NBER WORKING PAPER SERIES

SIMPLER BETTER MARKET BETAS

Ivo Welch

Working Paper 26105
http://www.nber.org/papers/w26105

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2019

The author has disclosed a financial relationship of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w26105.ack

Simpler Better Market Betas
Ivo Welch
NBER Working Paper No. 26105
July 2019
JEL No. G10,G11,G31

ABSTRACT

This paper proposes a robust one-pass estimator that is easy to code: Justified by the market-model itself and using a prior that market-betas should not be less than –2 and more than +4, the market-model is run on daily stock rates of return that have first been winsorized at –2 and +4 times the contemporaneous market rate of return. The resulting "slope-winsorized" estimates outperform (all) other known estimators in predicting the future OLS market-beta (on R^2 metrics). Adding reasonable age decay, suggesting a half-life of about 3 to 5 months, to observations entering the market-model further improves it. The estimates outpredict the Vasicek estimates by about half as much as the Vasicek estimates outpredict the OLS estimates.

Ivo Welch
Anderson School at UCLA (C519)
110 Westwood Place (951481)
Los Angeles, CA 90095-1482
and NBER
ivo.welch@anderson.ucla.edu

The most important moments in finance are expected rates of return, own volatilities, and factor exposures.

Most finance research is dedicated to estimating expected returns. They are both the most interesting and the most difficult moment to forecast. Lo and MacKinlay (1990), Harvey, Liu, and Zhu (2016), and others, have pointed out how historical data correlations have often been spurious. Mclean and Pontiff (2016) have pointed out how investors research can itself erode past significant associations.

Volatilities are important not only in the context of pricing financial derivatives but also because they seem to have an empirical associations with future rates of return. Volatilities are typically easier to estimate than expected returns because second moments improve with higher-frequency sampling. Engle (1982), Glosten, Jagannathan, and Runkle (1993), and others, have even shown how to predict time-varying volatility.

My paper assumes that factor exposures in general and market-betas in particular are the third intrinsically interesting moment worth estimating.

Interest in the estimation of market-beta began in earnest with portfolio theory in the 1950s. Even today, market-beta plays a special role.[1] The market portfolio is typically the first principal component and many investors start with market-like portfolios. Market beta is the primary measure of how individual assets contribute risk to the market portfolio. It informs investors how tilting assets in and out of market portfolios changes their risk exposures.

Market-betas can not only help investors assess the diversification contribution but can also influence the risk-return tradeoff through an associations with expected returns. The CAPM suggests a positive association, but the evidence against the CAPM is strong. Frazzini and Pedersen (2014) have claimed a negative empirical association, but there is now some disagreement.[2] Of course, even if there is zero association, risk-averse investors can still use market-beta estimates to tilt their portfolios towards low market-beta stocks

[1]The market excess rate of return is one among a number of factors of interest nowadays. Measuring the exposures of assets to other factor portfolios is an analogous general problem.

[2]The negative association has recently been disputed by Novy-Marx and Velikov (2018) and Han (2019), who show that the success of "betting-against-beta" was at least in part due to a mixing of time-varying volatility into the Frazzini-Pedersen estimated betas with accompanying market-timing. Section VI will show that the Frazzini-Pedersen market-beta estimator is a poor measure of the future OLS market-beta and of its future self.

and thereby increase their expected returns with lower risk. Finally, there is yet another category of market-beta users: managers and regulators (often required by law) may want reasonable estimates of market-beta when they seek to implement the CAPM—regardless of whether the model holds or not.

My paper does not explore the association of market-betas with average rates of return; much less does it take a stance on the CAPM. Instead, it explores only the measurement of market-betas and only in publicly-traded U.S. stocks. For validation purposes, it predicts future (OLS and other) market-betas.

Its recipe for a new type of robust estimator for market-beta, with emphasis on ease of implementation, is as follows:

1. Winsorize each rate of return $r_{i,d}$ to $[(1 \pm \Delta_s) \cdot r_{m,d}]$, where r is a daily (net-of-risk-free) rate of return, i is a stock index, m is the stock market index, d is a day index, and Δ_s is a parameter. A good Δ_s is 3, suggesting bounds of $-2 \cdot r_{m,d}$ and $+4 \cdot r_{m,d}$.
2. Run an OLS market-model regression on these winsorized daily rates of return. I refer to this estimator as bsw (for "beta slope-winsorized"). A better version estimates a WLS market-model, in which the weight of each observation decays with age. A good decay parameter is ρ=2/252 per day, suggesting a half-life of about 90 trading days. I refer to this estimator as bswa (for "beta slope-winsorized aged").

The suggested parameters are low-dimensional, static, and not fragile. The bswa estimator does not need need two parameters per stock, per time unit, or per stock-time unit, but literally only two static fixed numbers for the entire CRSP sample, modestly searched from coarse integer choices only. Winsorization levels of Δ_s=2 or Δ_s=4, and half-lives from 75 to 120 days give nearly identical performance. The paper also shows that potential gains from further fine-tuning the two parameters are likely to be small.

Our estimators share with Bayesian counterparts the need to specify a prior. Here, it is that *stocks have market betas between* –2 *or* +4. Although our use of the prior is non-Bayesian, it is also not arbitrary. Its functional form is based on the linear specification of the market model itself. Most importantly, bswa requires minimal implementation effort (no memory—stock returns are treated as they enter the market-model regression) and produces superior empirical results.

In essence, bsw is just another robust estimator in the toolkit (Tukey (1960)). It applies its prior not to the market-beta but to each and every stock return. In a Bayesian analysis with a bounded beta prior, only market-betas beyond the bounds would be considered implausible. Here, each rate of return beyond these bounds (that could cause market-beta to exceed these bounds) is considered not necessarily implausible but uninformative for estimating beta.[3]

Consider two examples. First, if the excess market rate of return is zero, all stock returns are winsorized to zero. Yet, this is not because stock returns are viewed as excessive outliers. They are just useless—they do not influence the parameter of interest, market-beta. Slope winsorization takes into account not only the implausibility of outlier but also the resulting effect on the parameter of interest, market-beta. Second, if a firm has a single stock return outlier of +300% that occurs on a day on which the stock market happened to return +1.0%, in a Bayesian analysis, this single outlier would likely fix the posterior to its upper beta limit. (If the prior was not bounded but leptokurtic, the estimated market-beta could be quite large, regardless of other return observations.) In contrast, bsw works by "defusing" this single outlier stock return. It moderates its influence. Other stock returns remain more influential.

In a different sense, slope winsorization uses the prior in a way that is more akin to how one would use a functional form for a prior in a classical specification equation. For example, a classical prior that a variable has zero influence would suggest omitting the variable from the estimation; a classical prior that the beta is exactly 1.0 would suggest predicting $r_{i,t} - r_{m,t}$ as the dependent variable. If there are positive and negative market rates of return, the winsorization does force the estimated market-betas to obey the specified bounds.

Despite their ease of use, bsw and bswa perform not just as well as but better than other prominent market-beta estimators, when assessed by an ability to predict future realized

[3]The prior is on the prediction-relevant variable, market-beta, not on the "nuisance" parameter σ. (Indeed, the winsorizing estimators "destroy" the informativeness of the mean squared residuals.) Specifying market-beta bounds may be easier and more reliable. For example, many financial economists share the intuition that market-betas beyond 0 and +3 seem implausible, but have little intuition of whether (market-adjusted) rates of return beyond 5%, 10%, or 20% per day are unusual. Section III will suggest plausible bounds on returns. Section III discusses a band winsorizer that performs only modestly worse than the slope winsorizer. Section VI discusses Martin and Simin (2003), which works with winsorizations based on stock-specific assessed return variances.

OLS (or other) market-betas. The competitors include the market-beta estimators in Vasicek (1973), Dimson (1979), Martin and Simin (2003), Frazzini and Pedersen (2014), and Ait-Sahalia, Kalnina, and Xiu (2014), as well as monthly market betas, industry betas, peer betas, and many combinations of the aforementioned estimators. The slope-winsorized betas outperform them in nearly all data partitions—new firms and old firms; big firms and small firms; high-trading stocks and low-trading stocks; high-volatility stocks and low-volatility firms; high OLS-beta stocks and low OLS-beta stocks, and so on. Despite trying, I am unaware of other methods (variants, variables, or estimators) that can further improve on bswa in a meaningful way.[4]

Although "big data" estimators (such as neural nets) based primarily on empirical performance rather than theory have recently become popular in economics and finance, bsw and bswa achieve their superior performances with simpler rather than more complex methods. They are easier to code and their reasoning and effects are easily understood. Outlier handling through slope priors rather than band priors is just another robust way to winsorize inputs. Future research may examine whether slope-based winsorization has use in other application contexts.

[Insert Table 1 here: **Glossary of Market-Beta Estimators**]

Section I now describes the sample, the data, and OLS market betas. Section II discusses the Vasicek market beta and its variants, later confirmed to be the best estimators in the literature to date and thus used for our first comparisons. Section III explains level, band, and slope winsorizers in the context of market-beta. Section IV predicts the future OLS market beta using only winsorization (bsw). Section V adds age decay (bswa). It also describes bswa's performance in separate data partitions and attempts to improve bswa further using more information (mostly unsuccessfully). Section VI shows how bsw and bswa outperform other beta estimators predicting the OLS market-beta and their own future selves. And Section VII concludes with use recommendations and a 10(!)-line code snippet that implements bsw and bswa in R. A glossary of considered market-betas is in Table 1.

[4] Some variables can be used to improve bswa but only by very little and with much specification and fine-tuning search. The best candidate is trading volume.

I Data and OLS Market Betas

Our primary concern is measuring the true underlying market-beta of every stock. Because it is unobservable, we have to estimate an OLS market-beta, typically in a block sampler using one year[5] with *daily* stock returns (Foster and Nelson (1996) and Ghysels and Jacquier (2007)). All returns are quoted in excess over the risk-free rate from Ken French's website.

The paper works with various subsets of the data. The first analysis uses a mid-year to mid-year data set. Later analysis uses 12-month betas calculated each and every month. In this case, there are overlapping observations in the predictive "gamma regressions" (i.e., the regressions in which we forecast future market-betas with current market-betas). However, there are never observations in which the returns used to calculate the dependent predicted future market-beta have also been used to calculate the independent predicting current market-beta.[6]

We start with the CRSP data set from 1925/12/31 to 2018/12/31 with 94,402,370 valid stock returns. The key data selection constraints are that stocks have to have share code 10 or 11 (75,988,595 returns), sometimes that data starts in 1963/07 (66,814,795) or 1974/01 (61,599,045). To calculate market-beta, we require that a stock has 6 or 20 trading days of data in a year, and (often but not always) the availability of a future one-year OLS market-beta. The maximum number of valid overlapping one-firm-year bols estimates is 3,255,424, but only 2,962,015 firm-years have two consecutive bols's (necessary for validation testing). Using only non-overlapping July to June intervals from 1926 on, there are 271,861 one-year and 247,426 consecutive observations. Starting in July 1964, there are 239,923 and 216,126 equivalent observations. None of the data constraints seem to make any difference in the *relative* performance ordering of estimators.[7]

Because our analysis is equal-weighted and our stock market index is (always) the value-weighted CRSP market, the average market-beta across stocks in our sample is not

[5] A modestly better estimation window than 12 months would be 16-20 months. However, 12-month windows have a natural calendar relation. Moreover, in Section V, our independent variable uses all historical data decayed by age.

[6] For example, we may run a predictive regression, in which we have one observation predicting a Jan-2015 to Dec-2015 beta with a Jan-2014 to Dec-2014 beta, followed by another predicting a Feb-2015 to Jan-2016 beta with a Feb-2014 to Jan-2015 beta. TThere is one exception late in the paper, when bswa has overlap when it predicts the future bswa. This will be noted.

[7] This is also the case for selection biases. Although it is possible that there is systematic survivorship, Experimentation with different selection criteria suggested no *relative* effects.

1.0 (as it would be by definition if we used an equal-weighted market or a value-weighted regression). Instead, the average market-beta was about 0.90, 0.82, and 0.77 in samples beginning in 1926, 1963, and 1974, respectively. The average market-beta varied (slowly) with the composition of stocks on CRSP over the years. From 1939 to 1945, it reached levels of about 1.4.

Whenever we benchmark market-beta estimators against one another, we always hold the set of observations (stock-years) strictly the same.

A The Sample

[Insert Figure 1 here: **Quasi-CDF: Frequency of Winsorizations**]

Figure 1 plots the empirically observed pooled cumulative frequencies of returns and OLS market-betas over the entire sample.

Raw Excess Rates of Returns: The top plot shows that about 6.7% of the sample had *absolute* rates of return (always net of the risk-free rate) above 7%; 3.2% above 10%; and 0.6% above 20%. Not shown, the average daily rate of return was 6.1 basis points, with a standard deviation of 4.4%.

Market-Adjusted Excess Rates of Return: The middle plot shows that about 22% of the sample had *absolute* market-adjusted rates of return above 3%; 3.1% above 10%; and 0.6% above 20%. Not shown, the equal-weighted average daily rate of return net of the contemporaneous value-weighted market rate of return was 3.6 basis points, with a standard deviation of 4.3%.

OLS Market Betas: The bottom plot shows absolute distances from 1.0 for one-year OLS market-betas. In this panel containing both firms and years, the observed high kurtosis is more akin to the kurtosis of a log-normal distribution. However, the skewness is more akin to a plain normal distribution. Not shown, there is only modest skewness and kurtosis for any given firm. (The average firm's kurtosis is 2.4.) Instead, most of the kurtosis comes from the many different firms in the cross-section in any given year. (The average year's kurtosis is about 11.0.) The figure shows that few firm-years have betas more than Δ_S=2 units away from 1.0 (i.e., –1.0 and +3.0). Under an equivalent mean-variance matched normal distribution, we would expect about 0.5% of the sample. However, the actual number in the data was almost twice as high at 0.9%. At Δ_s=3 units distance from 1.0, there are almost no betas. Nevertheless, the actual number was 0.170% observed (two per thousand), compared to a matched Gaussian normal frequency of 0.002%. There are market-beta outliers!

B Current and Future OLS Market-Betas

Examining one-year market-betas, 0.6% are less than –1.0, 0.06% are less than –2.0, 0.4% are greater than +3.0, and 0.03% are greater than +4.0. Conditional on a market-beta less than –2.0, only 4.8% of next year's market beta remain less than –2.0 (30% are positive and and 2.0% are greater than +4.0!). Conditional on a market-beta greater than +4.0, only 1.1% of next year's market beta remain greater than +4.0 (35% are negative and and 3.4%(!) are less than –2). Market-betas beyond what is implied by $\Delta_s = 3$ are simply not very predictive of their future selves.

[Insert Figure 2 here: **Percentiles of Future 1-Year OLS Market-Betas (bols) By Current 1-Year OLS Market-Beta (bols)**]

In Figure 2, current year's OLS beta estimates, $bols_{i,y}$, are first binned into 100 percentiles in each year (not in the full sample!). The figure then plots means of the OLS market-betas in the subsequent year, $bols_{i,y+1}$, as well as the bins' one and two standard deviation ranges. It also plots $bols_{i,y}$, i.e., the sort criterion for the X-axis. The shallower slope of the $bols_{i,y+1}$ line compared to the $bols_{i,y}$ line reflects the mean reversion of the OLS beta estimates, due both to estimation error and mean reversion of the underlying true market-beta.

Even for the minimum current market-beta percentile (with its market-beta mean below –1), the subsequent average market-beta mean does not drop below 0.3, with a two standard deviation range reaching from about -1.0 to about +2.0. The highest average future market-beta is about +1.7, with a two standard deviation range reaching from about +0.0 to about +3.5.

The monotonic relationship between the current and future OLS market-beta breaks down below a current bols of just about 0.0. If, anything, more negative betas begin to forecast higher future betas. Similarly but less stark, the current beta becomes just about uncorrelated with the future beta for betas above 3.0. However, because the data is a panel and this plot shows one grand pooled association, it does not mean that there are no stocks which have persistent predictive market-betas higher or lower than –1 or greater than +3.[8]

Not shown, an equivalent plot for stock volatility shows a relationship between current and future volatility that is always monotonic. This relationship implies that firms with stock return outliers are likely to repeat outliers. It seems that high-volatility stocks have regular and persistent outlier events, but they do not systematically occur either on market up-days or down-days.

C Summary

OLS beta estimates below 0 and above 3 are rare and seem not to be predictive of future beta estimates. Reasonable a-priori bounds on what constitutes non-indicative market-beta value should allow for truncation that is no less than plus-and-minus Δ=1.5 from 1.0 (i.e., market-betas from –0.5 to +2.5). Market-betas beyond –2 and +4 are exceedingly rare aberrations, with consecutive-year incidences just slightly above zero-correlation levels.

[8]The lack of the monotonic relation towards both edges is even more apparent in Appendix Figure 12, where we plot the slope of the prediction line. It is negative for negative current betas, and just about zero for betas above 3.0. This figure also shows that although the details change modestly, the main inference does not. Below 0 and above 3, there is no meaningful positive monotonic relationship between current and future market-betas.

II The Vasicek Market-Beta Estimator

The evidence below will suggest that the best estimators in the literature hitherto were variants of the Vasicek (1973) estimator, bVCK, when used on daily data and ideally linearly debiased as in Levi and Welch (2017), bLW. Thus, we use them as exploratory benchmarks. Other prominent estimators will be explained and compared in Section VI.

Neither bVCK or bLW are justifiable by decision theory that reflects empirical reality. Most importantly, neither can account for the facts that the underlying market-betas are mean-reverting and that there are outliers. Moreover, there is no a-priori known process for how (underlying true) market-betas mean-revert, nor do we know the distribution of rare outliers (itself very difficult to estimate because of their rarity). Instead, these estimators are intuitive methods to adjust for estimation-caused mean reversion and outliers (in the case of bVCK), and true-beta mean reversion (in the case of bLW).

A The Unadjusted Vasicek Estimator

The Vasicek (1973) estimator, $\text{bVCK}_{i,y}$, can be viewed either as a Bayesian shrinkage estimator or as the random-effects panel estimator. It requires first computing the OLS market-beta $\text{bols}_{i,y}$ and standard error $\sigma^2_{i,y}$ for each stock within the desired unit of time ("σ^2_i"). Then it requires calculating cross-sectional statistics over all stocks, $\overline{\text{bols}}_t$ for the cross-sectional mean and "σ^2_t" for the cross-sectional heterogeneity. (Because the composition of stocks and with it the average market-beta changes every year, even this cross-sectionally constant part contributes to the good performance of bVCK.) For each stock i at time t, the Vasicek estimate is then

$$\text{bVCK}_i \equiv \left[\frac{\sigma^2_t}{\sigma^2_i + \sigma^2_t}\right] \cdot \text{bols}_i + \left[1 - \frac{\sigma^2_t}{\sigma^2_i + \sigma^2_t}\right] \cdot \overline{\text{bols}}_t \ . \tag{1}$$

On average, the current Vasicek market-beta seems to predict the future OLS market-beta better than any other known estimators. Nevertheless, Vasicek estimators have seen lower rates of adoption than inferior alternatives. The neglect may be for several reasons. First, Vasicek estimators unintuitively entangle beta estimates with those of unrelated stocks. Including or excluding unrelated stocks in the sample changes every other beta estimate.

Second, there is extra effort involved. Even the simplest versions require running a first-step OLS regression, calculating cross-sectional statistics, and then going back to readjusting the OLS estimates based on (cross-sectional and time-series) standard errors. The Vasicek-derived estimators in Karolyi (1992) and Levi and Welch (2017) require even more steps. Third, Vasicek betas have never become *the* standard. They seem to be rarely requested by referees as best practice. Fourth, in the absence of recent public performance benchmarks, their superior performance may not have been fully appreciated. Fifth, researchers often want to investigate another issue and not add effort by first investigating how to estimate betas. Their interest in market-beta may be perfunctory. Many papers investigate different issues and are content merely to add some (any) control for market-beta. Worse control may not necessarily be harmful. Sixth, diversified portfolio betas (rather than stock betas) are more forgiving.

B The Adjusted Vasicek (Levi-Welch) Estimator

Blume (1971) first suggested debiasing the OLS estimator using an empirically estimated linear correction. The Levi and Welch (2017) estimator, bLW, adopts this idea to debias bVCK. The first step is estimating a relation between lagged $\text{bVCK}_{i,y}$ and future $\text{bols}_{i,y+1}$ estimates, $\text{bols}_{i,y+1} = \gamma_0 + \gamma_1 \cdot \text{bVCK}_{i,y}$ estimates. The resulting linear empirical debiasing coefficients are quite stable. By decade, the predictive gamma coefficient estimates were

	1960s	1970s	1980s	1990s	2000s	2010s	Yr-Avg	Pooled
γ_0 (Intercept)	0.37	0.19	0.16	0.21	0.22	0.31	0.21	0.19
γ_1 (on $\text{bVCK}_{i,y}$)	0.71	0.73	0.67	0.65	0.77	0.69	0.73	0.74

Thus, a reasonable Levi-Welch estimator in this sample period was

$$\text{bLW}_{i,y} \equiv 0.20 + 0.75 \cdot \text{bVCK}_{i,y} \ . \tag{2}$$

The use of 0.20 and 0.75 as the two adjustment parameters does introduce a modest hindsight bias. However, within 5 years of the start of the sample, the biggest in-time-averaged distance (starting from 1960) from 0.75 was 0.04.[9]

Because bLW is merely a linear transformation of bVCK, bLW has the same predictive R^2 as bVCK in in-sample regressions predicting any other variable, such as any future market-beta estimates. However, bLW has lower RMSE than bVCK when used as a direct proxy of market-beta. A reader who considers the 0.20/0.75 parameters to be excessively data snooped can ignore bLW and focus solely on bVCK. bLW is included only for benchmarking purposes.

C Assessment Preview

My paper will show that bVCK and bLW have predicted so well for a reason hitherto not widely understood. Their effectiveness was not so much due to shrinkage of normally distributed panel observations but due to aggressive shrinkage of outliers. The bsw and bswa estimators can outperform them only because outlier reduction explains most of their performance. bLW often performs almost as well as bsw, albeit with much more coding effort. bswa outperforms it easily.

[9] The required 1–0.75 = 0.25 shrinkage is not the typical errors-in-variable problem (because the predictive Vasicek beta remains measured with error) under a normal distribution—the required shrinkage is far too much. Over-the-envelope calculations suggest that with a standard error of about 0.2 (based on one year of daily stock returns) and given the cross-sectional dispersion of betas of about 0.7, a one-year Vasicek beta should predict a future one-year OLS beta with an errors-in-variables bias of under 5%.

III Estimators Based on Winsorizing Stock Returns

All winsorization schemes must rely on an “aggressiveness” parameter, called Δ, which trades off type-I and type-II errors: correctly winsorizing unpredictive outliers vs. incorrectly winsorizing predictive outliers. Good delta choices trade off being too lax (thereby not having any effect) and being too strict (thereby pushing all beta estimates too close towards the same value). Figure 4 in Section IV will show that higher deltas become progressively more useless and lower deltas are progressively too restrictive to reflect non-noise realizations of market rates of return.

Like the shrinking estimators in Section II, the winsorizing estimators are also not based on decision theory with realistic features. The theoretical properties of all estimators in realistic samples are therefore largely unknown. However, we are never interested in the in-sample properties of our estimators. We are interested only in their out-of-sample forecasting power.

A Level Winsorization

Common “level” winsorization limits firms’ rates of return to absolute levels

$$\mathrm{rlw}_{i,d} \in \quad [\, -\Delta_l, +\Delta_l \,] \,, \tag{3}$$

which is (obvious) shorthand notation to winsorizing stock rates of return from $-\Delta_l$ to $+\Delta_l$. The delta here is a single choice parameter. It is static rather than indexed by firm or time. The “lw” following “r” abbreviates “level winsorization.”

As always, stock returns are daily and always quoted above the risk-free rate. The intercept can be zero because firm-heterogeneity in daily expected excess returns is negligibly small. The level winsorized market-beta is from a market-model regression, hitherto analogously denoted “blw”

$$\mathrm{blw}_{i,y} \equiv \frac{\mathrm{cov}\big[\ \mathrm{rlw}_{i,d}(\Delta_l)\ ,\ r_{m,d}\ \big]}{\mathrm{var}(\ r_{m,d}\ )} \ . \tag{4}$$

The typical estimation time period in this paper is one year (y), for any month from month-end to month-end.

[Insert Figure 3 here: **Winsorization Techniques**]

The top plot in Figure 3 illustrates level winsorization. By compressing the range of the dependent variable in the market-model regression, the beta estimate becomes biased towards zero. This bias is especially undesirable when reasonable hypotheses would center the prior not on zero.

It turns out that the "level-winsorized" blw is inferior, because there are days on which the overall stock market itself had exceptionally positive or negative rates of return. Level winsorization then incorrectly cuts off too many informative large positive or negative individual rates of return.

B Band Winsorization

Model-specific winsorization schemes can do better. Consider an example in which stocks follow a fat-tailed return distributions with large outliers. Researchers should not want to reduce such outliers when estimating volatility. However, when estimating market-beta, researchers would want to recognize that such outliers may randomly and unrelatedly occur on days when the rest of the stock market happens to go up or down. Thus, one may want to reduce outliers differently when estimating volatilities than when estimating market-betas.[10]

One model-specific estimator is a "band winsorizer" (bbw). It winsorizes stocks differently on days on which the market has moved a lot. Specifically, it limits firms' rates of return to a constant plus or minus the market rate of return (still always quoted above the risk-free rate):

$$\mathrm{rbw}_{i,d} \in \; \mathrm{r}_{m,d} + \; [\, -\Delta_b, +\Delta_b \,] \; . \tag{5}$$

[10]In the CRSP sample, large return outliers should not be winsorized the same when estimating stock volatility as when estimating market-beta. The R^2 of predicting next year's standard deviation with this year's standard deviation is 46%. With level-winsorized rates of return, this drops to 44%, with band-winsorized returns to 21%, and with slope-winsorized returns to 6%.

As in Martin and Simin (2003), the model is used to winsorize the largest standardized residuals. Unlike in Martin and Simin (2003), there is no first-stage OLS regression to estimate Δ_b.

The band winsorized market-beta is

$$\text{bbw}_{i,y} \equiv \frac{\text{cov}\left[\ \text{rbw}_{i,d}(\Delta_b)\ ,\ r_{m,d}\ \right]}{\text{var}(\ r_{m,d}\ )}\ , \tag{6}$$

The middle plot in Figure 3 illustrates "band" winsorization. By compressing the range of the dependent variable diagonally, each beta estimate becomes biased towards 1.

Band winsorization has to lean on specification priors about net-of-market rates of return. It is most aggressive with positive outliers when the market return is negative and with negative outliers when the market return is positive.

C Slope Winsorization

My paper focuses on a "slope-winsorized" market-beta estimator (bsw). It performs a little better than band winsorization and makes it possible to work with priors on market-beta itself. "Slope" winsorization limits firms' market-model slopes:

$$\text{rsw}_{i,d} \in \left(1.0 + [\,-\Delta_s, +\Delta_s\,]\right) \cdot r_{m,d}\ . \tag{7}$$

This linear structure in the winsorization is provided by the model itself. The slope winsorized market-beta is

$$\text{bsw}_{i,y} \equiv \frac{\text{cov}\left[\ \text{rsw}_{i,d}(\Delta_s)\ ,\ r_{m,d}\ \right]}{\text{var}(\ r_{m,d}\ )}\ . \tag{8}$$

The bottom plot in Figure 3 illustrates "slope" winsorization. Beta estimates become biased towards 1. Slope winsorization is most aggressive where it matters least for regression coefficient estimates—around market returns of about zero. It is useless to report the fraction of returns that are winsorized—there are always many returns that are near zero that are winsorized. However, these winsorized points matter little in the slope estimation.[11]

[11] If all market returns in the sample are positive or negative, pathological cases can arise in which the estimated bsw exceeds the prior bounds.

IV Predictive Performance of Undecayed bsw

We begin by describing the performance of winsorized market-beta estimates in predicting future OLS market-betas without age decay. Section V will add age decay.

The regressions in this section use a data set with mid-year to mid-year one-year rates of return from 1927 to 2018. Thus, not only are the market-beta measures never overlapping, neither are the observations in the predictive gamma regressions.

We consider two measures of estimation success, both based on the ability of different estimated current market-betas ($\hat{b}$) to predict the one-year-ahead realized OLS market-beta, $\text{bols}_{i,y+1}$. The first comes from predictive "gamma" regressions,

$$R^2\left(\text{bols}_{i,y+1} = \gamma_0 + \gamma_1 \cdot \hat{b}_{i,y} + \epsilon_{i,y+1} \right) . \tag{9}$$

The second measure is the equivalent RMSE

$$\sqrt{\sum_{i,y}(\text{bols}_{i,y+1} - \hat{b}_{i,y})/N} . \tag{10}$$

The regression R^2 is not affected by bias in the proxy, while the RMSE is.

The tables typically show the results of pooled panel regressions, but the nature of the data means that the regressions are mostly cross-sectional. Each year has thousands of observations in the cross-section and there are fewer than one hundred years. Not shown, results with Fama-Macbeth-like specifications are always similar.[12]

Realized OLS market-betas are a measure of realized diversification benefits for an investor holding the market portfolio. It could be argued that they are themselves reasonable estimation bogies. However, they are not the expected OLS market-betas. To the extent that realized betas differ from expected betas by noise, our estimates suffer from noise (e.g., lower R-squareds in the gamma regressions). Noise is a minor concern for the dependent variable, the future bols. We expect ex-ante estimators that perform better in predicting the future realized beta bols also to be able to predict better the unknown true and expected

[12] Fama-Macbeth regressions exploit the fact that stock returns should be uncorrelated over time. Market-betas are not uncorrelated over time.

market-betas. Put differently, errors in the dependent variable is a benign complication. It is what OLS was designed for.

Errors-in-Variables: It is more problematic that the independent variable is also measured with error. If the underlying model is stable, then the asymptotic bias in a slope coefficient of past on future values is $1/(1+\sigma_e^2/\sigma_b^2)$, where σ_e^2 is the squared standard error of the noise and σ_b^2 is the cross-sectional dispersion in the (market-beta) predictor. The average estimated market-beta standard error in the market-model regression is a rough estimate for σ_e^2. It is about 0.05 (per day). The average dispersion of estimated market-betas in the cross-section is a rough estimate for σ_e^2. It is about 0.40 (per firm). Thus, an over-the-envelope estimate for the γ_1 bias is about $1-1/(1+0.05/0.4^2) \approx 2\%$.[13] The empirical bias is much larger because the underlying market-beta is also not stable. In addition to time-varying betas, our sample also has time-varying heterogeneity in firm-size and with it time-variation in cross-sectional market-beta means and standard deviations.

Achievable Prediction: If both the dependent and the independent variable are proxies drawn with error from an unknown true but stable normal variable, the R^2 of a cross-sectional regression of one proxy on the other yields an R^2 that is the square of the R^2 in an (infeasible) regression of the true (unknown) beta parameter on a realized OLS beta. For example, if the OLS market-beta, bols, can explain 25% of its future self, it would suggest that bols could explain about $\sqrt{.25} \approx 50\%$ of the true unknown market-beta. If the underlying market-betas are changing, the estimated square root of the R^2 is a lower bound. The association with the true market-beta would be higher.

A Winsorization Deltas

Section III already outlined the tradeoffs in choosing an appropriate single static parameter for the winsorization cutoffs. Too strict a winsorization destroys the heterogeneity of betas. Too lax a winsorization becomes meaningless. There is a golden middle. Section I.B already hinted at reasonable ranges, for which winsorization would not seem too harsh, cutting into the main distribution of stock returns or market-betas.

[13] These approximations have ignored the panel nature of the data. However, unreported simulations suggest that they are reasonably applicable in our panel sample, too.

[Insert Figure 4 here: **Effect of Winsorizing Deltas on Predictive Performance** (R^2)]

Figure 4 plots the R^2s from gamma regressions, predicting $bols_{i,y+1}$ with winsorized market-beta estimates using different deltas. The plots show that variations in the winsorization choices are never knife-edge. The predictive R^2s have reasonable ranges. The top plot shows that level winsorization is better than no winsorization, but blw never reaches the predictive power of bVCK regardless of Δ_l. The middle plot shows that bbw reaches it when winsorizing net-of-market returns at around 2% to 3%. The bottom plot shows that bsw modestly exceeds it for Δ_s from 1.5 to 3.5.

B Dynamic Deltas?

The deltas used in my paper are fixed in advance. They consume at most one degree of freedom and have (nearly) no hindsight.[14] In principle, deltas could be chosen based on data analysis or (in-sample) estimations. They could also be different for the upper or lower bounds. However, the spirit of a specification-based[15] estimate is to keep the deltas as simple as possible. No first-stage regressions are needed, either. Section V shows that it is not impossible to improve the parameter estimates dynamically, but the potential gains are likely to be small for the variables considered.

C Synopsis

[Insert Table 2 here: **Gamma Regressions Predicting One-Year Ahead OLS Market-Beta (bols)**]

Table 2 describes the characteristics and performance numerically, using good winsorization parameters.

[14] When pinned down in the years before the estimation (e.g., before 1963 for 1963–2018 estimation), they do not even consume a single degree of freedom or use any hindsight. With modest use of hindsight, they could also be assessed from the frequency of extreme market-beta estimates. Only 0.17% of all unwinsorized firm-year OLS market-betas exceed Δ_s=3. In defense, note that other estimators also require parameter choices. For example, although the Vasicek estimator seems to require no extra parameters, it results in such badly biased beta estimates that it effectively requires the debiasing parameters as in Levi and Welch (2017). (Otherwise, it is unsuitable for use as a direct market-beta proxy.)

[15] The "specification-based" moniker is intended to suggest that bsw is not based on a multi-step model-parameter-based estimation, but that it is a much simpler estimator based just on a pre-specified prior.

The two left columns show the means and standard deviations of the independent variables. bols and bVCK have similar means (by construction) of about 0.79. There is both good heterogeneity and time-series variability in their realizations, but the distribution of the shrinkage bVCK is about 20 percent tighter (0.55 vs. 0.68). The 7%-level-winsorized beta has a lower mean of 0.71 and reduces the heterogeneity from 0.68 (for bols) to 0.41. Both the bbw($\Delta_B = 3\%$) and the bsw($\Delta_s = 3$) leave the mean at 0.79 but reduce the heterogeneity to a startling low of 0.44. The means of all estimators considered here are similar, which makes it a little easier to predict them with one another.

The columns to the right show the performance predicting the future bols.

- Model A is a "null" benchmark, in which we predict bols with the cross-sectional bols mean from the prior year. This mean reflects the time-varying composition of stocks on CRSP and the equal-weighting of stocks in our market-model regression with the value-weighted market index. The lagged mean target is also effectively part of the estimation of bVCK. Model A has modest bias (γ_1=0.842) but low R^2 (6.09%).
- Model B predicts with the standard OLS beta, bols. bols is a badly biased predictor of its future self. It has an auto-coefficient of only $\gamma_1 \approx 0.565$ and an R^2 of 27.97%. The bias is much too large for a simple error-in-variables problem and points to underlying mean reversion in bols over the one-year intervals considered here.
- Model C predicts with the Vasicek beta, bVCK. bVCK is less biased ($\gamma_1 \approx 0.756$) and has a higher R^2 of 33.38%. By construction, the bLW variant in Model D achieves a nearly unbiased coefficient of $\gamma_1 \approx 1.008$.
- Model E predicts with the level-winsorized market-betas, blw. It predicted better than bols but a little worse than bVCK. Its gamma coefficient was $\gamma_1 \approx 0.721$, its R^2 was 31.84%.

The specification-based winsorized market-betas perform well:

- Model F predicts with the band-winsorized market-beta, bbw. bbw had a predictive coefficient of 0.943 and an R^2 of 33.27%.
- Model G predicts with the slope-winsorized market-betas bsw. bsw had a predictive coefficient of 0.977 and an R^2 of 33.82%.

The last rows consider extended estimators:

- Model H predicts with an estimator which first slope-winsorizes and then Vasicek shrinks. It could outperform bsw, but not in an economically meaningful way. The R^2 of 33.97 improves by only 0.15% relative to Model G.
- Model I predicts with two independent variables, bsw and bVCK. Its R^2 increases to 34.51%. The coefficients are 0.56 on bsw and 0.34 on bVCK. Depending on the application, this 0.7% increase relative to Model G may or may not be a useful improvement.
- Model J predicts with all estimators from A through G simultaneously. The R^2 increases to 34.77%. (blw could be omitted.)

Where applicable, the RMSE column in Table 2 shows the combined effect of bias and noise. The RMSE is relevant if a beta estimator is used as a direct 1-to-1 proxy for the true beta. Because of its bias, bVCK performs poorly. bLW, bbw, and bsw perform better.

Not shown and well known, all regressions lose some statistical power predicting two-year-ahead OLS market-betas. It suggests that the underlying market-betas are themselves not stable but mean-reverting. This deterioration could have been due to outliers, but investigation shows that this is not the case. It appears as strongly with predictors that have winsorized the outliers as those that have not. Furthermore, the inference is similar when running on stocks that have two years of data vs. stocks that disappear in the second year.

D Performance Over The Years And Assessing Potential Improvements

[Insert Figure 5 here: **RMSE by (July-June) Year for bsw**]

Figure 5 plots the RMSE and best Δ_s year by year. bsw outperforms bVCK not just on average, but in 69 out of 92 years (75%). The exceptions are, in order, 1943 (ending June 1943, predicting 1943/07-1944/06), 1942, 1969, 1971, 1944, 2016, 1939, 1947, 2000, 1946, 1965, and 1953. The remaining 11 inferior years were just mildly worse. There is a modest time-varying pattern to the optimal delta, but the figure shows that even perfectly knowing the ex-post optimal Δ_s one year ahead would not have greatly improved the predictive performance relative to the fixed Δ_s=3 (except in 1942-3 and 1966). The black dotted line just below the blue line in the bottom plot shows the performance relative to bLW, essentially removing the bias disadvantage of bVCK. The advantage of bsw shrinks slightly, but the pattern remains the same, especially after World-War II.

Not shown, averaged on a year-by-year basis, a bsw with an infeasible (ex-post optimal) delta yielded a 2.40% average improvement in R^2 over bVCK, while the feasible bsw yielded a 1.52% average improvement. After 1974, the former yielded a 2.67% improvement while the latter yielded a 2.26% improvement. This 0.4% is the best delta optimization that could be obtained by a good model of the year-dimension related parameter. However, the plot also shows that reaching this 0.4% will be difficult: the variance with a simple time-series analysis prediction on the optimal Δ_s is likely to remain large, making it difficult for researchers to reach this upper bound. Together, it seems unlikely that modeling time drifts in Δ_s will easily yield much improvement when (a) it is difficult to estimate the best delta and (b) the R^2 improvements are modest even if one were to know the best deltas. Using the same two criteria, Section V.C will show why many other variables are unlikely to help in choosing good Δ_s and age-decay ρ parameters.

V Predictive Performance of Age-decayed bswa

It has been known for a long time that the underlying beta drifts, and thus that the usefulness of historical returns decays with age. The common way to handle this issue is to use block-samplers (Ghysels and Jacquier (2007)), which is also what we did by using 12-month intervals in the preceding section.

A smoother alternative is to progressively disregard older observations (i.e., without suffering "drop off") in the market-model regression itself. A weighted-least-squares regression (WLS) can smoothly incorporate age decay. Such a regression requires a second parameter, henceforth named ρ. (Block sampling also requires a window choice.) The age-decayed market-beta estimator is named bswa.

To keep the WLS implementation simple, we investigate only simple exponential age decay structures with only one single parameter regardless of firm and time. Like bsw, users can calculate bswa in time without the need of a first-stage regression. Each day should receive a weight of about 0.7% to 0.9% higher than the preceding one. These weights imply half-lives somewhere between 75 and 120 days, with 90 days a good middle $(1/(1+2/256)^{90} \approx 0.5$. Ergo, yesterday's stock returns should have about twice the weight of stock returns from four months ago, eight times the weight of those from one year ago, and sixty times the weight of those from two years ago. Three-year-old stock returns are effectively irrelevant.[16]

A Contour Plots of Performance by Parameters Delta and Rho

[Insert Figure 6 here: **bswa: Slope-Winsorization and Decay with Entire History**]

Figure 6 shows contour plots of the RMSE[17] predicting the future 1-month-ahead (typically 21 trading returns) and 1-year-ahead OLS market betas (252 returns).

Winsorization: The optimal winsorization parameter continues to hover around 3, with little RMSE performance differences for choices ranging from 2.75 to 4.25.

Age Decay: The optimal age decay ranges from $2/252 \approx\approx 0.8\%$ to $3.5/252 \approx 1.4\%$ per day when predicting the short-term market beta, and $1.5/252 \approx 0.6\%$ to $3/252 \approx 1.2\%$ per day when predicting a long-term market-beta.[18] Interestingly, different prediction horizons seem to prefer different patience: when forecasting shorter-term market-betas, the most recent observations should be considered relatively more important.

Not shown, R^2 improves and bias worsens with higher (more lax) deltas and lower (more patient) age decays.

Because we want to work only with parsimonious integer parameter estimates, we suggest $(\Delta_s{=}3, \rho{=}2)$ as our base case for bswa. The performance differences between our

[16] The rapid age decay helps explain why market-beta estimates using *monthly* rates of returns, which have to sample more than 12 months, are mostly noise (Section VI.B).

[17] When comparing market-beta performance within subsets, it is easier to work with RMSE's than R^2's. It makes little sense to work with subset gamma regressions. When subsets become small, as they may well in 100-by-100 contour grid cells, the required de-biasing cell-by-cell advance knowledge exceeds what can be expected to be known by investors or researchers in-time.

[18] One may want to estimate the instantaneous market-beta instead of the one-year market beta. In this case, one could inversely decay forward-looking stock returns.

(Δ_s=3, ρ=2) choice and the assessed optimal fractional choices, both in the training and in the test sample, are statistically but not economically significant.

B Different Sets

[Insert Table 3 here: **Gamma Regressions Predicting One-Year Ahead OLS Market-Beta (bols)**]

Table 3 shows the relative performance of bswa compared to earlier estimators. Using the improvement of bVCK over bols as a 100% baseline, After 1964, bsw predicted the future bols about 14% better than bVCK. bswa improved it by about 48%. (The data set barely matters.)

C Categorized Performance and Potential Conditional Parameter Improvements

In the following figures, stocks are first partitioned into separate sets. Each set can be viewed as its own "out-of-sample" test from the perspective of the other sets—they are not overlapping. If the association were random, spurious, or driven by outliers, the fact that an estimator holds in one percentile would not suggest that it should hold similarly in its neighbor. The figures show that there are clear patterns: nearer partitions have more similar results. Moreover, the superior performance of bswa is near-universal.

Similar to the logic in Subsection IV.D, we can assess whether better modeling along some X dimension is likely to improve the market-beta prediction. In particular, if there are strong patterns in the optimal parameters across X (i.e., and with little residual variance when predicting the optimal parameter), *and* if knowing the ex-post optimal parameter yields a good improvement in RMSE, then modeling the link from X to optimal parameters is likely to meaningfully improve predictive performance.

[Insert Figure 7 here: **Various Betas Predicting Future OLS Beta, Categorized By Year**]

Year: The gray top lines in the left panel of Figure 7 show the performance and loess-smoothed performance of an estimator that uses only the lagged average market-beta without cross-sectional discrimination. The gray areas show that there were some years (e.g., 1998-2000) in which bols was so poor a predictor that the unconditional average bols across all stocks from the preceding year predicted the future bols just as well or better than the firm's own bols (the black line above 0).

The zero horizontal line represents the normalized predictive performance of bVCK. Because the bols line is above zero, bVCK (almost always) predicted the future bols better than bols itself.

The blue lines below zero show the performances of the following:

1. bsw, the block-sampled 12-month undecayed slope-winsorized market-beta;
2. a one-year slope-winsorized beta with decay parameter ρ=2;
3. bswa, the age-decayed slope-winsorized market-beta (for which we also show the individual year RMSEs with points);
4. an infeasible hypothetical bswa which uses the ex-post optimal Δ_S and ρ in each partition (here year).

These four lines make it easier to understand how much of the performance is due to age decay, how much is due to winsorization, and how much one could potentially improve the parameters. Figure 7 shows that the winsorization and the decay were roughly equally important over the years. The improvement of bswa over bVCK was about half as much as the improvement of bVCK over bols. The improvement was fairly stable over the decades. However, there were multiple years before 1963 and a few years after, in which bVCK outpredicted bswa.

The right panel shows that there were time-varying patterns to the best parameters from year to year. Unfortunately the large residual year-to-year variability in ex-post optimal Δ_S and ρ suggests that it would be difficult to reach the performance of the dotted line in the left plot.

In the next set of figures, observations within each year are first ranked by different variables. Each observation is then assigned a within-year percentile rank. (This can be done in time by an investor.) The figures then plot the RMSEs predicting next year's bols within each percentile.

[Insert Figure 8 here: **Various Betas Predicting Future OLS Beta, Categorized By Market Capitalization**]

[Insert Figure 9 here: **Various Betas Predicting Future OLS Beta, Categorized By Trading Volume**]

Market-Cap: The left plot in Figure 8 shows that bols was a poor predictor for firms of any size. For the smaller half of stocks, $\text{bols}_{i,y}$ predicted no better than the indiscriminate cross-sectional bols average over all stocks $\overline{\text{bols}}_y$ from the prior year.

Our interest lies more with the relative performance of the slope-winsorized market-betas vs. bVCK. The plot shows that this also does not change greatly across market cap. bsw and bswa always outperform bVCK, except for the very largest percentile where the method of estimation becomes irrelevant. Even bols performs well. The right plot shows a modestly useful pattern. For small and large firms, higher (more lax) winsorization thresholds and slower age decay could improve performance. However, the right plot shows that this would yield a useful prediction improvement only for small firms.

Dollar Trading Volume: The equivalent figure for stocks ranked on (lagged) dollar trading volume suggests a useful estimation possibility: For the bottom tertile of rarely trading stocks, improved patience (lower ρ) with laxer winsorization could yield better estimates than bswa. The residual variance around predicting better parameters is low, and better parameters could improve prediction.

The left plot also shows the performance of the Dimson beta, bdim (as a solid orange line). Because the predicted target is bols, it is not surprising that bdim predicts poorly for stocks on the left. However, if bdim's efficiency cost were low, it should predict the future bols about as well as the plain bols estimator for the percentiles of stocks that trade the most. This is not the case. When non-synchronicity is not an issue, bdim significantly underperforms bols. The use of bdim comes with a high efficiency cost.

[Insert Figure 10 here: **Various Betas Predicting Future OLS Beta, Categorized By OLS Market Beta**]

[Insert Figure 11 here: **Various Betas Predicting Future OLS Beta, Categorized By OLS Market Beta Standard Error**]

Finally, we can investigate the performance categorized by stock return based statistics:

OLS Market-Beta: Figure 10 shows that when the OLS beta coefficient is near 1, all estimators perform about the same. Meaningful improvements can occur only when the OLS market-beta is different from 1. For stocks with large OLS market-betas, the bswa estimator performs almost as well as an estimator with perfect ρ and Δ_S foresight. For stocks with very low (negative) OLS market betas, more patience (smaller age decay $\rho\approx1.5$) and more aggressive winsorization ($\Delta_S\approx2$) might yield a small improvement.

OLS Beta Standard Error: The second critical variable used in bVCK is the standard error of bols. Presumably, any attempt to fine-tune the delta parameter with firm-specific estimated information would also attempt to estimate the standard error of the OLS market-beta, and then use this estimate to decay or winsorize differently.

The left plot in Figure 11 shows that even an ex-post optimal Δ_S and ρ can barely improve the RMSE performance for all but the noisiest decile of stocks. Unfortunately, as the right plot shows, these stocks are also the same ones where the optimal parameters are all over the place. There is no promising strong pattern here. Thus, it is unlikely that modeling uncertainty for the sake of a Δ_S parameter improvement is likely to work.

Not shown, the plots for stock return volatility and residual market-model volatility look similar to that for the standard error of bols.

In sum, it seems that it is more important to choose reasonable winsorization parameters than to attempt to fine-tune them. The most promising improvement would be to estimate market-betas with more patience (lower ρ) and laxer winsorization (higher Δ_S) when trading volume (and marketcap) is low.

D Improvements in the Market-Model (First-Pass) Regressions

I could not identify another variable that could improve the market-model estimation in a meaningful manner. I investigated only relatively obvious high-frequency daily time-series, such as daily trading volume, spreads, etc.

E CRSP and Compustat Variables To Improve the Gamma (Second-Pass) Prediction Regressions

Low-frequency variables could be useful as additional predictors in the gamma regressions.

[Insert Table 4 here: **Predicting Future 12-Month OLS Market Beta (bols)**]

Panel A of Table 4 shows that including a measure of one-year trading volume—specifically the within-year rank of total dollar trading volume—can improve the prediction accuracy from 34.74% to 35.89%. Depending on the application, the 1.2% gain in R^2 can be a meaningful improvement.

The prediction equation suggests that one should increase the estimated beta for firms with high trading volume rank, and decrease the estimated beta for firms with low trading volume. (Not shown, on the margin, the most-frequently traded stock should have a market-beta estimate about 0.2-0.3 higher than the lowest trading stock.) Marketcap could be used instead of trading volume, with only a minimal loss of prediction accuracy. Using a log marketcap may be easier than first percentile ranking trading volume, so we now use it as our "base model" in Panel B and Table 5. What other variables can usefully predict these residuals?

Panel B shows that the other CRSP variables in Table 4 did not greatly improve the prediction, even in these in-sample regressions. The next best variables are again transforms of dollar trading and marketcap.

[Insert Table 5 here: **Explaining Base Model Residuals With Compustat-Based Variables**]

Table 5 extends the analysis of base model residuals to include variables obtained from Compustat. Some of them greatly restrict the sample due to limited data availability. None of these variables matter.

F Conclusion

The age-decayed slope-winsorized bswa is a superior predictor of the future bols. The most promising (but not pursued) improvements relate to more use of trading volume. Stocks with low trading volume could benefit from more patience (slower decay ρ) and less aggressive winsorization (higher Δ_S), and/or be assigned somewhat lower market-betas than suggested by bswa. These improvements would require substantially more coding and data snooping than the simple bswa considered here.

VI Alternative Beta Estimators Post 1974

A Benchmark Performance

Having explored good Δ_S and ρ parameters, and having come to the conclusion that the one-size-fits-all values of Δ_S=3 and ρ=2 perform well enough, we now benchmark this simple bswa to other beta estimators.

In this section, we restrict the sample to years after 1974 (when Nasdaq was available) and now use overlapping observations in the gamma regressions. The maximum number of observations is thus 2,961,446 firm-years. The benchmark comparisons in each table will use the same set of observations (firm-years), because different estimates have different sample availabilities.

We also predict not only bols but also other beta measures. All market-beta estimators attempt to uncover the true beta signal in noisy stock return data. If the estimation error is iid, the best estimator should predict not only the OLS market-beta estimates better but also other noisy estimates of market beta.

This logic breaks down if there is persistence in the errors (e.g., a persistent alternative focus), too. The Dimson (1979) beta estimator falls into this category. If trading frequency is stable and the betas of infrequently trading stocks are downward biased, then the future Dimson estimate should retain the same bias as the current estimate. The persistent bias should make it easier for the Dimson beta to predict its future self in competition with other betas.

Investors care about future market-betas, not current market-betas. For an estimator to predict its future self well is a necessary but not a sufficient success criterion. If an estimator does not predict its future self better than bswa (and investors and researchers typically care about the future market-beta), then one may as well use bswa instead. Yet, self-prediction is not enough. For example, a "beta estimator" claiming that beta is the firm's perm number could predict itself perfectly well, but it could not predict the OLS or other market-beta estimates. This qualification is important, because it is sometimes argued that industry betas should be used because they are more stable than firm betas. Although they are indeed more stable, the evidence suggests that industry betas are very poor predictors of their individual constituent firm betas.

Not shown, when predicting market-betas over the next month instead of over the next *twelve* months, the dependent variable is much noisier. However, the ordering of estimator performance remains the same.

B Specific Alternatives

[Insert Figure 6 here: **Full Sample Post-1974 Benchmark Performance Comparisons (**$N = 2,961,446$**)**]

Table 6 shows the performance for all estimators for which the full 2,961,446 (overlapping) market-beta pairs (observations) could be used.

The OLS Market-Beta: When predicting bols, bsw with its R^2 of 43.7% modestly outperforms bVCK (R^2≈44.2%), but bswa outperforms it comfortably (R^2≈46.2%). As in Table 3, bswa outperforms bVCK by about half as much as bVCK outperforms bols.

The Vasicek (1973) Beta: When predicting bVCK, bswa with its R^2≈53.4% predicts better than own lagged bVCK with its R^2≈50.6%. Ergo, researchers interested in working with the future bVCK should use the current bswa rather than the current bVCK.

The Slope-winsorized Market-Beta: When predicting bsw, bswa with its R^2≈56.3% predicts better than own lagged bsw with its R^2≈53.9%. Ergo, researchers interested in working with the future bsw should use the current bswa rather than the current bsw. The 56.3% estimate suggests that bsw could predict the true unknown beta with an R^2 of about 75% (see Section IV). (bswa should have an even larger 80-90% correlation with the unknown true market-beta.)

The Aged Slope-winsorized Market-Beta: Some of the excellent predictive performance of bswa on itself ($R^2 \approx 62.4\%$) is due to overlapping stock returns in the calculation of market-beta (because bswa uses the entire history).

Dimson (1979): The Dimson beta estimator (bdim) corrects for non-synchronous trading by including leads and lags of the market rate of return in the market-model regression. (bdim is the sum-total of the coefficients on differently timed market rates of return.)

The Dimson estimator is in wide use. It is easier to implement in practice than the Vasicek estimator, because beta estimates are not entangled with those from other firms. No cross-sectional statistics are needed.

The original Dimson (1979) used its namesake estimator only on low-frequency (monthly) returns of (typically decile) portfolios. It also suggested Vasicek (1973) shrinkage and Blume (1971) regression inspired improvements. However, the subsequent literature has mostly ignored these enhancements and used only the summed coefficients. Most papers have also applied bdim in other contexts, such as in the context of individual stocks and/or daily stock returns. My paper calls bdim the "Dimson beta" not because the original papers used bdim in the daily-frequency individual stock context, but because the widely-used version had its origin in Dimson (1979).[19] Although bdim is intuitive and appealing, Figure 9 already showed that it suffers from a large efficiency loss. There is a large gap between bdim and bols even for the most liquid stocks.

Model A^1.4 of Table 6 showed that bdim predicted bols poorly (as it did also in Levi and Welch (2017)). This poor performance is not surprising, because the (synchronicity-biased) OLS beta is not its intended target. If the current OLS estimate is biased due to non-synchronicity, then the future OLS estimate should be biased, too. A future Dimson estimate would be a better bogie. However, Models C.1 and C.2 of the table show further problems. When predicting bdim, bswa with its $R^2 \approx 31.6\%$ predicts better than own lagged bdim with its $R^2 \approx 22.0\%$. Ergo, researchers interested in

[19]Fowler and Rorke (1983) point out that the Dimson (1979) estimator is incorrect. It does not yield inconsistent estimates. However, in practice, this is unlikely to be a big problem, and it is the original Dimson and not the Fowler-Rorke corrected estimator that remains in wide use.

working with the future bdim should use the current bswa rather than the current bdim. This was unexpected. bswa is so much better extracting the true beta signal that it can outpredict despite the persistent bias adjustment provided by bdim. Note also that bswa would predict bdim even better on the RMSE metric if one were to add the mean difference of 0.91 – 0.80≈0.11 to each bswa.

Other Robust Estimators: Robust estimation was pioneered in Tukey (1960). It was first used in the market-beta context by Martin and Simin (2003). The latter used a band-like shrinker based on first-stage OLS regressions and a non-linear two-equation shrinkage solution. Unfortunately, the Martin-Simin betas are difficult to compute. Fortunately, Timothy Simin generously shared two alternative robust estimators of monthly market-beta estimates. To be included, a firm-year had to have a CRSP share code 10 and 11 firms and at least 240 daily observations. His estimates begin in June 1966 and end in June 2018. They contains 2,023,162 firm-months.

The first set of betas are "maximum likelihood type estimators," abbreviated bmm. They are described in Rousseeuw (1984). These betas are similar to those in Martin and Simin (2003), but use a Huber (1964) loss function instead of an optimal one. According to Simin, the Huber estimates are typically nearly identical, but their computation is five times faster. Nevertheless, it took days running distributed Matlab on the Penn State super-computer to obtain the bmm estimates. The second set of betas are least-trimmed squares beta based on Yohai (1987), abbreviated blts.

[Insert Table 7 here: **Benchmark Performance Comparisons, Alternative Robust Estimators (Martin and Simin (2003),** $N = 2,018,080$)]

With a different set of observations, we require a new table. Table 7 shows that both robust estimators perform well in predicting bols, but not as well as bVCK, bsw, and bswa. bmm performs better than blts. Not shown, the difference in means suggests that a simple bias adjustment would further improve the RMSE when bswa predicts either bmm or blts.

When predicting bmm, bswa with its R^2≈52.3% predicts better than own lagged bmm with its R^2≈49.7%. Ergo, researchers interested in working with the future bmm should use the current bswa rather than the current bmm. When predicting blts, bswa with its R^2≈49.7% predicts better than own lagged blts with its R^2≈45.7%.

Ergo, researchers interested in working with the future blts should use the current bswa rather than the current blts.

Frazzini-Pedersen: The Frazzini and Pedersen (2014) beta estimator (bfp) uses five years of data to calculate correlations and one year of data to calculate standard deviations. The estimator is then the correlation multiplied by the ratio of the two standard deviations, shrunk with a Blume (1971)-like linear equation, $0.6 \cdot \hat{b} + 0.4$. The origin of the estimator is not clear, because Frazzini and Pedersen (2014) did not validate or benchmark it. Nevertheless, perhaps due to the success of the paper in predicting future average rates of return, bfp has become prominent.[20]

[Insert Table 8 here: **Benchmark Performance Comparisons, Frazzini and Pedersen (2014) BaB Market-Betas** ($N = 1,440,636$)]

Again, requiring five years of returns to calculate correlations leading to the different set of observations, we need a new table. When predicting bfp, bswa with its $R^2 \approx 31.1\%$ predicts better than own lagged bfp with its $R^2 \approx 20.6\%$. Ergo, researchers interested in working with the future bfp should use the current bswa rather than the current bfp. However, because the bfp mean is much higher than the bols mean, bfp has a lower RMSE explaining itself than bols or bswa. The RMSE when bswa predicts bfp (in Model H.3) can be improved by 0.19 by adding $0.99 - 0.80 \approx 0.19$ to each bswa estimate.

Intra-Day Betas: Ait-Sahalia, Kalnina, and Xiu (2014) also generously shared their beta estimates, named btaq1. Because their intra-day TAQ based beta estimates are for one month each, we also consider a moving average of the most recent 12 months, named btaq12.

[Insert Figure 9 here: **Benchmark Performance Comparisons, Ait-Sahalia, Kalnina, and Xiu (2014) TAQ Market-Betas**]

Again, we need a new table due to the different set of observations. Panel A of Table 9 briefly shows that btaq1 does not predict either a one-month bols or itself well. When

[20] Novy-Marx and Velikov (2018) show how this estimator is biased in the time-series, picking up firm-specific time-changing volatility patterns. Han (2019) shows that most of the average return performance comes from the time-series component.

predicting btaq12, bswa with its $R^2 \approx 50.6\%$ predicts better than own lagged btaq12 with its $R^2 \approx 48.1\%$. Ergo, researchers interested in working with the future btaq12 should use the current bswa rather than the current btaq12.

A researcher interested in future market betas based on monthly stock returns is still better off estimating the predicting beta variable using ordinary daily stock returns. However, it may be possible to incorporate the slope-based winsorization technique directly into the intra-day estimation to improve it. (Another disadvantage of btaq12 vs. bswa is that it will remain more difficult to code and compute, likely relegating it to niche applications only.)

Monthly-Stock Return-Based OLS Market-Betas: Table 10 considers market-betas obtained from monthly-frequency stock returns. The window was expanded to 60 months, empirically a good (self-predictive) window. (Not shown, the following conclusions are robust to window lengths from 36 to 60 months.)

[Insert Figure 10 here: **Benchmark Performance Comparisons, Monthly-Return-Frequency Market-Betas** ($N = 1,715,377$)]

The table shows that future monthly-frequency market-betas are mostly noise. The predictive R^2's in the gamma auto-regressions are 1.6% for a monthly-frequency OLS market-beta bmols and 3.7% for its Vasicek equivalent bmvck. The own historical average, with no attempt at discrimination across stocks, does not perform much worse.

When predicting bmols, bswa with its $R^2 \approx 3.3\%$ predicts better than own lagged bmols with its $R^2 \approx 1.6\%$. Ergo, researchers interested in working with the future bmols should use the current bswa rather than the current bmols. When predicting bmvck, bswa with its $R^2 \approx 6.1\%$ predicts better than own lagged bmvck with its $R^2 \approx 3.7\%$. Ergo, researchers interested in working with the future bmvck should use the current bswa rather than the current bmvck.

Now shown, in this sample, bswa also has a mean of 0.80. Ergo, the RMSE when bswa predicts bmols (in Model A^6.3) can be improved by 0.27 by adding 1.07–0.80 to each bswa estimate, and by 0.24 in Model K.3 by adding 0.24 to each bswa estimate.

Industry Betas: Not shown, I also experimented with industry betas. Industry market-betas are poor predictors of the market-betas of the stocks in their industries.

Peer Betas: Not shown, I also experimented with averaging in the beta of the nearest industry-marketcap peer. Unfortunately, this did not improve predictive performance.

C Distinctiveness of Estimates

Our final concern is the distinctiveness of market-beta estimates. Does it even make a difference which market-beta is used?

[Insert Table 11 here: **Distinctiveness of Market-Beta Measures**]

Table 11 shows the root-mean-squared-difference among estimators for the set of 432,426 stock-years where all market-betas were available. The typical difference ranged from 0.1 to 0.8. The robust estimators (Vasicek, winsorized, and Martin-Simin relayed) tend to be similar. Their typical distance from one another was about 0.1 to 0.2. The Vasicek market-beta was most similar to the OLS market-beta.

VII Conclusion

The slope-winsorized beta estimator (bsw) and its age-decayed version (bswa) are easy to implement. Using at least 12-18 months of historical *daily* stock return data in `ri` and `rm`, the R code is

```
_bswa <- function( ri, rm, Delta, rho ) {
   wins.rel <- function( r, rmin, rmax ) {
      rlo <- pmin(rmin,rmax); rhi <- pmax(rmin,rmax)
      ifelse( r<rlo, rlo, ifelse( r>rhi, rhi, r ) )   }
   wri <- wins.rel( ri, (1-Delta)*rm, (1+Delta)*rm )
   beta <- function(...) coef(lm(...))[2]
   # ri and rm must be increasing in time
   bsw <- beta( wri ~ rm, w=exp(-rho*(length(ri):1)) )
}

bsw <- function( ... ) _bswa( ... , Delta=3.0, rho=0.0 )
bswa <- function( ... ) _bswa( ..., Delta=3.0, rho=2.0/256.0 )
```

C code to calculate all bswa at the end of each month for each CRSP stock is available. Because returns can be added progressively, it takes less than one minute to calculate monthly market-betas from daily stock returns for the entire CRSP universe.

Using a computer software analogy, the predictive performance (and modest difference) suggests that bVCK and bLW are good enough not to require an upgrade for completed research projects. However, the upgrade is free. The resulting bswa performs much better and is easier to use. All other market-beta estimators seem inferior. Many are inferior even to plain OLS market-beta estimate. In their defense, their use can be harmless—mostly in situations in which one may as well use 1.0 as the unversal market-beta estimate and not even bother with estimation: (I) Market-betas for individual stocks estimated on monthly frequency stock returns are mostly noise. Although bswa remains the superior predictor, the gains (in predicting mostly noise) are small. (II) The estimates of market-betas for many sufficiently large portfolios do not vary greatly with estimation method. In the extreme, portfolios similar to value-weighted market portfolios should have market-betas of around 1, regardless of beta estimator.

References

Ait-Sahalia, Yacine, Ilze Kalnina, and Dacheng Xiu (2014). *The Idiosyncratic Volatility Puzzle: A Reassessment at High Frequency*. Tech. rep. University of Chicago.

Blume, Marshall E. (1971). "On the assessment of risk". In: *The Journal of Finance* 26.3, pp. 1–10.

Bollerslev, Tim, Sophia Zhengzi Li, and Viktor Todorov (June 2016). "Roughing up beta: Continuous versus discontinuous betas and the cross section of expected stock returns". In: *Journal of Financial Economics* 120.3, pp. 494–490.

Dimson, Elroy (June 1979). "Risk measurement when shares are subject to infrequent trading". In: *Journal of Financial Economics* 7.2, pp. 197–226.

Engle, Robert F. (1982). "Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation". In: *Econometrica* 50, pp. 987–1007.

Foster, Dean P. and Dan B. Nelson (Jan. 1996). "Continuous Record Asymptotics for Rolling Sample Variance Estimators". In: *Econometrica* 64.1, pp. 139–174.

Fowler, David J and C Harvey Rorke (1983). "Risk measurement when shares are subject to infrequent trading: Comment". In: *Journal of Financial Economics* 12.2, pp. 279–283.

Frazzini, Andrea and Lasse Heje Pedersen (2014). "Betting against beta". In: *Journal of Financial Economics* 111.1, pp. 1–25.

Ghysels, Eric and Eric Jacquier (2007). *Market Beta Dynamics and Portfolio Efficiency*. Tech. rep. University of North Carolina and HEC Montreal.

Glosten, Lawrence R., Ravi Jagannathan, and David E. Runkle (Dec. 1993). "On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks". In: *The Journal of Finance* 48.5, pp. 1779–1801.

Han, Xing (2019). *Understanding the Performance of Components in Betting Against Beta*. Tech. rep. University of Otago, NZ.

Harvey, Campbell R., Yan Liu, and Heqing Zhu (Jan. 2016). "...and the Cross-Section of Expected Returns". In: *Review of Financial Studies* 29 (1), pp. 5–68.

Huber, Peter J. (1964). "Robust Estimation of a Location Parameter". In: *Annals of Statistics* 53.1, pp. 73–101.

Karolyi, G. Andrew (1992). "Predicting Risk: Some New Generalizations". In: *Management Science* 38, pp. 57–74.

Levi, Yaron and Ivo Welch (Apr. 2017). "Best Practice for Cost-of-Capital Estimates". In: *Journal of Financial and Quantitative Analysis* 52 (2), pp. 427–463.

Lo, Andrew W and A Craig MacKinlay (1990). "Data-snooping biases in tests of financial asset pricing models". In: *The Review of Financial Studies* 3.3, pp. 431–467.

Martin, R. Douglas and Timothy T. Simin (Sept. 2003). "Outlier-Resistant Estimates of Beta". In: *Financial Analysts Journal* 59.5, pp. 56–69.

Mclean, R. David and Jeffrey Pontiff (Feb. 2016). "Does Academic Research Destroy Stock Return Predictability?" In: *The Journal of Finance* 71.1, pp. 5–32.

Novy-Marx, Robert and Mikhail Velikov (Nov. 2018). *Betting Against Betting Against Beta*. Tech. rep. University of Rochester and FRB of Richmond.

Rousseeuw, Peter J. (1984). "Least Median of Squares Regression". In: *Journal of the American Statistical Association* 79.388, pp. 871–881.

Tukey, John W. (1960). "A survey of sampling from contaminated distributions". In: *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling*. Ed. by I. Olkin et al. Stanford University Press, pp. 448–485.

Vasicek, Oldrich A. (1973). "A Note on Using Cross-Sectional Information in Bayesian Estimation of Security Betas". In: *The Journal of Finance* 28.5, pp. 1233–1239.

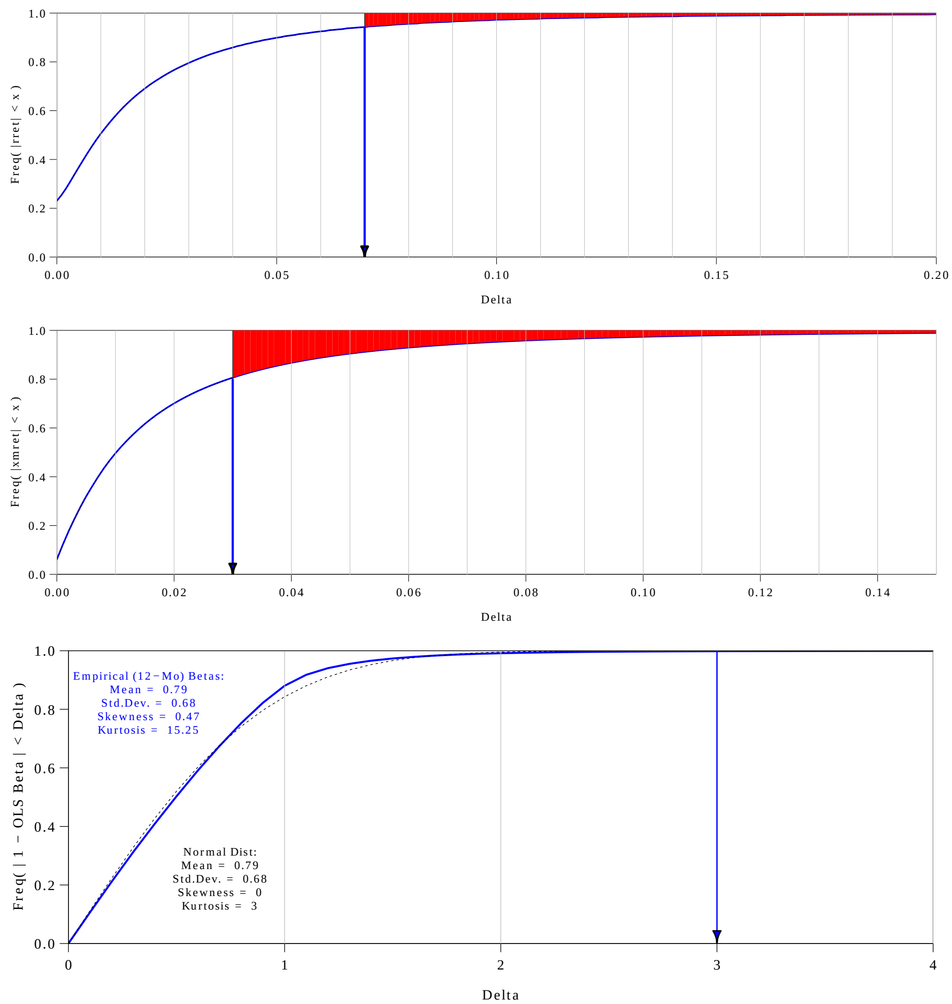
Yohai, Victor J. (1987). "High Breakdown-Point and High Efficiency Robust Estimates for Regression". In: *Annals of Statistics* 15.2, pp. 642–656.

Table 1: Glossary of Market-Beta Estimators

Abbrev	Long Name	Remarks
12 Months of Daily Stock Return Data		
bols	OLS	
bVCK	Vasicek	Random-Effects Panel Estimator. Vasicek (1973).
bLW	Levi-Welch	0.25 + 0.75·bVCK. Levi and Welch (2017).
blw	Level-Winsorized	Default Parameter: $\Delta_l = 5\%$.
bbw	Band-Winsorized	Default Parameter: $\Delta_b = 3\%$.
bsw	Slope-Winsorized	Default Parameter: $\Delta_s = 3$.
$\overline{\text{bols}}$	Average bols	Equal-weighted average over all stocks during preceding 12 months.
bdim	Dimson-like	Sum of three market coefficients. Dimson (1979).
bmm	Robust Maximum-Likelihood	Martin and Simin (2003), Rousseeuw (1984).
blts	Robust Least-Trimmed	Martin and Simin (2003), Yohai (1987).
btaq1	Intra-Day (1 Month!)	TAQ-based. Ait-Sahalia, Kalnina, and Xiu (2014).
btaq12	Intra-Day (12 Mos)	Average of 12 consecutive btaq1.
Extended (60 Months) of Stock Return Data		
bfp	Frazzini-Pedersen	Daily-frequency returns only, 12 months for variances, 60 for correlation. Frazzini and Pedersen (2014).
bmols	Monthly OLS	from 60 monthly stock returns.
bmvck	Monthly VCK	from 60 monthly stock returns.
Daily Stock Return Data As Long As Available		
bswa	Slope-Winsorized and Aged	Default Parameters: ρ=2, Δ_s=3. Rho implies $(1+2/252)-1 \approx 0.8\%$ decay per day past. Uses entire prevailing daily stock return history.

When self-computed, the input stock returns are always net of the prevailing risk-free rate from Ken French's website. The market is always the value-weighted CRSP stock index rate of return, also net of the risk-free rate of return. The bmm, blts, and btaq1 estimates were provided by the original authors. The bfp betas were checked against those estimated in Novy-Marx and Velikov (2018).

Figure 1: Quasi-CDF: Frequency of Winsorizations

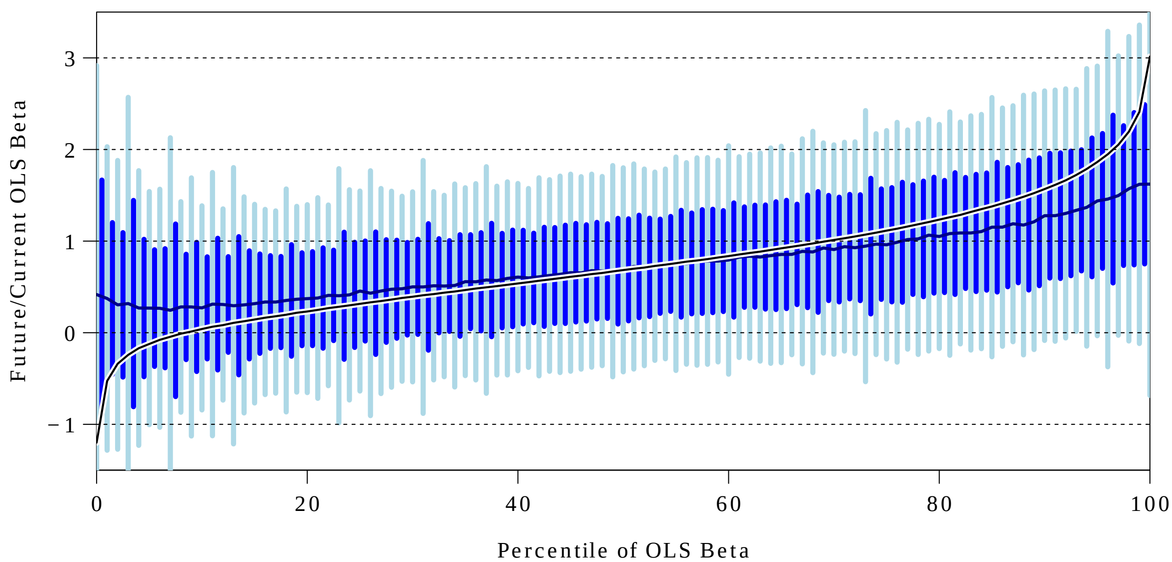


Explanations: In the top two panels, the sample are all valid stock returns (1926 to 2018). These plots show cumulative frequencies of *absolute* rates of return, either net of the risk-free rate (top) or net of the value-weighted market (middle). The red area shows the mass of returns winsorized at 7% and 3%, respectively. In the bottom panel, the sample are all annual OLS market-betas (bols). The blue line in the bottom panel shows the bols frequency that are at least a distance of x deltas away from 1.0. The thin dotted black line on the bottom panel is from a normal distribution with matched mean and variance. The arrows denote the preferred delta parameters in this paper. Absolute daily returns above 10% are not rare (3.2% frequency). Annual OLS market-betas below –1 and +3 are rarer (0.9%). OLS market-betas beyond a Δ of 3, i.e., below –2 and +4, are very rare (0.17%).

Interpretation: The preferred winsorization delta parameters effect different number of stock returns and market-betas. Warning: Δ_S may effect few overall market-betas but can influence many stock returns (although most will not matter much to the estimated regression slopes, bsw and bswa).

Source: fig-qcdf/qcdf-C.Rmd, forbeta-annual.csv.gz, crspdaily-clean.csv.gz → *cdf.pdf

Figure 2: Percentiles of Future 1-Year OLS Market-Betas (bols) By Current 1-Year OLS Market-Beta (bols)

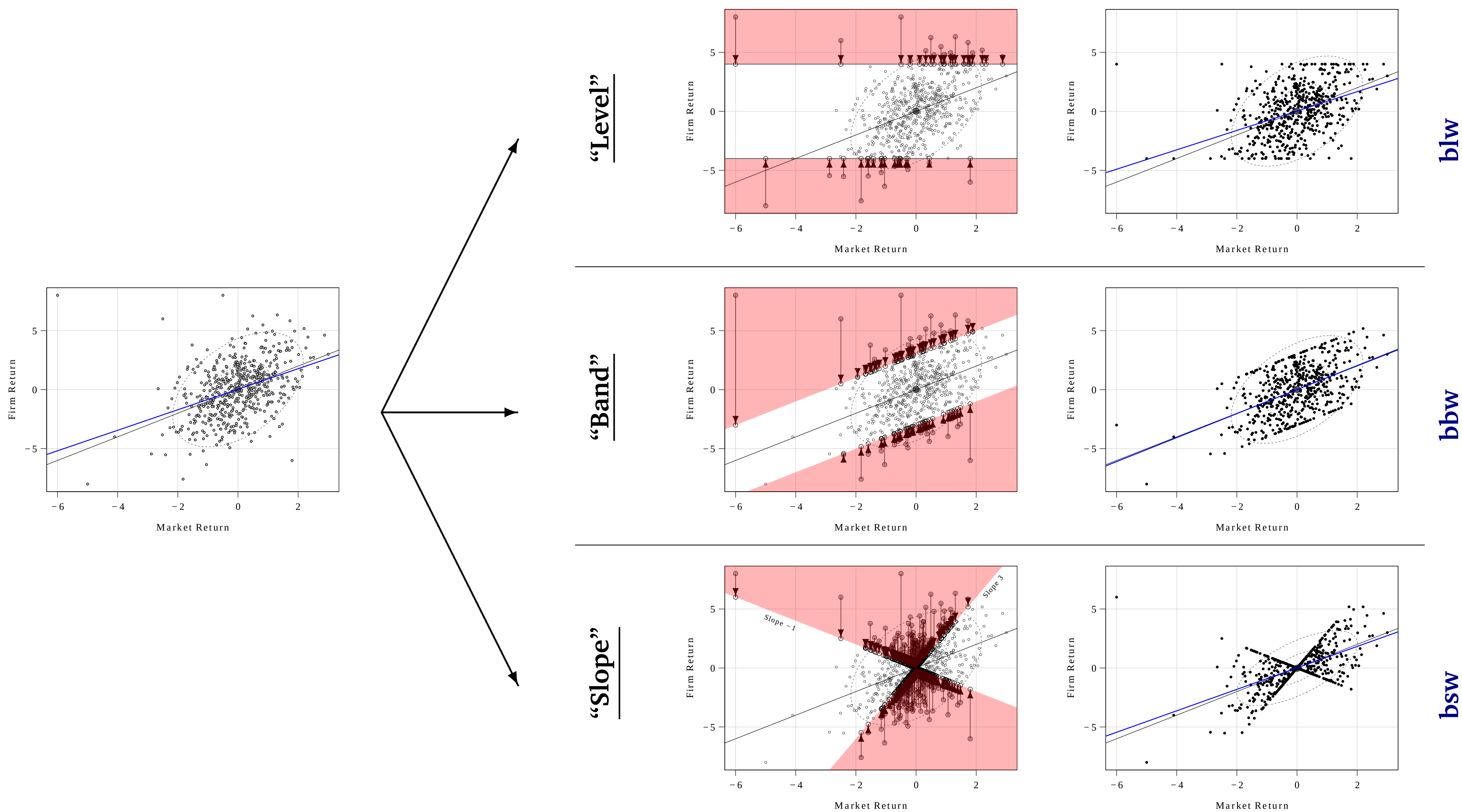


Explanations: The x-axis is the *within-year* percentile of the current estimated 12-month OLS market-beta (bols). (With about 2,000-4,000 stocks per year, this implies about 20-40 stocks per bin per year; over 55 years, about 1,000 stock-years.) The y-axis shows primarily the distributions of the equivalent future bols's. The thick blue line shows the mean of this next year's OLS market-beta. The blue bands show the one-standard-deviation and two-standard deviation bounds around this mean (based on realized future market-betas, not assessed from ex-ante standard errors). The black-white variable is the variable on which the binning was based. Its monotonic relation is mechanistic. The sample are all firm-years with valid current and future OLS market betas (bols) from 1926 through 2018.

Interpretation: Although about 5% of stock-years show negative market-betas, for negative market-beta estimates today, lower current market-betas no longer predict lower future market-betas *on average*. A prediction of a future bols of less than –2 is about 3 standard deviations below the mean prediction even for the most negative bols percentile today.

Source: fig-ols2fut/plotols2fut.Rmd, forbeta-annual.csv.gz → ols2fut*pdf (slopeols*pdf, olsbetahist.pdf)

Figure 3: Winsorization Techniques



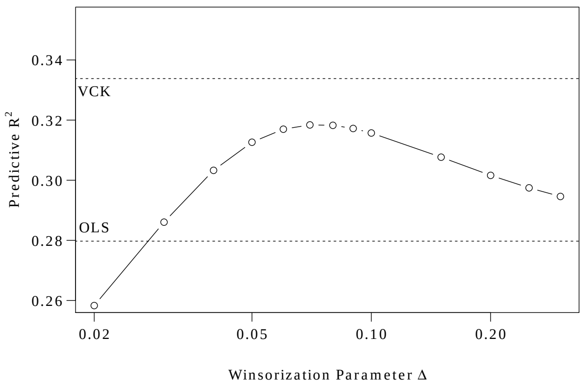
Explanations: The left column shows data drawn from a true $r_i = 0 + 1 \cdot r_m + \epsilon$ (black line). The estimated OLS line (in blue) happens to be a little flatter in this draw. The middle column illustrates the method of winsorization. Level winsorization truncates the dependent variable (the firm's own stock return). Band winsorization truncates based on the contemporaneous market rate of return. Slope winsorization truncates based on a beta prior. The right column shows the winsorized returns and the new OLS estimated line.

Interpretation: Level winsorization flattens the estimated slope. Band winsorization is harshest on positive return outliers when the market was bearish and on negative return outliers when the market was bullish. Slope winsorization is harshest where it matters least for the regression coefficient estimates, i.e., around r_m=0.

Source: fig-conceptual/conceptualplots.Rmd

Figure 4: Effect of Winsorizing Deltas on Predictive Performance (R^2)

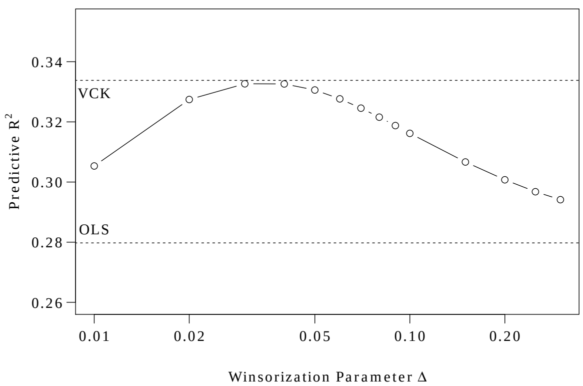
Level (blw)



$r_{it} \in [\, -\Delta_l \,,\; \Delta_l \,]$

Parameter Units:
Daily Net-of-Risk-Free Returns.
Reasonable Range: 4 – 10%.
Basecase: $\Delta_l = 7\%$.

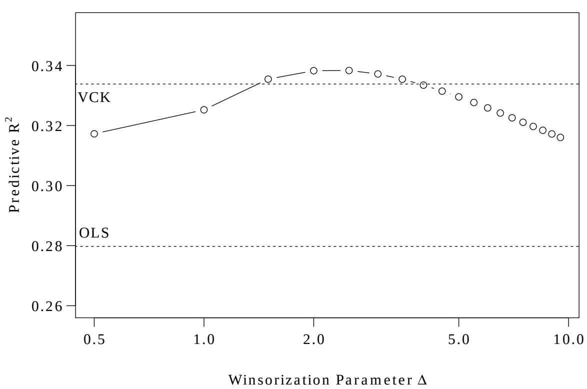
Band (bbw)



$r_{it} \in [\, r_{m,t} - \Delta_b \,,\; r_{m,t} + \Delta_b \,]$

Parameter Units:
Daily Net-of-Market Returns.
Reasonable Range: 2 – 5%.
Basecase: $\Delta_b = 3\%$.

Slope (bsw)



$r_{i,t} \in 1.0 + [\, (-\Delta_s) \,,\; (+\Delta_s) \,] \cdot r_{m,t}$

Parameter Units:
Market-Beta.
Reasonable Range: 1.5 – 4.0.
Basecase: $\Delta_s = 3$

Explanations: Variables are defined in Table 1, the sample in Section I. The 317,005 market-betas are for 1927/06 to 2019/12 (including the last half year). These plots show the R^2's in gamma regressions predicting future OLS market-betas (bols) with current market-betas when stock returns have first been winsorized, as indicated. The plots also show the predictive R^2's of the OLS and Vasicek/Levi-Welch market-beta estimators as horizontal lines. Variables are defined in Table 1, the sample in Section I.

Interpretation: Winsorization deltas are not overly sensitive but have to be reasonable. Level-winsorized betas predicted worse than band- or slope-winsorized betas.

Source: figs-annualbetas/1mkvarydeltas.Rmd, forbeta-xdaily.csv.gz → results.*/varydeltas.csv.gz, 2plotbydelta.Rmd → results.*/rsq-by-delta.pdf

Table 2: Gamma Regressions Predicting One-Year Ahead OLS Market-Beta (bols)

	Mean	SD	Abbrev	Predictor $b_{i,t}$	RMSE	γ_0	γ_1	R^2
A	0.80	0.21	$\overline{\text{bols}}$	Past Year Firm-Average OLS	0.700	0.111	0.842	6.09%
B	0.79	0.68	bols	(Own) OLS Market-Beta	0.680	0.332	0.565	27.97%
C	0.79	0.55	bVCK	Vasicek Market-Beta	0.604	0.184	0.756	33.38%
D	0.79	0.41	bLW	... Levi-Welch (0.75)	0.589	–0.017	1.008	–"–
E	0.71	0.56	blw	Level-Winsorized (Δ_l=7%)	0.621	0.271	0.721	31.84%
F	0.79	0.44	bbw	Band-Winsorized (Δ_b=3%)	0.590	0.033	0.943	33.27%
G	0.79	0.43	bsw	Slope-Winsorized (Δ_s=3)	0.587	0.008	0.977	33.82%
H	0.79	0.42		Slope-Wins Then Vasicek	0.586	–0.014	1.008	33.97%
I				Multivariate, bsw and bVCK				34.51%
J				Multivariate A to G				34.77%

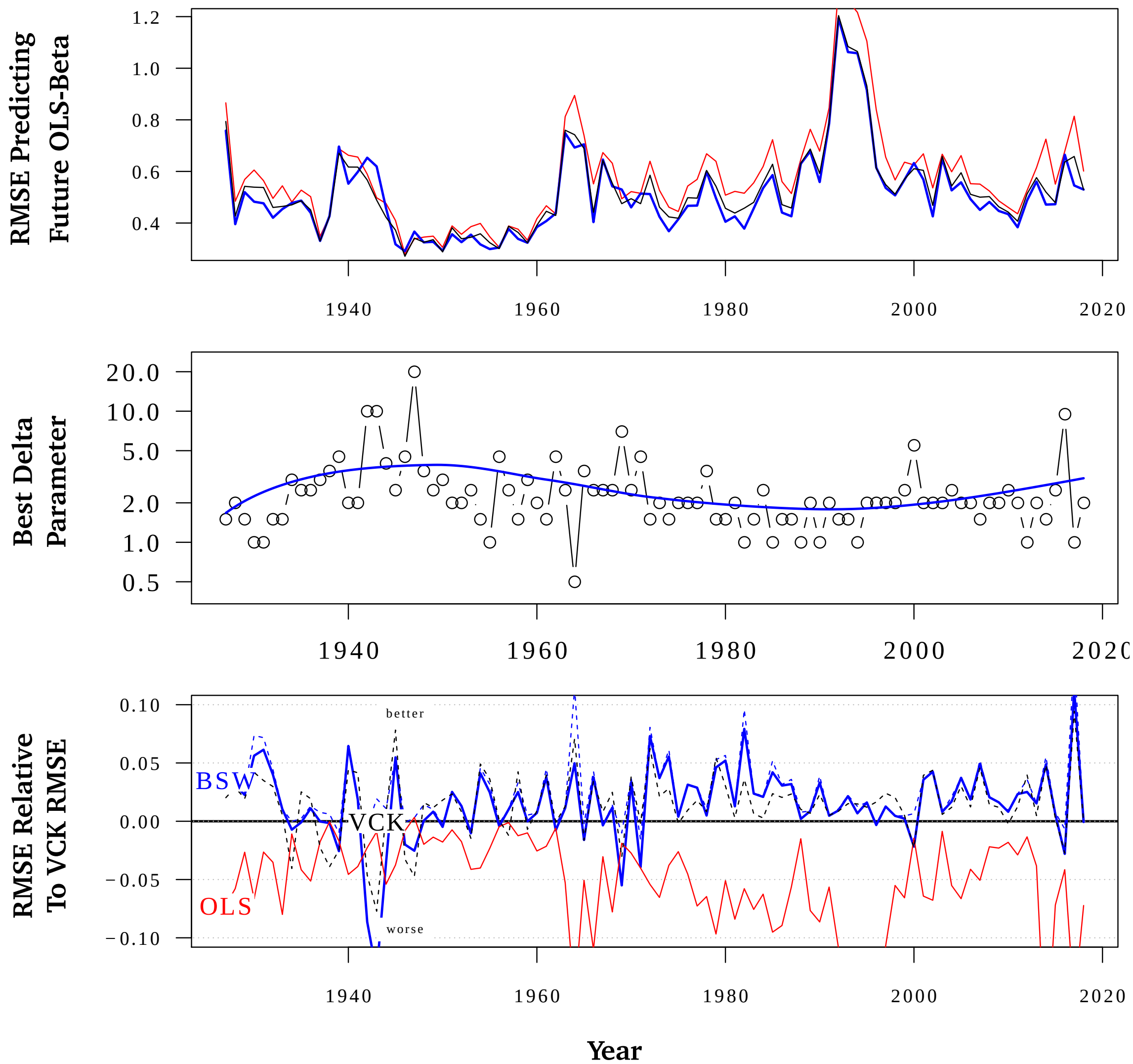
Explanations: Variables are defined in Table 1, the sample in Section I. The 317,005 market-betas are for 1927/06 to 2019/12 (including the last half year). The first two columns show the mean and standard deviation of the *in*dependent variable.

Each row shows the results of one grand pooled bivariate prediction. (Not shown, averaged year-by-year statistics (ala Fama-Macbeth) convey similar insights.) The RMSE data column is the direct proxy prediction error, as in RMSE $= \sqrt{\sum(\text{bols}_{i,t+1} - b_{i,t})^2}$. The final three columns are coefficient estimates and the R^2, as in $\text{bols}_{i,t+1} = \gamma_0 + \gamma_1 \cdot b_{i,t} + e_{i,t}$. Model H is a combination estimator, first slope-winsorized, then Vasicek-shrunk, which will not be used again elsewhere. (Unreported Newey-West coefficient standard errors are between 0.0021 and 0.0048, due to the hundreds of thousands of observations.) Models I and J use multiple independent variables.

Interpretation: The band- and slope-winsorized betas predicted the future OLS betas as well as the VCK/LW betas. There is not much gain to adding Vasicek shrinkage to the slope-winsorized market-beta.

Source: tbl-key-rsq.Rmd, forbeta-annual.csv.gz, mbols.csv.gz →

Figure 5: RMSE by (July-June) Year for bsw

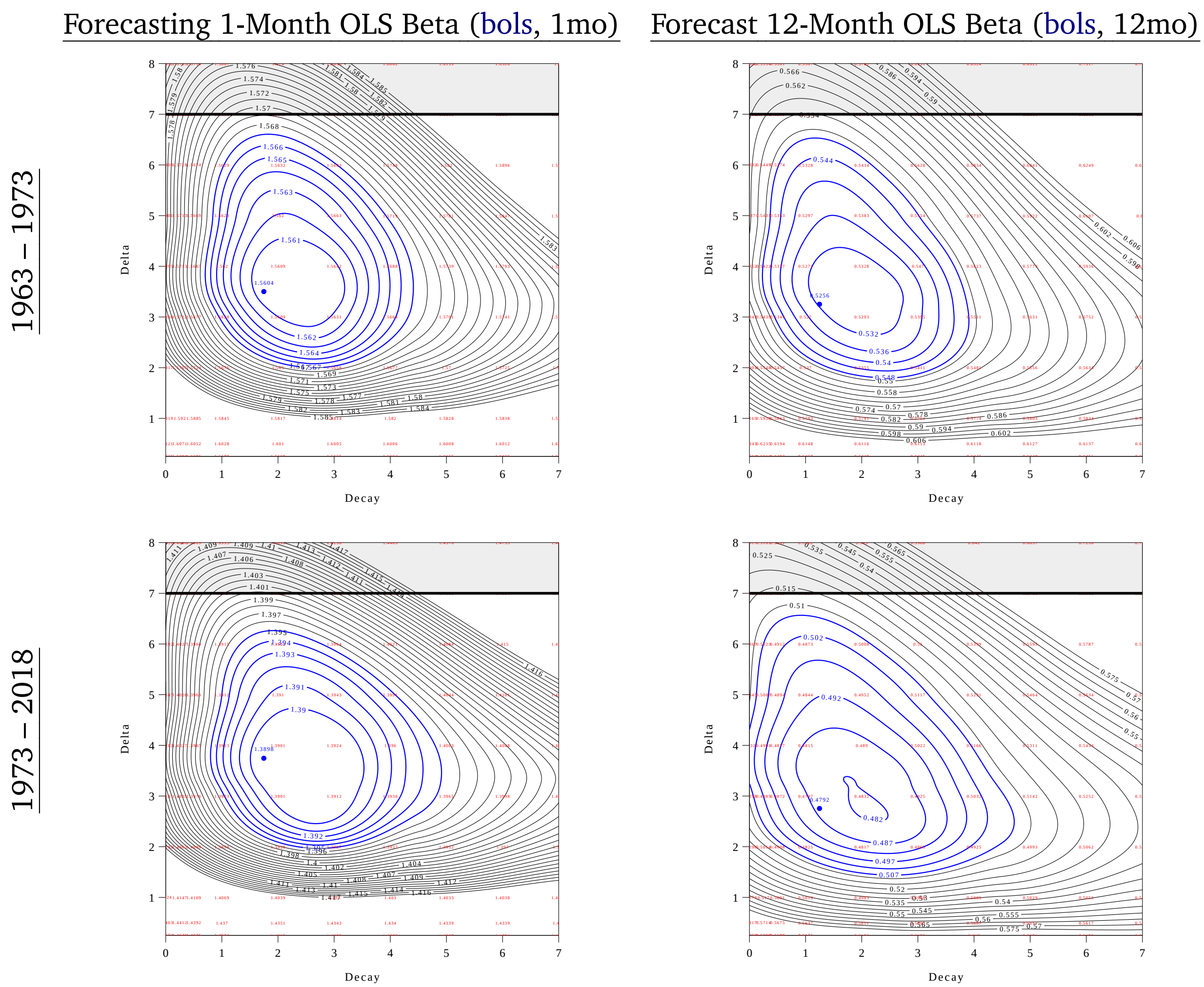


Explanations: Variables are defined in Table 1, the sample in Section I. This figure shows year-by-year detail about the predictive performance of the bsw estimator. This includes the absolute RMSE explanatory performance (top panel) and the (ex-post) best Δ_s winsorizer parameter (middle panel). The red line shows the performance of bols, the blue line of bsw, the black line of bVCK. In the top panel, higher is worse. In the bottom panel, higher is better. The dashed blue line in the bottom panel presumes foreknowledge of the best delta, as in the middle panel. The dashed black line is performance relative to bLW, rather than relative to bVCK.

Interpretation: The top plot shows that market-betas were unusually difficult to predict in the 1990s. The middle plot shows that there was only a mild time-pattern to the best Δ_s. The bottom plot shows that even ex-post knowledge of Δ_s would not have greatly improved the predictive performance (except in 1966, when Δ_s should have been much lower [stricter]).

Source: figs-annualbetas/1mkvarydeltas.Rmd, forbeta-xdaily.csv.gz → results.*/varydeltas.csv.gz, 2plotbycateg.Rmd → results.*/rmse-*.pdf

Figure 6: bswa: Slope-Winsorization and Decay with Entire History



Age Decay Factor ρ***:*** Use 2.0 to 3.5 for a predictor of the next one-month bols, and 1.5 to 3.0 for a predictor of the next twelve-month bols.

Slope Winsorization Delta Δ_S***:*** Use 2.75 to 4.25.

Explanations: These are contour-plots of the root-mean-squared errors, predicting either the 1-month-ahead or the 1-year-ahead OLS market-beta with age-decayed and slope-winsorized betas (bswa). Delta values plotted above 7 did not execute any (slope) winsorization; the unwinsorized RMSE values were placed as if $\Delta_S = 7$. The age decay was exp(–d/256), where d is the number of trading days. Contourlines in blue (black) show the range where the bswa estimator outperforms (underperforms) the 1-year bVCK estimator. R's contourplot function interpolates terrain optima to lie modestly east of the observed best sample point (indicated with a blue dot). The terrain surfaces are so flat that the optimal parameters are difficult to tell.

Interpretation: Comparing top and bottom figures, the best parameters seemed reasonably stable. There is no meaningful difference between a 1963–1973 "training" sample and a 1973–2018 "test" sample. However, this masks year-by-year variation shown in Figure 7.

Table 3: Gamma Regressions Predicting One-Year Ahead OLS Market-Beta (bols)

Panel A: 1927-2018, 2,962,015 Overlapping One-Year Samples

	Mean	SD	Abbrev	Predictor $b_{i,t}$	RMSE	γ_0	γ_1	R^2	Rel
A	0.81	0.64	bols	(Own) OLS Market-Beta	0.5632	0.304	0.618	38.64%	0.00
B	0.81	0.54	bVCK	Vasicek Market-Beta	0.4979	0.170	0.788	43.68%	1.00
C	0.81	0.49	bsw	Slope-Winsorized (Δ_s=3)	0.4860	0.097	0.880	44.15%	1.09
D	0.81	0.47	bswa	Decayed Slope-Winsorized (Δ_s=3, ρ=2)	0.4753	0.066	0.919	46.15%	1.49

Panel B: 1964-2018, 217,537 Non-Overlapping One-Year Samples

	Mean	SD	Abbrev	Predictor $b_{i,t}$	RMSE	γ_0	γ_1	R^2	Rel
A	0.79	0.65	bols	(Own) OLS Market-Beta	0.5847	0.315	0.590	35.53%	0.00
B	0.78	0.53	bVCK	Vasicek Market-Beta	0.5118	0.173	0.776	40.88%	1.00
C	0.78	0.48	bsw	Slope-Winsorized (Δ_s=3)	0.4985	0.097	0.869	41.62%	1.14
D	0.78	0.47	bswa	Decayed Slope-Winsorized (Δ_s=3, ρ=2)	0.4886	0.066	0.910	43.44%	1.48

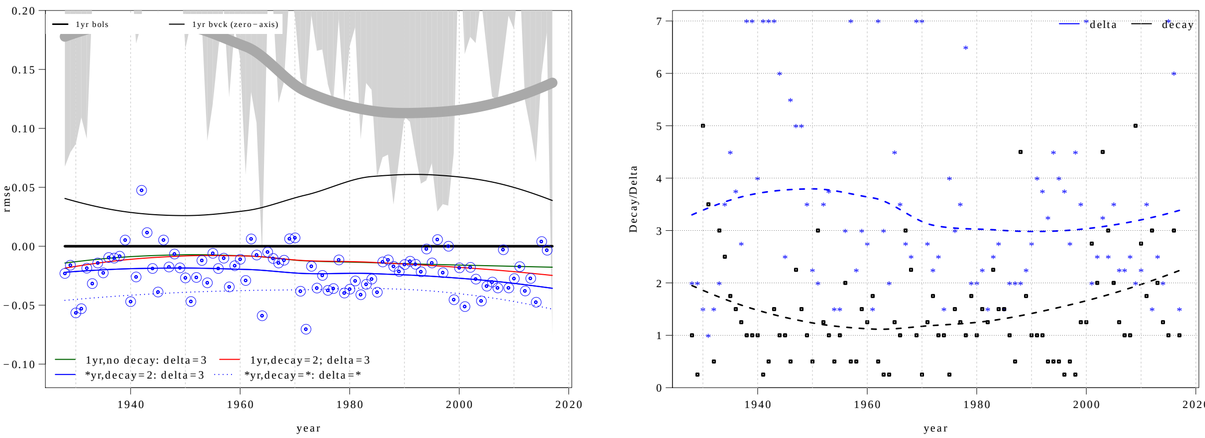
Explanations: Variables are defined in Table 1, the sample in Section I. This table largely mimicks Table 2, but adds bswa and illustrates the effects of using different data sets.

Interpretation: The slope-winsorized betas predicted the future OLS betas better than the VCK/LW betas.

Source: tbl-key-rsq-bswa.Rmd, benchmarks/

Figure 7: Various Betas Predicting Future OLS Beta, Categorized By Year

RMSE | Best Ex-Post Δ and ρ



Explanations: *The left panel* shows the RMSE predicting the future OLS market beta calculated either over 12 months of daily stock returns. Lower RMSE values are better. From top to bottom, the loess-smoothed (span=1) fitted lines are in the following order:

- The cross-sectional firm-averaged OLS market-beta over the last 12 months ($\overline{\text{bols}}$, in gray, both specific points and loess-fit);
- The firm-specific OLS market-beta, bols (in black, loess fit only);
- The Vasicek market-beta, bVCK (the fat black horizontal zero line, actual only);
- The undecayed slope-winsorized beta bsw (green, loess only);
- A 12-month age-decayed slope-winsorized beta (red, loess only);
- bswa, using data from inception of the firm on crsp (solid blue loess line), with the blue plotted points being the individual predictions; and
- a hypothetical equivalent beta estimate with (unavailable) knowledge of the ex-post best decay and slope delta parameters (dotted blue, loess only). This is useful to assess the max potential improvement from attempts to predict parameters better.

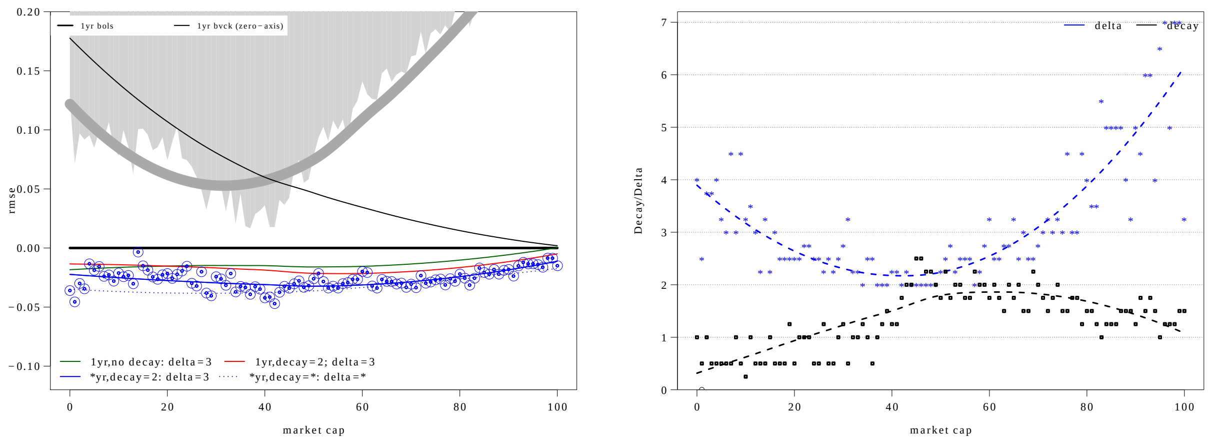
The right panel shows the ex-post best decay and delta parameter points (in each partition) that are used to obtain the hypothetical dotted blue estimate above (black and blue). In addition, they also show smoothed loess fits.

Interpretation: The improvement of bswa over bVCK was about half as large as the improvement of bVCK over bols. The pattern on the right shows that taking into account year-by-year drift in the optimal parameters could potentially improve the prediction, but only modestly.

Source: baseparms/by/1runstats.R(runbase.R), forbeta-annual.csv.gz, bsws/bsw-all.csv.gz, mean-ols-beta.csv.gz → by-[pctrank|year]-[dolvol|mcap|...]-w[1|12].csv.gz, plotstats.R, by-[pctrank|year]-[dolvol|mcap|...]-w[1|12].csv.gz

Figure 8: Various Betas Predicting Future OLS Beta, Categorized By Market Capitalization

RMSE | Best Ex-Post Δ and ρ



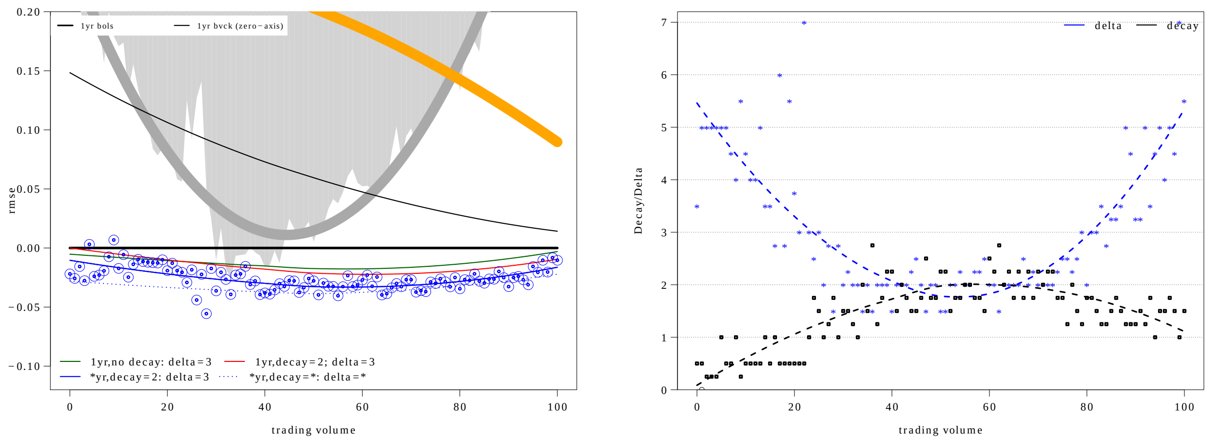
Explanations: This figure is identical to Figure 7, except that the x-axis is for 100 percentiles based on the market capitalization (from CRSP) at the end of the year over which the predictive betas were also estimated (i.e., in the 12 month immediately preceding the predicted $bols_{i,t+1}$).

Interpretation: The left plot shows that the R^2 ordering of estimators is stable across marketcap. bswa reliably outpredicts bVCK. The right plot shows that there are clear patterns to better parameters, but the left plot also shows that the resulting R^2 improvements would only be modest for the larger half of firms.

Source: baseparms/by/1runstats.R(runbase.R), forbeta-annual.csv.gz, bsws/bsw-all.csv.gz, mean-ols-beta.csv.gz → by-[pctrank|year]-[dolvol|mcap|...]-w[1|12].csv.gz, plotstats.R, by-[pctrank|year]-[dolvol|mcap|...]-w[1|12].csv.gz

Figure 9: Various Betas Predicting Future OLS Beta, Categorized By Trading Volume

RMSE | Best Ex-Post Δ and ρ



Explanations: This figure is identical to Figure 8, except that the x-axis is for 100 percentiles based on the dollar trading volume (during the same year at which the predictive betas are calculated).

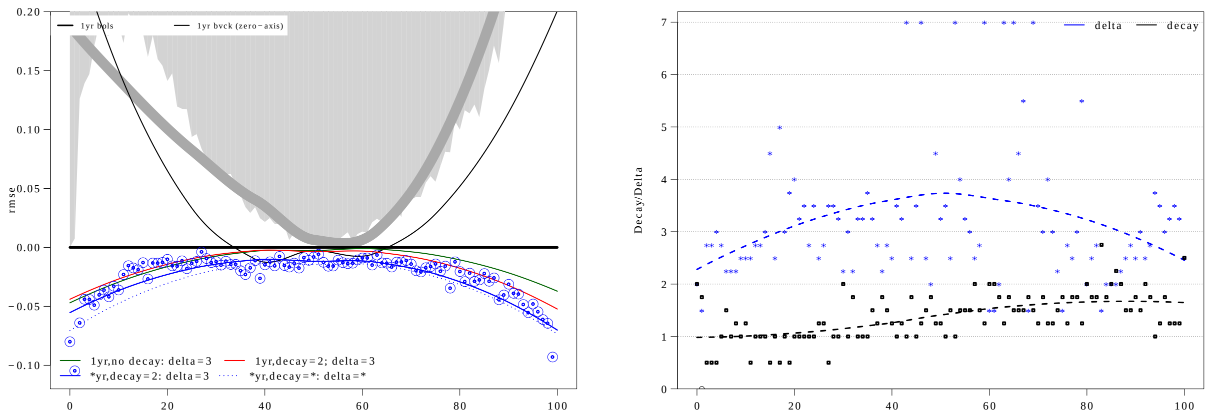
Interpretation: bswa reliably outpredicts bVCK. Modeling parameters for low-trading-volume stocks seems promising. The orange line on the top right is the OLS prediction error for bdim. It is large when non-synchronicity is not of concern (100th percentile).

Source: baseparms/by/1runstats.R(runbase.R), forbeta-annual.csv.gz, bsws/bsw-all.csv.gz, mean-ols-beta.csv.gz → by-[pctrank|year]-[dolvol|mcap|...]-w[1|12].csv.gz, plotstats.R, by-[pctrank|year]-[dolvol|mcap|...]-w[1|12].csv.gz

Figure 10: Various Betas Predicting Future OLS Beta, Categorized By OLS Market Beta

RMSE

Best Ex-Post Δ and ρ

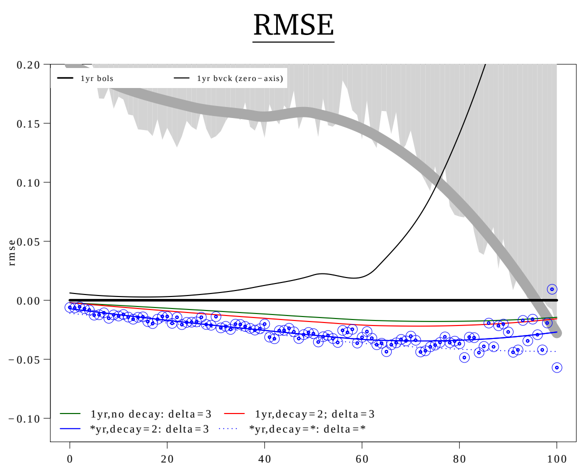


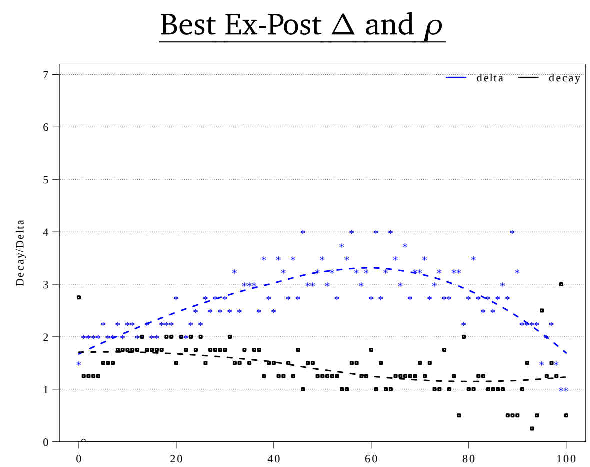
Explanations: This figure is identical to Figure 8, except that the x-axis is for 100 percentiles based on the (predictive) bols.

Interpretation: bswa reliably outpredicts bVCK. For market betas in the lowest tertiale, slower decay ($\rho\approx1$) and more lax winsorization ($\rho\approx2.5$) could improve the prediction.

Source: baseparms/by/1runstats.R(runbase.R), forbeta-annual.csv.gz, bsws/bsw-all.csv.gz, mean-ols-beta.csv.gz → by-[pctrank|year]-[dolvol|mcap|...]-w[1|12].csv.gz, plotstats.R, by-[pctrank|year]-[dolvol|mcap|...]-w[1|12].csv.gz

Figure 11: Various Betas Predicting Future OLS Beta, Categorized By OLS Market Beta Standard Error





Explanations: This figure is identical to Figure 8, except that the x-axis is for 100 percentiles based on the standard error of the (predictive) bols. Equivalent figures with OLS market-model residual standard errors or with plain stock volatility look very similar and are therefore omitted.

Interpretation: bswa reliably outpredicts bVCK. For stocks in the noisest 10%, harsh winsorization (or just using the past mean) could modestly improve prediction.

Source: baseparms/by/1runstats.R(runbase.R), forbeta-annual.csv.gz, bsws/bsw-all.csv.gz, mean-ols-beta.csv.gz → by-[pctrank|year]-[dolvol|mcap|...]-w[1|12].csv.gz, plotstats.R, by-[pctrank|year]-[dolvol|mcap|...]-w[1|12].csv.gz

Table 4: Predicting Future 12-Month OLS Market Beta (bols)

Panel A: Standardized Coefficients, Base Regressions

Name of V	Standardized Coefficients: bswa	V	bswa×V	$\bar{R}^2$	N
(None)	0.59			34.74	228,754
Log(Trading Volume)	0.57	0.05		34.96	
... %Rank	0.53	0.12		35.89	
Log(Market Cap)	0.55	0.10		35.67	(Base Model)
... %Rank	0.55	0.11		35.71	

Panel B: Explaining Base Model Residuals With CRSP-Based Variables

Name of V	Standardized Coefficients: bswa	V	bswa×V	$\bar{R}^2$	N
%Rank(ols.b)	–0.12	–0.07	0.19	0.23 %	228,754
%Rank(ols.xsgm)	–0.02	0.05	0.01	0.29 %	
%Rank(ols.b.se)	–0.01	0.05	0.01	0.30 %	
%Rank(sdy)	–0.03	0.04	0.03	0.28 %	
sdy	0.02	0.01	–0.03	0.01 %	
%Rank(mcap)	–0.11	–0.06	0.16	0.29 %	
%Rank(dolvol)	–0.12	–0.01	0.15	0.46 %	
Log Trading Volume	–0.07	–0.05	0.09	0.06 %	
log Market Cap	–0.20	–0.06	0.23	0.12 %	

Explanations: %Rank is the percent rank. The predicted variable is the future bols. The predicting variables are bswa, the named variable on the left, and the cross-variable. Most of the CRSP variables were used above. Reported coefficients have been standardized by multiplying them by the standard deviation of x and dividing by the standard deviation of y.

Interpretation: Although it would have been possible to improve the prediction statistically, such gains would have been economically small.

Source: tbl-acctg-xsect.Rmd, baseparms/bsws/bsw-2-3.csv.gz, forbeta-annual.csv.gz, compustat-monthly.csv.gz →

Table 5: Explaining Base Model Residuals With Compustat-Based Variables

	Standardized Coefficients				
Name of V	bswa	V	bswa×V	$\bar{R}^2$	N
Log(1 + BV AT)	–0.09	–0.09	0.12	0.18 %	213,470
log(1 + MV AT)	–0.12	–0.08	0.16	0.18 %	203,196
log(1 + EQ)	–0.09	–0.07	0.13	0.12 %	210,608
BV TL/AT	–0.01	–0.02	–0.01	0.07 %	210,596
MV TL/AT	–0.02	–0.03	0	0.11 %	202,174
BV FD/(FD+EQ)	–0.01	–0.01	0.01	0.01 %	204,851
MV FD/(FD+EQ)	–0.01	–0.01	0.02	0.01 %	202,083
Cash / AT	–0.02	0.03	0.01	0.22 %	210,848
BV EQ/MV EQ	–0.02	–0.04	0	0.13 %	208,767
Net Income/Sales	–0.01	–0.03	0.01	0.05 %	218,603
Fin CF/Sales	–0.01	0.04	0	0.16 %	133,420
Opr CF/Sales	–0.02	–0.04	0.03	0.03 %	133,541
Inv CF/Sales	0	0.02	–0.01	0.03 %	133,455

Explanations: BV means book value, MV means market value. The remaining mnemonics are largely as in the Compustat data base itself. They are lagged by six months.

Interpretation: There is no meaningful further predictive power in these variables.

Source: tbl-acctg-xsect.Rmd, baseparms/bsws/bsw-2-3.csv.gz, forbeta-annual.csv.gz, compustat-monthly.csv.gz →

Table 6: Full Sample Post-1974 Benchmark Performance Comparisons (N = 2,961,446)

	Mean	SD	(One-Period-Ahead) Dependent	(Lagged) Independent	g_0	g_1	$R^2_{(\%)}$	rmse
A[1]	0.81	0.655	bols	bols	0.30	0.62	38.8	0.562
				$\overline{\text{bols}}$	0.15	0.82	7.6	0.622
				bVCK	0.17	0.79	43.7	0.498
				bdim	0.38	0.46	27.9	0.685†
				bsw	0.10	0.88	44.2	0.486
				bswa	0.07	0.92	**46.2**	**0.475**
B	0.80	0.539	bVCK	bVCK	0.23	0.71	50.6	0.411
				bswa	0.14	0.83	**53.4**	**0.377**
C	0.91	0.731	bdim	bdim	0.48	0.47	22.0	0.756
				bswa	0.21	0.86	**31.6**	**0.617**
D	0.80	0.485	bsw	bsw	0.21	0.73	53.9	0.355
				bswa	0.19	0.76	**56.3**	**0.340**
E	0.80	0.474	bswa	bswa	0.17	0.78	62.4	0.308

† The rmse can be greater than the SD of the dependent variable, because standard deviations are de-biased unlike RMSEs.

Explanations: Variables are defined in Table 1, the sample in Section I. This sample starts in 1974, with 2,961,446 overlapping observations (one year of daily stock returns). These one-year betas were self-computed from CRSP data. (The predicted variable is always strictly ahead of the predictor variable, except bswa which uses stock return with unlimited history.) The best estimator for each predicted variable is boldfaced.

Interpretation: bswa is the best predictor for predicted market-beta of any kind. Researchers interested in the future one-year market-beta of any kind should still use bswa rather than own incarnations today.

Source: benchmarks/1mkbswall.R, mbols.csv.gz, dacheng-sahalia/dacheng-sahalia.csv.gz, simin/simin.csv.gz, frazzini-pedersen.csv.gz, monthlies/mobetawin60.csv.gz, monthlies/mobetawin12.csv.gz → dall.csv.gz, benchmarks/2runbenchmarks.R → 3benchtex.pl

Table 7: Benchmark Performance Comparisons, Alternative Robust Estimators (Martin and Simin (2003), $N = 2,018,080$)

	(One-Period-Ahead)			(Lagged)				
	Mean	SD	Dependent	Independent	g_0	g_1	$R^2_{(\%)}$	rmse
A^2	0.78	0.626	bols	bols	0.29	0.62	38.8	0.542
				bVCK	0.17	0.78	43.3	0.485
				bsw	0.09	0.88	44.0	0.472
				bswa	0.06	0.92	**46.0**	**0.461**
				bmm	0.30	0.69	42.5	0.514
				blts	0.33	0.68	40.5	0.533
F	0.70	0.589	bmm	bmm	0.21	0.70	49.7	0.453
				bswa	–0.02	0.92	**52.3**	**0.417**
G	0.66	0.587	blts	blts	0.21	0.68	45.7	0.472
				bswa	–0.04	0.89	**49.7**	**0.438**

Explanations: These estimators are fr om Martin and Simin (2003). They were available only for 2,018,080 (overlapping one-year daily market-beta) observations.

Interpretation: bmm and blts perform similar to but somewhat worse than shrunk and slope-winsorized market betas. A researcher who is interested in the best future bmm or blts should still rely on bswa and not on current bmm or blts.

Source: benchmarks/1mkbswall.R, mbols.csv.gz, dacheng-sahalia/dacheng-sahalia.csv.gz, simin/simin.csv.gz, frazzini-pedersen.csv.gz, monthlies/mobetawin60.csv.gz, monthlies/mobetawin12.csv.gz → dall.csv.gz, benchmarks/2runbenchmarks.R → 3benchtex.pl

Table 8: Benchmark Performance Comparisons, Frazzini and Pedersen (2014) BaB Market-Betas ($N = 1,440,636$)

	(One-Period-Ahead)			(Lagged)				
	Mean	SD	Dependent	Independent	g_0	g_1	$R^2_{(\%)}$	rmse
A^3	0.80	0.614	bols	bols	0.28	0.65	42.8	0.512
				bfp	–0.10	0.92	29.9	0.547
				bswa	0.07	0.93	**49.2**	**0.439**
H	0.99	0.370	bfp	bfp	0.54	0.46	20.6	**0.385**
				bols	0.74	0.31	27.1	0.564
				bswa	0.64	0.44	**31.1**	0.449

Explanations: This table explores 1,440,636 (daily stock returns, overlapping) observations in which Frazzini-Pedersen market-betas (bfp) based on daily stock returns could be computed. The dependent variable (future betas) had to be led by 60 months, because bfp has an ingredient requiring 60 months of stock returns.

Interpretation: bfp is a poor predictor either of future bols or future bfp.

Source: benchmarks/1mkbswall.R, mbols.csv.gz, dacheng-sahalia/dacheng-sahalia.csv.gz, simin/simin.csv.gz, frazzini-pedersen.csv.gz, monthlies/mobetawin60.csv.gz, monthlies/mobetawin12.csv.gz → dall.csv.gz, benchmarks/2runbenchmarks.R → 3benchtex.pl

Table 9: Benchmark Performance Comparisons, Ait-Sahalia, Kalnina, and Xiu (2014) TAQ Market-Betas

Panel A: One-Month Prediction: 940,801 Observations

	Mean	SD	(Dep)	(Indep)	g_0	g_1	$R^2_{(\%)}$	rmse
A[4]	0.96	0.608	bols (1 mo)	btaq1 (1 mo)	0.67	0.33	7.4	1.349
				bswa (1 yr)	–0.04	1.08	**17.1**	**1.110**
I	0.91	0.995	btaq1 (1 mo)	btaq1 (1 mo)	0.63	0.31	9.7	1.168
				bswa (1 yr)	0.01	0.97	**20.6**	**0.887**

Panel B: One-Year Prediction: 819,026 Observations

	Mean	SD	(Dep)	(Indep)	g_0	g_1	$R^2_{(\%)}$	rmse
A[5]	0.96	0.610	bols	btaq12	0.36	0.70	40.9	0.502
				bols	0.36	0.66	42.4	0.505
				bswa	0.14	0.90	**48.1**	**0.442**
J	0.94	0.560	btaq12	btaq12	0.30	0.70	47.5	0.440
				bols	0.33	0.62	44.7	0.476
				bswa	0.13	0.86	**50.6**	**0.399**

Explanations: This table explores the subset of months in which Ait-Sahalia-Xiu TAQ-based market-betas were available. The dependent variable (future daily-frequency one-month stock-return based market-betas) in Panel A were led only by one month, because the btaq1 independent variable was computed over one month only. btaq12 is averaged btaq1 over the past twelve months. There were 940,801 one-month and 819,026 one-year observations.

Interpretation: The btaq1and btaq12 variable predicted bols and themselves worse than bswa.

Source: benchmarks/1mkbswall.R, mbols.csv.gz, dacheng-sahalia/dacheng-sahalia.csv.gz, simin/simin.csv.gz, frazzini-pedersen.csv.gz, monthlies/mobetawin60.csv.gz, monthlies/mobetawin12.csv.gz → dall.csv.gz, benchmarks/2runbenchmarks.R → 3benchtex.pl

tbl:benchmark-monthly in `beta-exhibits.tex`*, July 16, 2019*

Table 10: Benchmark Performance Comparisons, Monthly-Return-Frequency Market-Betas ($N = 1,715,377$)

	Mean	SD	(One-Period-Ahead) Dependent	(Lagged) Independent	g_0	g_1	$R^2_{(\%)}$	rmse
A[6]	1.07	0.949	bmols	bmols	0.61	0.44	1.6	**0.955**
				bols	0.87	0.25	2.7	1.078
				bswa	0.77	0.37	**3.3**	1.013
K	1.04	0.504	bmvck	bmvck	0.46	0.57	3.7	**0.500**
				bols	0.90	0.18	4.8	0.745
				bswa	0.82	0.27	**6.1**	0.639

Explanations: This table explores the subset of months in which market-betas could be computed using monthly stock rates of return. The dependent variable (future monthly return based betas) had to be led by 60 months, because the independent variable was computed over 60 months. bmols is the plain OLS market-beta, bmvck its Vasicek equivalent. (bols and bswa continue to use daily-frequency stock returns.)

Interpretation: Monthly-return market-betas are almost unpredictable noise. Nevertheless, a researcher who is interested in a (bias-adjusted) future monthly-frequency market-beta should still use the daily-frequency-based bswa today.

Source: benchmarks/1mkbswall.R, mbols.csv.gz, dacheng-sahalia/dacheng-sahalia.csv.gz, simin/simin.csv.gz, frazzini-pedersen.csv.gz, monthlies/mobetawin60.csv.gz, monthlies/mobetawin12.csv.gz → dall.csv.gz, benchmarks/2runbenchmarks.R → 3benchtex.pl

Table 11: Distinctiveness of Market-Beta Measures

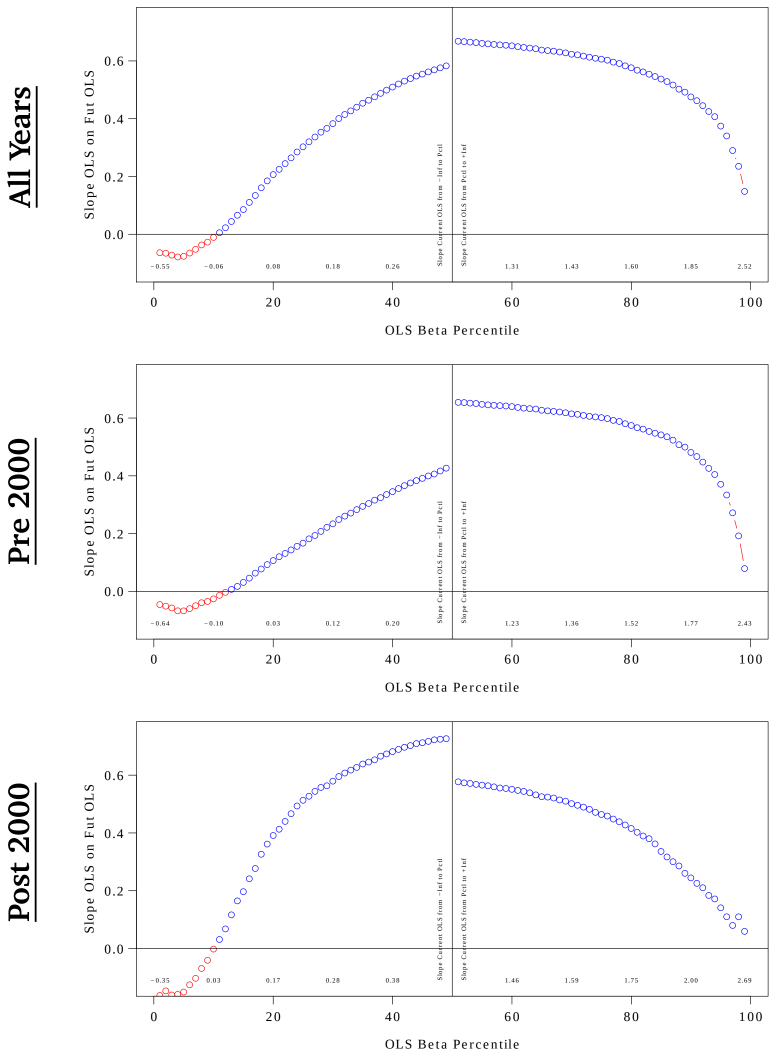
	"Null"	"Base"		"Robust"								"Monthlies"		"Intraday"	
	$\overline{\text{bols}}$	bols	bVCK	blw	bbw	bsw	bswa	bmm	blts	bdim	bfp	bmols	bmvck	btaq1	btaq12
$\overline{\text{bols}}$		0.58	0.52	0.49	0.41	0.48	0.47	0.55	0.55	0.67	0.34	0.33	0.26	0.81	0.53
bols	0.58		0.08	0.19	0.24	0.15	0.20	0.12	0.21	0.34	0.39	0.56	0.55	0.67	0.24
bVCK	0.52	0.08		0.16	0.18	0.09	0.15	0.12	0.19	0.35	0.34	0.51	0.50	0.66	0.23
blw	0.49	0.19	0.16		0.17	0.15	0.19	0.16	0.19	0.42	0.36	0.53	0.50	0.67	0.25
bbw	0.41	0.24	0.18	0.17		0.13	0.15	0.20	0.23	0.41	0.26	0.42	0.39	0.67	0.26
bsw	0.48	0.15	0.09	0.15	0.13		0.10	0.12	0.18	0.38	0.31	0.48	0.46	0.66	0.24
bswa	0.47	0.20	0.15	0.19	0.15	0.10		0.17	0.21	0.40	0.29	0.46	0.44	0.64	0.25
bmm	0.55	0.12	0.12	0.16	0.20	0.12	0.17		0.10	0.37	0.37	0.55	0.54	0.66	0.22
blts	0.55	0.21	0.19	0.19	0.23	0.18	0.21	0.10		0.41	0.39	0.57	0.55	0.67	0.25
bdim	0.67	0.34	0.35	0.42	0.41	0.38	0.40	0.37	0.41		0.48	0.62	0.62	0.75	0.43
bfp	0.34	0.39	0.34	0.36	0.26	0.31	0.29	0.37	0.39	0.48		0.33	0.29	0.72	0.37
bmols	0.33	0.56	0.51	0.53	0.42	0.48	0.46	0.55	0.57	0.62	0.33		0.13	0.81	0.53
bmvck	0.26	0.55	0.50	0.50	0.39	0.46	0.44	0.54	0.55	0.62	0.29	0.13		0.80	0.51
btaq1	0.81	0.67	0.66	0.67	0.67	0.66	0.64	0.66	0.67	0.75	0.72	0.81	0.80		0.64
btaq12	0.53	0.24	0.23	0.25	0.26	0.24	0.25	0.22	0.25	0.43	0.37	0.53	0.51	0.64	
Average	*0.46*	*0.30*	*0.27*	*0.29*	*0.27*	*0.26*	*0.27*	*0.28*	*0.31*	*0.44*	*0.35*	*0.45*	*0.44*	*0.66*	*0.33*

Explanations: This table is based on the sample of 432,426 firm-months with complete information on all market-beta estimates. Beta estimation windows can overlap. Variables are defined in Table 1, the sample in Section I.

Interpretation: The bswa estimate is most similar to bsw. All robust estimates (bsw,blw,bbw) and the Vasicek estimates (bVCK) are similar, with mutual RMSE differences of 0.2 or less. The constant ($\overline{\text{bols}}$), dimson (bdim), intraday (btaq1), and monthly estimates (bmols, bmvck) are all quite different from the average and from one another. The monthly-return-frequency bmols is far from its daily equivalent bols. (Any betas using daily returns over 12 months are mutually closer.)

Source: benchmarks/1mkbswall.R, mbols.csv.gz, dacheng-sahalia/dacheng-sahalia.csv.gz, simin/simin.csv.gz, frazzini-pedersen.csv.gz, monthlies/mobetawin60.csv.gz, monthlies/mobetawin12.csv.gz → dall.csv.gz, benchmarks/2runbenchmarks.R → 3benchtex.pl

Figure 12: Appendix: Predictive Slopes, Future OLS Market-Betas, By Current OLS Market-Beta Percentiles



Explanations: These are slopes from regressions predicting the future OLS market-beta with the current OLS market-beta in the set of betas that are above or below a given percentile (within each year). For example, a regression predicting future OLS betas (bols) with all stocks binned into the 60th percentile or higher had a slope estimate of about 0.6.

Interpretation: Firm-years with beta estimates below zero no longer monotonically predicted lower future OLS betas. The relationship between current and future market-beta estimates was just about zero for current market-beta estimates of three or higher.

Source: fig-ols2fut/plotols2fut.Rmd