LONG-RUN RISK-RETURN TRADE-OFFS*

Federico M. Bandi† Benoît Perron‡

First draft: *November 2006*
This draft: *June 2007*

Abstract

Excess market returns are correlated with *past* market variance. This dependence is statistically mild at short horizons (thereby leading to a hard-to-detect risk-return trade-off, as in the existing literature) but increases with the horizon and is strong in the long run (i.e., between 6 and 10 years). From an econometric standpoint, we find that the long-run predictive power of past market variance is robust to the statistical properties of long-horizon stock-return predictive regressions. From an economic standpoint, we show that, when conditioning on past market variance, conditional versions of the traditional CAPM and consumption-CAPM yield considerably smaller cross-sectional pricing errors than their unconditional counterparts.

JEL Classification: G12.
Keywords: Stock-return predictability, past market variance, long run.

*We thank Ron Gallant (the Editor), an anonymous referee, Alan Bester, Tim Bollerslev, Laurent Calvet, Tom McCurdy, Walter Distaso, Jeff Russell, Ross Valkanov, and seminar participants at the CIREQ time series conference (Montreal, December 8-9, 2006), New York Camp Econometrics II (Saratoga Springs NY, March 23-26, 2007), the Imperial College Financial Econometrics Conference (London, May 18-19, 2007) and the University of Chicago (finance lunch) for useful comments and discussions.

†*Graduate School of Business, The University of Chicago. Address: 5807 South Woodlawn Avenue, Chicago, IL, 60637, USA. Tel.: (773) 834-4352. E-mail: federico.bandi@gsb.uchicago.edu.* Financial support from the William S. Fishman Faculty Research Fund at the Graduate School of Business of the University of Chicago is gratefully acknowledged.

‡ *Dépt. de sciences économiques, Université de Montréal, CIREQ and CIRANO, C.P. 6128, Succ. centre-ville, Montréal, Québec, H3C 3J7, Canada. Tel. (514) 343-2126. E-mail: benoit.perron@umontreal.ca.* Financial assitance from FQRSC, SSHRC, and MITACS is gratefully acknowledged.

1 Introduction

Past long-run market variance is highly correlated with future long-run risk premia. While the dependence between past market variance and excess market returns is mild in the short run (thereby leading to a hard-to-detect classical risk-return trade-off, as reported in the existing literature), it increases with the horizon and is stronger in the long run. Consider regressions of the type

$$R_{t,t+h} = \alpha_h + \beta_h \sigma^2_{t-h,t} + \varepsilon_{t,t+h}, \tag{1}$$

where $R_{t,t+h}$ denotes excess market returns between months t and $t+h$, $\sigma^2_{t-h,t}$ denotes past market variance, and $\varepsilon_{t,t+h}$ is a forecast error. Assume $h = 1, ..., 120$ (1 month to 10 years). Using conventional (Newey-West style) methods of inference, we find a strongly significant correlation between $R_{t,t+h}$ and $\sigma^2_{t-h,t}$ for values of h equal to 72, 84, 96, 108, and 120 (6 to 10 years). In this range, the coefficients of determination $\left(R^2s\right)$ are between about 26% and 73%. This finding contrasts sharply with the short-horizon results. When focusing on horizons between one month ($h = 1$) and 4 years ($h = 48$), the corresponding R^2s are never larger than 1%. We show that the use of alternative inferential methods providing more accurate representations of the finite sample distributions of the relevant test statistics under the null of no dependence mildly mitigates, but by no means eliminates, the statistical significance of the reported long-run relations.[1]

The strong correlation between long-run excess market returns and past market variance is suggestive of an important correlation between past market variance and sources of time-variation in long-run risk premia.[2] Interestingly, we find that, in the long run, past market variance is a stronger predictor of excess market returns than both the classical dividend yield and the consumption-to-wealth ratio recently proposed by Lettau and Ludvigson (2001a). While the predictive ability of these variables is generally strongest at business cycle frequencies (mostly in the case of the consumption-to-wealth ratio), we show that the predictive ability of past market variance largely increases with the prediction horizon.

[1] Boudoukh et al. (2005) and Valkanov (2003), among others, have argued against the validity of standard econometric inference in long-run predictive regressions in finance. We accomodate their criticisms in what follows.

[2] A growing, recent literature in finance has studied the cross-sectional pricing implications of financial cash flows exposed to long-run macroeconomic risk (see, e.g., Bansal and Yaron, 2004, and Hansen et al., 2005). Here we focus on long-run market risk premia. We differ the reader to Section 9 for cross-sectional pricing results solely aimed at providing support for the predictive ability of past long-run market variance.

To further evaluate the economic relevance of the dependence between past market variance and excess market returns we provide an economic metric which complements our econometric findings. If past market variance tracks changes in expected excess market returns in an economically significant fashion, conditional or *scaled* versions of the classical CAPM and consumption-CAPM (C-CAPM) should deliver significantly smaller long-run pricing errors than their unconditional counterparts when conditioning on past market variance. While the literature has tested the validity of conditional versions of classical asset pricing models *given* predictors with a clear economic interpretation as risk proxies (see, e.g., Cochrane, 1996, and Lettau and Ludvigson, 2001b), we in effect evaluate the validity of past market variance as a risk proxy *given* the model. Using the 25 Fama-French size- and value-sorted portfolios we find that conditioning on past market variance translates into drastically smaller differences between cross-sectional long-run realized average returns and long-run average returns implied by the model(s). As an example, the difference in R^2 values between the classical CAPM and its conditional (on past market variance) version is striking. For aggregation levels $h = 84$, 96, 108, and 120 (7 to 10 years) the former delivers R^2 values equal to 11.4%, 3.4%, 1.3%, and .3% while the corresponding values for the latter are equal to 49.5%, 17.8%, 41.1%, and 61.6%.

Much recent work has been devoted to assessing the validity of the classical risk-return trade-off, namely the relation between short-run conditional expected excess returns on the market and the market's conditional variance.[3] The long-run implications of traditional short-term risk-return models have hardly been explored. Importantly, simple aggregation of short-term risk-return models under a classical (autoregressive) process for variance *cannot* imply our results. On the one hand, they would yield estimated regression slopes of long-run excess returns on *past* long-run variance which are decreasing with the level of aggregation. On the other hand, they would deliver estimated regression slopes of long-run excess returns on *contemporaneous* long-run variances that are (statistically) consistent for the true short-term trade-offs. We show that the regression slopes of long-run excess returns on past long-run variance have a tendency to increase with the level of aggregation. We also show that the

[3] The findings are mixed. Baillie and De Gennaro (1990), French et al. (1987), and Campbell and Hentschel (1992), for example, find a positive, but largely insignificant, relation. The results of Campbell (1987) and Nelson (1991), among others, point to a significantly negative relation. Glosten et al. (1989) and Turner et al. (1989), *inter alia*, report either a positive or a negative relation depending on the model used. Harrison and Zhang (1999), Scruggs (1998), Ghysels et al. (2005), Guo and Whitelaw (2006), Lundbland (2005), Maheu and McCurdy (2006), and Pástor et al. (2006), among others, find a risk-return trade-off. Interestingly, Harrison and Zhang (1997) also find a stronger relation between their estimated conditional excess market returns and conditional market variances at the longest horizons they consider (1 and 2 years). In recent work, Bollerslev and Zhou (2007) document a positive short-term relation between excess market returns and past variance risk premia (estimated as the difference between implied and realized variance measures).

regression slopes of long-run excess returns on *contemporaneous* long-run variance are neither significant nor do they converge to economically reasonable parameter values (corresponding, for example, to meaningful coefficients of relative risk aversion).

One could possibly reconcile these findings by invoking the misspecification of (short-run and, by aggregation, long-run) pricing models which do not allow for richer excess market return dynamics than those driven solely by time-varying market variance. However, the sense in which allowing for meaningful changes in investment opportunities to affect excess market returns (along the lines of Scruggs (1998) and Guo and Whitelaw (2006), among others) would yield our empirical results remains unclear. More importantly for our purposes (but in the same vein), our time-series and cross-sectional pricing results indicate that past long-run market variance has the potential to be an important proxy for a broader notion of long-run macroeconomic risk than time-varying variance risk. Understanding the nature of long-run market risk, and the economic channel through which past market variance proxies for it, are important challenges for future work.

The paper is structured as follows. Section 2 introduces the variance estimator in a fairly general continuous-time setting. Section 3 presents the data. Section 4 discusses the main empirical finding, i.e., the long-run correlation between excess market returns and past market variance. Section 5 provides simulations, a representation of the finite sample distributions of the relevant statistics under the null of no dependence, and two methods of robust inference. Section 6 further evaluates the robustness of our findings. Section 7 discusses the implications of our aggregation results for the classical short-term risk-return trade-off. In Section 8 we compare the long-run predictive ability of past variance to that of the dividend yield and of the consumption-to-wealth ratio. Section 9 further analyses the economic significance of the predictive ability of past market variance in the context of conventional cross-sectional asset pricing models. Section 10 concludes.

2 Variance estimator

We use *realized variance* to identify sample path variation in observed market returns. This estimator has a long history in finance. French et al. (1987), for instance, use it in the study of the risk-return trade-off at the monthly level ($h = 1$).

Consider a generic month t with n_t trading days. Denote by $r_{t+\frac{j}{n_t}}$ the j-th daily continuously-compounded return in month t. Realized variance in month t is given by

$$\sigma^2_{t,t+1} = \sum_{j=1}^{n_t} r^2_{t+\frac{j}{n_t}},$$

i.e., the sum of the squared daily returns over the period. When looking at horizon $h > 1$, the estimator is simply defined as

$$\sigma^2_{t,t+h} = \sum_{i=1}^{h} \sigma^2_{t+i-1,t+i}. \tag{2}$$

It is well-known that, under assumptions, $\sigma^2_{t,t+h}$ provides a consistent estimate of (increments in) the quadratic variation of the logarithmic price process in asymptotic designs allowing for $n_t \uparrow \infty$ for all t (i.e., as the number of observations in each month increases asymptotically without bound). For instance, assume the logarithmic price process is expressed as $\log p_t = \Phi^f_t + \Phi^l_t + \Phi^j_t$, where Φ^f_t is a continuous finite variation component, $\Phi^l_t = \int_0^t \sigma_s dW_s$ is a local martingale driven by Brownian shocks dW_t, $\Phi^j_t = \int_0^t \left(J_s dZ_s - \mu_j \lambda_s ds\right)$ is a compensated jump process with Z_t denoting a counting process with finite intensity λ_t, and J_t is a random jump size with mean μ_j and variance σ^2_j. Furthermore, assume the stochastic volatility process σ_s is càdlàg. This specification readily accommodates small and large shocks in the price's sample path as well as fairly unrestricted spot volatility dynamics. The quadratic variation of $\log p_t$ between t and $t + h$ is

$$[\log p]_{t,t+h} = [\log p]_{t+h} - [\log p]_t = \int_t^{t+h} \sigma^2_s ds + \sum_{t \leq s \leq t+h} (\log(p_s) - \log(p_{s-}))^2, \tag{3}$$

where $\log(p_{s-}) = \lim_{\eta \downarrow 0} p_{s-\eta}$, and is made up of two components, one associated with variation in the local martingale and one deriving from the presence of infrequent jumps in the sample path. Under this assumed model, the quantity $\sigma^2_{t,t+h}$ estimates $[\log p]_{t,t+h}$ consistently (for all h values) as $n_t \uparrow \infty$ for all t. Andersen et al. (2003) and Barndorff-Nielsen and Shephard (2002) have recently provided empirical and theoretical justifications for the use of realized variance in the presence of high-frequency asset price data under similar assumptions. As is customary in low frequency applications in finance, we do not take the asymptotics literally. Nevertheless, our use of daily data in the computation of long-run realized variance is bound to capture important variation in the market return's sample path.

To relate our measure more closely to traditional variance measures by allowing for residual autocorrelations left in the daily returns (due to a non-vanishing, in a finite sample, predictable component $d\Phi_t^f$), in Section 6 we also consider a HAC (Bartlett-type) modification of the monthly estimate $\sigma^2_{t,t+1}$, namely

$$\sigma^2_{t,t+1} = \sum_{j=1}^{n_t} r^2_{t+\frac{j}{n_t}} + 2\sum_{i=1}^{k}\left(1-\frac{i}{k+1}\right)\sum_{j=1}^{n_t-k} r_{t+\frac{j}{n_t}} r_{t+\frac{j+i}{n_t}} \tag{4}$$

for different choices of the number (k) of (realized) autocovariances. The original estimate obtains for $k = 0$. When running long-horizon regressions, we compute similar estimates for each aggregate period, i.e.,

$$\sigma^2_{t,t+h} = \sum_{j=1}^{n_{t,t+h}} r^2_{t+\frac{j}{n_{t,t+h}}} + 2\sum_{i=1}^{k}\left(1-\frac{i}{k+1}\right)\sum_{j=1}^{n_{t,t+h}-k} r_{t+\frac{j}{n_{t,t+h}}} r_{t+\frac{j+i}{n_{t,t+h}}}, \tag{5}$$

where $n_{t,t+h}$ is the total number of daily returns between month t and month $t+h$ and $n_{t,t+h} = \sum_{s=0}^{h-1} n_{t+s}$.

3 Data

We use the NYSE/Amex value-weighted index with dividends as our market proxy. The risk-free rate is the 30-day T-bill rate. The data are downloaded from CRSP for the post-war January 2, 1952 - December 29, 2006 period. To compute monthly continuously-compounded excess returns R we aggregate daily continuously-compounded excess returns $r - r^f$ by defining

$$R_{t,t+1} = \sum_{j=1}^{n_t}\left(r_{t+\frac{j}{n_t}} - r^f_{t+\frac{j}{n_t}}\right).$$

Continuously-compounded returns over $(t, t+h)$ are therefore expressed as

$$R_{t,t+h} = \sum_{i=1}^{h} R_{t+i-1,t+i}.$$

Table 1 presents descriptive statistics for all of the variables used in the paper. The first two columns contain empirical moments of the monthly excess returns and realized variances. As typically found in the literature, the monthly variances are considerably more skewed and fat-tailed than the monthly excess returns.

In what follows, we compare the predictive ability of past market variance to that of variables routinely used in forecasting excess market returns. The first variable is the dividend yield (d/p) obtained by subtracting logarithmic returns without dividends from logarithmic returns with dividends before averaging over the previous year to remove seasonal effects. The dividend yield is highly persistent and skewed to the left, but its tails are similar to those of a normal random variable. The second variable is Lettau and Ludvingson's consumption-to-wealth ratio (*cay*) (Lettau and Ludvigson, 2001a) derived as the residuals from a cointegrating regression of logarithmic consumption on logarithmic asset wealth and logarithmic labor income.[4] This variable is not available at the monthly frequency. We therefore report descriptive statistics for quarterly data (available through the end of 2005). *Cay* is not very skewed, has thin tails, and is also fairly persistent, albeit less than *d/p*.

Finally, the paper presents cross-sectional evidence about the predictive ability of past market variance in the context of conditional CAPM and C-CAPM models. We use the 25 size- and value-sorted portfolios of Fama and French (1996) as the relevant test assets. The monthly returns on these portfolios are updated annually and can be downloaded from Ken French's web site.[5] In the C-CAPM we use monthly growth of real per-capita consumption of non-durable goods and services. The consumption and population data are from the Bureau of Economic Analysis. The consumption data starts in February 1959. Monthly consumption growth is not highly skewed, has fairly thin tails, and low persistence.

4 Long-run risk-return trade-offs

To illustrate our findings, Figure 1 reports scatter plots of excess market returns $R_{t,t+h}$ and past market variance $\sigma^2_{t-h,t}$ at four levels of aggregation, namely $h = 1$, 12, 60, and 120. At the monthly frequency, the correlation between excess returns and past variance (i.e., a traditional form of the risk-return trade-off) is unclear and certainly not revealed by the use of conventional proxies for conditional expected excess returns and conditional variances, such as realized excess returns and realized past variances. As we increase the level of aggregation, an apparent correlation is revealed.

We provide an initial assessment of the extent of this dependence by running the regression in Eq. (1) for values of h between 1 (one month) and 120 (10 years). The estimated slopes, standard errors, and R^2s are reported in Table 2. In Table 2, and in all other tables below, the notation $h = 3$, for instance, signifies use of overlapping quarterly data. We correct the standard errors for the serial correlation induced by the overlapping nature of the data by using a kernel variance estimator with a quadratic spectral kernel, pre-whitening, and a bandwidth selected according to Andrews' (1991) data-based rule.

[4] We obtain the series from Sydney Ludvigson's web site: http://www.econ.nyu.edu/user/ludvigsons/

[5] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/

At the monthly level ($h = 1$), our results perfectly mirror the results of French et al. (1987). We find a slope coefficient equal to $-.79$ and insignificant. French et al. (1987) find a statistically insignificant coefficient equal to $-.349$ using data from 1928 to 1984. Except for the quarterly frequency, we obtain a positive, and significant, slope coefficient only when aggregating data for 72 months (6 years) and over. Our long-run slope estimates are between about 4 and 6.5. *If* past variance were an accurate predictor of future variance, *then* one might consider these values economically meaningful and consistent with structural interpretations relating the regressions' slopes to risk aversion. However, as we show below, past realized variance is not the best predictor of future variance. The long-run R^2s are large, i.e., between about 26.5% (at 6 years) and about 73% (at 10 years). These preliminary findings point to a substantial long-run dependence between excess market returns and past market variance.

For a clearer assessment, Figure 2 provides a graphical representation of the term structure of estimated slopes and corresponding 95% HAC confidence bands. On the one hand, the presence of a risk-return trade-off is hard to detect at short/medium horizons (1 month to about 5 years). In this range, the slopes' 95% bands easily include zero. This difficulty is well-known and has lead to a search for sophisticated methods of inference in the risk-return literature. On the other hand, the dependence between excess market returns and past market variance is pronounced at longer horizons. Classical statistical methods suffice to reveal it.

4.1 A useful restriction: zero intercept

In this subsection we constrain the intercept (in Eq. (1)) to be zero. This restriction is justifiable based on the insignificance (with only a few exceptions, in the very long run) of the estimated intercept coefficients in the previous regressions.[6] From a statistical standpoint, provided the restriction is true, the slope estimator is still estimated consistently but with increased precision.[7] Lanne and Saikkonen (2006) have recently made a similar point in a GARCH-in-mean model (Engle et al., 1987). Specifically, they have shown that the inclusion of an intercept term can lead to imprecise estimates of the variance-in-mean coefficient.

The results are in Table 3. We find that the restriction increases the statistical significance of our estimates at virtually all horizons. All estimates are positive and, with the exception of the 1-month and 6-month cases, statistically significant at all conventional levels. Importantly, the long-run slopes increase almost monotonically, as do their corresponding t-statistics (from

[6] The restriction would also be consistent with the classical short-term risk-return trade-off.

[7] In a univariate regression model with predetermined regressors x, the variance of the slope estimator goes from $\frac{\sigma_\varepsilon^2}{\sum_{i=1}^n (x_i - \overline{x})^2}$ to $\frac{\sigma_\varepsilon^2}{\sum_{i=1}^n {x_i}^2}$ when imposing the restriction. Clearly, $\sum_{i=1}^n (x_i - \overline{x})^2 < \sum_{i=1}^n {x_i}^2$.

3.7 at the 5-year horizon to 7.1 at the 10-year horizon). The dependence between long-run excess market returns and past long-run market variance is strong, and stronger than in the unrestricted case.

5 Inferential issues

5.1 Simulations

This section evaluates the accuracy of classical and HAC asymptotic inference in our framework. To this extent, we simulate monthly excess returns and an autoregressive variance process under the assumption of no dependence. Subsequently, we aggregate as previously done with data. The simulated process is:

$$R_{t,t+1} = \beta_1 \sigma^2_{t-1,t} + \varepsilon_{t,t+1}, \quad (6)$$

$$\sigma^2_{t,t+1} = \rho_0 + \rho_1 \sigma^2_{t-1,t} + u_{t,t+1}, \quad (7)$$

with $\rho_0 = 0$, $\rho_1 = 0.6$, $\sigma_\varepsilon = 1$, $\sigma_u = 1$, and $\rho_{\varepsilon u} = -0.3$, under $H_0 : \beta_1 = 0$ (i.e., the null of no dependence). The parameter values are meant to replicate the properties of our disaggregated monthly series (more on this later). Consistent with data, we simulate 660 monthly observations. The number of simulated paths is equal to $10,000$.

We run the regression in Eq. (1) for $h = 1, 3, ..., 120$. In addition, we consider regressions for which the intercept α_h is constrained to be 0. We test the null $\beta_h = 0$ for each choice of h at nominal level 5%. As above, we correct the standard errors by using a kernel variance estimator with a quadratic spectral kernel, pre-whitening, and a bandwidth selected according to Andrews' (1991) data-based rule.

The third and fourth row of Table 4, first panel, report the test sizes for the unconstrained intercept case when the standard t-statistic and the HAC statistic are used. As is well-known, the overlapping leads to severe size distortions of standard tests of the null $\beta_h = 0$. HAC corrections (applied earlier with data) reduce these distortions drastically, but the actual size is still beyond the nominal size. With an horizon of 120 months (10 years), for instance, the actual size of the HAC test is about 15% rather than 5%. The results from constraining the intercept are in the second panel of Table 4. The same pattern is observed, but size distortions are slightly less pronounced than in the unconstrained case. At the 10-year horizon, for example, HAC corrections yield a size of 13.4%.

The spuriously increasing slope estimates and R^2s, as well as the size distortions leading to over-rejections of the null of zero slope, are obvious concerns. Despite differences in aggregation, these results are reminiscent of similar findings in the context of classical predictive regressions of long-run returns on persistent (non-aggregated) financial ratios (Valkanov, 2003, and Boudoukh et al., 2005, among others). These behaviors have led to questioning the informational content of long-run regressions about the predictability of stock returns (Boudoukh et al., 2005). The next subsection provides tests of the null of no dependence with improved size properties.

5.2 An alternative asymptotic approximation

Write

$$R_{t,t+1} = \beta_1 \sigma^2_{t-1,t} + \varepsilon_{t,t+1}, \quad (8)$$

$$\sigma^2_{t,t+1} = \rho_0 + \rho_1 \sigma^2_{t-1,t} + u_{t,t+1}, \quad (9)$$

with $\rho_0 = 0$ and $\rho_1 = 1 + \frac{c}{T}$.[8] Assume the vector $[\varepsilon_{t,t+1}, u_{t,t+1}]$ is a vector martingale difference sequence with covariance matrix $[\sigma^2_\varepsilon, \sigma_{\varepsilon u}, \bullet, \sigma^2_u]$. The parameter c is a constant measuring deviations from unity that are decreasing in T. This framework is widely adopted in predictive regressions with persistent regressors (see, e.g., Campbell and Yogo, 2005, Valkanov, 2003, and Bandi, 2004, for a nonlinear approach). In our context, ρ_1 is smaller than in predictive regressions with persistent financial ratios (i.e., the negative parameter c is larger in absolute value). However, we will show by simulations that the convenient local-to-unity approach captures the salient finite sample features of our long-run regressions.[9]

Consider again the regression(s) in Eq. (1). As in Valkanov's asymptotic framework (Valkanov, 2003), we assume that $h = [\lambda T]$, i.e., the portion of the overlap is a constant fraction of the sample size ($[x]$ denotes, as always, the largest interval that is less than or equal to x).

[8] One could argue that the persistence properties of realized variance are more amenable to a fractional integrated process (see, e.g., Bandi and Perron, 2006, and the references therein). We use a near-unit root specification for two reasons. First, in light of the statistical properties of return data, fractional integration in variance would lead to *unbalanced* regressions. In ICAPM-style models, the balance could be restored by choosing suitable covariances between market returns and state variables that are fractionally-cointegrated with market variance. Second, long-run variance (as implied by autoregressive models) is virtually uncorrelated. Allowing for fractional integration in the disaggregated (monthly) series is not likely to change our results in important ways.

[9] Analogously to the case of a widely-used near-unity dividend yield process, technically the model does not restrict variance to remain positive. This is a well-known feature of this modelling choice which does not hinder its usefulness in deriving more accurate asymptotic approximations.

Differently from Valkanov's framework, however, regressor and regressand are aggregated over non-overlapping periods.

We are interested in the behavior of the slope estimates, R^2s, and test sizes for the null $\beta_h = 0$ (no dependence) under Eq. (8) and Eq. (9). Proposition 1 and Proposition 2 contain the relevant asymptotic approximations. Their proofs follow classical embedding methods in the unit-root literature (see, e.g., Phillips, 1991, and Cavanagh et al., 1995, among others).[10] For a thorough discussion of these methods in the context of long-run predictability issues in finance, we refer the reader to Valkanov (2003). In what follows the symbol $\Rightarrow$ denotes weak convergence as $T \to \infty$.

Proposition 1 (*The unrestricted regressions.*)

If the return and variance process follow Eq. (8) and Eq. (9), $\beta_1 = 0$, and the regression in Eq. (1) is run, then

1. $T\widehat{\beta}_{h=[\lambda T]} \Rightarrow \frac{\sigma_\varepsilon}{\sigma_u} \frac{\int_\lambda^{1-\lambda} \overline{W}(s,\lambda)\overline{J}_c(s,-\lambda)ds}{\int_\lambda^{1-\lambda} \overline{J}_c^2(s,-\lambda)ds}$,

2. $\widehat{\alpha}_{h=[\lambda T]}/\sqrt{T} \Rightarrow \frac{\sigma_\varepsilon}{1-2\lambda} \int_\lambda^{1-\lambda} \left(W(s+\lambda) - W(s)\right) ds - \frac{\int_\lambda^{1-\lambda} \overline{W}(s,\lambda)\overline{J}_c(s,-\lambda)ds}{\int_\lambda^{1-\lambda} \overline{J}_c^2(s,-\lambda)ds} \left(\frac{\sigma_\varepsilon}{1-2\lambda} \int_\lambda^{1-\lambda} \left(\int_{t-\lambda}^t J_c(s)ds\right) dt\right)$,

3. $\frac{t_{\widehat{\beta}_{h=[\lambda T]}}}{\sqrt{T}} \Rightarrow \frac{\int_\lambda^{1-\lambda} \overline{W}(s,\lambda)\overline{J}_c(s,-\lambda)ds}{\sqrt{\left(\frac{1}{1-2\lambda}\right)\left(\int_\lambda^{1-\lambda} \overline{W}^2(s,\lambda)ds \int_\lambda^{1-\lambda} \overline{J}_c^2(s,-\lambda)ds - \left(\int_\lambda^{1-\lambda} \overline{W}(s,\lambda)\overline{J}_c(s,-\lambda)ds\right)^2\right)}}$,

4. $R^2_{h=[\lambda T]} \Rightarrow \frac{\left(\int_\lambda^{1-\lambda} \overline{W}(s,\lambda)\overline{J}_c(s,-\lambda)ds\right)^2}{\int_\lambda^{1-\lambda} \overline{J}_c^2(s,-\lambda)ds \int_\lambda^{1-\lambda} \overline{W}^2(s,\lambda)ds}$,

where

$$\overline{W}(t,\lambda) = \{W(t+\lambda) - W(t)\} - \frac{1}{1-2\lambda} \int_\lambda^{1-\lambda} \left(W(s+\lambda) - W(s)\right) ds,$$

and

$$\overline{J}_c(t,-\lambda) = \left\{\int_{t-\lambda}^t J_c(s)ds\right\} - \frac{1}{1-2\lambda} \int_\lambda^{1-\lambda} \left(\int_{t-\lambda}^t J_c(s)ds\right) dt$$

with

$$dJ_c(s) = cJ_c(s)ds + dB(s) \qquad J_c(0) = 0$$

and $\{W(s), B(s)\}$ is a vector of standard Brownian motions with covariance $\frac{\sigma_{\varepsilon u}}{\sigma_\varepsilon \sigma_u}$.

[10] They are available from the authors upon request.

Proposition 2 (***The restricted regressions.***)

If the return and variance process follow Eq. (8) and Eq. (9), $\beta_1 = 0$*, and the regression in Eq. (1) is run with* $\alpha_h = 0$*, then*

1. $T\widehat{\beta}_{h=[\lambda T]} \Rightarrow \frac{\sigma_\varepsilon}{\sigma_u} \frac{\int_\lambda^{1-\lambda} W(s,\lambda) J_c(s,-\lambda) ds}{\int_\lambda^{1-\lambda} J_c^2(s,-\lambda) ds}$,

2. $\frac{t_{\widehat{\beta}_{h=[\lambda T]}}}{\sqrt{T}} \Rightarrow \frac{\int_\lambda^{1-\lambda} W(s,\lambda) J_c(s,-\lambda) ds}{\sqrt{\left(\frac{1}{1-2\lambda}\right)\left(\int_\lambda^{1-\lambda} W^2(s,\lambda) ds \int_\lambda^{1-\lambda} J_c^2(s,-\lambda) ds - \left(\int_\lambda^{1-\lambda} W(s,\lambda) J_c(s,-\lambda) ds\right)^2\right)}}$,

3. $R^2_{h=[\lambda T]} \Rightarrow 1 - \left(\frac{\int_\lambda^{1-\lambda} J_c^2(s,-\lambda) ds \int_\lambda^{1-\lambda} W^2(s,\lambda) ds - \left(\int_\lambda^{1-\lambda} W(s,\lambda) J_c(s,-\lambda) ds\right)^2}{\int_\lambda^{1-\lambda} J_c^2(s,-\lambda) ds \int_\lambda^{1-\lambda} \overline{W}^2(s,\lambda) ds}\right)$,

where

$$W(t, \lambda) = W(t + \lambda) - W(t),$$

and

$$J_c(t, -\lambda) = \int_{t-\lambda}^{t} J_c(s) ds$$

with

$$dJ_c(s) = cJ_c(s)ds + dB(s) \qquad J_c(0) = 0,$$

and $\{W(s), B(s)\}$ *is a vector of standard Brownian motions with covariance* $\frac{\sigma_{\varepsilon u}}{\sigma_\varepsilon \sigma_u}$.

The variance process is embedded in a mean-reverting Ornstein-Uhlenbeck process with mean parameter c. Under the null of no dependence, long-run returns are embedded in a standard Brownian motion. The correlation between the system's shocks makes the limiting OU process and the limiting Brownian motion correlated. The asymptotic distributions are stochastic functionals of these processes (and/or their demeaned versions). Differently from the results in Valkanov (2003), in light of the different aggregation method, the range of integration of the functionals is $(\lambda, 1 - \lambda)$ and the distance between the upper and lower limit of the stochastic integrands $W(t, \lambda)$ and $J_c(t, -\lambda)$ is 2λ.

Despite the autocorrelation of our regressor (market variance) being lower than in classical long-run predictive regressions (and arguably not a near-unit root), the asymptotic approximations capture the qualitative features of the simulations reported in the previous section. Under the null of no dependence, the slope estimator is super-consistent. However, its limiting distribution has a bias that is increasing (in absolute value) with the degree of overlap (i.e., with λ). If $\sigma_{\varepsilon u} < 0$, as in our data, the bias is positive, while it is negative with $\sigma_{\varepsilon u} > 0$. Similarly, the R^2 converges to a random variable whose mean increases with the overlap. Importantly, the standard t-statistic diverges with T, thereby determining likely over-rejections in the classical asymptotic framework. As in Valkanov (2003), we will rely on the pivotal (given the parameters c and $\frac{\sigma_{\varepsilon u}}{\sigma_\varepsilon \sigma_u}$) statistic $\frac{t_{\widehat{\beta}_{h=[\lambda T]}}}{\sqrt{T}}$ to test the null of no dependence. The last rows of Table 4 (first and second panel) report the rejection probabilities when using $\frac{t_{\widehat{\beta}_{h=[\lambda T]}}}{\sqrt{T}}$ in our simulations. The critical values are generated assuming $c = (\rho_1 - 1)T$ with $\rho_1 = 0.6$ and $\sigma^2_{\varepsilon u} = -0.3$. In other words, we assume the parameters are known. In the relevant region of the parameter space, however, we find that the distribution is not very sensitive to these values (this will be evident in Table 5 below). Hence, the results appear to be a good indication of what can be achieved in practice. All rejection probabilities lie between 5% and 3.5% and are therefore close to the nominal size of 5%.

5.3 Revisiting the long-run dependence between excess returns and past variance

Table 5 contains inference based on the $\frac{t_{\widehat{\beta}_{h=[\lambda T]}}}{\sqrt{T}}$ statistic. We report the statistic's values as well as the 5% right-tail critical values for the limiting distributions in Proposition 1 and Proposition 2 obtained using autoregressive parameters equal to 0.2 and 0.6. The former corresponds to the estimated autoregressive parameter of the monthly variance series over the full sample. The latter corresponds to the estimated autoregressive parameter of the monthly series when excluding the 1987 crash.

The statistical significance of the dependence between excess returns and past variance has now decreased, but only slightly. This dependence is still very significant at the 5% level over 7, 8, 9, and 10 years. It is also significant at the 10% level at 6 years. As pointed out earlier, restricting the intercept to be zero (as implied by a classical short-term risk-return trade-off, for example) leads to more significant slope estimates. In this case, we find again a strong long-run dependence at all conventional levels (Table 5, panel 2).

In order to provide further evidence about the significance of our results at long horizons, we also report critical values based on the bootstrap. Our algorithm is similar to the one suggested by Kilian (1999) . It involves resampling the monthly excess returns and past variances using a simple VAR(1) model and then aggregating up the monthly data as done in our empirical analysis. In addition, we impose the null of no dependence and allow for heteroskedasticity through the use of the wild bootstrap. Specifically, we fit the following bivariate VAR(1) model to our data

$$\begin{pmatrix} R_{t,t+1} \\ \sigma^2_{t,t,+1} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} R_{t-1,t} \\ \sigma^2_{t-1,t} \end{pmatrix} + \begin{pmatrix} \varepsilon^1_{t,t+1} \\ \varepsilon^2_{t,t+1} \end{pmatrix}, \tag{10}$$

and generate bootstrap draws by i.i.d. resampling of the residuals $\begin{pmatrix} \hat{\varepsilon}^1_{t,t+1} \\ \hat{\varepsilon}^2_{t,t+1} \end{pmatrix}$. The bootstrap sample is constructed from

$$\begin{pmatrix} R^*_{t,t+1} \\ \sigma^{2*}_{t,t,+1} \end{pmatrix} = \begin{pmatrix} \hat{a}_{11} & 0 \\ \hat{a}_{21} & \hat{a}_{22} \end{pmatrix} \begin{pmatrix} R^*_{t-1,t} \\ \sigma^{2*}_{t-1,t} \end{pmatrix} + z_t \begin{pmatrix} \varepsilon^{1*}_{t,t+1} \\ \varepsilon^{2*}_{t,t+1} \end{pmatrix}, \tag{11}$$

where z_t is an independent random variable with expectation zero and variance 1 to allow for heteroskedasticity (we use a standard normal distribution). We start the recursion at the unconditional expectation of the VAR(1) process. The 0 in the North-East corner of the coefficient matrix imposes the null of no dependence.

Given a bootstrap sample $\left\{R^*_{t,t+1}, \sigma^{2*}_{t,t+1}\right\}_{t=1}^{660}$, we aggregate the data to order h, estimate the regression

$$R^*_{t,t+h} = \alpha_h + \beta_h \sigma^{2*}_{t-h,t} + \varepsilon^*_{t,t+h}, \tag{12}$$

and compute the ratio statistic $\frac{t_{\widehat{\beta}_{h=[\lambda T]}}}{\sqrt{T}}$. We repeat this process 999 times and report the right-tail critical values for an equal-tailed 5% bootstrap test to make results comparable with the previous asymptotic findings. As earlier, without constraining the intercept, the dependence between excess market returns and past market variance is significant at the 5% level at 7, 8, 9, and 10 years. With a constrained intercept, the results are significant at all relevant frequencies, with the sole exception of the shortest monthly frequency.

6 Robustness

6.1 The lag choice

Selecting the same lag h for both excess returns and past variance is natural. This is generally done when testing short-term notions of the risk-return trade-off using realized variance measures (see, e.g., French et al., 1987).[11] This said, it is of interest to re-evaluate our regressions when allowing for h values that differ across regressand and regressor. Table 6 reports parameter estimates, t-statistics, and R^2 values associated with regressions (inclusive of an intercept) of h_r-period returns on h_v-period past variances. The values on the diagonal correspond to

[11] Ghysels, Santa-Clara, and Valkanov (2005) provide an important exception. They have recently tested the short-term risk-return trade-off using variance measures constructed on the basis of a longer window of daily return data than that used to construct the regressand (excess market returns). They have also allowed for unequally-weighted daily squared returns in the construction of their variance estimates.

the values reported in Table 2 for the case $h_r = h_v = h$. As expected, we find that the South-East regions of the tables (those associated with large h_r and h_v values) deliver statistically significant parameter estimates and large R^2 values. For large values of h_r and h_v, statistical significance is maximal right around the diagonal but, in general, not exactly on the diagonal.

6.2 HAC variance

Predictability induces correlation in observed returns. If one acknowledges that the asymptotic arguments leading to the consistency of realized variance (through a drift component vanishing to zero at speed dt) might not be satisfied in practice, one might want to correct the classical realized variance estimator for the correlation in the observed returns. In Table 7 we evaluate the forecasting ability of past market variance by using the estimator in Eq. (5) with different values of the number k of autocovariances. Our conclusions are not altered by these alternative specifications. The statistical significance of the slope estimates decreases mildly at 6 and 7 years. This decrease might be simply due to extra noise (resulting from the addition of statistically insignificant return autocorrelations) in the resulting variance estimates. As always, constraining the intercept to zero reinforces our findings.

6.3 The sample period

Choosing the sample period poses a standard trade-off between robustness and efficiency. The longer the sample, the more accurate the estimates. However, longer samples increase the likelihood of misleading estimates due to potential regime shifts. Having made this point, while it seems natural to use post-war data to analyze the dependence between past market variance and market risk premia as done previously, learning about long-run behavior with 40 years of monthly observations can be a complicated task. Here we re-assess the robustness of our findings by considering the longer 1934-2006 sample. Table 8 contains the corresponding results. With an unconstrained intercept, the inclusion of pre-war data decreases the numerical value of the slope estimates. Statistical significance is preserved in the long-run with a slight decrease at 6 years and, when employing the statistic $\frac{t_{\widehat{\beta}_{h=[\lambda T]}}}{\sqrt{T}}$ with asymptotic and bootstrap critical values, at 7 years. Constraining the intercept to zero results, as always, in more significant slope estimates.

7 Implications for the classical risk-return trade-off

Martingale prediction, i.e., the use of past market variance to predict future market variance, is of course only justifiable when variance is highly persistent. Table 9 presents the variance first-order autocorrelations at all levels of aggregation. We report estimates from a linear regression of realized variance on itself h periods in the past. In other words, we use non-overlapping realized variances at each aggregation level. Variance is highly positive dependent at short horizons. This is well-known. The relatively low autocorrelation coefficient (0.2) for the 1-month variances should not be surprising. It is simply a by-product of the 1987 crash. If we confined ourselves to pre-crash data (or if we used the extended sample in the previous section), we would find an autocorrelation value equal to about 0.6 for $h = 1$. Using only post-crash data yields an autocorrelation value equal to about 0.5.[12] As implied by standard autoregressive processes, the autocorrelations become quickly statistically insignificant with the level of aggregation. Realized variance is virtually uncorrelated in the long run.

This result deserves particular attention in our framework. Even if one were to neglect changes in investment opportunities (or, econometrically, covariances of asset returns with business cycle proxies) as sometimes done in the literature,[13] testing long-run notions of the classical risk-return trade-off using past long-run market variance is certainly inappropriate. As pointed out above, past long-run variance is likely not the best predictor of future long-run variance.

While our regressions should *not* be viewed as applications to the long-run of tests generally applied in the short-run, it is natural to ask whether our long-run results can be compatible with classical short-term risk-return trade-offs. To this extent, we ask the question: Are the reported long-run results compatible with a disaggregated risk-return model with $\beta_1 = \beta$ as in Eq. (8) and Eq. (9)? We argue that they are likely not. Proposition 3 discusses the asymptotic properties of the long-run regressions' slope estimates and R^2s when the disaggregated data-generating process implies a short-term trade-off between monthly returns and past monthly variance.

---

[12] Since the overall dynamics of the process conform more nicely with an autoregressive process with first-order autocorrelation equal to 0.6 than with an autoregressive process with first-order autocorrelation equal to 0.2 (induced by 1987 crash), the simulations (in Section 5) use the former value. However, the asymptotic critical values of the ratio statistic are obtained for *both* autocorrelation levels. Hence, our tests are robust to this choice.

[13] Recent papers that have forcefully emphasized the importance of accounting for changes in investment opportunities are Scruggs (1998) and Guo and Whitelaw (2006), among others.

Proposition 3

If the return and variance process follow Eq. (8) and Eq. (9), $\beta_1 = \beta$, and the regression in Eq. (1) is run with $\alpha_h = 0$, then

1. $\widehat{\beta}_{h=[\lambda T]} \Rightarrow \beta \frac{\int_\lambda^{1-\lambda} J_c(s,\lambda) J_c(s,-\lambda) ds}{\int_\lambda^{1-\lambda} J_c^2(s,-\lambda) ds}$,
2. $R^2_{h=[\lambda T]} \Rightarrow 1 - \left(\frac{\int_\lambda^{1-\lambda} J_c^2(s,\lambda) ds \int_\lambda^{1-\lambda} J_c^2(s,-\lambda) ds - \left(\int_\lambda^{1-\lambda} J_c(s,\lambda) J_c(s,-\lambda) ds \right)^2}{\int_\lambda^{1-\lambda} J_c^2(s,-\lambda) ds \int_\lambda^{1-\lambda} \overline{J}^2(s,\lambda) ds} \right)$,

where

$$J_c(t,-\lambda) = \int_{t-\lambda}^{t} J_c(s) ds$$

with

$$dJ_c(s) = cJ_c(s)ds + dB(s) \qquad J_c(0) = 0$$

and $B(s)$ is a standard Brownian motion.

The proposition implies that $\widehat{\beta}_{h=[\lambda T]}$ does not estimate β consistently. Neglecting asymptotic embedding arguments as in the proposition, simple aggregation of the model in Eq. (8) and Eq. (9) makes this statement obvious since, for a degree of overlap λ, the true slope coefficient is $\beta\rho_1^{[\lambda T]}$. Naturally, $\beta\rho_1^{[\lambda T]}$ becomes smaller and smaller with the level of aggregation. Interestingly, given our assumed (data-based) parameter values, the limiting distribution of $\widehat{\beta}_{h=[\lambda T]}$ has a negative bias which increases with λ. Notice that, in the data, the long-run slopes fail to decrease with the degree of overlap. Similarly, Proposition 3 implies that, for a large λ and our assumed parameter values, the limiting distribution of the R^2 should be more concentrated around zero. This is again contrary to our findings. Taken jointly, these observations imply that the disaggregated model in Eq. (8) and Eq. (9), and the type of temporal aggregation that would derive from it, are not supported by our data.

To reinforce this conclusion, we consider the case of *contemporaneous* aggregation under the null of a short-term trade-off, i.e., $\beta_1 = \beta$. In other words, we regress $\sum_{i=1}^{h} R_{t+i-1,t+i}$ on $\sum_{i=1}^{h} \sigma^2_{t+i-2,t+i-1}$. Theoretically, it is simple to show that in this case the slope estimator converges to β at speed T. Table 10 reports parameter estimates, t-statistics, and R^2 values associated with regressions (inclusive of an intercept) of $\sum_{i=1}^{h_r} R_{t+i-1,t+i}$ on $\sum_{i=1}^{h_v} \sigma^2_{t+i-2,t+i-1}$ (for potentially different values of h_r and h_v). We find that the relation between returns and "contemporaneous" variance is mild, and particularly mild for large h_r and h_v values.

The statistical and economic significance of the slope estimates does not increase with the aggregation level, as implied by theory under the null. In other words, the estimates do not converge to values that might be associated with meaningful levels of relative risk aversion.

In sum, our aggregation results are inconsistent with a classical notion of the short-term risk-return trade-off. Differently put, disaggregated asset pricing models which *solely* imply dependence between excess market returns and (autoregressive) conditional variance *can not* deliver our results upon aggregation. Arguably, the most widely employed time-series model implying dependence between conditional excess market returns and conditional variance is the classical GARCH-in-mean model. If we fit a GARCH-in-mean model to our data,[14] simulate from it, and aggregate as we do with data, we experience a decreasing (with the level of aggregation) pattern in the estimated regression slopes. This pattern is of course consistent with the implications of Proposition 3.

While conditional excess market returns can not solely depend on autoregressive conditional variance, the sense in which allowing for meaningful changes in investment opportunities (in the determination of market risk premia) could reconcile our empirical findings with theory ought to be determined. More generally, the pricing of short- and long-run market risk, and the economic channel through which past market variance proxies for the later, are fundamental issues for future work.

We now provide further time-series and cross-sectional evidence about the long-run predictive ability of past market variance.

8 Past market variance vs. predictors

Numerous variables have been found to display important degrees of correlation with long-run excess returns. Financial ratios are a well-known example (see, e.g., the discussion in Cochrane, 2001). This said, past market variance is admittedly a less natural predictor than predictors with prices in them, like the dividend yield (d/p). Let us focus on the price level (the denominator in d/p and other financial ratios). Prices are low (high d/p) when agents expect/demand higher returns. Prices are low, and expected returns are high, in recessions.

[14] The estimated model for our monthly data is:

$$\begin{aligned} R_{t,t+1} &= \underset{(3.9)}{3.61} h_t + \varepsilon_t, \\ \varepsilon_t &= \sqrt{h_t} u_t, \\ h_t &= \underset{(2.24)}{0.000082} + \underset{(24.67)}{0.86}\, h_{t-1} + \underset{(3.06)}{0.08} \varepsilon_{t-1}^2. \end{aligned}$$

Moving forward, low prices predict higher expected and realized returns (in expansions). In other words, prices are bound to go back up eventually. Alternatively, when taking the full d/p ratio seriously, classical present-value identities (see, e.g., Campbell and Shiller, 1998) imply that d/p ought to predict expected excess returns if it does not predict dividend growth, and it does not (Cochrane, 2006, thoroughly exploits this simple intuition).

In the case of past market variance, despite the documented higher volatility in recessions, the link between predictor and future long-run excess returns seems somewhat less mechanical. Yet, in the long run the forecasting ability of past variance is superior to that of d/p. Table 11 and Table 12 contain bi-variate regressions of long-run returns on past variance and d/p. We consider both sample periods examined earlier. As said, the dividend yield is obtained by subtracting logarithmic returns without dividends from logarithmic returns with dividends as in Cochrane (2006), for example.[15] For the longer sample, the regression results have a very familiar look.[16] The slopes associated with d/p largely increase with the horizon. This finding is consistent with dividend yield predictability. In particular, one can easily see that the result would be delivered by simple aggregation of a linear one-period model for expected excess market returns driven by a persistent dividend yield process. Importantly, the statistical properties of long-run predictive regressions make the result somewhat consistent with the null of *no* predictability as well (see, e.g., Valkanov, 2003, and Boudoukh et al., 2005). Regardless of whether one believes in dividend yield predictability or not, the sole point of these regressions is to illustrate that the long-run explanatory power of past market variance is hardly affected by the inclusion of a well-known, alternative predictor. Importantly, when controlling for past market variance, the incremental long-run return variation explained by d/p is rather minimal.

Lettau and Ludvigson (2001a) have recently advocated the consumption-to-wealth ratio (*cay*) as a measure capable of summarizing expectations of future excess returns to the market portfolio. The idea is that high consumption levels should be associated with expectations of high returns on wealth (or low consumption growth rates). Importantly, *cay* has been shown to have higher explanatory power for excess market returns than d/p. When running bivariate regressions of long-run excess market returns on past market variance and *cay* we again find that the long-run explanatory power of past market variance is hardly affected by the inclusion of *cay*. *Cay* has superior predictive ability at higher (business cycle) frequencies.

[15] Alternative dividend yield definitions have been used in the literature. Lewellen (2004), for example, employs the dividends paid over the prior year divided by the current value of the index. Our results (available upon request) do not change when using this alternative measure.

[16] The insignificance of the d/p slope estimates over the shorter sample starting in 1952 has been reported by other authors as well (see, e.g., Lettau and Ludvingson, 2003, Table 2).

The next session provides an important economic lens through which we further analyze the long-run predictive ability of past market variance.

9 An economic metric: cross-sectional asset pricing

If one believes in conditional asset pricing models, the use of conditioning variables should lead to "small" pricing errors *provided* the conditioning variables are suitable proxies for time-varying risk premia. This section evaluates the size of the pricing errors of the 25 Fama-French size- and value-sorted portfolios when conditioning on past market variance in the context of conditional versions of the classical CAPM and C-CAPM.

Write the fundamental pricing equation as

$$1 = E_t[M_{t,t+h}(1 + R^i_{t,t+h})],$$

where E_t denotes expectations conditional on time t information and $M_{t,t+h}$ is a stochastic discount factor. In the case of the CAPM, assume a conditional linear factor model $M_{t,t+h} = \theta^0_{t(h)} + \theta^1_{t(h)} R^M_{t,t+h}$, where $R^M_{t,t+h}$ denotes the h-period return on the market. In the case of the C-CAPM, assume $M_{t,t+h} = \theta^0_{t(h)} + \theta^1_{t(h)} \Delta c_{t,t+h}$, where $\Delta c_{t,t+h}$ is consumption growth between t and $t+h$.[17] If $\theta^0_{t(h)} = \delta_0 + \delta_1 \sigma^2_{t,t-h}$ and $\theta^1_{t(h)} = \vartheta_0 + \vartheta_1 \sigma^2_{t,t-h}$, the two specifications imply the unconditional models

$$1 = E[\left(\delta_0 + \delta_1 \sigma^2_{t,t-h} + \vartheta_0 R^M_{t,t+h} + \vartheta_1 \left(\sigma^2_{t,t-h} \times R^M_{t,t+h}\right)\right)(1 + R^i_{t,t+h})],$$

and

$$1 = E[\left(\delta_0 + \delta_1 \sigma^2_{t,t-h} + \vartheta_0 \Delta c_{t,t+h} + \vartheta_1 \left(\sigma^2_{t,t-h} \times \Delta c_{t,t+h}\right)\right)(1 + R^i_{t,t+h})],$$

respectively. Consistently, we run cross-sectional regressions of the form

$$\left(\frac{1}{T-h} \sum_{t=1}^{T-h} R^i_{t,t+h}\right) = \alpha + \lambda' \widehat{\boldsymbol{\beta}}^i + \varepsilon_i,$$

where the vector $\widehat{\boldsymbol{\beta}}^i = (\widehat{\beta}^i_1, \widehat{\beta}^i_2, \widehat{\beta}^i_3)'$ is obtained from

$$R^i_{t,t+h} = \kappa^i_0 + \beta^i_1 R^M_{t,t+h} + \beta^i_2 \sigma^2_{t-h,t} + \beta^i_3 \left(\sigma^2_{t-h,t} \times R^M_{t,t+h}\right)$$

[17] This approach is fairly general and certainly does not assume a potentially simplistic utility function, such as a time-separable power utility with constant relative risk aversion. It is, for example, consistent with the external habits of Campbell and Cochrane (1999), among other approaches. In Campbell and Cochrane's case, the parameter $\theta^1_{t(h)}$ is a function of the surplus-consumption ratio (see, e.g., the discussion in Lettau and Ludvingson, 2001b).

in the CAPM case and

$$R^i_{t,t+h} = \kappa^i_0 + \beta^i_1 \Delta c_{t,t+h} + \beta^i_2 \sigma^2_{t-h,t} + \beta^i_3 \left(\sigma^2_{t-h,t} \times \Delta c_{t,t+h}\right)$$

in the C-CAPM case.

The cross-sectional R^2s for conditional and unconditional versions of the models are in Table 14. Figure 3 through 6 contain graphical representations of the corresponding pricing errors. In the figures, as customary, each two-digit number identifies the relevant portfolio. The first digit represents the size quintile, while the second digit represents the book-to-market quintile. For example, 15 is the portfolio of small firms with high book-to-market values. Scaling by past market variance drastically reduces the pricing errors and increases the R^2s in both of our specifications. The improvements are impressive, particularly at short and long horizons. In the CAPM case the R^2s go from 37.3%, 19.9%, 23.6%, and 44.9% for $h = 1, 3, 6$, and 12 to 78.3%, 68.2%, 76.9%, and 82.3%, respectively. They go from 11.4%, 3.4%, 1.3%, .3% to 49.5%, 17.8%, 41.1%, 61.6% for $h =$ 84, 96, 108 and 120. Similar improvements occur in the C-CAPM case, largely in the short run.

In light of this paper's results, it is not surprising that conditioning on past market variance reduces the pricing errors at most aggregation levels. The classical CAPM clarifies this point. In the presence of a risk-free rate, the partial effect of $R^M_{t,t+h}$ on the stochastic discount factor (i.e., $\theta^1_{t(h)}$) is effectively a discounted Sharpe ratio, i.e., $\theta^1_{t(h)} = \frac{E_t\left(R_{t,t+h} - R^f\right)}{R^f Var_t(R_{t+h})}$. In the short term (for low h values) past market variance has predictive ability for $Var_t(R_{t+h})$. This is well-known. In the long run (for large h values), past market variance tracks predictable variations in $E_t\left(R_{t,t+h} - R^f\right)$, as shown earlier.

10 Conclusions

This paper illustrates (and provides *statistical* and *economic* support for) an interesting empirical phenomenon, i.e., the long-run dependence between expected excess market returns and past market variance.

When analyzing the relation between short-term risk premia and classical financial ratios, (return) aggregation is usually invoked to support predictability (see, for instance, the discussion in Cochrane, 2001). Our aggregation results, on the other hand, appear contrary to the traditional short-term risk-return trade-off. It is now important to evaluate disaggregated economic models which would deliver our findings upon aggregation. On a related issue, it is important to have a complete understanding of the economic mechanism through which past long-run market variance tracks predictable movements in risk premia at frequencies lower than business cycle frequencies. We leave these issues for future research.

A large amount of recent work has been devoted to the relevance of volatility in cross-sectional asset pricing tests. The existing work has largely focused on innovations in market volatility as a priced systematic factor (Ang et al., 2006, Bandi et al., 2006, and Moise, 2006, among others) as well as on idiosyncratic volatility (Ang et al., 2006, and Spiegel and Wang, 2005, *inter alia*). We suggest an additional channel through which volatility might play a pivotal role in helping our understanding of price formation in financial markets. As shown, scaling by past market variance drastically reduces the mispricing of the size and value-sorted portfolios (at most frequencies) in the context of well-known pricing paradigms. Further research on this issue is being conducted by the authors and will be reported in future work.

References

[1] Andersen, T.G., T. Bollerslev, F.X. Diebold, and P. Labys (2003). Modeling and forecasting realized volatility. *Econometrica* 71, 579-625.

[2] Andrews, D. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation, *Econometrica* 59, 817–858.

[3] Ang, A., R. Hodrick, Y. Xing, and X. Zhang (2006). The cross-section of volatility and expected returns. *Journal of Finance* 61, 259-299.

[4] Baillie, R.T., and R.P. DeGennaro (1990). Stock returns and volatility. *Journal of Financial and Quantitative Analysis* 25, 203-214.

[5] Bandi, F.M. (2004). On persistence and nonparametric estimation: with an application to stock-return predictability. Working paper.

[6] Bandi, F.M. and B. Perron (2006). Long memory and the relation between realized and implied volatility. *Journal of Financial Econometrics* 4, 636-670.

[7] Bandi, F.M., C. Moise, and J.R. Russell (2006). Market volatility, market frictions, and the cross-section of stock returns. Working paper.

[8] Bansal, R. and A. Yaron (2004). Risks for the long run: a potential resolution of asset pricing puzzles. *Journal of Finance* 59, 1481-1509.

[9] Barndorff-Nielsen, O.E., and N. Shephard (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society, Series B 64*, 253-280.

[10] Bollerslev, T. and H. Zhou (2007). Expected stock returns and variance risk premia. *Working paper.*

[11] Boudoukh, J., M. Richardson, and R.F. Whitelaw (2005). The myth of long-horizon predictability. *Review of Financial Studies,* forthcoming.

[12] Campbell, J.Y. (1987). Stock returns and the term structure. *Journal of Financial Economics* 18, 373-399.

[13] Campbell, J.Y. and J.H. Cochrane (1999). By force of habit: A consumption-based explanation of aggregate stock-market behavior. *Journal of Political Economy* 107, 205-251.

[14] Campbell, J.Y. and L. Hentschel (1992). No news is good news: an asymmetric model of changing volatility in stock returns. *Journal of Financial Economics* 31, 281-318.

[15] Campbell, J.Y. and M. Yogo (2005). Efficient tests of stock return predictability. Working paper.

[16] Campbell, J.Y. and R. Shiller (1988). The dividend-price ratio and expectations of future dividends and discount factors. *Review of Financial Studies* 1, 195-227.

[17] Cavanagh, C., G. Elliott and J. Stock (1995). Inference in models with nearly integrated regressors. *Econometric Theory* 11, 1131-1147.

[18] Cochrane, J.H. (1996). A cross-sectional test of an investment-based asset pricing model. *Journal of Political Economy* 104, 572-621.

[19] Cochrane, J.H. (2001). Asset Pricing. Princeton: Princeton University Press.

[20] Cochrane, J.H. (2006). The dog that did not bark: a defense of stock return predictability. Working paper.

[21] Engle, R., D. Lilien, and R. Robins (1987). Estimating time-varying risk-premia in the term structure: the ARCH-M model. *Econometrica* 55, 391-407.

[22] Fama, E. F. and K.R. French (1996). Multifactor explanations of asset pricing anomalies. *Journal of Finance* 51, 55-84.

[23] French, K.R., W. Schwert, and R.F. Stambaugh (1987). Expected stock returns and volatility. *Journal of Financial Economics* 19, 3-29.

[24] Ghysels, E., P. Santa-Clara, and R. Valkanov (2005). There is a risk-return trade-off after all. *Journal of Financial Economics* 76, 509-548.

[25] Glosten, L.R., R. Jagannathan, and D.E. Runkle (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance* 48, 1779-1801.

[26] Guo, H. and R. Whitelaw (2006). Uncovering the risk-return relation in the stock market. *Journal of Finance* 61, 1433-1464.

[27] Hansen, L.P., J.C. Heaton, and N. Li (2005). Consumption strikes back? Measuring long-run risk. Working paper.

[28] Harrison, P. and H.H. Zhang (1999). An investigation of the risk and return relation at long horizons. *The Review of Economics and Statistics* 81, 399-408.

[29] Kilian, L. (1999). Exchange rates and monetary fundamentals: What do we learn from long-horizon regressions? *Journal of Applied Econometrics* 14, 491-510.

[30] Lanne, M., and P. Saikkonen (2006). Why is it so difficult to uncover the risk-return trade-off in stock returns? *Economics Letters* 1, 118-125.

[31] Lettau, M. and S.C. Ludvigson (2001a). Consumption, aggregate wealth, and expected stock returns. *Journal of Finance* 56, 815-849.

[32] Lettau, M. and S.C. Ludvigson (2001b). Resurrecting the (C-)CAPM: a cross-sectional test when risk premia are time-varying. *Journal of Political Economy* 109, 1238-1287.

[33] Lettau, M. and S.C. Ludvigson (2003). Measuring and modeling variation in the risk-return trade-off. Handbook of Financial Econometrics, Elsevier Science. Forthcoming.

[34] Lewellen, J. (2004). Predicting returns with financial ratios. *Journal of Financial Economics* 74, 209-235.

[35] Lundbland, C. (2005). The risk-return trade-off in the long-run: 1836-2003. Working paper.

[36] Maheu, J. and T. McCurdy (2006). Components of market risk and return. Working paper.

[37] Moise (2006). Stochastic volatility risk and the size anomaly. Working paper.

[38] Nelson, D.B. (1991). Conditional heteroskedasticity in asset returns: a new approach. Econometrica 59, 347-370.

[39] Pástor, L., M. Sinha, and B. Swaminthan (2006). Estimating the intertemporal risk-return trade-off using the implied cost of capital. Working paper.

[40] Phillips, P.C.B. (1986). Understanding spurious regressions in econometrics. *Journal of Econometrics* 33, 311-340.

[41] Phillips, P.C.B. (1991). Regression theory for near-integrated time-series. *Econometrica* 56, 1021-1043.

[42] Schwert, G. W. (1990). Indexes of United States stock prices from 1802 to 1987. *Journal of Business* 63, 399-426.

[43] Scruggs, J.T. (1998). Resolving the puzzling intertemporal relation between the market risk premium and conditional market variance: a two-factor approach. *Journal of Finance* 52, 575-603.

[44] Spiegel, M. and X. Wang (2005). Cross-sectional variation in stock returns: liquidity and idiosyncratic risk. Working paper.

[45] Turner, C.M., R. Startz, and C.R. Nelson (1989). A markov model of heteroskedasticity, risk, and learning in the stock market. *Journal of Financial Economics* 25, 3-22.

[46] Valkanov, R. (2003). Long-horizon regressions: theoretical results and applications. *Journal of Financial Economics* 68, 201-232.

Table 1. Descriptive statistics : 1952-2006

	$R_{t,t+1}$	$\sigma^2_{t,t+1}$	d_t / p_t	cay_t	Δc_t
Mean	.0050	.0014	.0364	-6.987x10^{-6}	.0020
Variance	.0017	8.240x10^{-6}	4.127x10^{-4}	1.472x10^{-4}	1.486x10^{5}
Skewness	-.690	16.662	-.550	.006	.058
Kurtosis	5.891	359.314	2.523	2.647	4.340
Autocorrelation	.059	.200	.994	.848	-.256

We report descriptive statistics for monthly continuously-compounded excess market returns ($R_{t,t+1}$), monthly market variances ($\sigma^2_{t,t+1}$), per capita consumption growth (Δc_t), dividend yield averaged over the previous year (d_t / p_t), and Lettau and Ludvingson's consumption-to-wealth ratio (cay_t). All data is monthly except for *cay* which is quarterly. We use the NYSE/Amex value-weighted index with dividends as our market proxy. The risk-free rate is the 30-day T-bill rate. The data are downloaded from CRSP for the period January 2, 1952 - December 29, 2006. To compute monthly continuously-compounded excess market returns we aggregate daily continuously-compounded excess market returns $r - r^f$ by defining $R_{t,t+1} = \sum_{j=1}^{n_t} \left(r_{t+\frac{j}{n_t}} - r^f_{t+\frac{j}{n_t}} \right)$, where n_t is the number of trading days in month *t*. The monthly realized variances $\sigma^2_{t,t+1}$ are obtained by computing $\sigma^2_{t,t+1} = \sum_{j=1}^{n_t} r^2_{t+\frac{j}{n_t}}$. The consumption data is from the Bureau of Economic Analysis and is available starting in February 1959. The consumption-to-wealth ratio is taken from Sydney Ludvigson's web site (http://www.econ.nyu.edu/user/ludvigsons/) and is only available through the end of 2005.

Table 2. Risk-return estimates at different levels of aggregation: 1952-2006

	h=1	3	6	12	24	36	48	60	72	84	96	108	120
α	.006	.010	.022	.050	.124	.167	.162	.064	-.103	-.261	-.368	-.530	-.590
	(3.87)*	(1.43)	(1.40)	(1.20)	(1.47)	(1.48)	(.89)	(.28)	(-.42)	(-1.23)	(-.51)	(-3.32)*	(-2.69)*
β	-.790	1.277	.929	.585	-.406	-.282	.493	2.092	3.937	5.124	5.612	6.298	6.344
	(-1.56)	(2.54)*	(.92)	(.49)	(-.25)	(-.17)	(.26)	(1.16)	(2.04)*	(3.39)*	(4.35)*	(7.17)*	(5.73)*
R^2(%)	.3	1.0	.7	.3	.3	.2	.5	7.7	26.4	49.1	61.9	73.9	73.2

We run linear regressions (with an intercept) of *h*-period continuously-compounded excess market returns $R_{t,t+h} = \sum_{i=1}^{h} R_{t+i-1,t+i}$ on *h*-period past realized market variances $\sigma^2_{t,t-h} = \sum_{i=1}^{h} \sigma^2_{t-i+1,t-i}$ (the single period quantities are defined in Table 1). We consider values of *h* equal to 1 through 120 months (i.e., one month through 10 years). HAC *t*-statistics are in parenthesis.

Table 3. Risk-return slopes at different levels of aggregation (zero intercept): 1952-2006

	h=1	3	6	12	24	36	48	60	72	84	96	108	120
β	.065	2.071	2.106	2.252	2.017	2.140	2.419	2.747	3.023	3.103	3.113	3.105	3.134
	(.07)	(3.03)*	(1.665)	(2.12)*	(2.05)*	(2.20)*	(2.95)*	(3.78)*	(4.76)*	(5.53)*	(6.91)*	(7.25)*	(7.136)*

We run linear regressions (without an intercept) of *h*-period continuously-compounded excess market returns $R_{t,t+h} = \sum_{i=1}^{h} R_{t+i-1,t+i}$ on *h*-period past realized market variances $\sigma^2_{t,t-h} = \sum_{i=1}^{h} \sigma^2_{t-i+1,t-i}$ (the single period quantities are defined in Table 1). We consider values of *h* equal to 1 through 120 months (i.e., one month through 10 years). HAC *t*-statistics are in parenthesis.

Table 4. Simulation results - Comparison of size of 5% tests of null slope

DGP: $y_t = \varepsilon_t,\ \sigma_t^2 = 0.6\sigma_{t-1}^2 + u_t,\ corr(\varepsilon_t, u_t) = -0.3$

	h=1	3	6	12	24	36	48	60	72	84	96	108	120
						Estimated intercept							
β	.0014	.0014	.0019	.0026	.0061	.0101	.0144	.0186	.0229	.0274	.0314	.0357	.0404
R^2(%)	.2	.5	.7	1.4	2.8	4.2	5.7	7.3	9.0	10.6	12.3	14.1	16.0
Standard *t*-ratio	5.2	22.8	37.3	51.8	63.7	69.9	74.7	77.4	80.0	81.4	82.0	83.2	84.5
HAC	5.6	2.2	1.1	1.1	2.1	3.5	4.6	6.2	7.8	9.3	11.0	12.9	15.1
t/sqrt(T)	5.0	4.8	5.3	4.8	4.8	4.3	4.3	4.1	3.5	3.8	3.5	3.4	3.4
						Constrained intercept							
β	.0007	.0004	.0004	.0014	.0027	.0035	.0038	.0040	.0047	.0060	.0073	.0088	.0106
Standard *t*-ratio	5.3	23.1	38.0	52.3	64.1	70.2	73.9	76.3	78.2	79.7	81.1	82.5	83.6
HAC	5.7	2.1	1.1	1.3	1.9	3.0	4.0	5.4	6.6	8.0	9.3	11.3	13.4
t/sqrt(T)	5.5	5.1	4.8	5.0	4.2	4.2	3.5	3.1	2.3	2.6	2.3	2.0	2.0

We simulate continuously-compounded excess market returns and market variances under the assumption of no predictability. Subsequently, we aggregate excess returns and variances over *h* months and run regressions of excess market returns on past market variance as in the main text. We implement 10,000 replications. We set T = 660. The table reports rejection probabilities (of the null of no predictability) associated with the standard *t*-ratios, with HAC *t*-statistics (implemented using Andrew's 1991 data-driven method), and with the t/sqrt(T) statistic presented in the main text. The first panel reports regressions with an intercept, the second panel reports regressions without an intercept.

Table 5. Risk-return estimates – t/sqrt(T) statistic and critical values: 1952-2006

	h=1	3	6	12	24	36	48	60	72	84	96	108	120
							Estimated intercept						
t/sqrt(T)	-.055	.102	.082	.057	-.048	-.037	.065	.260	.529	.848*	1.072*	1.377*	1.316*
Right-tail c.v. (ρ=.2)	.076	.114	.154	.220	.314	.390	.452	.533	.602	.668	.736	.790	.866
Right-tail c.v. (ρ=.6)	.075	.124	.165	.232	.325	.402	.469	.534	.630	.663	.743	.809	.882
Boot. c.v.	.063	.113	.156	.225	.310	.451	.522	.591	.658	.702*	.837*	.857*	.873*
							Constrained intercept						
t/sqrt(T)	.005	.204*	.249*	.330*	.402*	.531*	.688*	.851*	1.086*	1.354*	1.532*	1.591*	1.608*
Right-tail c.v. (ρ=.2)	.077	.113	.163	.221	.324	.412	.492	.564	.626	.724	.805	.867	.985
Right-tail c.v. (ρ=.6)	.075	.122	.167	.235	.333	.413	.489	.568	.649	.723	.800	.901	.963
Boot. c.v.	.067	.121*	.173*	.272*	.388*	.481*	.569*	.657*	.744*	.847*	.937*	1.047*	1.100*

We run linear regressions (with and without an intercept) of h-period continuously-compounded excess market returns $R_{t,t+h} = \sum_{i=1}^{h} R_{t+i-1,t+i}$ on h-period past realized market variances $\sigma^2_{t,t-h} = \sum_{i=1}^{h} \sigma^2_{t-i+1,t-i}$ (the single period quantities are defined in Table 1). We consider values of h equal to 1 through 120 months (i.e., one month through 10 years). The quantity "t/sqrt(T)" is the statistic presented in the main text. "Right-tail c.v." is the right tail critical value of t/sqrt(T) for a 5% two-sided test with equal probability of rejection in the tails. The critical values are obtained by simulating the asymptotic distribution of t/sqrt(T) with autoregressive coefficients of .2 and .6, correlation between disturbances of -.3, and sample size of 660 observations. 10,000 replications have been used. * denotes significance at the 5% level based on a two-sided test. The bootstrap critical values are obtained using the wild bootstrap algorithm described in the text.

Table 6. Risk-return slopes at different levels of aggregation: 1952-2006

Point estimates

Degree of aggregation of excess returns	Degree of aggregation of variance												
	1	3	6	12	24	36	48	60	72	84	96	108	120
1	-.79	.32	.05	.04	-.06	.06	.06	.12	.07	.05	.05	.04	.00
3	1.24	1.28	.18	.14	-..17	.13	.16	.33	.23	.18	.15	.12	.01
6	2.09	1.89	.93	.20	-.26	.17	.28	.65	.50	.38	.29	.26	.06
12	2.15	1.67	.92	.58	-.46	.14	.56	1.25	1.08	.80	.51	.63	.22
24	3.93	3.07	1.54	.40	-.41	.19	.82	1.93	2.35	1.74	1.04	1.35	.65
36	.57	.86	.58	.06	-.58	-.28	.78	2.16	3.37	2.91	1.92	1.80	1.40
48	.80	.52	-.32	-.69	-.33	.18	.49	2.11	3.67	3.95	3.01	2.47	2.02
60	1.49	1.62	.96	.79	.81	.81	1.34	2.09	3.69	4.62	4.00	3.47	2.67
72	5.99	4.92	3.53	2.34	1.40	1.76	2.72	3.61	3.94	4.87	4.72	4.36	3.46
84	7.13	5.12	3.56	2.42	2.45	3.35	4.43	5.15	5.20	5.12	5.28	5.06	4.33
96	7.19	6.31	5.31	4.75	4.81	5.51	6.14	6.44	6.31	5.95	5.61	5.67	5.03
108	14.08	11.70	9.39	7.95	7.16	7.18	7.36	7.49	7.05	6.62	6.52	6.30	5.68
120	17.78	14.72	12.04	9é51	7.99	7.71	7.85	7.76	7.20	7.00	7.04	6.85	6.34

HAC *t*-statistics

Degree of aggregation of excess returns	Degree of aggregation of variance												
	1	3	6	12	24	36	48	60	72	84	96	108	120
1	-1.56	1.30	.31	.51	-.80	1.31	1.44	2.54	1.36	1.18	1.08	.83	.01
3	1.23	2.54	.36	.44	-.55	.80	1.03	1.81	1.11	1.08	.91	.71	.08
6	1.18	1.50	.92	.29	-.43	.48	.88	1.78	1.24	1.13	.89	.73	.17
12	1.13	1.19	.67	.49	-.50	.22	1.00	1.78	1.22	1.19	.69	.84	.30
24	1.53	1.42	.64	.22	-.25	-.15	.89	1.55	1.79	1.42	.74	1.04	.54
36	.20	.37	.24	.03	-.28	-.17	.57	1.91	2.46	2.81	1.32	1.16	.87
48	.24	.14	-.05	-.14	-.09	.07	.26	2.05	2.98	4.29	2.21	1.48	1.09
60	.36	.36	.23	.22	.27	.37	.64	1.16	2.25	4.01	3.04	2.27	1.51
72	.91	.88	.71	.58	.44	.72	1.02	1.58	2.04	3.00	3.05	2.74	1.92
84	.98	.90	.68	.63	.82	1.52	1.69	2.65	3.13	3.39	4.18	4.53	3.08
96	1.04	1.23	1.25	1.38	1.71	2.11	2.61	3.68	4.30	4.49	4.35	6.11	3.96
108	1.19	1.39	1.34	1.37	1.70	2.36	2.61	3.24	3.49	3.70	5.02	7.17	4.60
120	1.12	1.31	1.31	1.47	1.64	1.91	1.68	1.60	1.69	2.32	3.50	5.72	5.73

R^2 (%)

	Degree of aggregation of variance												
Degree of aggregation of excess returns	1	3	6	12	24	36	48	60	72	84	96	108	120
1	.3	.2	.0	.0	.1	.2	.2	1.0	.4	.3	.3	.2	.0
3	.2	1.0	.0	.1	.3	.3	.5	2.5	1.4	1.0	.8	.6	.0
6	.3	1.1	.7	.1	.3	.2	.8	4.8	3.5	2.3	1.6	1.4	.1
12	.2	.4	.3	.3	.6	.1	1.6	9.2	8.2	5.2	2.4	4.2	.6
24	.4	.9	.6	.1	.2	.1	2.1	13.0	21.5	14.0	5.7	10.8	3.0
36	.0	.1	.1	.0	.4	.2	1.4	12.9	35.5	30.9	15.4	15.0	10.7
48	.0	.0	.0	.2	.1	.0	.5	10.3	35.8	48.2	32.0	24.2	18.2
60	.0	.1	.1	.2	.5	.7	2.6	7.7	27.5	50.2	43.5	36.6	24.3
72	.4	1.0	1.3	1.4	1.2	2.7	8.5	18.4	26.4	47.9	51.4	49.7	34.8
84	.5	1.1	1.2	1.4	3.3	9,1	20.8	34.6	43.0	49.1	59.7	61.6	49.5
96	.5	1.5	2.5	4.8	11.2	22.1	36.4	49.7	57.8	59.9	61.9	71.2	60.8
108	1.6	4.2	6.5	11.2	20.7	31.6	44.0	56.3	59.9	61.0	68.4	73.9	65.6
120	2.3	6.0	9.6	14.4	23.0	32.1	43.7	52.1	53.9	59.0	68.5	74.5	73.2

We run linear regressions (with an intercept) of h_r-period continuously-compounded excess market returns on h_v-period past realized market variances $\sigma^2_{t,t-h_v} = \sum_{i=1}^{h_v} \sigma^2_{t-i+1,t-i}$ (the single period quantities are defined in Table 1). We consider values of h equal to 1 through 120 months (i.e., one month through 10 years). The first panel reports parameter estimates, the second panel reports HAC t-statistics, the third panel reports the coefficient of determination.

Table 7. Risk-return slope estimates at different levels of aggregation: 1952-2006
Variance computed with *k* autocorrelations and the Bartlett kernel

h=	1	3	6	12	24	36	48	60	72	84	96	108	120
	-.79	1.28	.93	.59	-.41	-.28	.49	2.09	3.94	5.12	5.61	6.30	6.34
k=0	(-1.56)	(2.54)*	(.92)	(.49)	(-.25)	(-.17)	(.26)	(1.16)	(2.04)*	(3.39)*	(4.35)*	(7.17)*	(5.73)*
	.3	1.0	.7	.3	.3	.2	.5	7.7	26.4	49.1	61.9	73.9	73.2
	-.79	1.20	.84	.48	-.51	-.56	.14	1.53	3.09	4.18	4.80	5.66	5.92
1	(-1.90)	(2.53)*	(.83)	(.44)	(-.36)	(-.42)	(.08)	(.90)	(1.63)	(2.62)*	(3.59)*	(6.77)*	(6.89)*
	.4	1.1	.7	.3	.5	.7	.0	5.1	20.7	41.5	56.6	73.3	77.2
	-.84	1.44	1.25	.48	-.70	-1.04	-.37	.97	2.32	3.32	4.15	5.21	5.66
5	(-1.77)	(2.54)*	(1.11)	(.39)	(-.50)	(-.73)	(-.22)	(.54)	(1.15)	(1.75)	(2.45)*	(4.50)*	(5.88)*
	.4	1.4	1.3	.2	.8	2.3	.3	2.0	12.1	27.6	44.3	65.0	73.6
	-.81	1.26	1.19	.55	-.75	-1.17	-.48	.91	2.28	3.29	4.14	5.36	5.95
10	(-2.35)*	(2.48)*	(1.16)	(.54)	(-.62)	(-.84)	(-.27)	(.47)	(1.15)	(1.88)	(2.42)*	(3.97)*	(6.24)*
	.5	1.3	1.3	.3	.9	2.7	.5	1.8	11.5	26.7	42.2	64.0	74.2

We run linear regressions (with an intercept) of *h*-period continuously-compounded excess market returns $R_{t,t+h} = \sum_{i=1}^{h} R_{t+i-1,t+i}$ on *h*-period past market variances $\sigma^2_{t,t-h} = \sum_{i=1}^{h} \sigma^2_{t-i+1,t-i}$. The single period returns are defined in Table 1. The single period variances are HAC variances constructed using Bartlett kernels and *k* auto-covariances (with *k* between 1 and 10). We consider values of *h* equal to 1 through 120 months (i.e., one month through 10 years). HAC *t*-statistics are in parenthesis.

Table 8. Risk-return estimates at different levels of aggregation: 1934-2006

	h=1	3	6	12	24	36	48	60	72	84	96	108	120
						Unconstrained intercept							
β	.-598 (-.94)	.759 (.90)	.500 (.59)	.297 (.51)	-1.178 (-1.24)	-.505 (-.52)	.746 (.62)	1.954 (1.83)	2.590 (1.95)	3.105 (2.37)*	3.790 (2.75)*	4.233 (2.42)*	4.760 (3.58)*
Boot. c.v.	2.18	2.35	2.04	1.87	2.16	2.64	2.81	2.64	2.72	3.03*	3.16*	3.23*	3.19*
t/sqrt(T)	-.040	.063	.047	.033	-.167	-.080	.128	.339	.440	.521	.633*	.708*	.827*
Right-tail c.v.	.066	.105	.148	.195	.281	.339	.410	.453	.508	.566	.592	.655	.707
Boot. c.v.	.071	.103	.128	.198	.294	.382	.476	.526	.572	.617	.627*	.657*	.689*
						Constrained intercept							
β	.353 (.42)	1.777 (2.18)*	1.921 (2.40)*	2.057 (3.28)*	1.639 (1.88)	2.195 (2.04)*	2.800 (3.25)*	3.315 (7.62)*	3.575 (5.62)*	3.727 (7.87)*	3.876 (7.00)*	3.886 (5.89)*	3.909 (8.00)*
Boot. c.v.	2.55	1.95*	1.72*	1.75*	1.83*	1.90*	2.03*	2.06*	2.15*	2.28*	2.34*	2.64*	2.75*
t/sqrt(T)	.027	.183*	.241*	.323*	.342*	.542*	.819*	1.091*	1.256*	1.410*	1.585*	1.691*	1.818*
Right-tail c.v.	.067	.107	.143	.203	.282	.347	.408	.470	.530	.579	.646	.688	.756
Boot. c.v.	.063	.130*	.147*	.213*	.323*	.404*	.499*	.582*	.633*	.709*	.761*	.846*	.928*

We run linear regressions (with and without an intercept) of *h*-period continuously-compounded excess market returns on *h*-period realized market variances (the single period quantities are defined in Table 1). We consider values of *h* equal to 1 through 120 months (i.e., one month through 10 years). The quantity "t/sqrt(T)" is the statistic presented in the main text. "Right-tail c.v." is the right tail critical value of t/sqrt(T) for a 5% two-sided test with equal probability of rejection in the tails. The critical values are obtained by simulating the asymptotic distribution of t/sqrt(T) with autoregressive coefficients of .6, correlation between disturbances of -.3, and sample size of 876. 10,000 replications have been used. * denotes significance at 5% level based on a two-sided test. The bootstrap critical values are obtained using the wild bootstrap algorithm described in the text. They are reported for the classical HAC t-statistic (second row) and for the t/sqrt(T) (fifth row).

Table 9. First-order autoregressive parameter for variance at different levels of aggregation: 1952-2006

	h=1	3	6	12	24	36	48	60	72	84	96	108	120
ρ	.200	.264	.262	.233	.110	.106	.073	-.028	-.140	-.088	.059	.175	.234
	(2.11)*	(2.87)*	(1.37)	(.71)	(.22)	(.13)	(.12)	(-.03)	(-.11)	(-.14)	(.13)	(.32)	(.32)

We run linear regressions of *h*-period realized market variances $\sigma^2_{t,t+h} = \sum_{i=1}^{h} \sigma^2_{t+i-1,t+i}$ on past *h*-period past realized market variances $\sigma^2_{t,t-h} = \sum_{i=1}^{h} \sigma^2_{t-i+1,t-i}$ (the single period quantities are defined in Table 1). The corresponding t-statistics are in parenthesis.

Table 10. Risk-return slopes at different levels of contemporaneous aggregation: 1952-2006

Point estimates

	Degree of aggregation of variance												
Degree of aggregation of excess returns	1	3	6	12	24	36	48	60	72	84	96	108	120
1	-.79	.31	.20	.09	.04	-.00	.01	.02	.03	.03	.03	.05	.07
3	-8.16	-2.69	-.60	-.21	-.04	-.12	-.06	-.01	.02	.05	.05	.09	.16
6	-9.88	-6.27	-3.07	-.92	-.30	-.31	-.19	-.12	-.02	.03	.04	.11	.26
12	-11.50	-7.53	-5.43	-3.07	-1.05	-.78	-.56	-.41	-.19	-.06	-.00	.11	.37
24	-12.98	-8.95	-6.69	-4.86	-2.99	-1.88	-1.43	-1.09	-.73	-.41	-.23	-.05	.32
36	-9.19	-6.50	-5.00	-3.94	-3.30	-2.88	-2.07	-1.65	-1.17	-.79	-.47	-.20	.24
48	-7.20	-4.68	-3.22	-2.48	-2.19	-2.54	-2.54	-2.01	-1.53	-1.10	-.75	-.38	.13
60	-6.87	-4.35	-3.11	-2.11	-1.43	-1.73	-2.19	-2.31	-1.77	-1.35	-.95	-.56	.03
72	-6.59	-3.61	-1.78	-.87	-.83	-1.02	-1.46	-1.87	-1.85	-1.39	-.99	-.53	.10
84	-9.66	-5.94	-4.17	-2.61	-1.08	-1.14	-1.32	-1.54	-1.64	-1.59	-1.14	-.64	.09
96	-4.14	-2.69	-1.78	-1.75	-1.44	-1.05	-1.23	-1.25	-1.19	-1.24	-1.27	-.72	.06
108	-6.79	-4.21	-2.37	-1.36	-.95	-1.19	-1.11	-1.16	-.95	-.84	-.96	-.86	-.05
120	-8.05	-4.94	-3.27	-2.33	-1.10	-1.03	-1.32	-1.18	-.97	-.71	-.67	-.63	-.39

HAC *t*-statistics

	Degree of aggregation of variance												
Degree of aggregation of excess returns	1	3	6	12	24	36	48	60	72	84	96	108	120
1	-1.56	1.01	1.12	.86	.67	-.09	.16	.43	.74	.98	.98	1.45	2.07
3	-4.73	-2.89	-.63	-.45	-.16	-.62	-.31	-.09	.18	.37	.39	.69	1.24
6	-2.37	-2.75	-3.56	-.80	-.44	-.67	-.43	-.31	-.07	.08	.14	.37	.84
12	-1.74	-1.30	-1.73	-3.59	-.85	-.85	-.72	-.54	-.25	-.10	-.00	.18	.65
24	-1.33	-1.02	-1.09	-1.41	-2.30	-1.46	-1.38	-1.08	-.69	-.40	-.25	-.06	.40
36	-1.00	-.73	-.71	-.70	-1.05	-1.65	-1.56	-1.43	-.98	-.65	-.40	-.20	.28
48	-1.22	-.70	-.47	-.42	-.52	-.80	-1.59	-1.35	-1.04	-.75	-.52	-.27	.12
60	-.94	-.57	-.41	-.44	-.36	-.53	-1.04	-1.71	-1.28	-1.04	-.72	-.43	.02
72	-.81	-.42	-.19	-.11	-.18	-.32	-.60	-1.04	-1.25	-.95	-.76	-.38	.08
84	-1.59	-.92	-.67	-.46	-.19	-.30	-.50	-.70	-.92	-.99	-.73	-.46	.07
96	-.78	-.37	-.26	-.27	-.32	-.27	-.46	-.59	-.63	-.68	-.76	-.45	.05
108	-.86	-.45	-.31	-.24	-.20	-.33	-.37	-.50	-.47	-.43	-.52	-.52	-.03
120	-1.09	-.58	-.33	-.39	-.22	-.29	-.51	-.50	-.49	-.38	-.39	-.38	-.25

R^2 (%)

		Degree of aggregation of variance												
		1	3	6	12	24	36	48	60	72	84	96	108	120
Degree of aggregation of excess returns	1	.3	.2	.2	.1	.1	.0	.0	.0	.1	.1	.1	.3	.7
	3	10.4	4.6	.6	.2	.0	.3	.1	.0	.0	.1	.1	.4	1.3
	6	7.4	12.1	7.4	1.7	.4	.8	.4	.2	.0	.0	.0	.3	1.7
	12	5.0	8.7	11.4	9.3	2.7	2.4	1.8	1.3	.3	.0	.0	.1	1.8
	24	3.6	7.0	9.8	13.0	12.3	7.8	6.8	5.4	2.9	1.1	.4	.0	.8
	36	1.4	2.8	4.2	6.5	11.3	13.9	10.5	9.5	6.0	3.1	1.2	.2	.3
	48	.7	1.3	1.5	2.2	4.2	9.2	12.8	11.0	8.3	5.0	2.6	.7	.1
	60	.5	.9	1.1	1.3	1.4	3.3	7.3	10.6	7.9	5.5	3.2	1.2	.0
	72	.4	.2	.3	.2	.4	1.0	2.7	5.8	6.9	4.6	2.9	.9	.0
	84	.8	1.2	1.5	1.5	.6	1.1	2.0	3.5	4.8	5.3	3.2	1.1	.0
	96	.1	.2	.3	.6	1.0	.9	1.6	2.1	2.4	2.9	3.4	1.2	.0
	108	.3	.5	.5	.3	.4	1.0	1.1	1.6	1.3	1.2	1.6	1.4	.0
	120	.4	.6	.7	.9	.5	.6	1.4	1.5	1.2	.7	.7	.7	.3

We run linear regressions (with an intercept) of h_r-period continuously-compounded excess market returns $R_{t,t+h_r} = \sum_{i=1}^{h_r} R_{t+i-1,t+i}$ on h_v-period realized market variances $\sigma^2_{t-1,t+h_v-1} = \sum_{i=1}^{h_v} \sigma^2_{t+i-2,t+i-1}$ (the single period quantities are defined in Table 1). We consider values of *h* equal to 1 through 120 months (i.e., one month through 10 years). The first panel reports parameter estimates, the second panel reports HAC *t*-statistics, the third panel reports the coefficient of determination.

Table 11. Estimates with dividend yield at different levels of aggregation: 1952-2006

	h=1	3	6	12	24	36	48	60	72	84	96	108	120
α	.010	.016	.040	.105	.260	.340	.399	.329	.218	.149	-.028	-.383	-.592
	(-1.56)	(.60)	(.72)	(1.04)	(1.44)	(1.58)	(1.09)	(.62)	(.47)	(.49)	(.10)	(-1.18)	(-1.11)
β	-.823	1.267	.863	.435	-.579	-.495	.1639	1.639	3.371	4.442	4.898	6.008	6.348
	(-1.70)	(2.38)*	(.77)	(.34)	(-.36)	(-.34)	(.08)	(.862)	(2.00)*	(3.77)*	(4.24)*	(5.96)*	(4.21)*
γ	-.082	-.135	-.383	-1.167	-2.911	-3.646	-4.798	-5.099	-5.960	-7.519	-6.831	-2.306	.027
	(-.59)	(-.25)	(-.33)	(-.54)	(-.90)	(-.84)	(-.85)	(-.66)	(-.78)	(-1.74)	(-1.82)	(-.53)	(.00)
R^2(%)	.4	1.1	.9	1.2	3.5	4.5	6.7	13.0	32.6	57.8	67.7	74.3	73.2

We run linear regressions (with an intercept) of *h*-period continuously-compounded excess market returns $R_{t,t+h} = \sum_{i=1}^{h} R_{t+i-1,t+i}$ on *h*-period past realized market variances $\sigma^2_{t,t-h} = \sum_{i=1}^{h} \sigma^2_{t-i+1,t-i}$ (the single period quantities are defined in Table 1) and the time *t* dividend-to-price ratio. We consider values of *h* equal to 1 through 120 months (i.e., one month through 10 years). HAC *t*-statistics are in parenthesis.

Table 12. Estimates with dividend yield at different levels of aggregation: 1934-2006

	h=1	3	6	12	24	36	48	60	72	84	96	108	120
α	-.001	-.011	-.017	-.022	.006	-.067	-.174	-.349	-.470	-.634	-.832	-.895	-.932
	(-.18)	(-.47)	(-.04)	(-.18)	(.03)	(-.30)	(-.72)	(-1.11)	(-1.32)	(-2.03)*	(-2.83)*	(-2.31)*	(-2.45)
β	-.592	.826	.511	.285	-1.174	-.659	.448	1.543	2.187	3.022	3.965	4.229	4.619
	(-.90)	(.96)	(.56)	(.47)	(-1.20)	(-.70)	(.34)	(1.46)	(1.87)	(4.22)*	(5.54)*	(3.38)*	(5.88)*
γ	.185	.571	1.119	2.014	3.981	7.129	9.577	12.652	14.766	16.811	18.683	18.800	17.721
	(1.48)	(1.07)	(.96)	(.77)	(1.01)	(1.58)	(1.88)	(2.29)*	(2.27)*	(2.81)*	(3.13)*	(2.94)*	(2.90)*
R^2(%)	.4	1.2	1.6	2.2	7.5	11.7	17.8	32.9	41.7	49.6	59.3	63.1	67.0

We run linear regressions (with an intercept) of h-period continuously-compounded excess market returns $R_{t,t+h} = \sum_{i=1}^{h} R_{t+i-1,t+i}$ on h-period past realized market variances ${\sigma^2}_{t,t-h} = \sum_{i=1}^{h} {\sigma^2}_{t-i+1,t-i}$ (the single period quantities are defined in Table 1) and the time t dividend-to-price ratio. We consider values of h equal to 1 through 120 months (i.e., one month through 10 years). HAC t-statistics are in parenthesis.

Table 13. Estimates with *cay* at different levels of aggregation: 1952-2005

	h=1	3	6	12	24	36	48	60	72	84	96	108	120
α	-	.008	.018	.044	.110	.191	.235	.152	.020	-.137	-.263	-.476	-.532
		(1.30)	(1.18)	(1.07)	(1.81)	(2.32)*	(2.65)*	(1.08)	(.11)	(-1.07)	(-1.28)	(-2.43)*	(-2.00)*
β	-	1.558	1.318	.813	-.216	-1.075	-.901	.794	2.402	3.819	4.630	5.817	5.970
		(1.68)	(1.17)	(.45)	(-.14)	(-.67)	(-.90)	(.54)	(1.44)	(3.69)*	(3.15)*	(4.98)*	(4.04)*
γ	-	1.426	2.683	4.841	8.616	11.401	12.475	12.511	13.443	10.332	7.818	4.323	3.017
		(3.46)*	(3.04)*	(2.68)*	(3.59)*	(4.18)*	(5.04)*	(4.37)*	(3.69)*	(2.71)*	(1.89)	(1.43)	(.73)
R^2(%)	-	6.3	9.2	13.9	25.7	39.0	37.1	34.3	48.9	58.7	65.4	74.7	72.5

We run linear regressions (with an intercept) of *h*-period continuously-compounded excess market returns $R_{t,t+h} = \sum_{i=1}^{h} R_{t+i-1,t+i}$ on *h*-period past realized market variances $\sigma^2{}_{t,t-h} = \sum_{i=1}^{h} \sigma^2{}_{t-i+1,t-i}$ (the single period quantities are defined in Table 1) and Lettau and Ludvingson's consumption-to-wealth ratio (*cay*) at time t. Because *cay* is provided quarterly (and only through the end of 2005), we consider values of *h* equal to 3 through 120 months (i.e., one quarter through 10 years). HAC *t*-statistics are in parenthesis.

Table 14. R^2 (%) from cross-sectional regressions on Fama-French 25 size- and value-sorted portfolios: 1959-2006

	h=1	3	6	12	24	36	48	60	72	84	96	108	120
$R^m_{t,t+h}$	37.3	19.9	23.6	44.9	68.4	73.5	54.7	35.5	23.6	11.4	3.4	1.3	.3
$\Delta c_{t,t+h}$	1.7	5.9	2.5	.4	.2	2.5	6.4	19.8	32.4	47.9	59.2	65.6	60.6
$R^m_{t,t+h}, \sigma^2_{t-h,t}, R^m_{t,t+h} \times \sigma^2_{t-h,t}$	78.3	68.2	76.9	82.3	72.9	76.6	55.5	39.5	27.4	49.5	17.8	41.1	61.6
$\Delta c_{t,t+h}, \sigma^2_{t-h,t}, \Delta c_{t,t+h} \times \sigma^2_{t-h,t}$	25.8	21.6	67.0	76.6	44.7	43.4	61.3	73.2	44.7	54.2	63.9	67.1	73.2

The table reports cross-sectional R^2s from a linear regression of average returns on the Fama-French 25 portfolios sorted on size and book-to-market: $\left(\frac{1}{T-h}\sum_{t=1}^{T-h} R^i_{t,t+h}\right) = \alpha + \lambda' \hat{\beta}_i + \varepsilon_i$. The vector $\hat{\beta}$ is obtained from time-series regressions of $R^i_{t,t+h}$ on the variables in the first column.

Figure 1. Excess returns and realized variance at different levels of aggregation

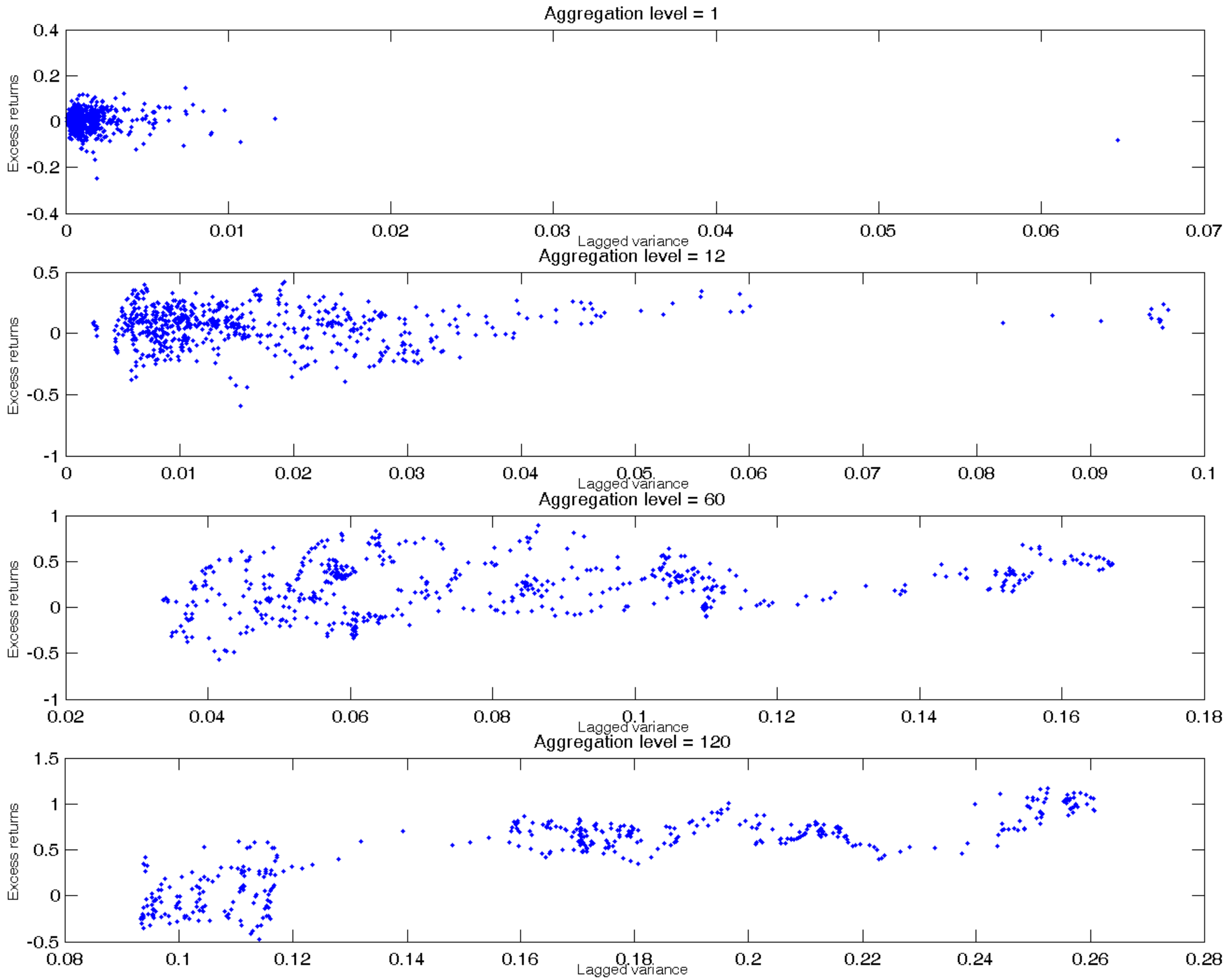

Figure 2. Slope estimates at different levels of aggregation and 95% confidence intervals

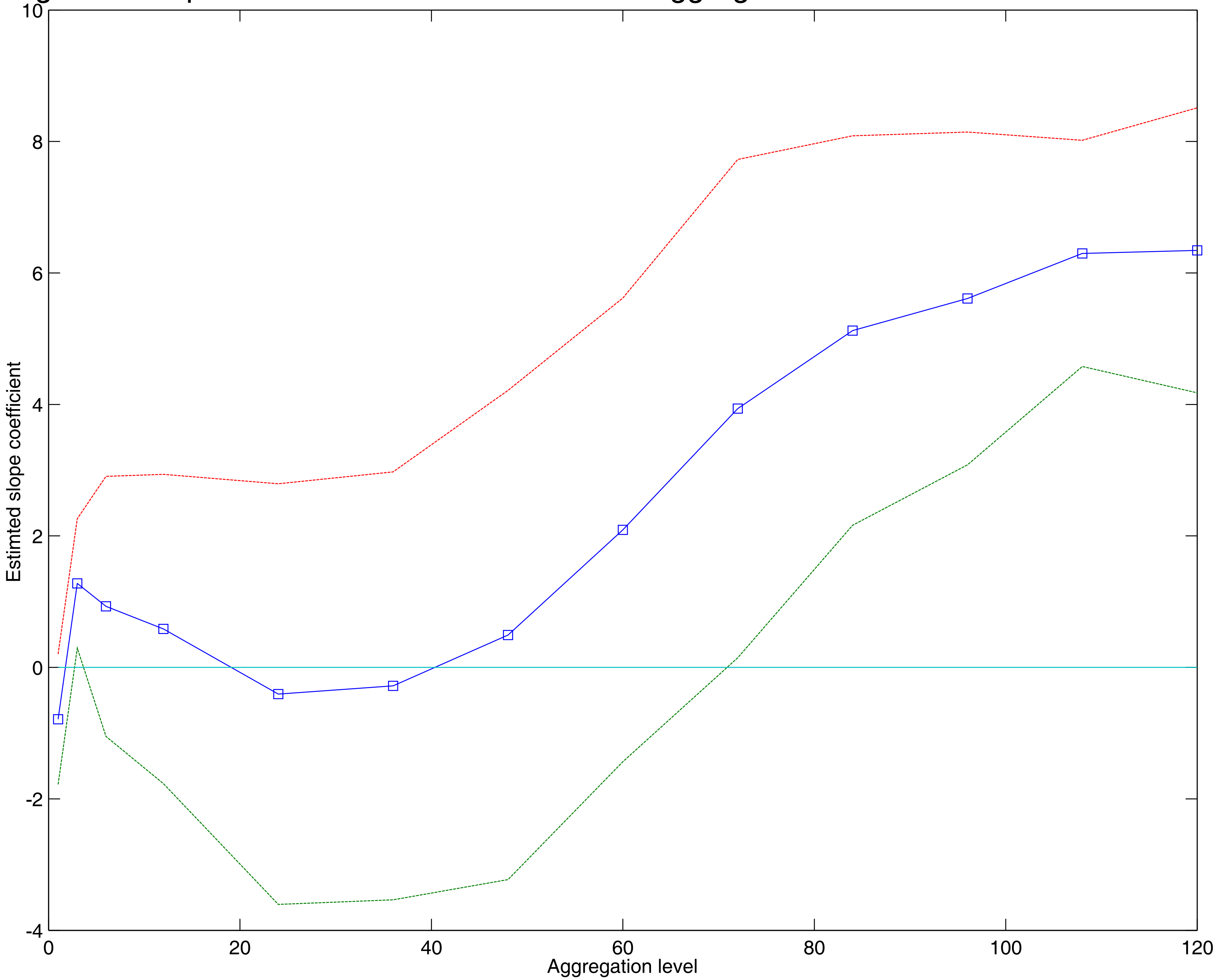

Figure 3. Actual vs. fiitted excess returns on FF portfolios - market as factor

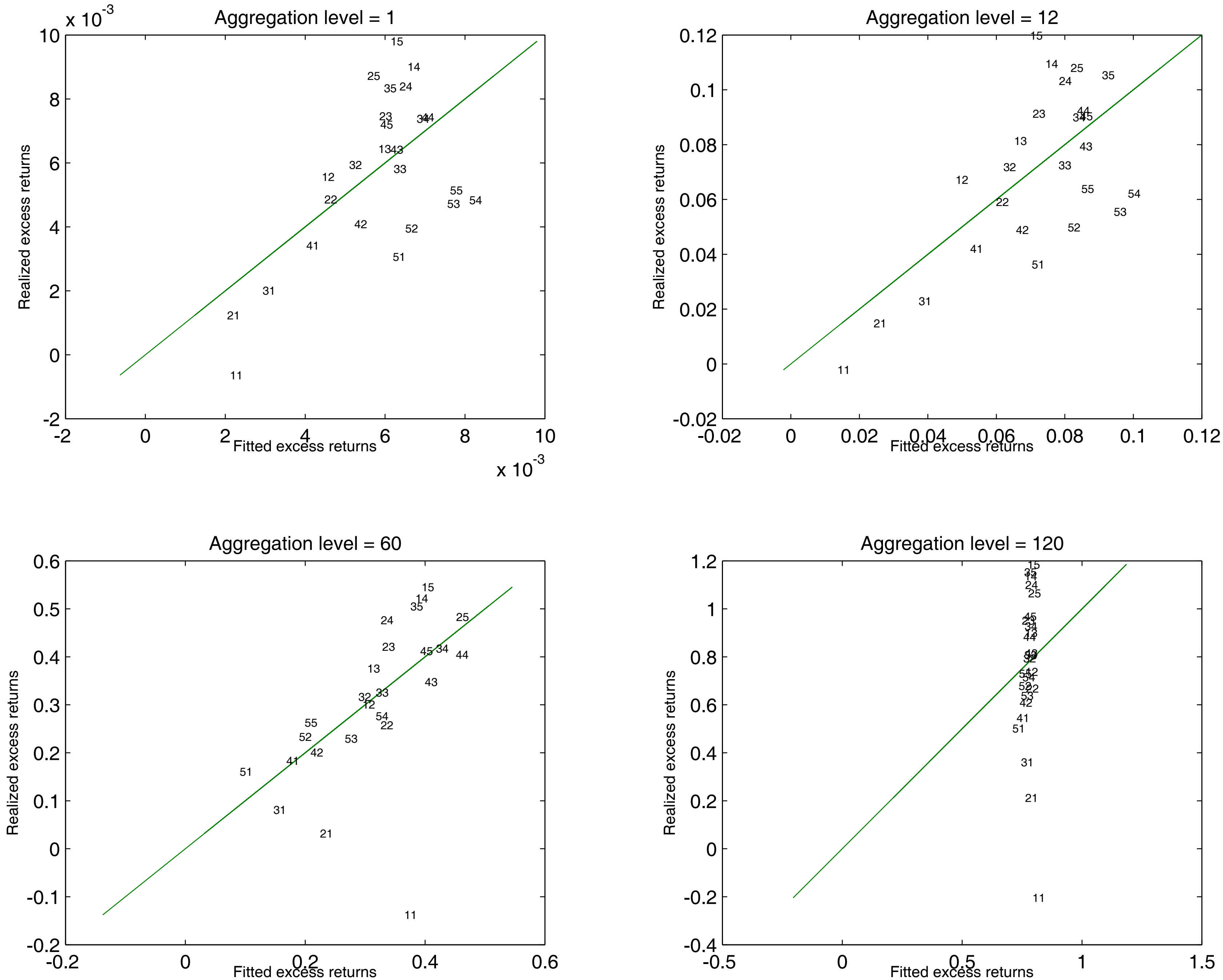

Figure 4. Actual vs. fitted excess returns on FF portfolios - market as factor and past variance as conditioning variable

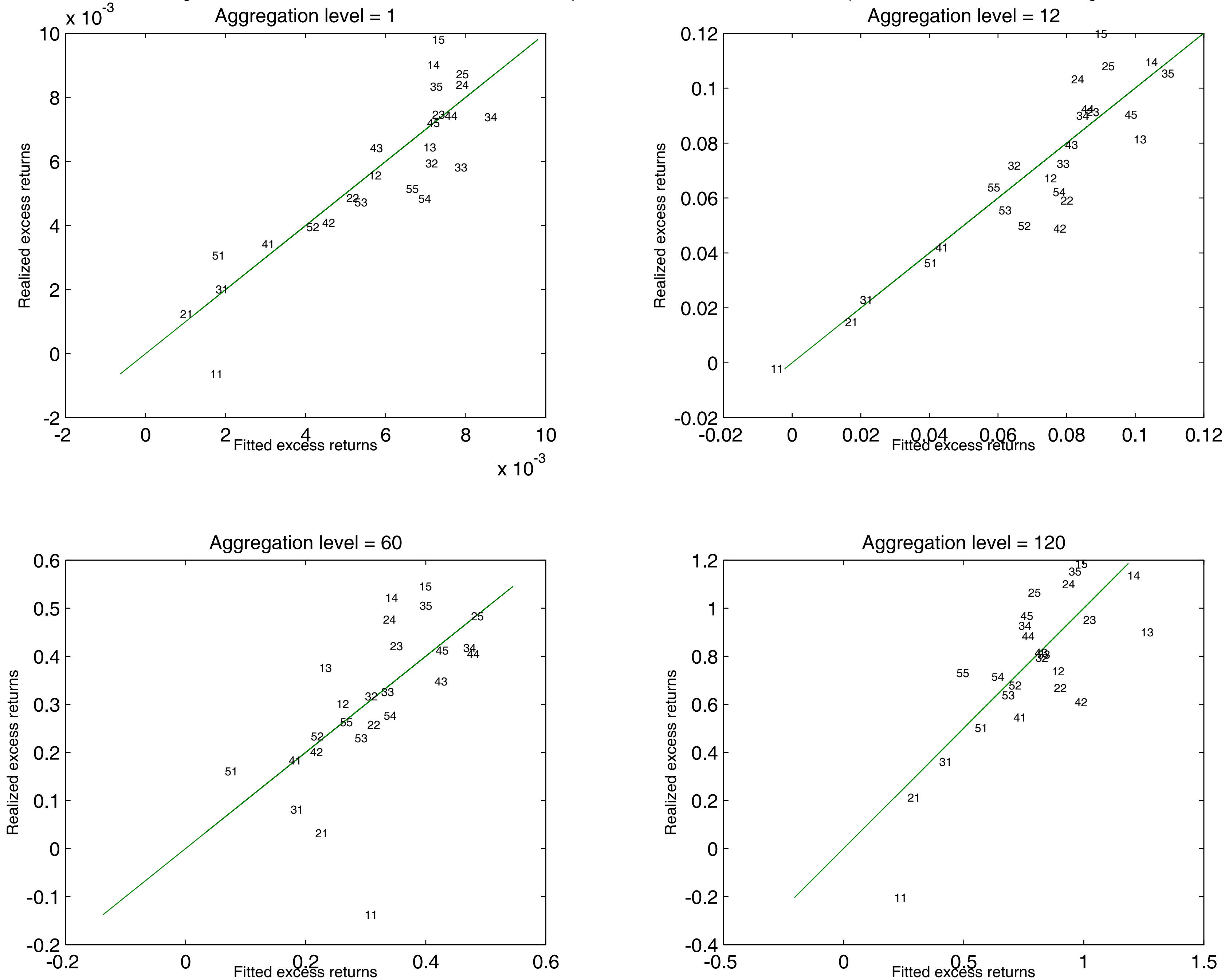

Figure 5. Actual vs. fitted exces returns on FF portfolios - consumption as factor

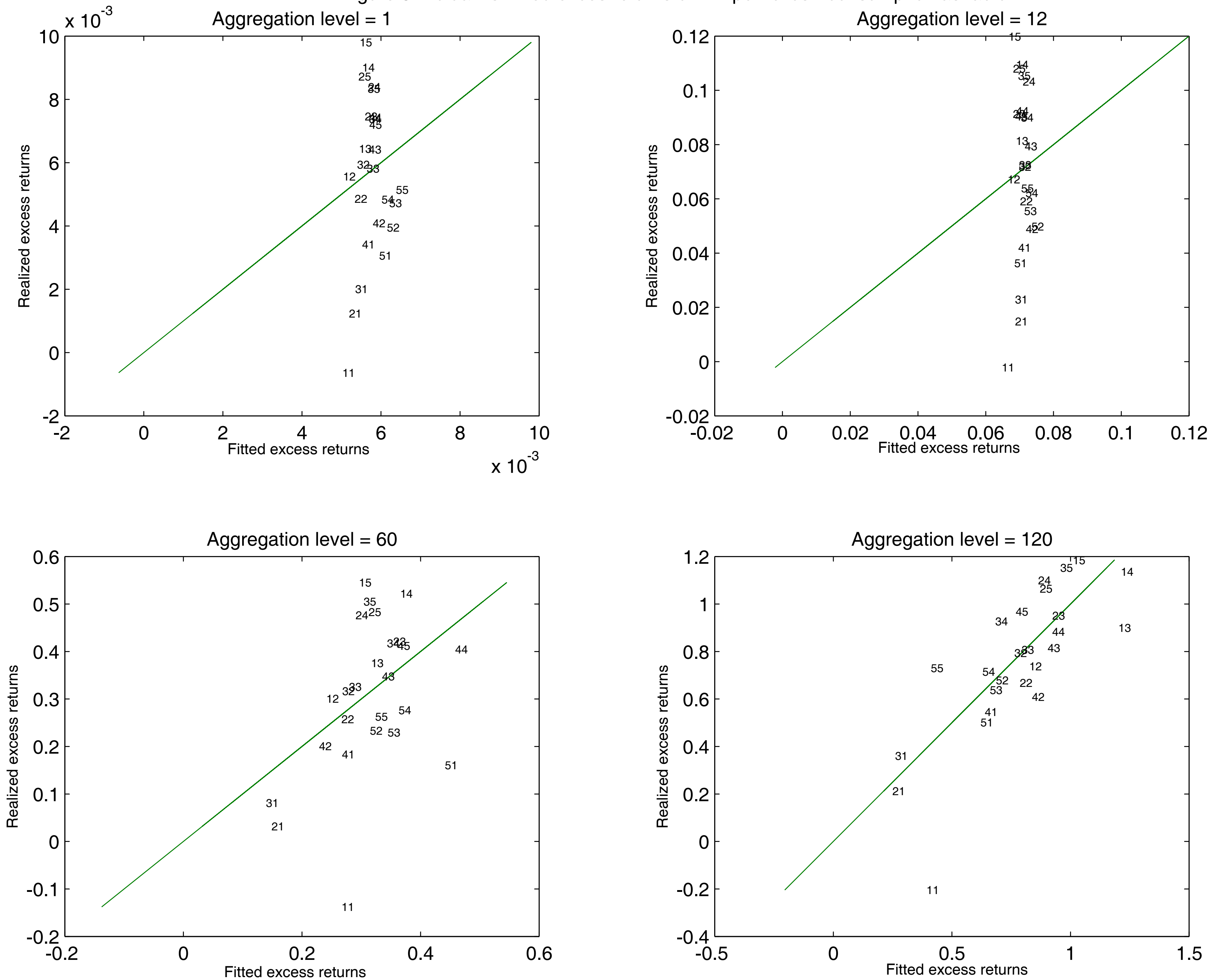

Figure 6. Actual vs. fitted excess returns on FF portfolios - consumption as factor and past variance as conditioning variable