

# Best Practice for Cost-of-Capital Estimates

Yaron Levi and Ivo Welch\*

## Abstract

Cost-of-capital assessments with factor models require quantitative forward-looking estimates. We recommend estimating Vasicek-shrunk betas with 1–4 years of daily stock returns and then shrinking betas a *second* time (and more for smaller stocks and longer-term projects), because the underlying betas are themselves time-varying. Such estimators also work well in other developed countries and for small-minus-big (SMB) and high-minus-low (HML) exposures. If own historical stock returns are not available, peer betas based on market cap should be used. Historical industry averages have almost no predictive power and should never be used.

## I. Introduction

Given the copious asset pricing literature, it is not difficult to find papers arguing that the capital asset pricing model (CAPM) or the Fama–French model (FFM) could be true (e.g., Levy and Roll (2010)), or that the models (perhaps with some modifications) still suggest good-practice benchmarks (e.g., Da, Guo, and Jagannathan (2012)), or that the models have performed so poorly even in the context for which they were developed (for 1-month-ahead stock return prediction; see, e.g., Lewellen (2015)) that they may as well not be used.<sup>1</sup>

Our paper does not contribute to the asset pricing literature debates. Instead, our paper investigates the inputs to these models in a context in which they have remained dominant: corporate cost-of-capital assessments. Graham and Harvey (2001) report that 70% of their firms use the CAPM, whereas 30% use related multifactor models (e.g., the FFM (1992)). Jacobs and Shivdasani (2012) find an even higher number, with 90% claiming to use the CAPM.

---

\*Levi, ylevi@marshall.usc.edu, Marshall School of Business, University of Southern California; Welch (corresponding author), ivo.welch@anderson.ucla.edu, Anderson School of Management, University of California at Los Angeles. We thank Zhi Da (the referee), Jarrad Harford (the editor), Susan Huot (the business and preproduction manager), many colleagues, and seminar participants at California State University at Fullerton, the University of California at Los Angeles, Cornell University, the Federal Reserve Board, and Brigham Young University.

<sup>1</sup>Even though the high-minus-low (HML) factor has performed well, the full multifactor linear expected-return prediction of the FFM has not performed well.

The three most common MBA corporate finance textbooks (in alphabetical order) prescribe the following cost-of-capital estimates:

- Berk and DeMarzo (2014) cover the CAPM in chapter 12, concluding that “the CAPM approach is very robust. While perhaps not perfectly accurate, when the CAPM does generate errors, they tend to be small. . . . [T]he CAPM is viable, especially when measured relative to the effort required to implement a more sophisticated model. Consequently, it is no surprise that the CAPM remains the predominant model used in practice to determine the cost of capital.” Berk and DeMarzo use an equity-premium estimate of 4%–6% over short-term and 3%–5% above long-term Treasuries. They discuss estimating betas with 2–5 years of weekly or monthly returns. Moreover, they state that industry betas reduce estimation error and improve accuracy. In their appendix, they mention shrinkage, pointing out that Value Line shrinks betas, whereas Reuters, Yahoo!, and Capital IQ do not, and Bloomberg offers both.
- Brealey, Myers, and Allen (2016) cover the CAPM in chapter 9. They are more skeptical but note that there are no CAPM alternatives and that three-quarters of all managers are using the CAPM. They use an equity-premium estimate of 7%. In their beta-calculation example, they use 5 years of monthly stock returns. They then propose that “the estimation errors . . . [are] why financial managers often turn to industry betas.”
- Ross, Westerfield, Jaffe, and Jordan (2016) cover the CAPM in chapter 11, concluding that it is “the Nobel prize-winning answers to these questions [that] form the basis of our modern understanding of risk and return.” They “settle on an estimate of 7 percent for the market risk premium, though this number should not be interpreted as definitive” and emphasize its large standard errors. Their first subsection in Section 13.3 discusses how they estimate beta: “[W]e use five years of monthly data for this plot . . . in line with calculations performed in the real world.” Section 13.4 describes changes in beta estimates over time as ever-present noise but largely innocuous. The final subsection explains that “people can better estimate a firm’s beta by involving the whole industry,” suggesting that executives are better off using industry beta if they believe the operations of the firm are similar to those of others in the industry.

That is, as of 2016, these three corporate finance textbooks lead with the use of the CAPM as the primary model, suggest equity premia in excess of 4%–6% per annum, and illustrate betas calculated with unshrunk monthly stock returns. All describe industry average betas as the superior alternative in common cases. (Welch (2014) is largely based on an earlier draft of this paper and thus is more skeptical.)

Our paper endorses neither the CAPM nor the FFM. Instead, it helps corporate managers assess quantitative inputs *if/when* they have decided to use models that require forward-looking factor means and factor exposures. Most of our paper is dedicated to good-practice beta estimation, but our last section shows

how equity-premium estimates are time-period and duration sensitive. Finally, in the conclusion, our paper suggests a specific viable alternative capital budgeting model that we name the “pragmatic” model.

There are two key problems in estimating good betas: The first is measurement error; the second is that underlying market betas are not constant over long windows. Because the Vasicek (1973) shrinkage estimator was developed only to address the former, it is insufficient. That underlying changes should have an impact has been recognized at least since Blume (1971) and Fama and MacBeth (1973) and developed more rigorously by Jagannathan and Wang (1996), but it has mostly been ignored. Indeed, the literature has been fairly quiet, and some of it seems to have been forgotten.<sup>2</sup> The Internet Appendix (available at [www.jfqa.org](http://www.jfqa.org)) reviews all 2013–2015 papers in the *Journal of Finance*, the *Journal of Financial Economics*, and the *Review of Financial Studies* that calculate first-stage factor exposures. These academic betas are highly naïve. Approximately two-thirds calculate their exposures with monthly frequency stock returns (even for individual stocks), with estimation windows ranging from 12 months to 60 months. Approximately one in three assumes time-invariant betas and works with one full-sample-data beta per stock. Only one paper (Frazzini and Pedersen (2014)) shrinks beta estimates at all. The Dimson (1979) corrections for nonsynchronous trading are more popular, used in approximately one in six papers.

Our paper recommends more shrinkage that unambiguously improves forward exposure estimates. If a firm wants to rely on factor exposure estimates for its cost of capital estimates, it should do as follows:

*Recommendation.* Cost-of-capital estimates should be shrunk far more than is common practice. To estimate a good 1-year-ahead forward beta (BETA) for individual firms (LW\_BETA), we recommend starting with the popular Vasicek (VCK) (already-shrunk) beta estimator (VCK\_BETA), calculated with between 1 and 4 years of *daily* frequency historical stock returns, and then shrinking it a second time by another 20%–30%:

$$(1) \quad \text{LW\_BETA} \approx 0.75 \times \text{VCK\_BETA} + 0.25 \times \text{TARGET.}$$

If only the ordinary least squares (OLS) 1-pass beta estimate (OLS\_BETA) is available, shrinking it by 30%–40% offers acceptable predictions:

$$(2) \quad \text{LW\_BETA} \approx 0.65 \times \text{OLS\_BETA} + 0.35 \times \text{TARGET.}$$

A better 1-pass estimator replaces the fixed coefficient. It multiplies the OLS\_BETA by a transform of its standard error (SE),  $1/(1 + \text{SE}(\text{OLS\_BETA}))$  and reduces the second-pass shrinkage to 20%. If the stock is small, then another 5%–10% shrinkage can be used. If the goal is to predict the beta not over the next 1 year but over the next 10 years, then another 10% shrinkage is appropriate.

As to target, smaller firms have *lower* average market betas. Consequently, TARGET should vary with firm market cap. The smallest market cap tertial has a

<sup>2</sup>Much of this early literature is not focused on individual firm (and long-term) stock-specific cost-of-capital estimates but on improved testing of the CAPM on 1-month-ahead returns for *portfolios*, often with monthly frequency stock returns.

mean market beta of approximately 0.5, the middle of approximately 0.7, and the largest of approximately 0.9.

Although some improvements to this basic “recipe” are possible (which is perhaps not surprising, given that we start with millions of stock returns), most of these improvements are rarely meaningful enough to matter in *our* intended application, which is assessing the corporate cost of capital. For example, we consider intraday data, monthly data, financial statement data, leverage, industry membership, industry betas, value-growth adjustments, more and optimal lags, peer betas, and more sophisticated beta estimators. For individual stocks, the best improvements increase the in-sample predictive  $R^2$  only by up to 3% on a baseline of approximately 50% and/or a root mean square error (RMSE) in the market-beta estimate of approximately 0.02 on a baseline standard error of approximately 0.40. Contrary to common impression (perhaps based on academic papers, having worked largely with portfolios and 1-month returns), stock-specific factor exposures remain noisy and unstable (even with our improvements), but they predict with useful accuracy. Important corporate cost-of-capital uncertainty is not just in the factor premia and the model.

Stocks’ own prevailing 1-year betas, estimated on daily data and suitably shrunk, can predict market beta with  $R^2$ s of approximately 50%. Some common beta estimators produce far inferior forecasts. Monthly betas based on stock return data have approximately half the explanatory power, even with the necessary far-longer estimation windows and much more shrinkage. Moreover, the commonly recommended capital budgeting practice of replacing (proxying) stock market betas with industry betas (because stock betas are “too noisy”) destroys all predictive usefulness: In explaining own stock beta, the industry beta has an  $R^2$  of only 4% (instead of 50%). It is simply a common misinterpretation of Fama and French (1997) that industry betas are better predictors than own betas. (Their paper never recommends this!) If necessary (i.e., if own historical stock returns are not available), the best alternative is to use peer stocks with similar market caps (not necessarily in the same industry or with similar book-to-market ratios). Similar-size firms can predict firm-specific stock betas with  $R^2$ s of up to approximately 35%. Moreover, nonsynchronous trading adjustments, as in Dimson (1979), are detrimental in forecasting betas even for small stocks, although not as badly.

Although the use of the CAPM should remain controversial, our suggestions for better input estimates should be uncontroversial. Moreover, underlying beta changes have an effect even on short-term market-beta forecasts, but they become more important for longer-term projects. Few academic treatments have been guiding corporate managers in the use of financial cost-of-capital assessment models for such long-term projects. The only paper explicitly concerned with costs of equity for capital budgeting purposes is the seminal Fama and French (1997), but it focuses on shorter-term and industry predictions. Even in those contexts, they rightfully lament the great uncertainty in estimates. (Equity-premia estimates often still anchor on the Mehra and Prescott (1985) puzzle estimates, although other estimation methods and samples have yielded much lower estimates for decades.) Our paper emphasizes pragmatic shrinkage improvements and

quantitative forward-looking estimates for the financial cost of capital of individual stocks and over many years and decades.

Because of the intended use in a corporate scenario, our paper restricts itself to relatively simple and mostly linear exposure estimators. At times, it assesses, at least implicitly, when increases in estimation complexity seem (or do not seem) justified by their improvements in predictive accuracy. The small increase in calculation complexity for our estimators will, we hope, contribute to more academic use, too.

The cross-sectional cost-of-capital estimates that arise from better input estimates are quite different from the textbook recommendations. *They are best summarized by our in-text table in the conclusion.* Instead, the estimates are not greatly different from a pragmatic model in which managers would base cost-of-capital estimates only on i) cash flow timing, ii) financing characteristics, and iii) imperfect market aspects. Discrimination based on estimates of historical equity market betas seems misguided. The pragmatic model is a direct and viable practical alternative to the CAPM. In fact, all its inputs need to be considered in best-practice application of the CAPM, too. These three cost-of-capital premia are easy to defend. The Treasury term structure is evidence that cash flows further in the future can often require higher expected rates of return. There is evidence that firms have lower *tax-adjusted* costs of debt. And imperfect market adjustment has always been considered a necessity for firms that are not among the largest stocks in the public markets.

These three project characteristics themselves reflect much heterogeneity in risk. Again, they are already known to warrant appropriate informal consideration even in CAPM usage. The pragmatic model merely advises against the current practice of using market-beta estimates for expected-return forecast discrimination. If betas are used, then one should do so only after much greater shrinkage than is common. Cost-of-capital estimates from a CAPM with more homogeneous equity-beta estimates and smaller equity premia would often not be very different from this pragmatic alternative. We emphasize that the pragmatic model does *not* recommend that all projects' cash flows be treated as if they have equal risks or equal expected costs of capital.

Of course, even if the CAPM fails, estimating the market betas and equity premia is not intrinsically useless. These CAPM inputs can also be quantitatively important in many other applications. For example, they matter when a manager wants to form a hedged low-variance portfolio. The simplicity of our linear second-stage exposure adjustments helps to make this quite easy.

Our paper proceeds as follows: In Section II, we describe our data and methods. In Section III, we investigate market-beta predictions over the next calendar year. In Section IV, we show how these estimates deteriorate with longer forecast horizons, and thus how longer-term projects require more beta shrinkage. In Section V, we describe the market betas of (corporate) bonds, exposures of domestic stocks to the Fama–French small-minus-big (SMB) and high-minus-low (HML) factors, and exposures of foreign stocks to their home-market-index returns. In Section VI, we briefly describe the historical equity premium relative to the risk-free rate. This section points out that the most commonly used Ibbotson

equity premium is large compared with plausible alternatives. In Section VII, we conclude with a comparison of costs of capital obtained using different methods.

## II. Data and Methods

### 1. Data Sources

The Center for Research in Security Prices (CRSP) and Compustat provide standard data for our study. Kenneth French's Web site ([http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data\\_library.html](http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html)) provides the daily factor premia, risk-free Treasury, and 49-industry portfolio returns. Compustat also provides the foreign stock return data.

### 2. Underlying Process

The fact that underlying betas are slowly time-varying has the following estimation consequences: First, recent returns are more relevant to the estimation than older returns; second, the estimation uncertainty does not vanish over longer-term estimation windows, and thus it could become better to shrink relatively more; and third, the degree of shrinkage depends on the length of time over which the intended forward beta is to be used. We do not assume an exact underlying econometric process specification (e.g., as in Jostova and Philipov (2005) or Mamaysky, Spiegel, and Zhang (2007)).<sup>3</sup> Instead, our approach is to consider the performance of (mostly) linear estimators using the historical data experience. This can be justified either by the lack of confidence in any specific underlying process or by the corporate finance context in which we need simple estimators to facilitate adoptability. We do not mean to imply that the specific-econometric-process assumption approach is intrinsically better or worse.

### 3. Beta Calculations

Our benchmark specification considers betas calculated over each calendar year using daily stock returns. Absent other econometric issues, higher-frequency daily returns are known to be better than lower-frequency monthly returns (e.g., Kothari, Shanken, and Sloan (1995)) when estimating second moments and comoments. For each firm and year, we first calculate the OLS\_BETA estimates from the standard market-model time-series regressions as

$$(3) \quad \text{OLS\_BETA}_{i,d} = \frac{\text{cov}(r_{i,d} - r_{f,d}, r_{m,d} - r_{f,d})}{\text{var}(r_{m,d} - r_{f,d})},$$

where  $r_{i,d}$  is the rate of return for stock  $i$  on day  $d$ ,  $r_{f,d}$  is the 1-year Treasury rate of return, and  $r_{m,d}$  is the value-weighted CRSP market rate of return. We required 220 days (out of approximately 252 days) of daily stock return data within a year and winsorized betas at  $-3$  and  $+5$  (in approximately 100 out of approximately 200,000 firm-years).

Blume (1971) first points out the mean reversion of betas over time and suggests a linear adjustment toward the grand beta. Vasicek (1973) proposes a Bayesian estimator that is identical to the frequentist random-effects panel model

<sup>3</sup>Mamaysky et al. (2007) are interested in short-term forecasts. In 1-year forecasts, the Kalman-filtered OLS\_BETA forecast does not improve on the shrunk VCK.BETA.

(Greene (2008), p. 202f). We refer to this as the VCK market-beta estimator. It is calculated as

$$(4) \text{ VCK\_BETA}_{i,d} = \left[ \frac{\text{SD}(\text{OLS\_BETA})^2}{\text{SD}(\text{OLS\_BETA})^2 + \text{SE}(\text{OLS\_BETA})^2} \right] \times \text{MEAN}(\text{OLS\_BETA}) \\ + \left[ \frac{\text{SE}(\text{OLS\_BETA})^2}{\text{SE}(\text{OLS\_BETA})^2 + \text{SD}(\text{OLS\_BETA})^2} \right] \times \text{OLS\_BETA}_{i,d},$$

where  $\text{SD}(\text{OLS\_BETA})$  is the cross-sectional heterogeneity in  $\text{OLS\_BETA}$  estimates,  $\text{SE}(\text{OLS\_BETA})$  is the standard error of the time-series regression coefficient estimate, and  $\text{MEAN}(\text{OLS\_BETA})$  is the cross-sectional average  $\text{OLS\_BETA}$  estimate.<sup>4</sup> Karolyi (1992) introduces alternative and multiple shrinkage targets based on industry and firm size. From 1963 to 1985,  $\text{OLS\_BETAs}$  estimated from daily frequency over 1 year have a predictive mean squared error (MSE) of 0.319, whereas  $\text{VCK\_BETAs}$  have a predictive MSE of 0.242. Multiple shrinkages toward industry and size portfolio can further reduce the MSE from 0.242 to 0.227. This is a more meaningful improvement in asset pricing tests than in corporate finance applications, in which other issues can swamp this 0.015 improvement.

The Vasicek (1973) and Karolyi (1992) estimators are optimal *if* the underlying beta processes are not themselves time-varying. Yet this may not be the correct assumption in our longer-term application. This is pointed out by Johnson, Bennett, and Curcio (1979) and by Barry ((1980), p. 88), who states, “[W]hen beta is suspected to vary over time, a technique that explicitly recognizes the variation (be it systematic or random) should be used. No technique that assumes stationarity should be expected to perform well under nonstationarity...” In a context in which betas move only slowly and users need *some* beta estimator, this seems harsh. The  $\text{VCK\_BETA}$  uses heterogeneity and uncertainty information in a way that is potentially more efficient than the simpler  $\text{OLS\_BETA}$  or the linear Blume (or Merrill Lynch) beta;<sup>5</sup> there is no consensus about the exact nonstationary time-series process, and there is no widely known closed-form alternative beta estimator.

Jostova and Philipov (2005) offer one good structural alternative: a mean-reverting beta process. They derive an estimator to test the CAPM with *monthly* stock return data on industries. Their estimator is clearly superior to our own *if* their underlying process specification is correct. Unfortunately, even if it is, it is still unlikely that their approach could be widely adopted for cost-of-capital budgeting. This is because their estimator lacks a closed form (even for the instant beta). It is for this reason that our own paper restricts itself to variants of Blume and Vasicek estimators.<sup>6</sup>

<sup>4</sup>The  $\text{MEAN}(\text{OLS\_BETA})$  has a (time-series) mean of 0.87 and not 1, because the betas are not value-weighted.

<sup>5</sup>Merrill Lynch no longer advertises its market betas. Our previous draft investigated secondary descriptions and found them to perform no better than those investigated here. We were under the common but mistaken impression that Yahoo! Finance shrank its market-beta estimates. Instead, as of mid-2016, it reports monthly data unshrunk 36-month estimates.

<sup>6</sup>Both Fisher and Kamin (1985) and Jostova and Philipov (2005) also investigate heteroskedasticity-adjusted betas. They find that allowing for heteroskedasticity improves the estimation only modestly. Thus, such methods seem more suitable to more sophisticated asset pricing tests. In our work, we ignore further market heteroskedasticity adjustments.

#### 4. Time Changes and Estimation Window

In the absence of time variation in the underlying data-generation process, betas should be estimated with as much historical return data as possible. In the presence of time variation in the underlying betas, the best historical estimation interval (or the best time-decaying estimation weights) becomes a trade-off between the desire to have more days (to reduce the estimation noise) and the desire to predict the future beta with more relevant recent data.

#### 5. Predictive Regressions

Our approach is to let the historical relation between the lagged prevailing beta and the future beta (in the cross section) guide *linear* predictions. Our typical inference is from a regression that predicts future beta  $BETA_{i,y}$  with prevailing beta  $BETA_{i,y-1}$ :

$$(5) \quad BETA_{i,y} = \gamma_0 + \gamma_1 \times BETA_{i,y-1} + \epsilon,$$

where  $y$  is (typically) the calendar year. The auto-coefficient  $\gamma_1$  helps assess by how much one should shrink the prevailing beta to obtain the lowest MSE forecast of the future beta. The parameter  $\gamma_0$  adjusts for time changes in the mean of the betas, which are caused by the fact that our regressions are usually not value-but equal-weighted. We occasionally refer to the auto-coefficient  $\gamma_1$  as a shrinkages or gamma coefficient. The common use of the prevailing beta estimate as a direct proxy forecast for the future beta is equivalent to imposing the “one-to-one” restriction that  $\gamma_0 = 0$  and  $\gamma_1 = 1$ .

Our 1-year base scenario predicts nonoverlapping (in the dependent variable) annual regressions with their convenient natural calendar-year breakpoints. With approximately 90 years of stock returns and between several hundred and several thousand stocks per year, a typical predictive regression is based on several hundred thousand data points, each of which itself aggregates hundreds of daily stock returns. Standard errors are always very small, and models are always statistically significantly different, allowing us to focus on estimate magnitudes. In some regressions, we begin in 1973 because CRSP’s NASDAQ data came online in late 1972. NASDAQ stocks greatly increase the sample heterogeneity in market cap.

Note the contrast between market betas and stock returns (as in Fama and MacBeth (1973)). A typical cross-sectional regression that explains next year’s stock returns (among firms) may have an  $R^2$  of less than 5%, whereas a regression that explains next year’s beta may have an  $R^2$  of more than 50%. This makes the task of understanding persistence in betas considerably easier than the task of understanding expected rates of return and asset pricing models. Again, because a data set this large results in statistical reliability far beyond ordinary magnitudes, our discussions can focus on economic meaning.

#### 6. Errors in Variables

The typical error-in-variables (EIV) concern in tests of asset pricing models does not apply in our context. Errors in any dependent variable (i.e., the fact that the future beta is not the true beta) are handled gracefully by the (predictive) OLS regressions. Errors in the independent variable (i.e., in the prevailing beta estimate) bias the gamma coefficient estimates downward but only relative



to the relation that would obtain between the *true* beta and the future beta. However, corporate managers in our setting are assumed to have at their disposal only estimated prevailing betas. Their inputs could thus not be the true-but-unknown betas. In this sense, our gamma estimates merely spell out how a user of prevailing historical beta estimates should shrink his or her estimates to obtain forecasts of future betas. Our estimators should be seen as providing good-practice inputs given the data we have rather than as attempts to provide good ingredients for asset-pricing-model tests given unknown true parameters.

## 7. Variations

There is a multitude of possible estimation variations (and complications), both for the beta estimations and for the predictive regressions, many of which are discussed in the next few sections.

## III. Empirical Estimates of Future Market Betas

### A. A Single-Year Illustration: 2013

For illustration, let us begin with a typical year. In calendar years 2012 and 2013, 4,750 stocks have enough daily stock return data to calculate both 1-year betas. The typical time-series standard error of the (1-year) OLS\_BETA or VCK\_BETA estimates is approximately 0.2. The typical cross-sectional heterogeneity is approximately 0.62 for OLS\_BETAs and 0.54 for VCK\_BETAs. To predict the OLS\_BETA in 2013 ( $\text{OLS\_BETA}_{i,2013}$ ) in this cross section, the one-to-one direct predictions (i.e., (badly) guessing that the estimated prevailing 2012 betas would be the 2013 betas) are as follows:

To predict  $\text{OLS\_BETA}_{i,2013}$ , use either of the following:

$$\begin{aligned} 0.00 + 1.00 \times \text{OLS\_BETA}_{i,2012} &\Rightarrow R^2 \approx 50\% \quad \text{RMSE} \approx 0.38, \text{ or} \\ 0.00 + 1.00 \times \text{VCK\_BETA}_{i,2012} &\Rightarrow R^2 \approx 57\% \quad \text{RMSE} \approx 0.35. \end{aligned}$$

The  $R^2$  is calculated as  $1 - \text{SSE}/\text{SST}$  with the 2013 betas. The dependent variable can be viewed as measured with error or as the variable of interest itself. (Hedging the *realized* market beta would be as useful as hedging the *expected* market beta.) The best linearly shrunk OLS regression predictions are as follows:

To predict  $\text{OLS\_BETA}_{i,2013}$ , use either of the following:

$$\begin{aligned} 0.26 + 0.69 \times \text{OLS\_BETA}_{i,2012} &\Rightarrow R^2 \approx 63\% \quad \text{RMSE} \approx 0.33, \text{ or} \\ 0.21 + 0.75 \times \text{VCK\_BETA}_{i,2012} &\Rightarrow R^2 \approx 64\% \quad \text{RMSE} \approx 0.32. \end{aligned}$$

These four lines suggest that i) OLS\_BETAs predict worse than VCK\_BETAs, and ii) more aggressive shrinkage is even more important than the use of VCK\_BETAs instead of OLS\_BETAs. The reduction in RMSE is meaningful but limited. More importantly, it is obtainable with minimal effort.

Instead of predicting the 1-year OLS\_BETA, we can predict the 1-year VCK\_BETA. (With a different dependent variable,  $R^2$ 's cannot be compared to those noted previously.)

To predict  $VCK\_BETA_{i,2013}$ , use either of the following:

$$0.00 + 1.00 \times VCK\_BETA_{i,2012} \Rightarrow R^2 \approx 56\% \quad RMSE \approx 0.32, \text{ or}$$

$$0.25 + 0.70 \times VCK\_BETA_{i,2012} \Rightarrow R^2 \approx 69\% \quad RMSE \approx 0.27.$$

Thus, the best linear predictor of the 1-year-ahead  $VCK\_BETA$  is the lagged  $VCK\_BETA$ , shrunk by another 30%. Although specific to 2013 and not always appropriate, the 30% shrinkage rule-of-thumb generally provides good estimates for the entire sample and most methods. The envelope theorem assures us that modest deviations from 30% have only second-order effects.

## B. The Basic Pooled-Panel Predictive Regression

Although we are interested in the cross-sectional beta predictions in any given year, the results from a pooled-panel regression over all firms and years provide almost identical inference and are easier to summarize.

In Table 1, we report the ( $R^2$  and  $\gamma$ ) performance of predictions for 1-year future betas for individual stocks with prevailing betas, either OLS or VCK, and calculate either with 1 year, 3 years, or 5 years of daily stock return data. In consecutive panel rows, the observations in the dependent variable are the same, thus permitting the direct comparison of  $R^2$ s. Each regression is bivariate. Although values are not reported because there are so many observations (here and in subsequent tables), all models are statistically significantly different from one another.<sup>7</sup> The reader can focus on the economic meaning of the reported values.

In Panel A of Table 1, we predict the OLS\_BETA. The best predictive performance (with an  $R^2 \approx 50\%$ ) comes from the use of a  $VCK\_BETA$  estimated in a 1-year window and then shrunk again by approximately 20%:

$$(6) \quad LW\_BETA \approx 0.15 + 0.81 \times VCK\_BETA_{i,t}.$$

There is little economic difference between prevailing  $VCK\_BETAs$  calculated over 1 year or 3 years, but  $VCK\_BETAs$  calculated over 5 years are worse. A reasonable alternative to the  $VCK\_BETA$  is an OLS\_BETA calculated with 3 years of daily stock returns and shrunk by 20% ( $R^2 \approx 48\%$  and  $\gamma_1 \approx 0.79$ ). If a 3-year OLS\_BETA is not available, then a 1-year OLS\_BETA shrunk by 30% can be used.

In Panel B of Table 1, we predict the  $VCK\_BETA$ . Again, the best predictive performance comes from the use of a  $VCK\_BETA$  estimated in a 1-year window and then shrunk again by approximately 25% ( $\gamma_1 \approx 0.74$ ). A 3-year beta-estimation window produces reasonably similar predictions.

In both panels of Table 1, the RMSE prediction error remains an economically large 0.4–0.5, but this overstates the problem. Individual stock betas are noisy. This is because the past beta is measured with error, the underlying beta is changing, and the future beta is measured with error. This last issue is not of

<sup>7</sup>We establish the statistical distribution under the null hypothesis by randomly switching prediction errors between any two models and recomputing the metrics. We then compare the location of the actual empirical  $R^2$  (RMSE) estimate to the distribution of simulated  $R^2$  (RMSE) estimates under the null distribution.

TABLE 1  
Predicting 1-Year-Ahead Future Market Betas, Pooled Regressions

In Table 1, the statistics are from pooled bivariate predictive regressions that explain (in-sample) 1-year-ahead calendar stock market betas (computed from daily stock returns and winsorized at  $(-3, 5)$ ) with prevailing stock market betas. The prevailing betas (independent variables) are as described in the column headers and are calculated over 1 year, 3 years, or 5 years. The  $R^2$ 's are comparable in each row because the dependent observations are identical (thus requiring availability of 5-year prevailing betas even when only a 1-year beta is used). When  $\gamma_0 = 0$  and  $\gamma_1 = 1$ , the  $R^2$  is  $1 - \text{SSE/SST}$ , equivalent to the predictive regression with the constraints  $\gamma_0 = 0$  and  $\gamma_1 = 1$ .  $R^2$ 's are quoted in percentages. The best predictor in each row is highlighted in boldface. Although values are not reported because there are so many observations, all models are statistically significantly different from one another. *Data Set:* Stocks with 1-year, 3-year, and 5-year prevailing market-beta availability, requiring 220, 400, and 1,000 daily stock returns, respectively, from CRSP 1927–2014 data. *Interpretation:* i) The best estimation windows are 1 to 3 years. ii) VCK\_BETAs predict better than OLS\_BETAs. iii) Predicted betas should be shrunk. In the case of VCK\_BETA, they should be shrunk a second time.

Metric	Independent Variable: BETA Calculated over Last x Years						SD(y)	N
	1 Year		3 Years		5 Years			
	OLS	VCK	OLS	VCK	OLS	VCK		
<i>Panel A. Dependent Variable: OLS_BETA</i>								
RMSE	0.50	0.45	0.46	0.45	0.47	0.46	0.62	199,783
$R^2$	35.6	47.4	44.9	<b>47.8</b>	42.3	43.9		
$\gamma_1$	Forced to 1		Forced to 1		Forced to 1			
$\gamma_0$	Forced to 0		Forced to 0		Forced to 0			
RMSE	0.46	0.44	0.45	0.44	0.46	0.46	0.62	199,783
$R^2$	46.1	<b>50.0</b>	48.3	49.4	45.0	45.5		
$\gamma_1$	0.68	0.81	0.79	0.85	0.80	0.84		
$\gamma_0$	0.25	0.15	0.16	0.12	0.16	0.13		
<i>Panel B. Dependent Variable: VCK_BETA</i>								
RMSE	0.45	0.38	0.40	0.38	0.40	0.39	0.54	199,783
$R^2$	30.3	49.5	44.8	<b>50.2</b>	42.7	45.9		
$\gamma_1$	Forced to 1		Forced to 1		Forced to 1			
$\gamma_0$	Forced to 0		Forced to 0		Forced to 0			
RMSE	0.38	0.35	0.37	0.36	0.38	0.38	0.54	199,783
$R^2$	50.7	<b>56.2</b>	53.3	55.1	49.7	50.6		
$\gamma_1$	0.61	0.74	0.71	0.77	0.73	0.77		
$\gamma_0$	0.29	0.20	0.21	0.17	0.21	0.18		

concern because for capital budgeting, we can wish to predict the *expected* covariation, not the *realized* covariation.<sup>8</sup>

Because Table 1 suggests that the VCK\_BETA better predicts not only itself but also the future OLS\_BETA, it also suggests that the VCK\_BETA is a better measure of the true beta than the OLS\_BETA. The direct-prediction row shows that most of the VCK's advantage is due to its 10% lower bias. (It is, after all, already shrunk once.) However, some of its advantage is due to the VCK's more efficient use of the uncertainty information from the time-series regressions.<sup>9</sup>

<sup>8</sup> A simple experiment helps illustrate this. Assume that beta is Gaussian normally distributed  $\beta_i \sim N(1, 0.5)$  with  $i \in 5,000$  firms. Assume there are no underlying changes in beta, so  $\beta_{i,t} = \beta_{i,t-1}$ . Assume that the measured beta is  $b_i \equiv \beta_i + \epsilon_i$ , where  $\epsilon_i \sim N(0, 0.3)$ . The resulting predictive cross-sectional regression on observables yields  $b_t = 0.26 + 0.74 b_{t-1}$  with an  $R^2$  of 0.55 and a standard error of 0.4, which is roughly what we observe. If the association between the true expected beta and the prediction could be computed, then  $\beta_{i,t} = 0.26 + 0.74 b_{i,t-1}$  with an  $R^2$  of 0.74 and a standard error of 0.25. An error of 0.25 is still economically meaningful in a cost-of-capital context but not as bad as an error of 0.4.

<sup>9</sup> The first conclusion is from the fact that the linear OLS\_BETA increases 15%–20% over the one-to-one prediction, whereas the VCK\_BETA increases by only 2%–7% over the one-to-one prediction. The second conclusion is from the fact that the VCK prediction  $R^2$  remains above the OLS  $R^2$  even when the shrinkage is linear.

The VCK estimator requires a 2-pass calculation, an inconvenience relative to the 1-pass OLS estimator. Although not shown in Table 1, the 1-pass OLS estimator's performance can be improved by multiplying the independent variable by a transform of the OLS\_BETAs standard error,  $1/(1 + \text{SE}(\text{OLS\_BETA}))$ . In predicting the OLS\_BETA,  $\gamma_0 \approx 0.21, \gamma_1 \approx 0.86, R^2 \approx 48.7\%$ ; in predicting the VCK\_BETA,  $\gamma_0 \approx 0.26, \gamma_1 \approx 0.78, R^2 \approx 44.2\%$ :

$$(7) \quad \text{LW\_BETA} \approx 0.80 \times \left[ \frac{1}{1 + \text{SE}(\text{OLS\_BETA})} \right] \times \text{OLS\_BETA} \\ + 0.20 \times \text{TARGET.}$$

### C. The Intercept and Approximations

By definition, the value-weighted average market beta is 1. Thus, the intercept  $\gamma_0$  is  $1 - \gamma_1$  in a value-weighted regression.

In our equal-weighted regressions in Table 1 (with respect to a value-weighted market), the average market beta is less than 1, thus lowering the  $\gamma_0$  intercept. In Table 1, the estimates are  $\gamma_0 + \gamma_1 \approx 0.95$ . Therefore, the best forward estimates using both intercept and slope shrink toward lower average market betas of, for example, approximately 0.83 ( $\text{LW\_BETA} \approx 0.20 + 0.74 \times \text{VCK\_BETA}$ ) and not toward 1. As noted in the Introduction, the smallest tertial of firms had VCK\_BETAs averaging 0.5, and the largest averaged 0.9.

It depends on the corporate use context and assumptions about estimation accuracies whether equal-weighted or value-weighted analysis is more suitable. Although results are not reported in Table 1, if the predictive regressions are also value-weighted, the  $\gamma_1$  coefficient remains almost the same: Predicting the OLS\_BETA (VCK\_BETA) with the lagged 1-year VCK\_BETA changes  $\gamma_1$  from 0.81 (in Table 1) to 0.80 (VCK 0.74 to 0.76), that is, just a tiny bit. It is the  $\gamma_0$  intercepts that increase from 0.15 to 0.20 and from 0.20 to 0.24, respectively.

Even in the equal-weighted regressions, the sum of the gammas remains reasonably close to 1. The envelope theorem suggests such modest deviations from the exact gamma coefficients are reasonable. Although results are not reported in Table 1, if we constrain  $\gamma_0 + \gamma_1 = 1$ , the coefficients increase from  $\gamma_0 = 0.20$  and  $\gamma_1 = 0.74$  to  $\gamma_0 = 0.22$  and  $\gamma_1 = 0.78$ . Neither the RMSE (0.35) nor the  $R^2$  (56.2%) deteriorates to the second digit after the decimal point (the accuracies reported in Table 1). It is often inconsequential whether the user shrinks the coefficient by 22% toward 1 or by 26% toward 0.8. Table 2 notes that the year-by-year annual gamma estimates have fairly high standard deviations (approximately 0.15). Thus, more pseudo-precision than reported seems misguided. A reasonable guide is to shrink very small firms toward 0.5, midsize firms toward 0.75, and large firms toward 1.

### D. Specification Variations

The previous beta estimator, in which VCK\_BETAs are estimated on daily stock returns over 1 to 3 years and then shrunk 20%–30% (toward  $\approx 0.9$ ) a second time, is quite robust. In this section, we summarize the findings of a subset of our specification investigations.

TABLE 2  
 Predicting 1-Year-Ahead Future Market Betas, Time-Series Averages  
 of Annual Cross-Sectional Regression Statistics

In Table 2, the (yearly) average  $R^2$ s and gammas are from cross-sectional annual predictive regressions (similar to those of Fama and MacBeth (FM) (1973)). For other basic explanations, refer to Table 1. The best predictor in each row is highlighted in boldface. *Data Set:* As in Table 1. *Interpretation:* The FM regressions are almost identical to the pooled regressions. <sup>†</sup>The standard deviation of the  $\gamma_1$  mean of 0.83 is 0.15, with an autocorrelation of  $-0.24$ .

		Independent Variable: BETA Calculated over Last x Years							
Dependent Variable	Metric	1 Year		3 Years		5 Years		SD(y)	N
		OLS	VCK	OLS	VCK	OLS	VCK		
OLS_BETA	$R^2$	49.9	<b>51.9</b>	50.4	51.0	47.4	47.7	0.62	199,783
	$\gamma_1$	0.70	<b>0.83<sup>†</sup></b>	0.80	0.85	0.82	0.85		
VCK_BETA	$R^2$	52.5	<b>55.7</b>	53.5	54.6	50.5	51.2	0.54	199,783
	$\gamma_1$	0.63	0.70	0.72	0.77	0.73	0.77		

1. Cross Section

In Table 2, we confirm that the results from averaging cross-sectional calendar-year regressions are almost identical to the results of the pooled regression. With the 1-year window (which predicts better than the 3-year window), the VCK\_BETA predicts its future self with an  $R^2$  of 55.7% (rather than 56.2% in the pooled specification), and the predictive  $\gamma_1$  coefficient is 0.70 (rather than 0.74).<sup>10</sup>

The RMSE in the beta predictions remains an alarmingly large 0.40. Standard errors may be low in large portfolios, but they are not low for individual stocks. Nevertheless, as we show in Section III.E, it is incorrect to conclude that it is therefore better to use industry portfolios.

2. Nonlinearities

When stocks are sorted into quintiles based on OLS\_BETAs (31,473 firm-years per bin, post-1973, 5 years of data), they show no obvious nonlinearities in mean reversion:

Lagged OLS_BETA Interval	-3, 0.2	0.2, 0.5	0.5, 0.8	0.8, 1.2	1.2, 5
Lagged OLS_BETA mean	-0.01	0.36	0.66	1.00	1.61
(Next) OLS_BETA mean	0.23	0.43	0.67	0.94	1.33
(Next) VCK mean	0.27	0.45	0.67	0.92	1.27

The same is the case for mean reversion of the VCK\_BETAs:

Lagged VCK_BETA Interval	-2.9, 0.26	0.26, 0.52	0.52, 0.80	0.80, 1.14	1.14, 3.77
Lagged VCK_BETA mean	0.08	0.39	0.66	0.96	1.49
(Next) VCK_BETA mean	0.25	0.46	0.67	0.90	1.29
(Next) VCK standard deviation	0.31	0.32	0.35	0.37	0.45

<sup>10</sup>Note that we cannot use fixed-effects regressions (the auto-coefficient drops to 0.44, implying shrinkage of more than 50%) because such a specification primarily answers a time-series question and not a cross-sectional question. A fixed effect effectively assumes knowledge of each firm's (long-run) sample beta mean. Yet the long-run beta mean is only realized ex post and thus not necessarily known to the managers. A manager who better knows his or her beta (and thus the consequence of deviation of 1 year's beta from the long-term mean) can of course do better. Note also that there are years in which the predictive regression  $\gamma_1$  is higher or lower. We assume that managers do not have foreknowledge of such changes, especially over longer horizons, and thus we assume that they rely on the average forecast coefficients instead. Our key question is how to adjust stocks with low betas differently from stocks with high betas.

The mean reversion in the two extreme quintiles is similar if we remove the  $\approx 100$  winsorized extreme firm-years from the data set.

Although the decline of higher betas over time could be explained by firms' projects declining in risk as they mature, the increase in lower betas suggests that beta changes are not primarily a firm-aging effect.

### 3. Return-Measuring Frequency

There is good a priori reason to believe that betas calculated from higher-frequency stock returns are superior proxies for the underlying true beta (e.g., Kothari, Shanken, and Sloan (1995)). This can also be empirically assessed.

In Panel A of Table 3, we still predict future 1-year daily return betas. Measurement noise in the dependent variable is handled well by the OLS method. The  $R^2$ s and  $\gamma_1$ s show that (even the best 5-year) prevailing monthly return betas predict worse ( $R^2 \approx 20\% - 30\%$ ) than the (1-year) prevailing daily return betas ( $R^2 \approx 50\%$ ).

The opposite is not the case. In Panel B of Table 3, we predict future 1-year monthly return betas. Although these dependent variables are noisier (especially over 1 year), we can still compare the relative predictive power of different independent variables. In predicting future monthly return betas, 1-year daily return betas predict as well as or better than lagged betas, which are themselves estimated with monthly stock returns (including betas calculated over much longer horizons).<sup>11</sup> Although not reported here, this also remains the case for dependent variables (monthly return betas) that are estimated over longer horizons.

For users of OLS market betas calculated with 3 to 5 years of monthly stock returns, which was the most common academic approach in 2013–2015, rather than with daily return betas, the historical evidence suggests that the prediction is improved by shrinking beta estimates toward the cross-sectional mean by approximately 50%.

If daily data produce a better estimator than monthly data, do intraday data then produce a better estimate than daily data? Ait-Sahalia, Kalnina, and Xiu (2014) develop an intraday estimator and apply it to New York Stock Exchange (NYSE) Trade and Quote (TAQ) data for 18,335 firm-years from 2007–2013. They shared with us the resulting market-beta estimates. In results not tabulated here, we find a 12-month<sup>12</sup> average of their intraday market beta has broadly similar performance as the daily market beta in predicting future Vasicek market betas. The intraday market beta always has its RMSE within 0.02 of the daily market-beta RMSE, often but not always better. Second-time shrinkage again remains highly advisable. The intraday market beta is unambiguously better in predicting its own future realization, with an RMSE lower by approximately 0.10 than the comparable daily estimator.

<sup>11</sup>The fact that low-frequency monthly betas are difficult to predict by either low-frequency or high-frequency betas is consistent with the hypothesis that low-frequency betas are less informative. A thought experiment clarifies this: A random statistic is not better forecastable by a good high-frequency daily return beta than by itself.

<sup>12</sup>A 1-month market beta produces worse estimates than a 12-month average.

TABLE 3  
Monthly Stock Return Frequency

Some of the independent variables in Panel A of Table 3 and the dependent variables in Panel B are calculated not from daily but from monthly stock returns. For other basic explanations, refer to Table 1. The best predictor in each row is highlighted in boldface. The "10 Years" statistics are not based on the same sample as the shorter statistics and therefore are not directly comparable. *Data Set:* As in Table 1. In Panel B, a 12-month return data requirement results in dropping 18% of the firm-years (instead of 15%). *Interpretation:* Betas calculated from monthly stock returns predict worse than betas calculated from daily stock returns.

Dependent Variable	Metric	Independent Variable: BETA Calculated over Last x Years										Daily Beta (≤ 5 Years)	
		1 Year		3 Years		5 Years		10 Years		Over 1 Year		SD(y)	N
		OLS	VCK	OLS	VCK	OLS	VCK	OLS	VCK	OLS	VCK		
<i>Panel A. Predicting 1-Year-Ahead Betas Based on Daily Return Frequencies with Betas Based on Monthly Return Frequencies</i>													
OLS_BETA (daily)	$R^2$	11.9	17.0	21.0	26.2	22.8	26.7			46.1	50.0	0.62	199,783
	$\gamma_1$	0.19	0.38	0.36	0.58	0.43	0.63			0.68	0.81		
VCK_BETA (daily)	$R^2$	12.9	19.3	22.8	29.6	24.9	30.1			50.7	56.2	0.54	199,783
	$\gamma_1$	0.17	0.35	0.32	0.53	0.38	0.58			0.61	0.74		
<i>Panel B. Predicting 1-Year-Ahead Betas Based on Monthly Return Frequencies</i>													
OLS_BETA (monthly)	$R^2$	5.1	6.8	9.4	11.0	10.4	<b>11.4</b>	9.4	9.5	10.7	<b>11.4</b>	1.15	197,068
	$\gamma_1$	0.23	0.44	0.45	0.70	0.54	0.76	0.59	0.73	0.61	0.72		
VCK_BETA (monthly)	$R^2$	7.0	10.9	13.0	17.0	14.5	17.6	13.2	14.3	16.2	<b>18.0</b>	0.68	197,068
	$\gamma_1$	0.16	0.33	0.31	0.51	0.37	0.56	0.43	0.54	0.44	0.53		

4. Multivariate Prediction

In previous tables, we entertained only one independent variable at a time. Table 4 predicts the 1-year VCK\_BETA with multiple variables for the set of firms with 5 years of market cap and beta information.

TABLE 4  
Multivariate Prediction of 1-Year-Ahead Daily VCK Betas

Table 4 predicts 1-year-ahead VCK\_BETAs after 1973, using more predictors. Panel A includes OLS\_BETA and the OLS standard deviation. Panel B includes multiple lags of the 1-year VCK\_BETAs. The intercept is  $\gamma_0$ . Although not reported, the cross-sectional heterogeneity (standard deviation) of SD(VCK\_BETA) (from the time-series regression) is approximately 0.2. For other basic explanations, refer to Table 1. *Data Set:* As in Table 1, except all 5 years have to be individually available, too, and the data set begins in 1973. This results in 157,367 firm-year observations. *Interpretation:* In Panel A, shrinking more aggressively based on the estimated time-series error of beta (toward the cross-sectional mean) rather than that suggested by VCK is modestly better. In Panel B, using 1 or 2 more years of historical VCK\_BETAs is (modestly) better, with progressively less weight on more distant market betas.

Panel A. Predicting with Lagged OLS\_BETA, VCK\_BETA, and Time-Series Uncertainty Estimates

	1	2	3	4
$R^2$	53.7%	54.8%	53.7%	54.8%
Intercept	0.19	0.16	0.20	0.17
VCK_BETA lagged	0.73	1.15	0.73	1.16
OLS_BETA lagged		-0.37		-0.38
SE(OLS_BETA) lagged			-0.03	-0.06

Panel B. Predicting VCK\_BETA with Many Lags of Past VCK\_BETAs

	1	2	3	4	5
$R^2$	53.7%	56.7%	57.4%	57.6%	57.6%
Intercept	$\gamma_0$ 0.19	0.15	0.13	0.12	0.12
1-year lagged VCK_BETA	$\gamma_1$ 0.73	0.54	0.51	0.50	0.50
2-year lagged VCK_BETA		0.25	0.18	0.17	0.17
3-year lagged VCK_BETA			0.13	0.10	0.09
4-year lagged VCK_BETA				0.06	0.04
5-year lagged VCK_BETA					0.03

In Panel A of Table 4, we include both OLS\_BETAs and VCK\_BETAs. In effect, this relaxes the Vasicek uncertainty-weighting function. The gamma coefficients on the two betas suggest that shrinking with the standard deviations more aggressively than suggested by the VCK formula is better, but the predictive  $R^2$  increases only modestly, from 53.7% to 54.8%.

Elton, Gruber, and Urich (1978) suggest that betas are principally proxies of their standard deviations, but the table shows that stocks' own residual market-model standard deviations add little above and beyond the beta in predicting the future beta.

In Panel B of Table 4, we show that including further lags of the VCK\_BETA is more useful, increasing the  $R^2$  from 53.7% to 56.7% with 1 more lag (and to approximately 57.6% with 5 lags). Moreover, the gamma decay with age is reasonable, suggesting that we are not overfitting the data.

5. Time Decay

Figure 1 plots the prediction of  $x$ -month VCK\_BETAs with lagged  $y$ -month VCK\_BETAs.<sup>13</sup> The figure shows that the best estimation window to calculate

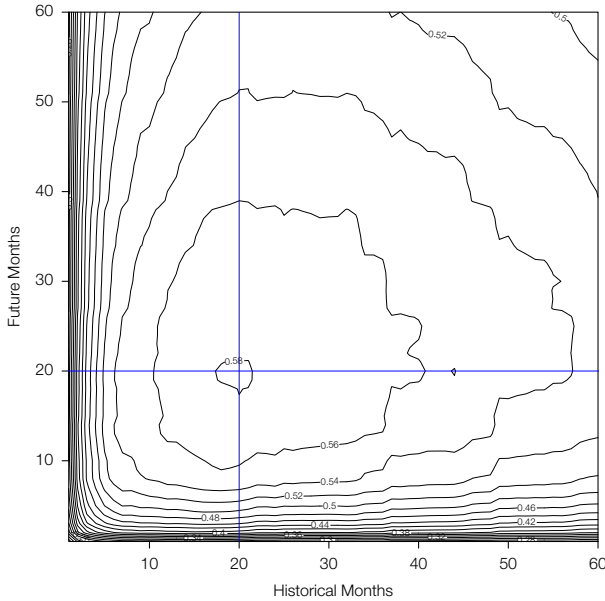
<sup>13</sup>The regressions are run every January using daily stock return data. Thus, for periods greater than 12 months, there are overlaps in the beta calculations for consecutive regression observations. (The regressions are always predictive, in that calculation periods for the dependent and independent variables never overlap.)



FIGURE 1

Predicting  $y$ -Month VCK\_BETA with  $x$ -Month VCK\_BETA

In Figure 1, the variable on the  $x$ -axis ( $y$ -axis) is a VCK\_BETA computed over an event window of  $x$  ( $y$ ) months. The contour plot graphs the  $R^2$  in bivariate predictive regressions. For other basic explanations, refer to Table 1. *Data Set:* Unlike most other tables, we allow the data availability to vary with the horizon. For 1-month forecasts, we have 230,860 observations; for 60-month forecasts, we have 153,762 observations, from CRSP 1927–2014 data. *Interpretation:* The best estimation window is 20 months, but 12- to 36-month windows also yield reasonably good predictions.



the VCK\_BETA is not 12 months but approximately 20 months, increasing the predictive  $R^2$  in this variation from approximately 56% to approximately 58%.

The multivariate (2-year) prediction improvement from Table 4 could thus have come either from the fact that it includes another 8 months of relevant stock returns or because it weights older days less. Both contribute nearly equally to the improvement. The  $R^2$  improvements are as follows:

Using the last 12 months (baseline)	55.2%
Using the last 20 months (improved)	56.0%
Using two separate 12 months (multivariate)	58.2%

In turn, this suggests another possible predictive improvement: weighting past days less in the market-model time-series regression. For example, a 0.001/day decay weighted least squares (WLS) time-series regression reduces the weight of the first return of the year to approximately  $0.999^{252} \approx 78\%$ . At 0.002/day and 0.003/day, these first-day weights decrease to 60% and 47%, respectively. We try these weighting functions with up to 1 year and up to 5 years of historical return data. However, when predicting the same unweighted VCK forward betas, the time-decayed market-model betas yield no improvement:

Market Model Time-Series Weight Decay	1 Year	5 Years
Undecayed	55.2%	50.8%
Using a 0.001/day decay	55.2%	47.5%
Using a 0.002/day decay	55.0%	44.0%
Using a 0.003/day decay	54.8%	40.9%

In sum, we can recommend that instead of shrinking the 1-year VCK\_BETA to approximately 75%, a corporate manager could shrink the prevailing 1-year beta to 50% and add 25% of the shrunk beta from the year before.

### 6. Selection

In Table 5, we confirm that selection concerns are not responsible for the main findings. With the ex ante 5-year data requirement (which we use in Table 1 to keep the dependent variable identical for different time intervals), the gamma predictive coefficient is  $\gamma_1 \approx 0.74$ , and the  $R^2$  is 56.2%. Imposing a completely different selection scheme instead, an ex post selection criterion that insists on the full year of data and requires only 1 year of lagged data, we find that the coefficient is  $\gamma_1 \approx 0.70$ , and the  $R^2$  is 54.5%. Other variations produce similar results. From the perspective of a corporate executive, the main recommendation (i.e., 1–3 years of daily return betas, Vasicek shrunk and then shrunk another 20%–30%) remains intact.

TABLE 5  
Selection Biases

In Table 5, the second set of regressions drops the requirement of having longer historical prevailing betas and imposes an ex post survival requirement of at least 1 year. For other basic explanations, refer to Table 1. *Data Set:* First rows: 1-year-ahead (220 daily stock returns) and 5-year-ahead (1,000 daily stock returns) beta availability. Second rows: Only 1-year-ahead prevailing (220 daily stock returns) forward calendar stock returns, from CRSP 1927–2014 data. *Interpretation:* Even starkly different stock selection rules have only modest influence on the second-level shrinkage recommendation.

Dependent Variable	Set	Metric	Independent Variable: Daily BETA over Last Year				SD(y)	N
			OLS	VCK	OLS	VCK		
OLS_BETA	Ex ante	$R^2$	46.1	50.0			0.6	199,783
	5 years available	$\gamma_1$	0.68	0.81				
VCK_BETA	Ex post	$R^2$			43.2	47.5	0.6	280,181
	1 year available	$\gamma_1$			0.66	0.81		
VCK_BETA	Ex ante	$R^2$	50.7	56.2			0.5	199,783
	5 years available	$\gamma_1$	0.61	0.74				
VCK_BETA	Ex post	$R^2$			48.2	54.5	0.5	280,181
	1 year available	$\gamma_1$			0.60	0.70		

### 7. Pre- versus Post-1980

In Table 6, we confirm that the inference does not change greatly after 2000. The best estimator remains the VCK\_BETA, the best estimation window remains 1 to 3 years, and the best predictive  $\gamma_1$  coefficient remains around 0.75.

### 8. Market Cap

In Table 7, we divide the sample by market cap, beginning in 1973 (when NASDAQ came online). The table confirms that the gamma and  $R^2$  estimates are similar in the two tertials with large market capitalization ( $\gamma_1 \approx 0.8$ ,  $R^2 \approx 50\%$ ). For the small tertial, however, the prediction is worse. The best predictive coefficient is only  $\gamma_1 \approx 0.6$  using the 3-year estimation window (and  $\gamma_1 \approx 0.5$  using

TABLE 6  
Time Period

Table 6 breaks out pre-2000 and post-2000 observations. For other basic explanations, refer to Table 1. The best predictor in each row is highlighted in boldface. *Data Set:* As in Table 1. *Interpretation:* Shorter estimation windows are relatively better after 2000. The shrinkage coefficients remain about the same, but the explanatory power improves.

		Independent Variable: BETA Calculated over Last x Years						SD(y)	N
		1 Year		3 Years		5 Years			
Dependent Variable	Metric	OLS	VCK	OLS	VCK	OLS	VCK		
<i>Panel A. Pooled Post-2000 (daily return frequency)</i>									
OLS_BETA	$R^2$	56.6	<b>58.5</b>	54.2	54.6	48.4	48.6	0.62	64,505
	$\gamma_1$	0.74	0.82	0.80	0.83	0.79	0.81		
VCK_BETA	$R^2$	59.0	<b>61.6</b>	56.8	57.5	50.8	51.1	0.56	64,505
	$\gamma_1$	0.69	0.77	0.74	0.77	0.73	0.76		
<i>Panel B. Pooled Pre-2000 (daily return frequency)</i>									
OLS_BETA	$R^2$	41.1	45.6	45.5	<b>46.9</b>	43.8	44.6	0.62	135,278
	$\gamma_1$	0.64	0.81	0.78	0.85	0.81	0.86		
VCK_BETA	$R^2$	46.4	52.9	51.5	<b>53.9</b>	49.8	51.1	0.52	135,278
	$\gamma_1$	0.57	0.73	0.70	0.77	0.72	0.77		

TABLE 7  
By Lagged Stock Market Cap (1972–2014)

In Panel A of Table 7, stocks are grouped into tertials based on market cap at the final day over which the x variable (the prevailing beta) is calculated. In Panel B, stocks are value-weighted. For other basic explanations, refer to Table 1. The best predictor in each row is highlighted in boldface. *Data Set:* As in Table 1. However, the data begin later because NASDAQ is added in Dec. 1972 (increasing the number of stocks on CRSP from 2,606 to more than 5,000). With 5 years of required data availability, the first predicted beta is effectively for 1978. This time period is sometimes referred to as "post-NASDAQ." *Interpretation:* Smaller-stock betas benefit from a longer estimation window, are harder to predict, and require more second-stage shrinkage.

		Independent Variable: BETA Calculated over Last x Years						SD(y)	N	
		1 Year		3 Years		5 Years				
Set	Dependent Variable	Metric	OLS	VCK	OLS	VCK	OLS	VCK		
<i>Panel A. Pooled Tertial-Specific Prediction</i>										
Small cap	OLS	$R^2$	9.6	14.6	15.3	<b>17.3</b>	15.4	16.5	0.59	41,451
		$\gamma_1$	0.32	0.56	0.52	0.64	0.56	0.63		
	VCK	$R^2$	15.1	25.1	23.7	<b>28.3</b>	24.2	26.6	0.41	41,451
		$\gamma_1$	0.28	0.51	0.45	0.57	0.14	0.56		
Mid cap	OLS	$R^2$	45.5	<b>47.8</b>	44.6	45.4	39.6	40.2	0.61	51,834
		$\gamma_1$	0.68	0.79	0.77	0.82	0.78	0.81		
	VCK	$R^2$	48.8	<b>52.5</b>	47.9	49.3	42.3	43.2	0.53	51,834
		$\gamma_1$	0.61	0.73	0.70	0.75	0.70	0.74		
Large cap	OLS	$R^2$	55.6	<b>56.0</b>	53.3	53.4	47.1	47.1	0.53	70,087
		$\gamma_1$	0.73	0.80	0.79	0.83	0.78	0.80		
	VCK	$R^2$	55.6	<b>56.8</b>	53.6	54.1	47.4	47.6	0.48	70,087
		$\gamma_1$	0.67	0.74	0.73	0.76	0.72	0.74		
<i>Panel B. Pooled Value-Weighted Prediction</i>										
All cap	OLS	$R^2$	54.8	<b>55.2</b>	51.4	51.4	45.9	45.8	0.60	163,372
		$\gamma_1$	0.73	0.79	0.78	0.80	0.76	0.77		
	VCK	$R^2$	54.7	<b>55.6</b>	51.6	51.8	46.1	46.2	0.51	163,372
		$\gamma_1$	0.68	0.74	0.73	0.75	0.71	0.73		

the 1-year window). When predicting future betas for small-cap firms, managers should use longer stock return histories and shrink beta estimates more.<sup>14</sup>

In Table 8, we investigate whether the prediction can be improved by using low trading-frequency nonsynchronicity corrections. To do this, we select the smallest 1,000 stocks (again, beginning in 1973) and apply the Dimson (DMS) adjustment (which is the sum of the coefficients of the 1-day-past, current, and 1-day-ahead market-model time-series coefficients). We find that the Dimson adjustment is not helpful. It cannot even predict its own future realization better than the VCK\_BETA. In predicting the future OLS\_BETA or VCK\_BETA, it performs considerably worse.

TABLE 8  
Smallest 1,000 Stocks, Post-NASDAQ, With and Without Dimson Adjustment

In Table 8, we use only the smallest 1,000 stocks *before/after* the prevailing beta formation period. The Dimson-adjusted market beta (DMS) is the sum of three market-model time-series coefficients as explained in detail by Dimson (1979). For other basic explanations, refer to Table 1. The best predictor in each row is highlighted in boldface. *Data Set:* The data in this table first select the smallest 1,000 stocks and then require them to have 5 years of data. CRSP data begin after NASDAQ's introduction in 1972. Over 40 years, this leaves only approximately 14,000 observations. *Interpretation:* The Dimson adjustment worsens the prediction. The DMS\_BETA cannot even predict itself well.

Dependent Variable		Independent Variable: BETA Calculated over Last x Years									SD(y)	N
		1 Year			3 Years			5 Years				
		OLS	VCK	DMS	OLS	VCK	DMS	OLS	VCK	DMS		
OLS_BETA	$R^2$	4.9	10.1	4.1	9.3	<b>11.9</b>	8.3	10.2	11.7	9.8	0.6	14,574
	$\gamma_1$	0.23	0.52	0.16	0.45	0.62	0.33	0.52	0.61	0.40		
VCK_BETA	$R^2$	9.8	22.3	7.7	17.9	<b>24.7</b>	15.8	20.1	23.8	19.0	0.4	14,574
	$\gamma_1$	0.21	0.50	0.14	0.40	0.57	0.29	0.47	0.56	0.36		
DMS_BETA	$R^2$	4.3	8.4	3.5	7.9	<b>9.7</b>	7.4	8.8	<b>9.7</b>	8.9	0.9	14,556
	$\gamma_1$	0.29	0.64	0.19	0.56	0.75	0.41	0.65	0.75	0.51		

## 9. Financial Statement Information

A number of commercial data vendors sell risk metrics (including beta estimates) that are based on “refining” their estimates with financial statements and other information. Some skepticism is in order. This is not only because vendors, such as Morgan Stanley Capital International (MSCI), do not make their historical estimates available for benchmark testing. MSCI, which acquired Barra in 2004, no longer makes beta or risk metrics available to academics. Thus, and raising suspicion, it is impossible to assess the quality of their estimates.

We also try to replicate a number of academic papers that advocate accounting models to calculate implied costs of capital. Our investigations find that these

<sup>14</sup>In the post-1973 sample, both the mean and the standard deviation of the market betas are smaller for smaller firms. This explains why the  $R^2$  of the predictive regression does *not* improve greatly when we include a cross-dummy that allows the VCK gamma to vary with quantile. The standard deviation of the VCK\_BETA for the bottom tertial market cap is approximately 0.4, whereas the standard deviation in the other two tertials is approximately 0.5. With fewer data points, the range of betas for the smaller firms is also smaller. This means that there are relatively few cases in which small-firm predictions are far off the center. It would be these cases in which the 0.56 predictive gamma coefficients (for the smaller firms) rather than the 0.73 (for the larger firms) would make a difference in predicted values. In turn, this means that including the market cap cross-product variable does not greatly improve the overall regression  $R^2$ .

accounting models perform no better than simpler earnings–price (EP) models, and EP itself performs modestly better only in a subset of firms (for up to 5 years). Moreover, these models are sensitive to some seemingly innocuous specification changes, and *they work well only in the first 2 decades of the sample, but no longer in the second 2 decades*.<sup>15</sup>

Although conditioning exposures on other variables could be useful in the context of asset pricing models (Jagannathan and Wang (1996)), they are unlikely to perform better in our long-term cost-of-capital context. Managers need to predict the cost of capital for future cash flows for long-term projects *today*. Thus, cost-of-capital estimates can only depend on variables known *today*, not on future conditioning variables.<sup>16</sup>

In Table 9, we investigate some variables from the financial statements (assuming a 3-month publication delay) by using predictive regressions of the form

$$\begin{aligned} \text{VCK\_BETA}_{i,t} = & \gamma_0 + \gamma_1 \times \text{VCK\_BETA}_{i,t-1} + \gamma_2 \times F_{i,t-1} \\ & + \gamma_3 (F_{i,t-1} \times \text{VCK\_BETA}_{i,t-1}) + \epsilon_{i,t}, \end{aligned}$$

where  $F_i$  is one of the variables defined in the table.

Table 9 shows that none of these variables offers economically large improvements.<sup>17</sup> Using more than 20 variables together improves the  $R^2$  in predicting 1-year-ahead market beta only from 44.74% to 45.60%. In predicting 5-year total betas for the same observations, the predictive power increases from 40.27% to 41.67%. Imposing (economic) restrictions on sensible coefficients would further reduce this optimistic empirical improvement in  $R^2$ .

## 10. Industries

In Table 10, we predict the betas for the 49 industry portfolios defined by Fama and French (1997).<sup>18</sup> Given the smaller standard deviations among industry returns, the explanatory power between OLS\_BETAs and VCK\_BETAs is more similar, although OLS\_BETAs still require approximately 5% more shrinkage than VCK\_BETAs. The conclusion is again that 1-year daily stock-return-based

<sup>15</sup>The accounting models are not *standard* cost-of-capital models that are either nowadays recommended in corporate finance textbooks or in wide use. Thus, a more detailed investigation of such models is beyond the scope of our paper.

<sup>16</sup>For the same reason, the momentum factor is not useful for predicting long-run cost of capital beyond a few years. Managers cannot meaningfully predict whether their projects will be “momentum stocks” in a few years.

<sup>17</sup>This is from a corporate capital budgeting perspective, in which these tiny improvements are swamped by other uncertainties. From a corporate capital budgeting perspective, it is further undesirable that these extended regressions require comprehensive historical estimations to obtain the extra coefficients.

<sup>18</sup>Fama and French (1997) state on page 154 that their rationale is that aggregated industry portfolios should favor the models: *Industries give an understated picture of the problems that will arise in estimating risk loadings for individual firms and investment projects*. Analogously, if the models cannot make solid recommendations for individual stocks, giving recommendations for individual projects would seem even more fruitless. This is the case not only because of the further disaggregation but because individual projects rarely have price histories to allow researchers even to investigate (much less confirm) the performance of these models. There is also simply no reference data set for individual projects on which academic research could test models. Unfortunately, this does not absolve executives from the need to assign hurdle rates, and they have to do so not to industries, not even to their overall firms, but to their individual projects.

TABLE 9  
Predicting 1-Year-Ahead VCK\_BETA with Lagged VCK\_BETA and Financial Statement Data

In Table 9, additional prevailing financial statement variables are used to improve the predictive regressions. When we include any such variables, we always include both the variable itself and the cross product of this variable with the prevailing market beta. The 5-year-ahead estimate predicts the total forward beta (not the marginal 1-year forward beta) over a 5-year window. BLIABRATIO is total liabilities (Compustat: LT) divided by total assets (AT). MLIABRATIO is the asset AT denominator converted into market value by subtracting the book value of equity (CEQ) and adding the (CRSP) market value of equity (MVE). BLEVRATIO is the book value of debt (DLC + DLTT), divided by debt plus equity (CEQ). MLEVRATIO is like BLEVRATIO, but with MVE instead of CEQ. CASH is cash (CHE) normalized by assets (AT). EARN\_GROWTH is the growth in operating income after depreciation (OIADP). EARN\_VOL is the absolute earnings growth change,  $\text{abs}(\text{EG}(t) - \text{EG}(t-1))$ . NEG\_EARN is a dummy for whether earnings growth is negative,  $\text{EG}(t) < 0$ . BOOK-TO-MARKET is book value divided by market value,  $\text{CEQ}/\text{MVE}$ . TAXES\_ONEARN is income taxes (TXT) divided by earnings (EBIT). TAXES\_ONSLS is income taxes (TXT) divided by sales (REV). CAPEX is capital expense (CAPX) divided by book assets (AT). ISSUING is the sale of common and preferred stock (SSTK) divided by the market value of equity (MVE). For other basic explanations, refer to Table 1. *Data Set:* As in Table 1, except that all prevailing financial statement variables must be available. *Interpretation:* These financial statement variables do not greatly improve the prediction of future betas.

Independent Variable(s)	$R^2$	
	1-Year	5-Year
VCK_BETA (only)	44.74	40.27
VCK_BETA, BLIABRATIO, and VCK_BETA × BLIABRATIO	44.77	40.59
VCK_BETA, MLIABRATIO, and VCK_BETA × MLIABRATIO	45.10	40.93
VCK_BETA, BLEVRATIO, and VCK_BETA × BLEVRATIO	44.74	40.27
VCK_BETA, MLEVRATIO, and VCK_BETA × MLEVRATIO	44.85	40.57
VCK_BETA, CASH, and VCK_BETA × CASH	45.01	40.58
VCK_BETA, EARN_GROWTH, and VCK_BETA × EARN_GROWTH	44.74	40.27
VCK_BETA, EARN_VOL, and VCK_BETA × EARN_VOL	44.74	40.27
VCK_BETA, NEG_EARN, and VCK_BETA × NEG_EARN	44.83	40.31
VCK_BETA, BOOK-TO-MARKET, and VCK_BETA × BOOK-TO-MARKET	44.74	40.27
VCK_BETA, TAXES_ONEARN, and VCK_BETA × TAXES_ONEARN	44.75	40.27
VCK_BETA, TAXES_ONSLS, and VCK_BETA × TAXES_ONSLS	44.75	40.27
VCK_BETA, CAPEX, and VCK_BETA × CAPEX	44.74	40.27
VCK_BETA, ISSUING, and VCK_BETA × ISSUING	44.74	40.28
VCK_BETA, ALL_PRECEDING, and VCK_BETA × ALL_PRECEDING	45.56	41.67
No. of firm-years	112,391	69,097

TABLE 10  
Predicting 49-Industry VCK\_BETAs

In Table 10, the observation units are no longer individual stocks but the 49-industry portfolios defined by Fama and French (1997). For other basic explanations, refer to Table 1. The best predictor in each row is highlighted in boldface. *Data Set:* Kenneth French's 49-industry portfolio returns, 1927–2013. The first prediction is for 1932. *Interpretation:* The 49-industry market betas can predict 1-year-ahead 49-industry market betas. The 1-year event windows remain best, and OLS\_BETAs are almost as good as VCK\_BETAs.

Dependent Variable		Independent Variable: BETA Calculated over Last x Years							
		1 Year		3 Years		5 Years		SD(y)	N
		OLS	VCK	OLS	VCK	OLS	VCK		
OLS_BETA	$R^2$	<b>64.7</b>	64.4	58.6	58.7	51.0	51.0	0.34	3,827
	$\gamma_1$	0.80	0.85	0.82	0.85	0.81	0.81		
VCK_BETA	$R^2$	64.2	<b>64.7</b>	58.0	58.4	50.6	50.8	0.32	3,827
	$\gamma_1$	0.75	0.80	0.77	0.79	0.74	0.76		

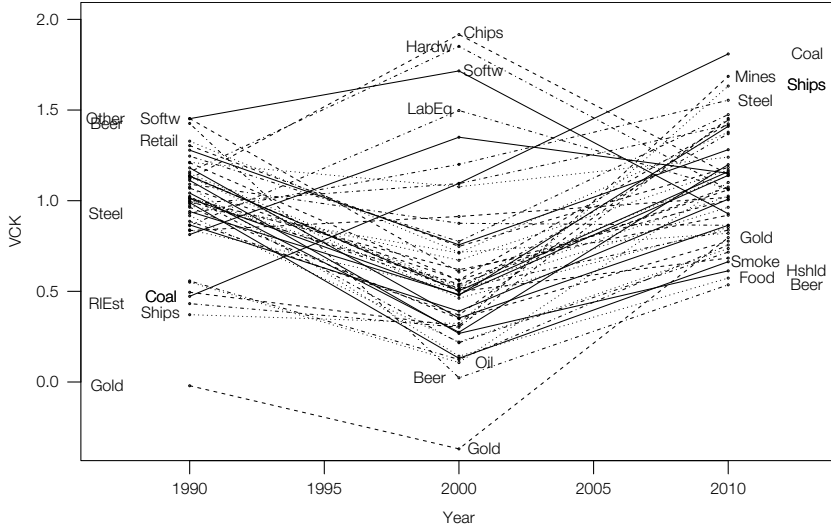
estimation windows remain best, now with a higher  $R^2$  of approximately 65% and a higher gamma of approximately  $\gamma_1 \approx 0.8$ .

With only 49 industries, it is possible to visualize the mean reversion in betas. Figure 2 plots industry market betas in 1990, 2000, and 2010. The figure shows that industry beta rankings are not stable. It is *not* difficult to invent ex post rationalizations. For example, betas could have changed because measurement error is still a factor, even for portfolios as highly aggregated as industries and even

FIGURE 2

The 1-Year Daily 49-Industry OLS\_BETAs in 1990, 2000, and 2010

Figure 2 breaks out the beta estimates for 1990, 2000, and 2010 for the 49 industries from Table 10. Years in between are ignored. Interpretation: Market betas show strong but imperfect mean reversion over 5- to 10-year horizons.



after VCK shrinkage; some industries experienced ups; others experienced downs; the financial crisis of 2008 changed leverage and firm size, especially in the financial services industries; and so on. But it is difficult to come up with ex ante predictions for how betas will change in the future.

### E. Substituting with Peer Exposures

It is often the case that corporate managers have no history of stock returns for the new projects for which they want to assess the cost of capital. In such cases, one common prescription is to rely on the market betas of peers or the market beta of the industry instead. In Table 11, we investigate how well such substitutes perform.

In Panel A of Table 11, we predict stocks' own betas not only with their own lagged betas but also with their own 49-industry betas. The table shows that whereas own stock beta has an  $R^2$  of more than 50% in predicting stock beta, lagged industry beta has an  $R^2$  of only 4%. If the stock's own beta and the industry beta are both included, the predictive power of the trivariate model does not surpass that of using only the stock's own beta. *Industry betas are practically useless for individual stocks.* When own market betas are available, the corporate finance textbook suggestion of using industry market betas "because they are less noisy" is not justifiable.

In Panel B of Table 11, we describe a variance decomposition. Neither industry nor market-value groups can explain much of the total variance across market betas. Most of the variance remains idiosyncratic. The overall variation of the Vasicek 1-year beta in the full sample is 0.55. The cross-industry variation is the standard deviation of the 49-industry beta means. For the within-industry

TABLE 11  
Predicting 1-Year-Ahead 49-Industry and Peer Proxied Betas

In Panel A of Table 11, we predict beta not only with the own lagged beta but also with 49-industry-proxied betas. To be included, stocks must have data on current and lagged betas, market caps, and book-to-market ratios, and they must be in 1 of the (Fama–French) 49 industries with at least 5 firms. In Panel B, we decompose the standard deviation into across and within variations. In Panel C, the sample is post-NASDAQ, resulting in a first prediction for 1978, and the stocks must have 5 years of data plus data on market caps and book-to-market ratios and at least 5 peers in 1 of the 49 industries. The proxies are similar-size stocks in the same 49 industry sectors. The peer betas are designated averages of individual stock VCK\_BETAs. For other basic explanations, refer to Table 1. The best predictor in each row is highlighted in boldface. *Data Set:* Panel A uses 5-year-ahead prevailing beta (1,000 daily stock returns) and 1-year-ahead beta (220 daily stock returns) for 1927–2014. Panel B requires 220 trading days and data to classify industry or market cap. Panel C requires availability of lagged market cap and at least 5 firms in the 49-industry-year set for peer firm identification. The data are from CRSP after the addition of NASDAQ in 1972. *Interpretation:* Panel A shows that the 49-industry market betas should not be used as substitutes for own stock betas. Panel B shows that most of the variance is idiosyncratic, not specific to industry or market value. Panel C shows that using betas from stocks with similar market caps within the same 49 industries provides worse but more reasonable substitutes for firms' own betas.

*Panel A. With 49-Industry Proxies*

Dependent Variable	Independent Variable	$R^2$		SD(y)	N
		OLS	VCK		
OLS_BETA	Bivariate, own OLS_BETA only	43.2	<b>47.5</b>	0.64	277,286
	Bivariate, 49-industry OLS_BETA only	3.9	3.9	0.64	277,286
	Trivariate, both OLS_BETAs	43.3	<b>47.5</b>	0.64	277,286
VCK_BETA	Bivariate, VCK_BETA only	48.2	<b>54.5</b>	0.54	277,286
	Bivariate, 49-industry VCK_BETA only	4.1	4.1	0.54	277,286
	Trivariate, both VCK_BETAs	48.3	<b>54.5</b>	0.54	277,286

*Panel B. Standard Deviation Decompositions*

Pooled, Daily Data, VCK_BETA	Total	Across	Within	N
49-industry (1 year)	55%	15%	52%	305,865
49-industry (5 year)	50%	16%	47%	244,103
Market-value quintiles (1 year)	52%	21%	49%	264,441
Market-value quintiles (5 year)	48%	20%	44%	205,688
Market-beta quintiles (1 year)	54%	37%	43%	284,868

*Panel C. With Individual Firm Market Peers, Same 49-Industry, VCK\_BETAs Prediction Only*

Independent Variable(s)	$\gamma_{own}$	$\gamma_{peer}$	$R^2$	SD(y)	N
Own VCK_BETA	0.71		51.2%	0.36	82,158
Market cap peer firm VCK_BETA, nearest 1		0.46	21.1%	0.46	82,158
Market cap peer firms VCK_BETA, nearest 2		0.60	27.8%	0.44	82,158
Market cap peer firms VCK_BETA, nearest 4		0.71	32.4%	0.42	82,158
Trivariate: VCK_BETA, own and nearest 1	0.66	0.11	52.0%	0.36	82,158
Multivariate: VCK_BETA, own and nearest 2	0.63	0.16	52.4%	0.36	82,158
Multivariate: VCK_BETA, own and nearest 4	0.60	0.21	52.8%	0.35	82,158
Peer firm VCK_BETA, nearest book-to-market		0.29	8.0%	0.49	85,963
Peer firm VCK_BETA, nearest 2 book-to-market		0.42	12.0%	0.48	85,963

variation, we subtract the industry beta mean and calculate the standard deviation of the full sample. The cross-industry beta standard deviation is 0.15; the within-industry standard deviation is 0.55. Both industry and market cap are only modest correlates. The final row presents an upper limit on explanatory power, computing within versus across variation by lagged market beta. The quintiles have 37%/54% residual across-group variation. (It is not suitable as an instrument when market beta is unknown. When market beta is known, it is better used itself.)

In Panel C of Table 11, we proxy the market beta of one stock with those of other stocks. The best proxies that we can identify are the betas of stock(s) that have similar market caps. However, it takes the average of two larger and two smaller firms' betas to reach a similar shrinkage recommendation. Even then, the average peer beta has only roughly half the explanatory power (i.e., 30% rather than 50%) in the predictive regression. This table also shows that very small



improvements can be obtained by including both own betas and peer betas (52.8% rather than 52.0%). Finally, the table shows that firms with similar book-to-market ratios (in the industry) are not good peers.

Moreover, we suspect, but we cannot test, that project costs of capitals suffer from similar beta nonrepresentability problems relative to their stock (aggregates), in the same way that stocks suffer from beta representability problems relative to their industry (aggregates). We know of no canonical data set of corporate projects with rates of returns on which we could examine this. If our suspicion is correct, then it only further reinforces our conclusions.

#### IV. Time Decay

Few corporate projects that require cost-of-capital estimates are as short as 1 year. Interesting cost-of-capital concerns arise more in the context of longer-term projects. The time-varying mean-reverting nature of betas suggests that managers should not only limit the estimation window for the historical stock return input data and shrink their estimates but also take into account what they need the beta for: a near or far-off expected cash flow. Over longer forward horizons, the beta is more likely to revert.

Panel A of Table 12 shows that the problem is not merely estimation error but underlying beta mean reversion. If the underlying true beta process does not change, then the prediction of any 1-year market beta should be as good for the 1-year beta in 1 year as it would be for the 1-year beta in 2 years or 3 years. The gammas and  $R^2$ s should remain constant. Instead, Panel A shows that the  $\gamma_1$ s decline by approximately 7%–9% for every year further in the future.

Panel B of Table 12 shows the effect on predicting a market beta over a fixed interval beginning today. To predict the market beta over the next 1 month of daily stock returns, the estimation noise in the dependent variable is more important than the time decay, suggesting a gamma estimate of only  $\gamma_1 \approx 0.67$ . (But recall that the needed input is not the realized beta; rather, it is the expected beta!) To predict the market beta beyond the next 3 years, the time decay is more important than the estimation noise, suggesting gammas under 0.70. Realized beta estimates are best forecast when working with 1- to 2-year cash flows (with a gamma of approximately 0.7). For 10-year cash flows, users should shrink market betas by approximately 40% rather than just 30%.

In an earlier version of our paper, we fit the observed empirical moments (in particular, the decay in gamma when future months are predicted) to a data-generating model. Although this is an interesting exercise, it is ultimately not useful in our context. We are not trying to assess an underlying model (or derive an optimal estimator in the context of a specific underlying model assumption). We are merely trying to advise users about linear forward-looking beta estimates, given access only to the prevailing empirical data and not to the true underlying exposures. Thus, this moment-matching model yields no further insights of great application use: The estimated beta is noisy, and our empirical gamma coefficient suffers from an EIV problem (i.e., downward bias) relative to the true gamma (i.e., if the inputs could be true market betas). The prediction quality decreases with time-varying underlying betas and uncertainty and estimation windows that are too short or too long (or weighting functions that decay too steeply or not

TABLE 12  
Time Decay of Prediction: Daily VCK\_BETAs

In Panel A of Table 12, the 1-year beta in  $x$  years is predicted by the prevailing 1-year daily return frequency VCK\_BETA. The leftmost column describes estimation variations. In Panel B, the total beta over the period indicated in the column header is predicted by the prevailing 1-year daily return frequency VCK\_BETA. For other basic explanations, refer to Table 1. *Data Set:* As indicated in the left column, beginning in 1927. *Interpretation:* The underlying beta process is time-varying, so a 10-year market-beta forecast should be shrunk by another 10% relative to a 1-year market-beta forecast.

*Panel A. Predicting 1-Year VCK\_BETA in  $x$  Years*

Method and Set	Metric	1 Year	2 Years	3 Years	5 Years	10 Years
Any stocks $N \in [284,868, 129,794]$	$R^2$	54.4%	44.5%	38.4%	28.2%	16.6%
	$\gamma_1$	0.73	0.66	0.61	0.52	0.39
Same stocks: 5 years $N = 195,575$	$R^2$	54.2%	43.7%	38.0%	28.3%	—
	$\gamma_1$	0.74	0.66	0.62	0.52	—
Same stocks: 10 years $N = 127,621$	$R^2$	55.6%	45.1%	38.7%	28.2%	16.8%
	$\gamma_1$	0.74	0.67	0.62	0.53	0.40
Fama–MacBeth stocks daily $T \in [87, 78]$	$R^2$	54.2%	43.7%	38.0%	30.1%	20.5%
	$\gamma_1$	0.74	0.66	0.61	0.54	0.42
Fama–MacBeth same stocks daily VCK $T = 83$	$R^2$	54.8%	44.0%	38.2%	30.2%	—
	$\gamma_1$	0.74	0.67	0.62	0.54	—
Fama–MacBeth same stocks daily $T = 87$	$R^2$	55.7%	45.1%	39.3%	31.0%	20.6%
	$\gamma_1$	0.75	0.68	0.63	0.55	0.42

*Panel B. Predicting Total VCK\_BETA over  $x$  Years*

Method and Set	Metric	1 Month	6 Months	1 Year	2 Years	3 Years	5 Years	10 Years
Any stocks, pooled	$R^2$	29.5%	50.7%	54.4%	54.6%	54.3%	50.4%	51.3%
	$\gamma_1$	0.67	0.73	0.73	0.70	0.71	0.69	0.62
	$N$	3,737,170	589,366	284,868	135,270	85,315	40,960	13,308
Same stocks: 10 years $N = 11,964$	$R^2$	26.6%	55.3%	54.6%	58.1%	57.5%	56.7%	52.6%
	$\gamma_1$	0.72	0.77	0.77	0.78	0.77	0.74	0.63
Fama–MacBeth: 10 years $T = 7$	$R^2$	28.6%	53.7%	58.8%	59.7%	57.9%	55.0%	52.4%
	$\gamma_1$	0.71	0.76	0.78	0.76	0.74	0.70	0.62

steeply enough). When we use the model's simulated data, we can fit the empirical data well, and the best shrinkage coefficient quite closely matches the one deduced from the empirical coefficients.

## V. Other Exposures and Models

### A. Bond and Other Exposures to the U.S. Stock Market

The Federal Reserve Economic Data (FRED) provides daily yield data based on the Bank of America Merrill Lynch bond indexes beginning in 1996. This makes it possible to calculate (approximate) daily bond returns and hence bond stock market betas for each calendar year.

Table 13 first confirms the Campbell, Sunderam, and Viceira (2017) result that the 10-year Treasury bond market beta declines by approximately 0.015 every year, ending up near  $-0.2$  in 2015. Without a clear reason, this points out how difficult it is to intuitively assess market betas even for large asset classes.<sup>19</sup> Investment-grade corporate bonds suffer smaller annual market-beta declines (approximately  $-0.003$  per year), but their market betas are also not positive by

<sup>19</sup>Unlike for equity market betas, we would not expect reversion and thus would not recommend shrinkage toward 1.0 for bonds. We have evidence that investment-grade bond betas have declined over the last decades, but without a good reason why (as is also the case for Campbell et al. (2017)), we cannot recommend projecting this to continue.

TABLE 13  
Corporate Bond Betas with Respect to the Stock Market

On the left-hand side of Table 13 are year-end yields (12/31) from FRED. For the right-hand side, we first interpolate daily bond returns from reported FRED yields. Days on which a bond or stock market return is not available (e.g., Sept. 11–16, 2011) are omitted, yielding extra multiple-day return inputs into the market-model regressions. Annual market betas are with respect to the CRSP value-weighted market (VWRETD) and net of the 1-year Treasury (DGS1), each from approximately 250 daily returns. For the constant quality bonds, we assume an equivalent 5-year zero-bond maturity. Changes in calculation assumptions can make a difference of approximately 0.1. The last column is from a simple time-series regression on 19 annual market betas in which the independent variable is a year index. (The time trend in market betas happens to be highly multicollinear with the beta auto-coefficient. Thus, we can report only one or the other, and the trend variable fits better.) Except for the 1- to 7-year portfolios, the time trend is statistically significant, despite being based only on 19 observations each. *Data Set:* The series are from FRED. (The series name is in the second column.) Although not reported, the results for the 10-year Treasury (DGS10) are similar but not identical if we use the CRSP 10-Year Constant Maturity Treasury Bond Series (Series ID 2000007). CRSP uses the single nearest bond series, whereas the Treasury interpolates a spline across the curve. *Interpretation:* Investment-grade corporate bonds (BBB or better) have 0 market betas with the stock market as of 2015. Noninvestment-grade bonds have betas of approximately 0.1–0.3.

Bond Category	FRED Series ID	Yield		OLS_BETA		
		1996	2015	1997	2015	Time Trend
1-year Treasury	DGS1	5.51%	0.65%	—	—	—
10-year Treasury	DGS10	6.43%	2.27%	0.11	-0.18	-0.015
AAA	C0A1CAAAY	6.72%	2.86%	0.04	-0.06	-0.004
AA	C0A2CAAAY	6.82%	2.82%	0.03	-0.05	-0.003
A	C0A3CAAY	6.97%	3.16%	0.03	-0.05	-0.003
BBB	C0A4CBBBEY	7.26%	4.42%	0.04	-0.03	-0.003
B	H0A2HYBEY	10.01%	8.96%	0.03	0.13	+0.007
CCC	H0A3HYCEY	13.97%	18.15%	0.02	0.19	+0.011
1–3 years	C1A0C13YAY	6.38%	2.16%	0.01	-0.01	+0.000
3–5 years	C2A0C35YAY	6.67%	2.94%	0.03	-0.03	-0.002
5–7 years	C3A0C57YAY	6.88%	3.70%	0.03	-0.05	-0.003
7–10 years	C4A0C710YAY	7.04%	4.11%	0.03	-0.04	-0.004
10–15 years	C7A0C1015YAY	7.20%	4.61%	0.08	-0.10	-0.012
15+ years	C8A0C15PYAY	7.43%	5.17%	0.10	-0.16	-0.024

2015. Only noninvestment-grade bonds still have positive (and possibly increasing) market betas in 2015, with CCC bond betas in the range of +0.1 to +0.3.

Some asset holdings are often deemed relevant and worth breaking out for managerial judgments about project market betas. We also briefly look at (U.S.) market betas for some exchange rates, foreign stock market indexes, and commodities; the untabulated results are summarized as follows:

*Currencies.* From 1970 to 2015, the U.S. dollar index has market betas ranging between -0.2 and +0.3 (approximately +0.1 in 2015), the yen between -0.1 and 0.2 (+0.2), the euro between -0.1 and 0.2 (-0.1), and the British pound between 0.0 and 0.1 (0.0).

*Foreign Stocks.* As of late 2015, the Hong Kong and Nikkei stock markets seem to have U.S. market betas of approximately 0.3. This is counter to common wisdom that foreign stock markets currently offer little diversification. The FTSE and DAX are more strongly correlated with the U.S. market, with betas of approximately 0.8 (and steadily drifting up).

*Commodities.* Our findings are in line with those of Bhardwaj, Gorton, and Rouwenhorst (2015). For nearly a decade, the oil price has had a strong positive market beta. Platinum, gold, silver, soybeans, and other commodities have low or time-varying market betas.<sup>20</sup>

<sup>20</sup>However, prevailing common correlations seem unreliable insurance in a strong downturn. For example, on Oct. 15, 1987, when the market declined by approximately 10%, Platinum also declined by 7%, and soybeans by 4%.

## B. International Market Betas

In Table 14, we estimate the *local* market betas of individual stocks from the United Kingdom, Germany, France, Australia, Japan, China, and India with respect to a self-constructed market-cap-weighted market index composed of the firms in the Compustat Global Vantage sample. In the first four countries, the best estimates are VCK\_BETAs estimated over 1-year windows. In Japan, a 3-year estimation window performs modestly better. In China, OLS\_BETAs perform modestly better. And in India, OLS\_BETAs perform better in predicting OLS\_BETAs, and VCK\_BETAs perform better in predicting VCK\_BETAs. But in all countries, the predictive  $\gamma_1$  coefficients range from 0.60 to 0.80, suggesting 20%–40% shrinkage.

Table 15 shows that U.S. firms seem to have modestly more stable market betas than firms in the other countries. This finding should not be overread: It

TABLE 14  
International Own-Market Betas

In Table 14, the regressions are for other countries. The market indices are self-computed value-weighted market cap portfolios (with individual stock returns truncated at  $-0.5$  and  $+2.0$  and with maximum firm size truncated to 5% of the total country market capitalization). For other basic explanations, refer to Table 1. The best predictor in each row is highlighted in boldface. *Data Set:* Compustat Global Vantage. All stocks with 5-year (1,000 daily stock returns) and 1-year (220 daily stock returns) market-beta availability. Lagged market cap is used to calculate the local market index. For India, daily stock return data start in 1991; for China, they start in 1993; and for all other countries, they start in 1986. *Interpretation:* Forward market betas are more difficult to predict, especially in the Pacific Rim countries. Yet the basic recipe remains the same: 1–3 years of daily return data, VCK shrunk first, and then VCK shrunk again.

Country	Dependent Variable	Metric	Independent Variable: BETA Calculated over Last x Years						SD( $\gamma$ )	N
			1 Year		3 Years		5 Years			
			OLS	VCK	OLS	VCK	OLS	VCK		
United Kingdom	OLS	$R^2$	49.0	<b>52.5</b>	48.5	49.5	44.9	45.4	0.5	14,725
		$\gamma_1$	0.71	0.84	0.80	0.86	0.81	0.84		
	VCK	$R^2$	53.3	<b>58.5</b>	53.4	55.0	49.7	50.5	0.4	14,725
		$\gamma_1$	0.64	0.76	0.72	0.78	0.73	0.77		
France	OLS	$R^2$	49.3	<b>52.8</b>	46.3	46.7	41.9	41.9	0.5	6,371
		$\gamma_1$	0.71	0.82	0.78	0.82	0.79	0.81		
	VCK	$R^2$	54.0	<b>57.8</b>	49.9	50.6	44.7	44.8	0.4	6,371
		$\gamma_1$	0.65	0.76	0.71	0.75	0.72	0.74		
Germany	OLS	$R^2$	21.8	<b>30.4</b>	24.0	28.2	21.2	23.1	0.6	8,559
		$\gamma_1$	0.47	0.75	0.63	0.76	0.63	0.71		
	VCK	$R^2$	31.7	<b>48.5</b>	38.3	45.1	33.7	36.8	0.5	8,559
		$\gamma_1$	0.42	0.69	0.58	0.71	0.58	0.66		
Australia	OLS	$R^2$	17.0	<b>23.5</b>	20.0	22.3	17.6	19.3	0.7	8,328
		$\gamma_1$	0.43	0.68	0.64	0.77	0.68	0.80		
	VCK	$R^2$	25.2	<b>35.3</b>	28.6	32.3	26.0	28.2	0.5	8,328
		$\gamma_1$	0.38	0.60	0.54	0.66	0.59	0.70		
Japan	OLS	$R^2$	39.3	39.9	40.1	<b>40.5</b>	39.9	40.2	0.4	47,801
		$\gamma_1$	0.62	0.69	0.72	0.76	0.78	0.81		
	VCK	$R^2$	40.6	41.8	41.6	<b>42.3</b>	41.6	42.1	0.4	47,801
		$\gamma_1$	0.57	0.64	0.67	0.70	0.72	0.75		
China	OLS	$R^2$	<b>36.3</b>	35.5	34.6	34.2	27.7	27.6	0.3	10,696
		$\gamma_1$	0.61	0.68	0.71	0.76	0.70	0.74		
	VCK	$R^2$	<b>35.3</b>	34.9	33.5	33.4	26.8	26.8	0.3	10,696
		$\gamma_1$	0.53	0.59	0.62	0.66	0.61	0.64		
India	OLS	$R^2$	<b>35.4</b>	35.2	33.0	32.8	30.2	30.2	0.5	9,327
		$\gamma_1$	0.64	0.71	0.79	0.83	0.87	0.91		
	VCK	$R^2$	36.0	<b>36.6</b>	33.1	33.3	31.0	31.2	0.5	9,327
		$\gamma_1$	0.57	0.65	0.71	0.75	0.79	0.82		

TABLE 15  
Predicting Longer-Term VCK\_BETAs

Table 15 extends the dependent variable in Table 14 up to 5 years. Thus, to be included, a stock must have both a 1-year-ahead and a 5-year-ahead beta. There are fewer firm-years than there are in Table 14 because we require at least 1,000 nonmissing stock returns in the historical beta calculation. The independent variable is always the prevailing 1-year VCK\_BETA and is identical in each row. *Data Set:* See Table 14, but with the additional restriction of 5-year-prevailing and 5-year-ahead beta availability. (The U.S. data are post-NASDAQ.) Lagged market cap is used to calculate the local market index and must be available. *Interpretation:* U.S. local market betas are the most predictable betas. Decay patterns are similar in other countries.

Country	Auto-Coefficient $\gamma_1$			$R^2$			N
	In 1 Year	In 5 Years	Over 5 Years	In 1 Year	In 5 Years	Over 5 Years	
United States	0.74	0.52	0.65	54.7	28.3	50.4	195,498
United Kingdom	0.73	0.50	0.64	54.1	24.1	46.4	10,851
France	0.71	0.54	0.61	52.2	30.0	46.7	5,509
Germany	0.64	0.41	0.52	43.2	20.4	38.7	7,547
Australia	0.54	0.36	0.44	28.1	10.7	25.1	5,377
Japan	0.64	0.45	0.52	40.5	19.1	35.2	39,878
China	0.58	0.26	0.36	32.6	7.2	20.7	8,899
India	0.52	0.44	0.46	27.4	12.0	27.3	7,749

could well be a result of the mix of stocks in different countries covered by the database vendor.

### C. Fama–French Factor Exposures

We can also investigate exposures to the Fama–French factors with multivariate OLS time-series regressions, as follows:

$$(8) \quad r_{i,t} - r_{f,t} = a_i + b_i \times \text{XMKT}_t + s_i \times \text{SMB}_t + h_i \times \text{HML}_t + \epsilon_{i,t},$$

where SMB and HML are the two additional Fama–French factors.<sup>21</sup>

Table 16 shows that the auto-coefficient estimates and annual decay rates suggest that  $s$  and  $h$  are less stable than  $b$ . The best estimators to predict 1-year exposures are again prevailing VCK exposures, computed from daily stock returns and shrunk again by 30%–40% (albeit toward 0). To predict 5-year exposures,  $s$  should be shrunk by more than 50% and  $h$  by 40%. (This is subject to the further qualification that smaller firms tend to attrition more.)

It should be noted that with higher residual standard deviations in the predictive cross-sectional regressions (i.e., of approximately 1 instead of 0.5) and with less certain economic perspectives about the meaning of SMB and HML factors vis-à-vis an MKT factor, the use of historical  $s$  and  $h$  estimates in forward-looking capital budgeting applications of the FFM could suffer more from total expected return uncertainty than the CAPM model.

### D. Other Models

Corporate cost-of-capital models assume that alpha should be 0. However, in the context of investment funds, it is not uncommon to estimate alphas empirically, too. Harvey and Liu (2016) now suggest that alphas, like betas, are best shrunk aggressively. This is beyond the scope of our paper.

<sup>21</sup>In untabulated results, we find that in predicting future realized stock returns (not betas!), the HML spread performs well after Fama and French (1992) was published, but the complete FFM does not. Moreover, it is often difficult to convince executives that riskier growth projects should have lower rather than higher hurdle rates.

TABLE 16  
Fama–French Factor Exposures (*b*, *s*, and *h*) for U.S. Stocks

In Table 16, the dependent and independent variables are the estimated coefficients from multivariate Fama–French (1992) (rather than bivariate market-model) time-series regressions. The reported coefficients are the analogs of those in Table 1. The table does not report the included predictive cross-coefficients. In Panel B, the independent variables are always the prevailing equivalent 1-year VCK exposure estimates. The best predictor in each row is highlighted in boldface. *Data Set:* As in Table 1. *Interpretation:* Factor exposures on SMB and HML are more difficult to predict than market exposures. The same recipe works well: 1–3 years of daily stock returns, VCK shrunk, and then VCK shrunk again. Over a 1-year-ahead period, we recommend approximately 40% shrinkage. Over 5 years, we recommend approximately 60% shrinkage.

Panel A. Predicting Exposure in 1 Year

Dependent Variable	Metric	Independent Variable: BETA Calculated over Last <i>x</i> Years						SD( <i>y</i> )	<i>N</i>	
		1 Year		3 Years		5 Years				
		OLS	VCK	OLS	VCK	OLS	VCK			
SMB exposure ( <i>s</i> )	OLS	<i>R</i> <sup>2</sup>	35.0	<b>38.5</b>	29.4	29.9	25.6	26.0	1.0	199,783
		$\gamma_1$	0.59	0.72	0.63	0.69	0.68	0.73		
	VCK	<i>R</i> <sup>2</sup>	39.3	<b>44.8</b>	32.3	33.5	27.8	28.7		
		$\gamma_1$	0.53	0.66	0.56	0.62	0.60	0.65		
HML exposure ( <i>h</i> )	OLS	<i>R</i> <sup>2</sup>	38.0	<b>41.4</b>	35.9	37.0	34.3	34.7	1.2	199,783
		$\gamma_1$	0.61	0.74	0.65	0.70	0.65	0.68		
	VCK	<i>R</i> <sup>2</sup>	42.2	<b>47.8</b>	40.5	42.5	39.0	39.9		
		$\gamma_1$	0.56	0.68	0.59	0.65	0.60	0.63		

Panel B. Predicting Longer-Term VCK-Style Exposures for FFM Exposures

Coefficient	Auto-Coefficient $\gamma_1$			<i>R</i> <sup>2</sup>			<i>N</i>
	In 1 Year	In 5 Years	Over 5 Years	In 1 Year	In 5 Years	Over 5 Years	
MKT exposure ( <i>b</i> )	0.74	0.52	0.65	54.7	28.3	50.4	195,498
SMB exposure ( <i>s</i> )	0.64	0.33	0.44	38.9	9.6	23.8	195,498
HML exposure ( <i>h</i> )	0.72	0.41	0.57	48.5	16.8	32.4	195,498

## E. Alternative Capital Budgeting Models

There are also alternative cost-of-capital estimators. Our paper is about model estimates based on historical stock returns. Estimates based primarily on theoretical considerations (such as real options, as discussed by Da, Guo, and Jagannathan (2012)) could in principle perform better. In practice, such methods are not used in a manner that we can identify, however, perhaps due to their complexity and lack of consensus. Assessing their relative merits is beyond the scope of our paper.

Theory suggests that financial cost-of-capital assessments should determine corporate project discount rates. Unfortunately, our knowledge of firms' actual mapping from their financial cost-of-capital estimates to their discount rates remains tenuous. Although most firms claim they use the CAPM, Jagannathan, Matsa, Meier, and Tarhan (2016) find hints that this may be lip service, at least for some firms. They investigate 86 firms that agreed to participate in a survey in 2003 and that provided their discount rates (but not their own cost-of-capital estimates). Their sample is small, and replication remains impossible because of confidentiality agreements, but some evidence is better than theory alone. In their Table 2, they find that their calculated adjusted-CAPM betas explain approximately 40% of the variation in firms' self-reported discount rates (with an inferable equity-premium intercept of approximately 5%–6%). But firms also seem to use discount rates that are significantly higher than the academic estimates of

costs of capital, given their assumptions about equity premia and risk-free rates. Jagannathan et al. ((2016), Table 9) find that some of this bias can be explained by operational constraints and cash.

## VI. Equity-Premia and Duration-Premia Estimates

As difficult as it is to predict factor exposures, which are second moments, they are easier to assess than factor premia, which are first moments.

There is no definitive method to assess expected factor premia. The two most common methods are historical averages and dividend-yield implied averages. (Fama and French (2002) suggest that the equity premium suffers a “Peso problem,” as noted by Rogoff (1977) and Krasker (1980). The true market risk premium may well be lower than its ex post realization, e.g., as noted by Welch (2016).) Jacquier, Kane, and Marcus (2005) and others have pointed out that a historical geometric average, lower than the arithmetic average, is more suitable for assessing equity premia. There is no agreed-upon best estimation period. It is unknown whether the 30-year, 60-year, or 120-year historical experience is the most accurate guide for the future.

Rather than defend a specific perspective, our paper tries to help the reader assess his or her own equity-premium estimates. To do so, we construct reasonable time series of a longer duration. For stocks, we use both the Wilson and Jones (1987) series from 1870 to 1926 and the Schwert (1990) series from 1802 to 1925. For corporate bonds, we use a combination of FRED and Macauley National Bureau of Economic Research (NBER) data back to 1900 and spread-adjusted railroad bond data before 1900. For government bond and bill data, we use FRED data back to 1925, Jones–Wilson data back to 1870, and Officer (2014) data and Macauley’s New England Bond data before 1870. Of course, over longer time horizons, asset classes change as well. Our series are reasonable but not definitive. Note also that the U.S. Treasury short-term rate is not widely considered to be a risk-free rate of return before 1913.

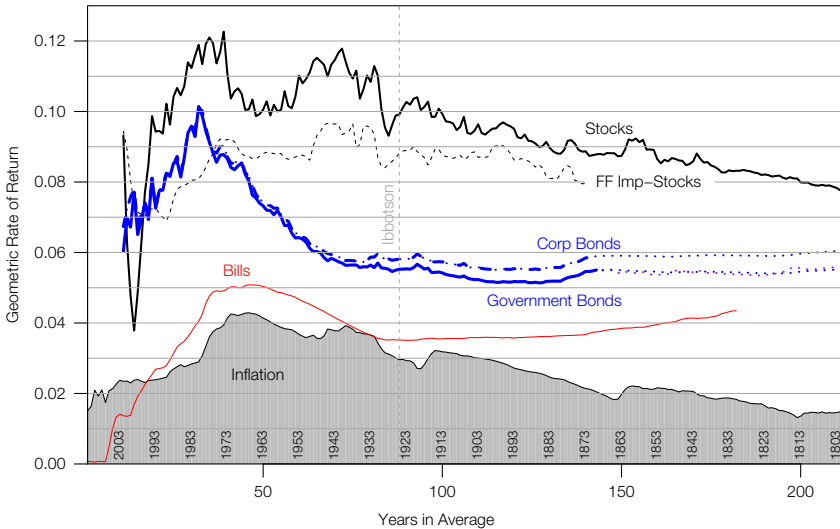
Figure 3 plots the geometric performance of the standing of our asset classes in early 2014 and looking back for up to 200 years. Any one individual annual rate of return that enters our geometric average rates of return becomes less reliable the further back in time one observes. However, in this perspective, longer-term observations also contain and thus are still averaged with subsequent returns, too. For example, the points at 1900 do not contain only 1-year unreliable rates of return but all 114 years since then.

Figure 3 shows that for a user in need of an equity premium above *short-term* bonds, unless the accumulation begins with the 2000 bear market, a 20-year perspective suggests an equity premium of 5% to 7%. Perspectives that are longer than the Ibbotson sample, which begins in 1926, suggest smaller equity premia. For a user in need of an equity premium above *long-term* bonds, perspectives shorter than 40 years (beginning around 1973) or longer than 150 years (beginning around 1830) suggest geometric equity premia of approximately 2%. A perspective of the approximately 100-year Ibbotson sample suggests a larger spread of approximately 4%.

FIGURE 3

## Backward-Looking Geometric Average Rates of Return

Figure 3 shows geometric averages beginning with the year on the x-axis and ending in 2013. The "FF Imp-Stocks" compound the prevailing Fama and French (2002) implied stock return estimates based on prevailing dividend yield and dividend growth. Government and corporate bonds are 10- to 20-year bonds. *Data Set:* The series is patched together from multiple data sources. The construction is explained in the text and in great detail on Ivo Welch's Web site (<http://ivo-welch.info>). *Interpretation:* With estimation windows reaching back either only to the 1970s or to the early 19th century, the historical geometric equity premium relative to long-term bonds is approximately 2%. However, reaching back to 1926 (i.e., as in the Ibbotson sample), the historical geometric equity premium reaches approximately 4%.



## VII. Conclusion

It is not controversial that users of factor models should adopt not a backward-looking but a forward-looking perspective. The historical experience suggests that longer-term projects require i) greater shrinking of factor exposures and ii) lower stock market equity premia (due to a higher risk-free-rate component). These are merely suggestions for better estimators of model-input parameters, not arguments about whether the models themselves hold or are empirically useful. Like Fama and French (1997), our paper should not be interpreted as endorsing the validity or use of either the CAPM or the FFM. We remain skeptics. In our paper, we even avoid tests of the models' powers to predict future expected rates of return.

The sophistication of beta estimators can reflect their use. In some contexts, even unshrunk beta estimates are good enough (e.g., when the goal is merely to hold market exposure constant, any linearly transformed control may do); in other contexts, sophistication that can tease out small improvements in beta prediction can be important (e.g., when the goal is to hedge out liquid market risk). In our corporate cost-of-capital context, however, it is principally the approximate magnitudes themselves that matter.

Now consider a manager who has decided to use the CAPM for cost-of-capital budgeting purposes and considers long-term buy-and-hold investments



into one of two traded projects/stocks. The first stock has an estimated daily 1-year OLS\_BETA of 0.0. The second has one of 1.5. These are roughly the means of the first and fifth quintiles in the distribution of individual-stock market betas. Sketching their cost-of-capital differentials suggests the following:

Equity Premium ( $R_m - R_f$ )		Beta Differential		Cost-of-Capital Differential
Estimation Period	Risk-Free Treasury	Beta Estimation	Second Shrinkage	$\Delta\text{BETA} \times (R_m - R_f)$
Ibbotson	1-Year	OLS_BETA	No	$1.5 \times 6.5\% \approx 10\%$
Ibbotson	1-Year	VCK_BETA	No	$1.35 \times 6.5\% \approx 9\%$
Ibbotson	1-Year	VCK_BETA	Yes	$1.0 \times 6.5\% \approx 7\%$
Ibbotson	10-Year	VCK_BETA	Yes	$1.0 \times 4.0\% \approx 4\%$
1973–2013	10-Year	VCK_BETA	Yes	$1.0 \times 2.0\% \approx 2\%$
1850–2013	10-Year	VCK_BETA	Yes	$1.0 \times 2.0\% \approx 2\%$

Managers may then need to reduce their spread estimates further in the event of any of the following issues:

- Market betas are estimated not from the projects themselves but only from noisily related proxy stocks (e.g., the rates of returns for the firm as a whole, peer stocks, or industry averages).
- Betas are calculated from monthly rather than daily stock returns.
- Corporate bonds have to pay a liquidity spread above Treasuries.
- There is uncertainty in parameter estimates (Armstrong, Banerjee, and Corona (2013)).

All of these concerns would apply even if the CAPM cost-of-capital linear equation were, in fact, the correct underlying true model predicting expected stock returns. This adds the important caveat: The spread could be even lower if there is some uncertainty about whether the CAPM (with its reliance on market beta) is the correct model.

Yet the empirical evidence is not favorable. Frazzini and Pedersen (2014) point out forcefully that betas have not even been reliably positively related to average return performances in 1-month-ahead predictions.<sup>22</sup> Baker, Hoeyer, and Wurgler (2016) describe an extensive literature that remains puzzled about the fact that market betas seem to have no marginal impact on expected rates of return. Figure 4 shows, in the simplest possible way, how badly even the most quintessential CAPM prediction fails: High-beta stocks have not even outperformed low-beta stocks over the last half century.

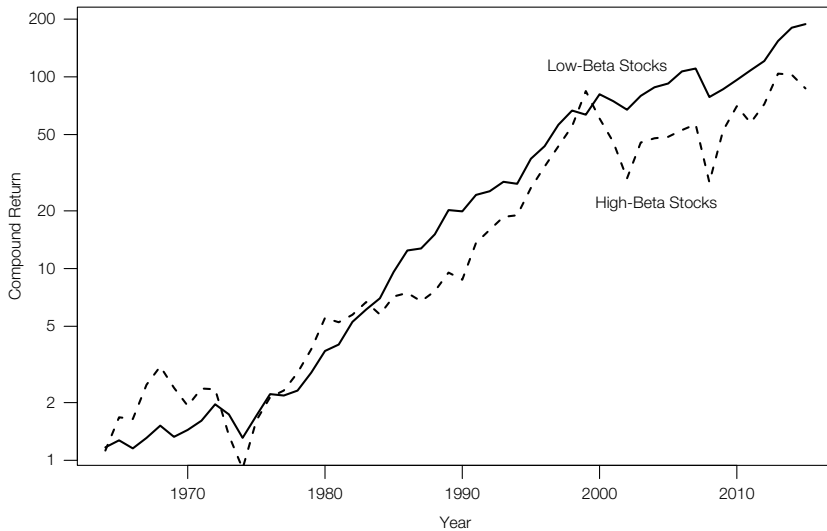
Finally, if the stock projects have leverage, the asset cost-of-capital differential between the two projects would be smaller than the equity cost of capital differential. To the extent that leverage is stable (e.g., biotech companies have consistently lower leverage than utilities), differently leveraged projects would have lower cost-of-capital differential spreads, although these spreads would then also create a more predictable time-series persistence in their asset-beta spreads.

<sup>22</sup> Anecdotal evidence suggests that a (not yet mentioned) other real-world objection to the use of the models is their vulnerability to “cost-of-capital gaming” among different constituents of the firm. This concern could suggest further spread reduction.

FIGURE 4

## Performance of High-Beta and Low-Beta Stocks

Figure 4 shows the compounded rates of return of portfolios formed based on the two extreme market-beta quintiles. Data Set: Kenneth French's Web site ([http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data\\_library.html](http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html)). Interpretation: Even the quintessential prediction of the CAPM, that high-beta stocks should outperform low-beta stocks, does not hold over the last 50 years.



We can speculate where overestimating spread differentials could have the most severe consequences. Firms that rely on a naïve textbook application of the CAPM or FFM would likely overestimate the heterogeneity of their projects. This perceived spread differential would be more pronounced among small firms and longer-term projects. The worst decision problems would arise for low-market-beta long-term projects that are financed primarily with equity, in which small changes in the *perceived* equity cost-of-capital rate can induce large changes in the *perceived* net present values. (For high-market-beta projects, cash flows in the distant future matter relatively less, so cost-of-capital misestimation would be relatively less problematic.) It also suggests that the weighted average cost of capital (WACC) fallacy (e.g., Krüger, Landier, and Thesmar (2015)) may not be as problematic as commonly taught.

We conclude that *even if factor models truly hold and corporate managers know which is correct*, they would have to greatly attenuate their estimates of the model-implied differential costs for *equity*, especially for long-term projects. In many cases, this attenuation reaches the point where they may as well ignore the beta estimates altogether. This does not mean that market betas are unimportant. They have other purposes, such as in variance reduction and market hedging. Betas would merely no longer be presumed to be positive discriminants for assessments of expected return hurdles.

One obstacle to abandoning current CAPM practice may be the lack of a clearly articulated alternative cost-of-capital recommendation. We suggest that one managerial alternative would be a *pragmatic, empirical cost-of-capital model*,

which would recognize that the expected rate of return on leveraged equity is nearly unpredictable, especially over longer horizons. Such a model would not assume equal costs of capital but continue to discriminate among projects based on two perfect-market criteria, i) project leverage and ii) project duration, and based on imperfect-market criteria, such as iii) limited access to capital markets, asymmetric information, transaction costs, and taxes (and possibly even behavioral mispricing). Risk in the economy would still matter but primarily in its manifestations through these premia.

This simpler pragmatic model would still leave it to the user to specify cost-of-capital differences between debt- and equity-financed projects, between short- and long-term projects, and between perfect- and imperfect-market aspects. Assessing these premia is not easy. Equity-premium estimates are controversial but indispensable. The duration premium is visible for debt financing, but it is not settled for equity financing. (Van Binsbergen, Brandt, and Koijen (2012) argue that the equity-premium term structure may be downward sloping, but Schulz (2016) argues that it is not.) Imperfect-market premia are so specific that modern academic finance has explored them only on their edges.

To its credit, the pragmatic model is still a simplification of currently recommended best practice. Any problems the pragmatic model faces are also faced by CAPM best practice and its common ad hoc adjustments. This is because application of the pragmatic model requires the same inputs that CAPM best-practice application requires. Consequently, the pragmatic model is largely identical, except for its one simplification: It suggests that it is reasonable to abandon attempts to discriminate among equity costs of capital (expected rates of return) based on estimated factor exposures. For assessments of common corporate equity costs of capital, given projects and leverage, it is good practice to assume that factor exposures are reasonably indistinguishable.<sup>23</sup>

## References

- Ait-Sahalia, Y.; I. Kalnina; and D. Xiu. "The Idiosyncratic Volatility Puzzle: A Reassessment at High Frequency." Working Paper, University of Chicago (2014).
- Armstrong, C. S.; S. Banerjee; and C. Corona. "Factor-Loading Uncertainty and Expected Returns." *Review of Financial Studies*, 26 (2013), 158–207.
- Baker, M.; M. F. Hoeyer; and J. Wurgler. "The Risk Anomaly Tradeoff of Leverage." Working Paper, Harvard Business School, University of Oxford, and New York University (2016).
- Barry, C. B. "Bayesian Betas and Deception: A Comment." *Journal of Financial Research*, 3 (1980), 85–90.
- Berk, J., and P. DeMarzo. *Corporate Finance*, 3rd ed. Boston, MA: Pearson (2014).
- Bhardwaj, G.; G. Gorton; and G. Rouwenhorst. "Facts and Fantasies about Commodity Futures Ten Years Later." Working Paper, NBER (2015).
- Blume, M. E. "On the Assessment of Risk." *Journal of Finance*, 26 (1971), 1–10.
- Brealey, R.; S. Myers; and F. Allen. *Principles of Corporate Finance*, 12th ed. New York, NY: McGraw-Hill/Irwin (2016).
- Campbell, J. Y.; A. Sunderam; and L. M. Viceira. "Inflation Bets or Deflation Hedges? The Changing Risks of Nominal Bonds." *Critical Finance Review*, 6 (2017).
- Da, Z.; R.-J. Guo; and R. Jagannathan. "CAPM for Estimating the Cost of Equity Capital: Interpreting the Empirical Evidence." *Journal of Financial Economics*, 103 (2012), 204–220.

<sup>23</sup>Baker et al. (2016) push this even further, suggesting that firms should change their capital structures (and possibly projects) when increases in leverage cause commensurate increases in beta but not in expected rates of return.

- Dimson, E. "Risk Measurement When Shares Are Subject to Infrequent Trading." *Journal of Financial Economics*, 7 (1979), 197–226.
- Elton, E. J.; M. J. Gruber; and T. J. Urich. "Are Betas Best?" *Journal of Finance*, 33 (1978), 1375–1384.
- Fama, E. F., and K. R. French. "The Cross-Section of Expected Stock Returns." *Journal of Finance*, 47 (1992), 427–465.
- Fama, E. F., and K. R. French. "Industry Costs of Equity." *Journal of Financial Economics*, 43 (1997), 153–193.
- Fama, E. F., and K. R. French. "The Equity Premium." *Journal of Finance*, 57 (2002), 637–659.
- Fama, E. F., and J. D. MacBeth. "Risk, Return, and Equilibrium: Empirical Tests." *Journal of Political Economy*, 81 (1973), 607–636.
- Fisher, L., and J. H. Kamin. "Forecasting Systematic Risk: Estimates of 'Raw' Beta That Take Account of the Tendency of Beta to Change and the Heteroskedasticity of Residual Returns." *Journal of Financial and Quantitative Analysis*, 20 (1985), 127–149.
- Frazzini, A., and L. H. Pedersen. "Betting against Beta." *Journal of Financial Economics*, 111 (2014), 1–25.
- Graham, J. R., and C. R. Harvey. "The Theory and Practice of Corporate Finance: Evidence from the Field." *Journal of Financial Economics*, 60 (2001), 187–243.
- Greene, W. H. *Econometric Analysis*, 6th ed. Boston, MA: Pearson (2008).
- Harvey, C. R., and Y. Liu. "Rethinking Performance Evaluation." Working Paper, Duke University and Texas A&M University (2016).
- Jacobs, M. T., and A. Shivdasani. "Do You Know Your Cost of Capital?" *Harvard Business Review*, 90 (2012), 118–124.
- Jacquier, E.; A. Kane; and A. J. Marcus. "Optimal Estimation of the Risk Premium for the Long Run and Asset Allocation: A Case of Compounded Estimation Risk." *Journal of Financial Econometrics*, 3 (2005), 37–55.
- Jagannathan, R.; D. A. Matsa; I. Meier; and V. Tarhan. "Why Do Firms Use High Discount Rates?" *Journal of Financial Economics*, 120 (2016), 445–463.
- Jagannathan, R., and Z. Wang. "The Conditional CAPM and the Cross-Section of Expected Returns." *Journal of Finance*, 51 (1996), 3–53.
- Johnson, D. J.; R. E. Bennett; and R. J. Curcio. "A Note on the Deceptive Nature of Bayesian Forecasted Betas." *Journal of Financial Research*, 2 (1979), 65–69.
- Jostova, G., and A. Philipov. "Bayesian Analysis of Stochastic Betas." *Journal of Financial and Quantitative Analysis*, 40 (2005), 747–778.
- Karolyi, G. A. "Predicting Risk: Some New Generalizations." *Management Science*, 38 (1992), 57–74.
- Kothari, S.; J. Shanken; and R. Sloan. "Another Look at the Cross-Section of Expected Returns." *Journal of Finance*, 50 (1995), 185–224.
- Krasker, W. S. "The Peso Problem in Testing the Efficiency of Forward Exchange Markets." *Journal of Monetary Economics*, 6 (1980), 269–276.
- Krüger, P.; A. Landier; and D. Thesmar. "The WACC Fallacy: The Real Effects of Using a Unique Discount Rate." *Journal of Finance*, 70 (2015), 1253–1285.
- Levy, M., and R. Roll. "The Market Portfolio May Be Mean/Variance Efficient After All." *Review of Financial Studies*, 23 (2010), 2464–2491.
- Lewellen, J. "The Cross Section of Expected Stock Returns." *Critical Finance Review*, 4 (2015), 1–44.
- Mamaysky, H.; M. Spiegel; and H. Zhang. "Estimating the Dynamics of Mutual Fund Alphas and Betas." *Review of Financial Studies*, 21 (2007), 233–264.
- Mehra, R., and E. C. Prescott. "The Equity Premium: A Puzzle." *Journal of Monetary Economics*, 15 (1985), 145–161.
- Officer, L. H. "What Was the Interest Rate Then? A Data Study." Available at <http://www.measuringworth.com/interestrates/> (2014).
- Rogoff, K. "Rational Expectations in the Foreign Exchange Market Revisited." Working Paper, Massachusetts Institute of Technology (1977).
- Ross, S.; R. Westerfield; J. Jaffe; and B. Jordan. *Corporate Finance*, 11th ed. New York, NY: McGraw-Hill/Irwin (2016).
- Schulz, F. "On the Timing and Pricing of Dividends: Comment." *American Economic Review*, 106 (2016), 3185–3223.
- Schwert, G. W. "Stock Returns and Real Activity: A Century of Evidence." *Journal of Finance*, 45 (1990), 1237–1257.
- Van Binsbergen, J.; M. Brandt; and R. Koijen. "On the Timing and Pricing of Dividends." *American Economic Review*, 102 (2012), 1596–1618.

- Vasicek, O. A. "A Note on Using Cross-Sectional Information in Bayesian Estimation of Security Betas." *Journal of Finance*, 28 (1973), 1233–1239.
- Welch, I. *Corporate Finance*. 3rd ed, Los Angeles, CA: Ivo Welch (2014).
- Welch, I. "The (Time-Varying) Importance of Disaster Risk." *Financial Analysts Journal*, 72 (2016), 14–30.
- Wilson, J. W., and C. P. Jones. "A Comparison of Annual Common Stock Returns: 1871–1925 with 1926–85." *Journal of Business*, 60 (1987), 239–258.