]	[4.1]	[6.0]	[5.8]	[-3.2]	[2.1]	[-2.3]	[0.3]	[5.2]	[7.5]	[-0.2]	[-3.0]	[-0.2]	[-0.1]	[4.5]	[6.1]	[1.3]	[-3.8]	[3.1]	[3.2]	
	L	-0.51	-0.22	0.47	0.75	-0.35	-3.90	2.76	0.25	0.24	0.42	3.13	0.39	-48.07	0.96	0.00	0.31	4.33	0.57	-6.51	-4.11	86.1
		[-3.6]	[-1.2]	[4.9]	[5.5]	[-0.2]	[-1.0]	[0.5]	[1.9]	[2.6]	[4.8]	[1.2]	[0.1]	[-2.0]	[4.3]	[0.0]	[2.3]	[1.5]	[0.1]	[-0.2]	[-0.9]	
	WL	0.87*																				

Notes: This table presents the results for the single-stage instrumental variables conditioning method under the 6-0-6 momentum strategy using the Fama-French three-factor model. Reported are the estimates, t -statistics, and adjusted R^2 from the regression $R_{i\tau} = \alpha_i^{IV1} + \Sigma_j(\beta_{ij}F_{j\tau}[1 - Z_{j,\tau-1}]) + \eta_{i\tau}$, where $R_{i\tau}$, $i \in \{W, L\}$, and $F_{j\tau}$, $j \in \{MKT, HML, SMB\}$, are monthly portfolio excess and factor returns, and instruments $Z_{\tau-1}$ include 6- and 36-month LC betas for the three factors, and the standard instruments (DY, TS, TB, DS). Alphas are in percent. Conditional winner minus loser IV alphas that are significantly smaller than UC alphas at the 5% level are marked with an asterisk. The sample period is from January 1930 to December 2005.

Table A3. CAPM MOMENTUM ALPHAS: FILTERING METHOD

		Stage 1					Stage 2			
		γ_0	γ_1	γ_2	γ_3	R^2	α^{FI}	δ_0	δ_1	R^2
A. 6-0-6 Strategy										
(1)	W	0.03	1.00			41.8	0.50	-0.03	1.00	77.0
		[0.6]	[26]				[4.1]	[-0.4]	[14]	
	L	-0.02	0.99			36.9	-0.09	0.02	1.01	64.1
		[-0.3]	[23]				[-0.4]	[0.1]	[9.0]	
	WL						0.59			
(2)	W	0.07	0.88	0.10		43.2	0.57	-0.10	1.05	77.6
		[1.3]	[19]	[4.8]			[4.8]	[-1.1]	[15]	
	L	0.01	0.91	0.08		37.6	-0.07	-0.18	1.13	64.9
		[0.1]	[18]	[3.3]			[-0.3]	[-1.0]	[10]	
	WL						0.64			
(3)	W	0.03	1.00		0.18	41.8	0.45	-0.04	1.00	76.9
		[0.6]	[26]		[0.9]		[3.7]	[-0.4]	[14]	
	L	-0.02	0.99		-0.07	36.9	-0.08	0.04	1.00	64.0
		[-0.3]	[23]		[-0.3]		[-0.3]	[0.2]	[8.9]	
	WL						0.52			
(4)	W	0.07	0.87	0.12	0.43	43.4	0.45	-0.10	1.04	77.2
		[1.4]	[19]	[5.2]	[2.2]		[3.7]	[-1.1]	[14]	
	L	0.01	0.91	0.08	-0.05	37.5	-0.06	-0.17	1.13	64.8
		[0.1]	[18]	[3.3]	[-0.2]		[-0.3]	[-0.9]	[10]	
	WL						0.51			
B. 6-1-1 Strategy										
	W	0.02	0.90	0.12	0.52	42.4	0.52	-0.09	1.05	74.4
		[0.3]	[18]	[5.2]	[2.4]		[4.1]	[-1.0]	[14]	
	L	-0.20	1.03	0.08	-0.21	42.4	0.30	0.05	0.99	61.9
		[3.0]	[20]	[3.4]	[-0.7]		[1.2]	[0.3]	[10]	
	WL						0.22			
C. 6-1-6 Strategy										
	W	0.07	0.87	0.11	0.37	43.8	0.58	-0.12	1.07	78.4
		[1.3]	[19]	[4.9]	[1.9]		[4.8]	[-1.3]	[15]	
	L	0.09	0.86	0.08	0.01	36.0	-0.30	-0.27	1.19	65.1
		[1.5]	[18]	[3.1]	[0.0]		[-1.4]	[-1.5]	[10]	
	WL						0.88			

Notes: This table reports the results of the filtering methods. First stage regression is

$$\beta_{\tau}^{FC} = \gamma_0 + \gamma_1 \beta_{\tau}^{IV} + \gamma_2 \beta_{\tau}^{CP} + \gamma_3 \varepsilon_{M\tau} + u_{\tau},$$

where β_{τ}^{FC} is the 12-month forward component beta; β_{τ}^{CP} is computed using one month of daily data; β_{τ}^{IV} is calculated using the one-step instrumental variables approach with LC6, LC36, RU6, RU36, DY DY, DS, TB, and TS as instruments; and unexpected market return $\varepsilon_{M\tau}$ is the residual from regressing $R_{M\tau}$ on the same set of instruments. The second stage performance regression uses the fitted beta $\hat{\beta}_{\tau}^{FC}$ from the first stage:

$$R_{\tau} = \alpha^{FI} + \left(\delta_0 + \delta_1 \hat{\beta}_{\tau}^{FC} \right) R_{M\tau} + \nu_{\tau}.$$

Regressions are run separately for winners and losers, and t -statistics are in square brackets. The sample period is 1930-2005.

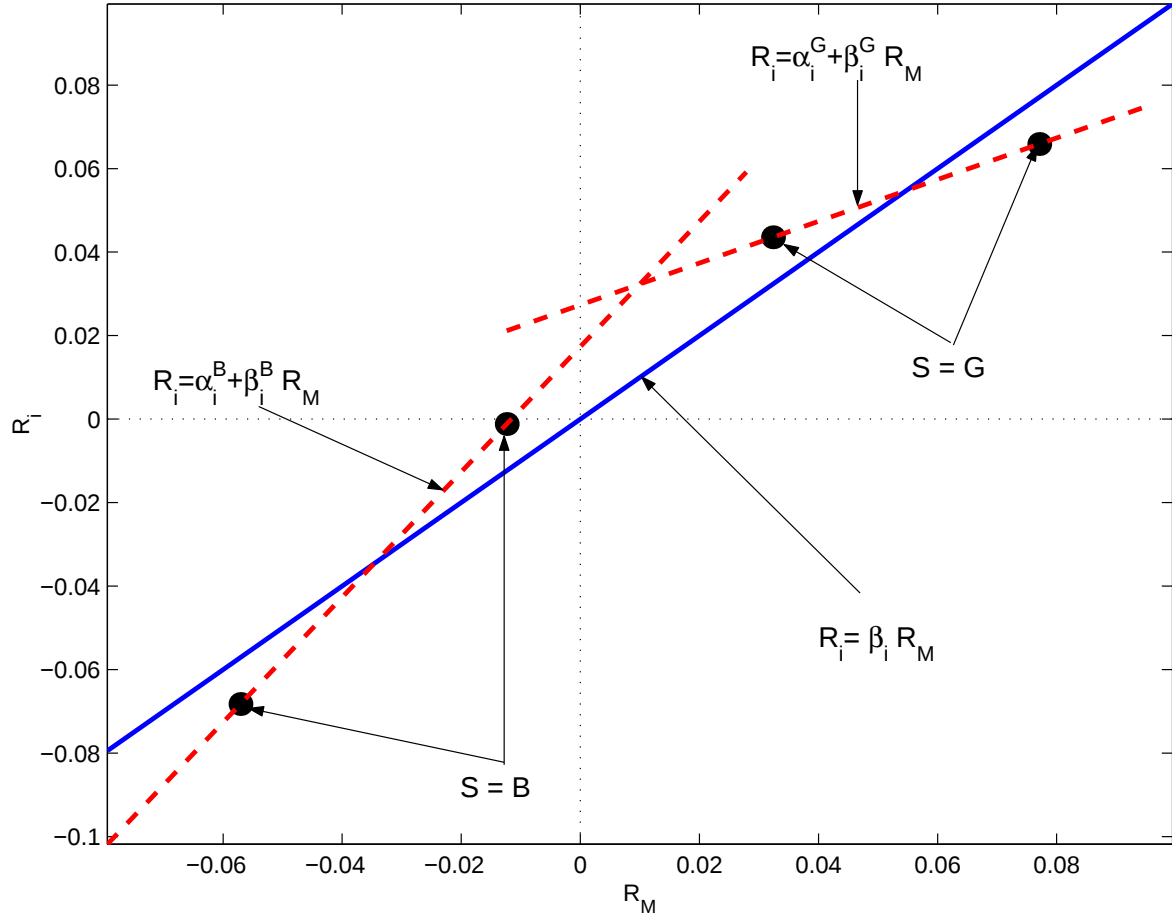


Figure 1. Overconditioning in a 4-state Example. This figure plots portfolio returns against the market return to illustrate overconditioning in a 4-state example. The solid line passing through the origin shows the investor-conditioned pricing relation, while the dashed lines represent the nonlinearity in payoffs, or the overconditioned pricing relations. Returns are $R_M(\Omega) \equiv [-0.057, -0.012, 0.032, 0.077]$ and $R_i(\Omega) = [-0.068, -0.001, 0.044, 0.066]$, and conditional betas are $\beta_i^B = 1.5$ and $\beta_i^G = 0.5$.

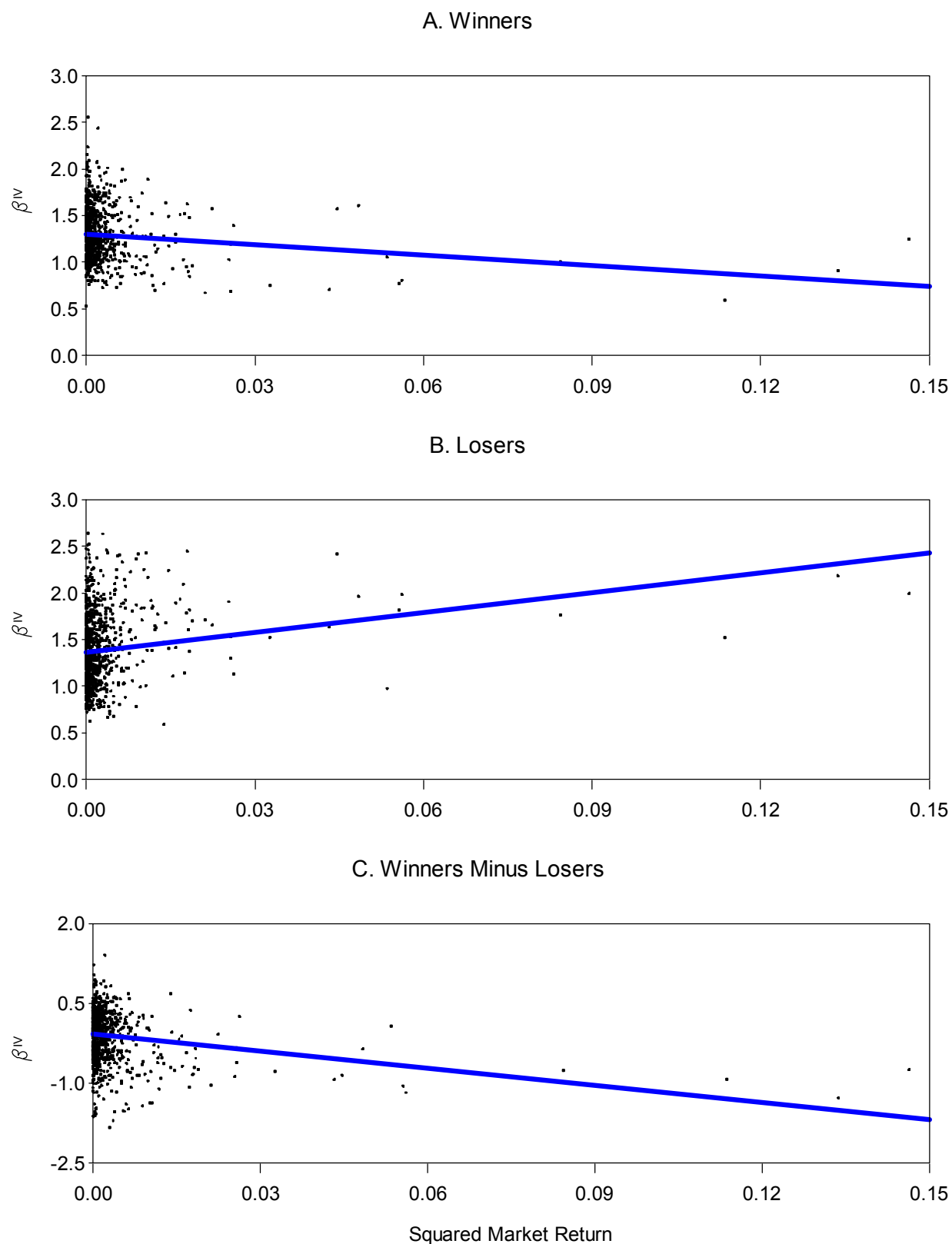


Figure 2. Volatility-Timing in Momentum. This figure shows the scatterplot of the IV betas of the 6-0-6 momentum strategy vs. squared monthly market return. DY, TB, DS, TS, 6- and 36-month market runup, and 6- and 36-month LC betas are used to obtain IV betas. The sample period is from January 1930 to December 2005.

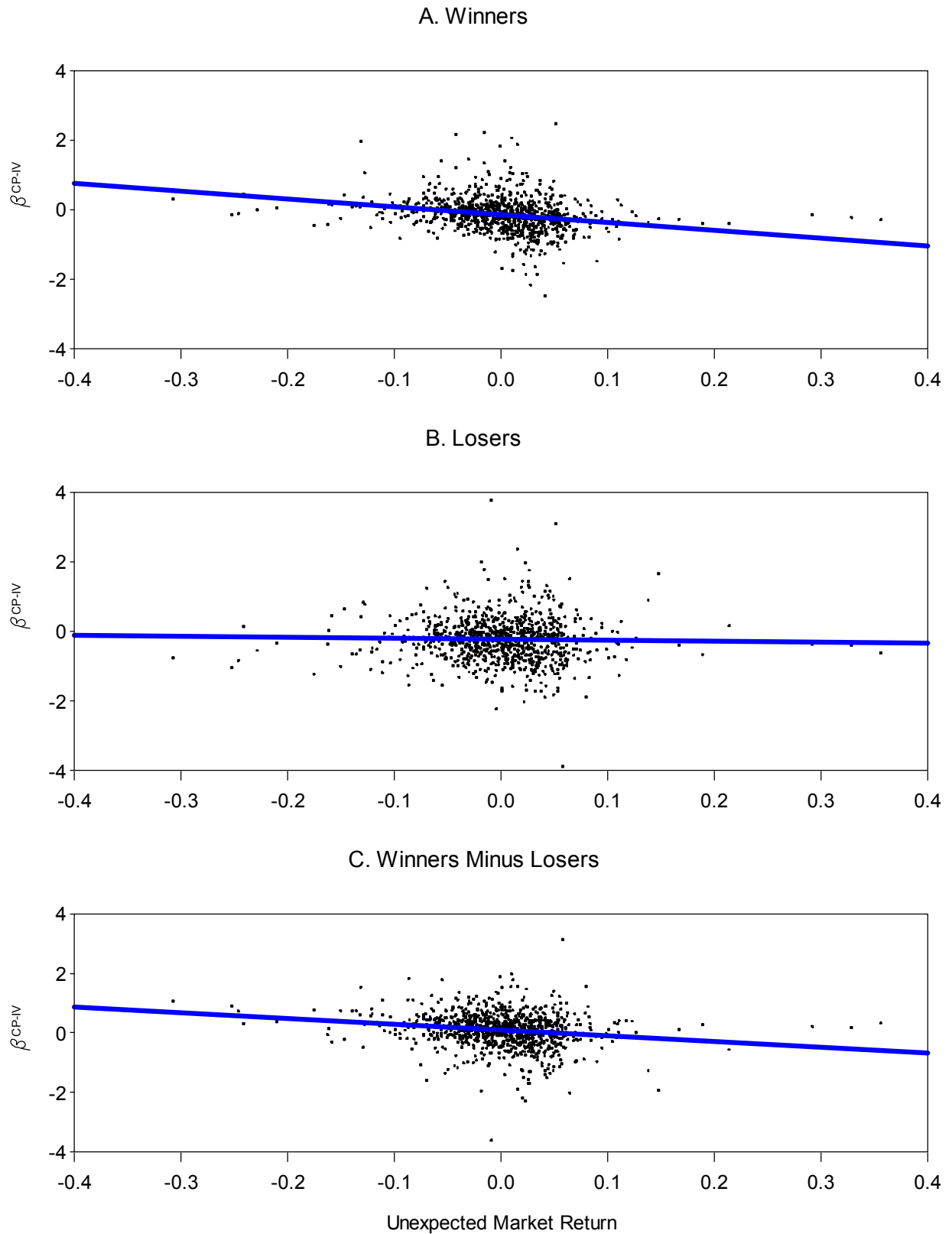


Figure 3. Overconditioning in Momentum. This figure shows the scatterplot of the difference between CP and IV betas of the 6-0-6 momentum strategy vs. unexpected monthly market return. DY, TB, DS, TS, 6- and 36-month market runup, and 6- and 36-month LC betas are used to obtain IV betas and unexpected market return. The sample period is from January 1930 to December 2005.