Real risk or paper risk? Mis-measured factors, granular measurement errors, and empirical asset pricing tests *

Sung Je Byun† Lawrence D.W. Schmidt‡

This version: March 2019
First version: April 2016

Abstract

Given that the size distribution of publicly traded firms has fat tails, large idiosyncratic returns on large stocks can have nontrivial effects on the returns of value-weighted portfolios. We study effects of "granular measurement errors"– which are present when the law of large numbers fails and idiosyncratic returns are not fully diversified away – on standard empirical asset pricing tests. We construct an empirical proxy for the granular measurement error and demonstrate that it contributes substantially to the observed volatility of the CRSP value-weighted index and other market proxies. Unpriced granular measurement errors lead to downward-biased estimates of the intertemporal risk-return relationship and generate biased estimates of systematic risk exposures in cross-sectional asset pricing tests. After making simple corrections to eliminate the effects of granular measurement errors, we find much stronger evidence of an intertemporal risk-return relationship for the market index. In the cross section, betas for most portfolios – especially portfolios of small stocks – are severely biased downwards. After correcting estimated betas for the granular residual, the size anomaly disappears, and we find evidence of an expected return-beta relationship consistent with basic CAPM/APT theory. Finally, we use instrumental variables estimates to provide direct evidence that the granular residual is less informative about current and future real activity, suggesting an economic rationale for it having a lower or even zero risk price.

* *JEL classification*: C15; C58; G12; G17

* *Keywords*: Risk-return trade-off; Idiosyncratic risk; Empirical asset pricing

---

*We are grateful to John Campbell, John Cochrane, Xavier Gabaix, Stefano Giglio, Daniel Greenwald, and Leonid Kogan for helpful discussions related to the paper. We thank Maziar Kazemi for outstanding research assistance. The views expressed are those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Dallas or the Federal Reserve System.

†Financial Industry Studies Department, Federal Reserve Bank of Dallas, 2200 N. Pearl Street, Dallas, TX. SungJe.Byun@dal.frb.org.

‡MIT Sloan School of Management, 30 Memorial Dr, Cambridge, MA. ldws@mit.edu.

1 Introduction

The asset pricing literature wrestles to reconcile two seemingly contradictory empirical results about the riskiness of the stock market. On one hand, risk premia on broad market indices are quite large, even puzzlingly high relative to predictions of many macroeconomic models. This robust empirical result suggests that market returns must be very risky on average. On the other, measures of the quantity of market risk – volatility of market returns and covariance with the market return across assets – vary considerably yet are only tenuously linked with expected returns in the data. In the time series, evidence for a tradeoff between expected index returns and volatility is somewhat mixed.[1] In the cross-section, estimates of the risk-premium for covariance with broad market returns are often insignificant or even negative. Unlike the equity premium puzzle, these findings seem to suggest that the market return – or at least a sizable component of the market return – is not very risky.

While it is possible to reconcile these empirical results by introducing additional sources of risk and/or time-variation in risk prices, this paper considers a simple and complementary explanation which closely relates to the "granular hypothesis" of Gabaix (2011) and is operative even in simple linear factor models such as the Capital Asset Pricing Model (CAPM). Our starting point is that common benchmark portfolios used in empirical asset pricing tests are weighted averages of individual stock returns where, in most cases, a relatively small number of very large stocks are collectively associated with a substantial fraction of market value. Under these conditions, the law of large numbers can fail and a portion of the market return is driven by firm-specific shocks experienced by the largest stocks in the index which are not diversified away; we refer to this average as the "granular residual".[2]

We show that a non-negligible granular residual introduces non-classical measurement error into data used in conventional empirical asset pricing tests. Qualitatively, the associated econometric biases are capable of reproducing some of the seemingly inconsistent results discussed above. To assess their quantitative importance, we pro-

[1] Results are sensitive to choices of stock indices, sample periods, volatility forecasting models and/or approaches adopted by researchers.

[2] For example, in November 2018 *The Economist* magazine recently reported that "some 37% of the rise in the value of all firms in the S&P 500 index since 2013 is explained by six of its members: Alphabet, Amazon, Apple, Facebook, Microsoft, and Netflix. About 28% of the rise in Chinese equities over the same period is owing to two firms: Alibaba and Tencent." Source: "Big Tech's sell-off", *The Economist*, November 1, 2018.

pose and implement simple corrections to alleviate them using data from the US since 1928. We find that the effect of granular measurement errors is quite substantial, and that, after correcting these biases, results are more consistent with basic theory. Beginning with the time series, we find stronger evidence of an intertemporal relationship between the conditional mean and variance of the market excess return. In cross-sectional tests, we find evidence of substantial biases in estimates of market risk (beta) exposures and and a tight link between the size "anomaly" and the econometric biases associated with the granular residual.

We develop analytical results in a standard linear asset pricing framework in which each stock return in the portfolio is a linear combination of a small number of common, priced factors (which affect all stocks and linearly enter the stochastic discount factor) and an independent mean-zero error term. When the market value distribution has sufficiently thin tails and the number of traded stocks is large, the law of large numbers approximately holds and the market return is only driven by common factors. When idiosyncratic shocks are not diversified away, value-weighted portfolio returns also feature an additional term: the difference between the realized market return and its projection onto the priced common factors. We refer to this linear projection as the "true market factor". We characterize econometric biases for the case this extra term – the granular residual – is assumed to be orthogonal to the stochastic discount factor, and, accordingly, not associated with any risk compensation.[3]

The intuition for the bias induced for the intertemporal risk-return tradeoff under these assumptions is straightforward. When the granular residual is unpriced, the

[3]Since the granular residual term is driven by "idiosyncratic" shocks experienced by the very largest firms in the economy, it has an aggregate component and could in principle be priced. Also, given that we estimate the granular residual by aggregating stock-level returns that cannot be explained by a low-dimensional factor model, our estimates may also reflect specific news about industries which are overrepresented among very large firms that are unrelated to broader macroeconomic conditions. However, our empirical findings are consistent with the granular residual having a small (or zero) risk price. Such a result could be sensible in light of the Roll (1977) critique of tests of the CAPM related to using an incorrect market proxy. In a world where US public equity is only one of many sources of macroeconomic risk – such as those associated with other traded financial assets, private equity, real estate, uninsurable idiosyncratic risk, and human capital returns – that can enter the stochastic discount factor, changes in the value of public equity induced by idiosyncratic shocks to large public firms do not command high risk premia. See, e.g., Mayers (1973), Fama and French (1996), Campbell (1996), Heaton and Lucas (2000), Malloy et al. (2009), Campbell et al. (2016), and Schmidt (2016), among many others. To this end, we show that the industry distribution of "megacap" firms (those in the top 2.5% of the market value distribution, using NYSE market cap breakpoints) is not representative of the industry distribution of the rest of the market and exhibits much more pronounced and volatile time-variation.

benchmark index has the same risk-premium as the true factor. Under constant risk prices, only changes in variance of the true factor change the market risk-premium. Yet a non-negligible granular residual implies that estimates of market variance also reflect a forecast of the squared granular residual. Empirically, both the true factor and granular residual feature time-varying volatilities with different dyanamics. This generates a disconnect between the conditional variance of the benchmark portfolio and the true factor and leads to attenuation biases in estimates of the risk-return tradeoff in the time series. We present empirical and simulation evidence that this is sufficient to explain the weak evidence for such a tradeoff in extant studies.

We construct monthly estimates of granular residuals from value-weighted idiosyncratic returns of individual stocks using the same individual weights applied to construct the Center for Research in Security Prices (CRSP) value-weighted index. In the data, the distribution of market capitalization of listed stocks has extremely fat tails.[4] We estimate the standard deviation of the granular residual to be large, at about 3/5 of the standard deviation of excess returns on the CRSP value-weighted index. In contrast, granular residuals are negligible in the equal-weighted index, in which idiosyncratic returns are almost perfectly diversified away.

Next, we estimate models for the relationship between the conditional mean and variance of the market excess return.[5] We first replicate a familiar result from the literature; namely, the coefficient relating the expected market return to the conditional variance is positive yet statistically insignificant.[6] Then, we repeat the exercise using an "adjusted" market index return which subtracts our proxy for the granular residual. Across a wide variety of specifications, the risk-return tradeoff coefficient is larger in magnitude and highly statistically significant. Further, there is very strong evidence of a risk-return tradeoff using the CRSP equal-weighted portfolio, to which no adjustments are required because granular residuals are very small. We obtain similar results for other definitions of the market portfolio. Moreover, time-variation in the volatility of granular measurement errors can also have important quantita-

[4] At the end of our sample, the largest 10 and 25 firms collectively were associated about 18% and 30% of the total market capitalization of the index, respectively, and we present a number of examples in which large, event-driven price movements of a single stock moved the index by between 30 and 100 bp over short periods of time.

[5] We estimate the variance using different models in the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) family of Engle (1982) and Bollerslev (1986).

[6] This result is similar to findings of either negative or statistically weak coefficient estimates (Nelson 1991, Glosten et al. 1993, Koopman and Hol Uspensky 2002).

tive implications for the inference about the relative importance of time-varying risk quantities (volatility) versus risk prices.[7] After making our correction, estimates of a time-varying risk return tradeoff coefficient (price of risk) from a model with time-varying parameters are much less volatile relative to the unadjusted index.

In the cross section, we propose a simple instrumental variables method to address potential biases in market betas, where we instrument for the CRSP value-weighted return with a linear combination of principal components constructed using a rolling window of the cross-section of stock returns. We estimate the projection of the market index onto a small number of common factors and use the projection in place of the index return in subsequent asset pricing tests.[8] Empirically, we find similar results for various choices of fitting windows and numbers of principal components, suggesting that the projection of the market portfolio onto the principal components is fairly easy to estimate in practice.

We find a tight link between granular measurement errors and biases in estimated betas for sorted portfolios, an effect which is closely linked with firm size. In the presence of non-trivial granular measurement errors, CAPM betas for small stocks are severely biased downwards. This happens for two reasons. First, the use of a mis-measured factor generates a multiplicative attenuation bias which, as would happen with classical measurement error in an OLS regression, pushes estimates towards zero. This bias is particularly pronounced for small firms which, empirically, tend to have larger betas. Second, very large firms have returns which are directly correlated with the (non-classical) granular measurement error, which pushes their OLS betas upwards. When we investigate portfolios formed on size and beta, we find evidence that both effects are present and quantitatively large. Per our estimates, it is not uncommon for portfolios of mostly small stocks to have market betas that are biased downwards by as much as 20%.

After making these corrections, we find that the size premium becomes substan-

[7] Granular measurement errors introduce a multiplicative bias in the risk-return tradeoff coefficient, a bias which is time-varying when the ratio of the variances between the true factor and the granular residual is time-varying. This bias, through the lens of a regression of the market return on its conditional variance with time-varying parameters, would appear to reflect time-variation in risk attitudes. A similar bias can appear in cross-sectional Fama-Macbeth regressions.

[8] Our procedure is closely related to using two-stage least squares to correct for measurement errors in a standard regression setting. In our theory framework (and simulations), this procedure correctly addresses the attenuation bias due to granular measurement errors, because principal components are asymptotically uncorrelated with idiosyncratic shocks of individual firms, making them valid instruments.

tially smaller. The annualized CAPM alpha on the Small minus Large (1 - 10) long short decile portfolio shrinks from 3.5% to 0.67% for the post 1963 sample and from 5% to 2.2% over the full sample. Once we control for exposure to the component of the Fama-French value factor (HML) orthogonal to our instrument, OLS alphas on these portfolios remain substantial at (2.3% and 3.1% for the two samples, respectively), whereas alphas associated with our IV estimator are only -0.56% and 0.49% over the two samples. Empirically we find that pricing errors from a two-factor model which uses our instrument plus this orthogonal value factor are quite similar to those from the Fama-French (1993) three factor model.

We perform our corrections on a large number of characteristic sorted portfolios and consistently find that betas are more strongly biased for small stocks, and that OLS estimates tend to understate the variation of "true" betas across portfolios in the cross section.[9] At first blush, this effect would seem to exacerbate the familiar result that the relationship between expected returns and market betas is flatter than one would predict from the CAPM. However, our estimate of the slope of the expected return-beta relationship is steeper for IV – almost twice as large at 3.2% relative to 1.8% for OLS – despite the fact that the cross-sectional distribution of betas widens. This result is likely due to the fact that beta is measured with less error. Once we control for exposures to the value factor which is orthogonal to the market, our estimates of the expected return-beta slope are squarely in line with theory.[10]

The above results are all consistent with our assumption that the granular residual has a risk price near zero. Such a result could naturally arise if, for example, the granular residual conveys little to no information about future real activity. In the final section of the paper, we provide direct evidence consistent with this conjecture by following Schmidt (2016) and estimating model-free impulse response functions similar in spirit to the Jordà (2005) local projection method. The object of interest is how a stock return surprise of a given magnitude would cause an econometrician to change her forecast of future macroeconomic variables (e.g., consumption, GDP, employment, etc). If the granular residual is less informative than the true factor

[9]The standard deviation of IV betas (0.26) is about 35% larger than the standard deviation of OLS betas (0.19), and the standard deviation of the estimated *bias* in beta is 0.09, just shy of half the standard deviation of OLS betas. The average portfolio's beta biased downward by about 0.07.

[10]This result suggests that the "flatness" of the relationship between expected return and market beta is partially driven by the fact that stocks with low market betas tend to have above-average exposures to the value factor, which generates an omitted variable bias.

about these outcomes, we would expect to see an attenuation bias in these estimated impulse responses and, accordingly, infer that correlations between the stock market and the real economy are lower than they actually are for the true factor. When we use principal components (PCs) to construct an instrument for the true factor which is orthogonal to the granular measurement error, we consistently obtain IV estimates of these impulse responses which are considerably larger in magnitude. In many economic models, this finding would naturally imply that the granular residual would have a smaller or even zero risk price relative to the true factor.

Our results relate to two literatures – both of which are quite mature and which we will not attempt to survey here – which have found mixed results when testing for an intertemporal risk-return tradeoff as well as a relationship between betas and expected returns in the cross-section. Many papers in the former literature have tended to emphasize various approaches to form conditional variance forecasts (e.g., functional form for the conditional variance forecasts, distributional assumptions for the innovation process, etc.) and results have been found to be sensitive to these modeling choices. In contrast, our critique relates to a fundamental source of mis-specification and applies even when a researcher has access to a "perfect" measure of the volatility of the market index. Accordingly, we find that addressing it leads to much stronger results, where magnitudes line up with predictions from simulation exercises. We also find quite strong evidence of a risk-return tradeoff for various equal-weighted portfolios; to our knowledge, this result is new to the literature.

In the cross-sectional asset pricing literature, the potential challenges associated with a mis-measured market proxy dates back to Roll's (1977) critique, namely that the correct market proxy is unobservable and tests of the univariate CAPM may fail without it. More recently, Giglio and Xiu (2017) emphasize the potential importance of measurement errors in cross-sectional asset pricing tests and also propose a resolution which differs along several dimensions but also involves the use of principal components. Their empirical results, which complement ours, also suggest that using these adjustments can yield stronger evidence the market portfolio has a risk price which is consistent with its time series mean. Our paper additionally contributes by proposing a specific, empirically testable source of measurement error in the market portfolio, which lends itself to simple adjustments that can be integrated into other tests. The resulting changes in factor exposures obtained via these corrections are remarkably consistent with our analytical expressions for biases generated by the granular residual.

The remainder of the paper is organized as follows. Section 2 describes a simple linear asset pricing framework and provides a formal definition of the granular residual within it, and provides some motivating nonparametric evidence of the lack of diversification of the market portfolio in the US. Next, in sections 3 and 4, we develop analytical results to characterize the econometric biases introduced by the granular residual and propose and implement tests to correct them in the time series and cross section, respectively. Finally, section 5 demonstrates a similar attenuation bias in estimates of the degree of comovement between stock returns and measures of real activity, and section 6 concludes.

2 Theory and definition of granular residual

Here, we describe a simple linear asset pricing framework to motivate our empirical tests and provide a formal definition of the granular residual. Let $r_{i,t+1}$ denote the monthly *excess* return on individual stock i from month t to $t+1$. The market index return of interest is a weighted average of individual excess returns,

$$r_{t+1} = \sum_{i=1}^{N_t} w_{i,t} r_{i,t+1},$$

where $w_{i,t}$ is a predetermined weight of a stock i's return and N_t is a number of individual stocks in the stock index return in period t.[11] Unless otherwise noted, we assume that the index is value-weighted with weights w_{it} proportional to firms' market capitalization.

2.1 Basic framework

We assume that the excess return of each stock is described by a linear factor model

$$r_{i,t+1} - E_t[r_{i,t+1}] = \beta_i \cdot [f_{t+1} - E_t f_{t+1}] + \omega_i'[g_{t+1} - E_t g_{t+1}] + \eta_{i,t+1}, \tag{1}$$

[11] For example, $w_{i,t} = 1/N_t$ for an equal weighted index and $w_{i,t} = p_{i,t} s_{i,t} / \sum_{k=1}^{N_t} p_{k,t} s_{k,t}$ for a value weighted index where $p_{i,t}$ is stock i's price and $s_{i,t}$ is its number of outstanding shares at t.

where f_{t+1} is a scalar random variable which we will refer to as the "true factor", β_i is a stock i's exposure to f_{t+1}, g_{t+1} is a set of additional common factors (and ω_i the corresponding loadings) where $E(g_{t+1}|f_{t+1}) = 0$, and $\eta_{i,t+1}$ is a mean-zero idiosyncratic shock which has mean zero conditional on f_{t+1} and g_{t+1}.[12] Next, we assume that no arbitrage holds, implying the existence of a stochastic discount factor (m_{t+1}) satisfying the law of one price: $E_t[m_{t+1}r_{i,t+1}] = 0$ for $\forall i$ in the economy. If we further assume that stochastic discount factor takes the linear function of the form (as would be innocuous taking the continuous time limit with Brownian shocks)

$$m_{t+1} - E_t[m_{t+1}] = -\gamma \cdot (f_{t+1} - E_t[f_{t+1}])E_t[m_{t+1}] - \Lambda(g_{t+1} - E_t g_{t+1})E_t[m_{t+1}], \quad (2)$$

we obtain the familiar CAPM expression for the risk premium

$$\begin{aligned} E_t[r_{i,t+1}] &= \gamma \cdot \beta_i \cdot Var_t[f_{t+1}] + \Lambda' E_t[(g_{t+1} - E_t g_{t+1})(g_{t+1} - E_t g_{t+1})']\omega_i \\ &\equiv \gamma \cdot \beta_i \cdot \sigma^2_{t+1} + \Lambda'\Omega_t \omega_i, \end{aligned} \quad (3)$$

where σ^2_{t+1} and Ω_t capture the conditional variance of the factors, and an asset's risk premium is a product of risk prices γ and Λ, risk quantities, and the asset's factor exposures β_i and ω_i.

Clearly, the same restriction holds for the market portfolio r_{t+1}, which is a linear combination of individual stocks. We will separately identify the first factor f_{t+1} via the following rotation/normalization restrictions: $\sum_{i=1}^{N_t} w_{i,t}r_{i,t+1} = 1$ and $\sum_{i=1}^{N_t} w_{i,t}\omega_i = 0$. In other words, f_{t+1} captures the common factor exposures of the market portfolio and the other factors g_{t+1} are orthogonal to the market, which implies that

$$E_t[r_{t+1}] = \gamma \cdot \sigma^2_{t+1}. \quad (4)$$

As is textbook material, under some standard conditions on the (myopic) portfolio choice problem, the above linear model holds with $f_{t+1} = r_{t+1}$ and $\sigma^2_{t+1} = Var_t[r_{t+1}]$ and $\Lambda = 0$, because the first order condition for investors' optimal portfolio choice implies that equations (1) and (3) jointly hold with the f_{t+1} equal to the tangency

[12] Specifically, this formula would follow if the linear projection of stock i's return on f_{t+1} is constant, a common assumption in empirical work. When constructing empirical estimates of the granular measurement errors in section 3.2, we accommodate time-variation in β_is, allowing individual stocks to respond to the market factor differently over time. Specifically, we use 6-months' of past daily observations to capture time variation in firm-specific characteristics.

portfolio (which equals the market portfolio by market clearing).[13] This argument provides a strong rationale for ubiquitous use of a value weighted index of all stocks as an empirical proxy for f_{t+1} in the literature. The expression for the market risk premium, originally derived in Merton (1973), also motivates a literature seeking to empirically test the time-series implication of the CAPM for broad market expected returns:

$$E_t\left[r_{t+1}\right] = \gamma \cdot Var_t\left[r_{t+1}\right], \tag{5}$$

where the risk price γ is expected to be positive.

In our linear factor model, we can make an alternative argument for using the market portfolio as an empirical proxy for f_{t+1}, which is closer in spirit to Ross' (1976) Arbitrage Pricing Theory (APT). Consider the case in which the true data generating process is as in (1), where $\eta_{i,t+1}$ is independently distributed across stocks. In other words, returns have a single factor structure, but assumptions which would be used to derive the CAPM from first principles (and would imply $\Lambda = 0$) may or may not hold. Let's normalize $E_t[f_{t+1}] = \gamma \cdot \sigma^2_{t+1}$ so that f_{t+1} can be thought of as the "true" market portfolio of the univariate CAPM/APT model. Then, the market index return is a natural proxy for the unobservable true factor (f_{t+1})

$$r_{t+1} = \sum_{i=1}^{N_t} w_{i,t} r_{i,t+1} = \sum_{i=1}^{N_t} w_{i,t}\left(\beta_i \cdot f_{t+1} + \omega_i' g_{t+1} + \eta_{i,t+1}\right) \equiv f_{t+1} + 0 \cdot g_{t+1} + \eta_{t+1},$$

given our assumptions that the factors are defined such that $\sum\limits_{i=1}^{N_t} w_{i,t}\beta_i = 1$ and $\sum\limits_{i=1}^{N_t} w_{i,t}\omega_i = 0$. Since $E_t[\eta_{i,t+1}] = 0$, we know that r_{t+1} will be an *unbiased* estimator of f_{t+1}. Since the number of stocks N_t is large, under some conditions, r_{t+1} is also *consistent* for f_{t+1}, i.e.,

$$\operatorname*{plim}_{N_t \to \infty} r_{t+1} = f_{t+1} \iff \operatorname*{plim}_{N_t \to \infty} \eta_{t+1} = 0, \tag{6}$$

in which case the market portfolio is an ideal test asset, exactly revealing the true factor.

[13] For example, investors only consume financial wealth, have no short sales constraints, and maximize a mean-variance utility function.

2.2 Defining the granular residual

Whether or not (6) approximately holds depends on the extent to which the law of large numbers applies, and some necessary conditions (see, e.g., Gabaix (2011)) are easy to test. For example, if $\underline{\kappa} < Var[\eta_{i,t+1}] < \overline{\kappa} < \infty$ for all stocks and $\eta_{i,t+1} \perp \eta_{j,t+1}$ for all $i \neq j$, a necessary condition for $\eta_{t+1} \xrightarrow{p} 0$ is $\lim_{N_t \to \infty} \sum_{i=1}^{N_t} w_{i,t}^2 = 0$. As we demonstrate and discuss further (see Figure 1) below, this condition does not hold for the CRSP universe of stocks in the US, the dominant data source for most empirical asset pricing tests. For example, there have been two instances in which large numbers of individual stocks are added to the database, but the weights on the largest stocks barely change due to the highly skewed distribution of market values across firms.

We will refer to $\eta_{t+1} \equiv \sum\limits_{i=1}^{N_t} w_{i,t}\eta_{i,t+1}$ as the "granular residual", as it captures the contribution to the overall market return from idiosyncratic shocks which aren't fully diversified away. Our empirical estimates suggest that, due to the fat-tailed distribution of firm market values, the granular residual is quite large for the market portfolio, as is also likely the case for other test assets (usually value-weighted portfolio). When this is the case, our proxy for f_{t+1} is measured with error. This implies that the conditional variance of the market portfolio is larger than that of the true factor; $\sigma_{m,t+1}^2 = \sigma_{t+1}^2 + Var_t[\eta_{t+1}] \equiv \sigma_{t+1}^2 + \sigma_{\eta,t+1}^2$.[14] In the model, only the true factor is priced, and the risk-premium on r_{t+1} is proportional to the true σ_{t+1}^2, not $\sigma_{m,t+1}^2$.

One practical reason to favor the value-weighted portfolio is that, due to our normalizing assumption that $\sum_i w_{it}\omega_i = 0$, the market portfolio provides an unbiased (but inconsistent when the law of large numbers fails) estimator of f_{t+1} without needing to estimate any parameters. Alternative estimators might require obtaining direct estimates of nuisance parameters such as ω_i and β_i. While this need increases the statistical complexity of the analysis, methods such as Principal Components Analysis (PCA) are intended to exploit the "blessings of dimensionality" as the number of stocks and time periods are large. These methods treat the linear space spanned by f_{t+1} and g_{t+1} as unobserved and simultaneously recover consistent estimates of both factor loadings and weights. Therefore, it is fairly straightforward in practice to use these methods to estimate factors and factor loadings using data from a large cross-section of stocks. Most importantly, asymptotic properties of these estimators are well-understood, so, in contrast to the value-weighted portfolio when the mar-

[14]The lack of a cross term reflects our assumption that $\eta_{t+1} \perp f_{t+1}$.

ket value distribution has sufficiently fat tails, one can obtain consistent estimates of common factors with a large number of stocks and time periods.[15]

2.3 Discussion of key assumptions

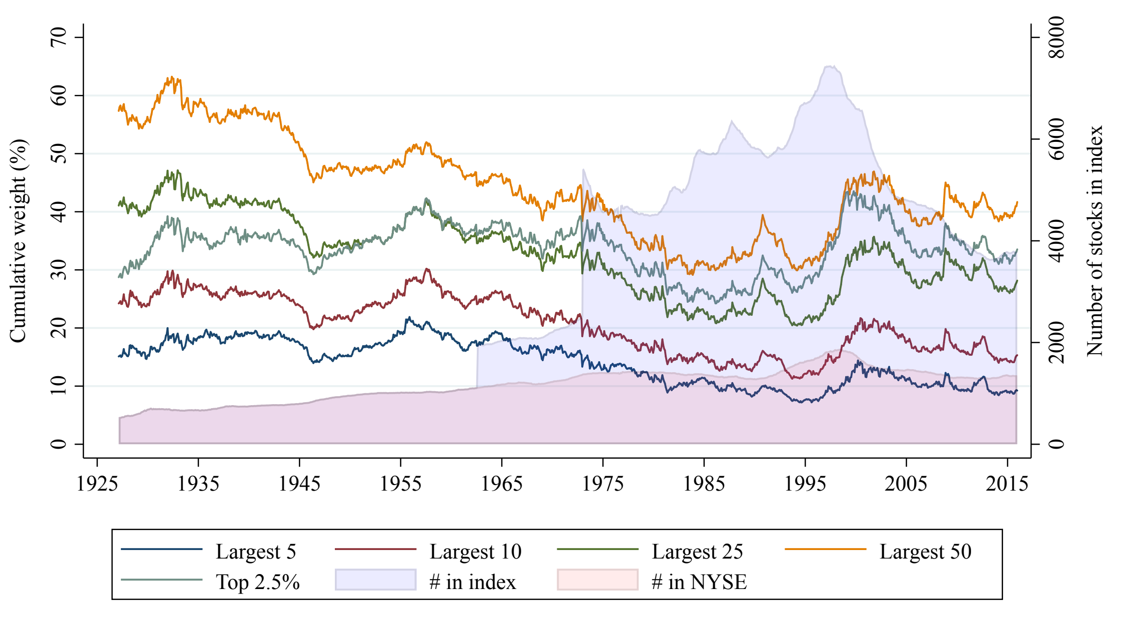
The extent to which the granular residual η_{t+1} is quantitatively large is an empirical question. To support our discussion and empirical tests that follow, we provide some empirical evidence from the U.S. stock market, suggesting this is likely the case.

Figure 1 highlights the market concentration among largest stocks in the CRSP value-weighted index. The figure plots the share of market value which is associated with the largest 5, 10, 25, and 50 stocks, respectively. Throughout the entire sample period from 1926 to 2014, the largest five stocks are associated with between 10 and 20 percent of the market capitalization of the index as a whole. The next 20 stocks are also associated with a substantial fraction of market value, so that the combined weight of the top 25 stocks ranges from almost 50% early in the sample to 30% at the present. Stocks 26-50 constitute another nontrivial percentage of total market value.

These numbers decline somewhat over the sample in part because the CRSP universe expands, as is clear from the number of stocks which is plotted on the right axis. Nonetheless, the market is highly concentrated throughout the sample, and it is striking that one can barely detect any changes in these weights even when the number of stocks included in the index jumps dramatically upward. These line plots hold the number of stocks fixed in each bin, which might suggest that granular measurement errors is smaller today than in the past. However, while the average weight per stock has declined, the fact that there are many more stocks in the upper tail of the market value distribution acts as a countervailing force.

[15]Bai and Ng (2006) proved that errors associated with estimating factors do not affect the limiting distribution of factor-augmented vector autoregression (FAVAR) estimators, which nests our model as a special case (with the minor caveat that our panel of individual stock returns is unbalanced, whereas their theoretical analysis assumes the panel is balanced). In other words, estimates of factors are consistent and subsequent regression estimates which use estimated principal components have the same asymptotic properties as if the true unobserved factors were used instead. This property holds when $\frac{T}{N} \to 0$ as $N, T \to \infty$. Sample sizes which are used to construct estimates of factors and loadings usually involve a number of stocks which is an order of magnitude larger than the number of time periods, so this condition approximately holds. Even when this condition fails, estimated factors can be treated as data and $\sqrt{T}$-consistent estimates of second stage regression parameters obtain when $\frac{\sqrt{T}}{N} \to 0$, but standard errors need to be corrected to reflect the fact that a generated regressor is used. Stock and Watson (2002) provide conditions under which factors are consistently estimable even when loadings exhibit moderate time-variation.

Figure 1: Combined weights of largest stocks in the CRSP value-weighted index



Note. This figure, on the left axis, plots combined weights of individual stocks for the 5, 10, 25, and 50 largest stocks in the CRSP value-weighted index respectively. The top 2.5% line, also on the left axis, plots the cumulative market value associated with stocks whose rank in the cross-sectional market value distribution is less that 2.5% times the number of stocks in the NYSE (using the NYSE to compute this breakpoint avoids generating mechanical increases in the series as many small AMEX and NASDAQ firms are added to the database). The right axis plots the total number of stocks included in the CRSP index as well as the NYSE only, where the former includes two large jumps as the CRSP sample coverage expands. One can visually confirm that combined weights from a very small number of large firms dominate the rest of firms in the CRSP value-weighted index.

We also compute an alternative measure of concentration which allows the number of stocks in the bin to grow as the number of listed stocks increases. Specifically, we compute the share of stocks in the top 2.5th percentile of the market value, where this measure sums up weights for the largest stocks whose rank in the CRSP market value distribution is less than 2.5% times the number of stocks in the NYSE for each month.[16] The concentration measure lacks a downward trend, hovering between 25 and 40 percent of the total value of the stock market.

In Appendix A, we provide additional pieces of supporting evidence, including several examples of situations in which drastic changes in a large company's stock price singlehandedly pulled down the market index by nontrivial amounts. We also demonstrate that the industry composition of the largest firms is quite volatile over

[16]Here we use the count of stocks in the NYSE (shaded in red on the right axis) to avoid mechanical increases in our concentration measure due to a large number of (mostly small) AMEX and NASDAQ stocks added during the second half of our sample period.

time and quite unrepresentative from the rest of the stock market.

Before moving on, a few comments are in order. First, the distinction between aggregate and idiosyncratic shocks becomes quite blurry when one considers a large shock experienced by one of the largest firms in the economy; to this end, Gabaix (2011) focuses on the role of large firms' idiosyncratic shocks in contributing to aggregate fluctuations. First, large changes in large firms' valuations (e.g., associated with information releases such as earnings announcements) may act as bellwethers for other firms. Second, large firms are connected to many other firms in the economy through customer-supplier relationships, suggesting that network effects will further amplify aggregate effects of these shocks (Acemoglu et al., 2012).

However, even if the word "idiosyncratic" may not describe these shocks perfectly, it is plausible that movements in the market index driven by idiosyncratic news about very large firms (coming from industries that change over time and may not necessarily be representative of the broader economy) are less likely to cause concern for investors relative to price movements which are common across a large number of individual stocks.[17] First, public firms are an important part, but not the only part, of the U.S. economy.[18] Thus, our argument is similar to Campbell and Vuolteenaho (2004), who find that ICAPM theory predicts that market declines driven by cash flow news are more painful for long-term investors than movements driven by discount rate news because the former leave investment opportunities unchanged whereas the latter are associated with improvements in future investment opportunities. In our case, more broad-based movements across many stocks are likely to be more tightly linked with these other risk factors. In the penultimate section of the paper, we provide additional direct evidence consistent with the granular residual being less informative than the true factor about future real activity.

A second challenge comes the fact that empirical asset pricing tests require estimates of the systematic exposure of an asset (or portfolio of assets). Thus, even if these shocks were priced to some extent, estimates of assets' loadings on these shocks are less likely to be precisely estimable or stable over time, given the nontrivial time

[17]Note that these more broad-based price movements would still likely reflect the systematic component of information received by the market about large firms other than the direct effects of these individual stock price changes on the value of the index of the whole.

[18]For example, Davis et al. (2006) estimate using Census data that public firms comprised about 30% of total employment in the U.S. over their sample. Moreover, for many investors, news about expected labor and/or entrepreneurial income (and potentially uninsurable risk associated with these factors) may be more important than dividends from investments in public firms.

variation in the composition of these megacap firms and extremely heterogeneous nature of the shocks in question.

3 Time-series: Reassessing the risk-return relationship for the market portfolio

In this section, we investigate effects of the granular residuals on the evaluation of the risk-return relationship in excess returns of the stock market index. We first characterize the econometric biases of unpriced granular measurement errors on empirical asset pricing tests below. We then correct for effects of the granular residuals directly by constructing an adjusted index return which subtracts an estimate the granular residual from the index return itself. We find substantial evidence that the granular residual generates an attenuation bias which weakens the evidence for an intertemporal risk-return relationship, whereas the risk-return relation is more robust in our adjusted index and in equal-weighted portfolios.

3.1 Econometric biases

Consider a researcher seeking to test implications of equation (3) for the market, $E_t[r_{t+1}] = \gamma \cdot \sigma^2_{t+1}$.[19] When a researcher estimates the expected return of the true market factor from the stock market index return, effects of the granular residual are mild. The historical average stock market index return is an unbiased, though not necessarily efficient estimator of the true market risk premium: $\gamma \cdot E[\sigma^2_{t+1}]$. Such an estimate implies a fairly high value of the price of risk γ, a lower bound for which (in the presence of a granular residual) can be obtained by dividing through by an estimate of the average variance of the market return.

However, in the presence of a non-negligible granular residual, much more substantial issues arise when attempting to test the relationship between expected returns and time-varying volatility. The reason is simple: *squared forecast errors of the stock market index return are not unbiased predictors of the conditional variance of the true market factor*. Granular measurement errors will lead to inaccurate (expected and realized) variance estimates which are larger than those of the true market factor and

[19] Recall that this expression follows from our normalization that the $\sum_{i=1}^{N_t} w_{i,t}\beta_i = 1$ and $\sum_{i=1}^{N_t} w_{i,t}\omega_i = 0$. Note that σ^2_{t+1} is measurable with respect to the time t information set.

likely exhibit very different time series dynamics.

To further simplify exposition, suppose that the researcher has access to a "perfect" conditional variance estimate for $\sigma^2_{m,t+1}$. If the data are generated according to the model above and the granular residual is absent, we recover the true market price of risk γ via an OLS regression of r_{t+1} on a constant and $\sigma^2_{m,t+1}$. When $\sigma^2_{\eta,t+1} \neq 0$, the OLS regression coefficient is biased and converges to

$$\operatorname*{plim}_{T\to\infty}\widehat{\gamma} = \frac{Cov\left[\sigma^2_{m,t+1}, r_{t+1}\right]}{Var\left[\sigma^2_{m,t+1}\right]} = \gamma \cdot \frac{Var\left[\sigma^2_{t+1}\right] + Cov\left[\sigma^2_{t+1}, \sigma^2_{\eta,t+1}\right]}{Var\left[\sigma^2_{t+1}\right] + Var\left[\sigma^2_{\eta,t+1}\right] + 2\cdot Cov\left[\sigma^2_{t+1}, \sigma^2_{\eta,t+1}\right]}. \quad (7)$$

The use of a mis-measured proxy for the priced factor is associated with a downward multiplicative bias provided that $Var\left[\sigma^2_{\eta,t+1}\right] \geq -Cov\left[\sigma^2_{t+1}, \sigma^2_{\eta,t+1}\right]$. This condition is guaranteed if σ^2_{t+1} and $\sigma^2_{\eta,t+1}$ are positively correlated or if $Var\left[\sigma^2_{\eta,t+1}\right] > Var\left[\sigma^2_{t+1}\right]$.[20] When variances are estimated from the data using a parametric model, things are slightly more complex because we have a generated regressor, but the intuition and basic mechanism is essentially the same.

In the empirically relevant case, the researcher would obtain a downward-biased coefficient estimate for γ even when using the correct stock market index variance as a predictor. As is well-known, measurement error in the key variable of interest (in this case, σ^2_{t+1}) often produces an attenuation bias. For instance, when $Cov\left[\sigma^2_{t+1}, \sigma^2_{\eta,t+1}\right] = 0$, heteroskedasticity in the granular measurement errors generates classical measurement errors in the regression equation, which attenuates the coefficient towards zero by a multiplicative factor of $Var\left[\sigma^2_{t+1}\right] / \left(Var\left[\sigma^2_{t+1} + \sigma^2_{\eta,t+1}\right]\right)$.[21] Before proceeding, note that the key source of attenuation bias in $\widehat{\gamma}$ is *time variation* in $\sigma^2_{\eta,t+1}$, which appears to be comparable in magnitude to the time variation in the variance of the true market factor. If a constant term is included in the regression,

[20]The intuition for why a negative correlation between σ^2_{t+1} and $\sigma^2_{\eta,t+1}$ can introduce an upward bias is as follows. A negative correlation between σ^2_{t+1} and $\sigma^2_{\eta,t+1}$ has two effects: the first is to reduce the covariance term in the numerator, which pushes the ratio downwards, but the second is to reduce the variance of $\sigma^2_{m,t+1}$ (the denominator). When the variance of $\sigma^2_{\eta,t+1}$ isn't very high relative to σ^2_{t+1}, the effect of reducing the denominator can dominate. In the extreme case for which $Var\left[\sigma^2_{\eta,t+1}\right]$ approaches $Var\left[\sigma^2_{t+1}\right]$ from below and the two variables are perfectly negatively correlated, the bias can become arbitrarily large. However, this is not the empirically relevant case.

[21]A correlation between the two generates an additional adjustment, but the attenuation bias dominates empirically. Using our empirical proxy for the granular residuals introduced in the following subsection, we estimate that $Var\left[\sigma^2_{t+1}\right] / \left(Var\left[\sigma^2_{t+1} + \sigma^2_{\eta,t+1}\right]\right)$ is 0.846, 0.868 and 0.811, respectively for total, the first sample period from 1928.1 to 1962.6, and the second sample period from 1962.7 to 2014.12.

$\sigma^2_{\eta,t+1} > 0$, is not sufficient to generate the attenuation bias in $\widehat{\gamma}$ unless $\sigma^2_{\eta,t+1}$ exhibits sizable variation over time. In Appendix D, we illustrate the potential quantitative importance of such attenuation bias using several Monte Carlo simulations.

Finally, consider the case in which the true market price of risk γ is constant but a researcher estimates a model which allows it to vary over time. When σ^2_{t+1} and $\sigma^2_{\eta,t+1}$ are both time-varying and not perfectly correlated, biases introduced by the granular residual are not constant over time. Therefore, if one allows for time-variation in γ, changes in the biases in equation (7) can lead a researcher to infer that the price of risk is changing even if it is constant. We examine this possibility in section 3.5.

3.2 Estimating the granular residual

Following our definition in Section 2, we construct monthly estimates for the granular residuals as a weighted average of idiosyncratic returns:

$$\eta_t \equiv \sum_{i=1}^{N_{t-1}} w_{i,t-1}\eta_{i,t}, \tag{8}$$

where $w_{i,t-1}$ is a stock i's (observable) weight in the stock market index and $\eta_{i,t}$ is a stock i's idiosyncratic return in a month t.

Our approach to estimating the granular residual is close to a recursive application of the cross-sectional regression, introduced in Fama and MacBeth (1973). We estimate idiosyncratic returns from a cross-sectional regression:

$$r_{i,t} = \lambda_{0,t} + \lambda_{1,t} \cdot \beta_{i,t-1} + \lambda'_{2,t}\omega_{i,t-1} + \eta_{i,t} \text{ for } i = 1, ..., N_t, \tag{9}$$

where $\beta_{i,t-1}$ and $\omega_{i,t-1}$ a stock i's estimated factor loadings, and $\lambda_{0,t}$, $\lambda_{1,t}$, and $\lambda_{2,t}$ are coefficients. We allow individual stocks' loadings on the factors to vary over time by estimating loadings using high frequency data, and use asymptotic PCA analysis (Connor and Korajczyk 1986, 1988) to jointly estimate $\beta_{i,t-1}$ and $\omega_{i,t-1}$ and high frequency analogs of f_{t+1} and g_{t+1}.[22] These estimated loadings are used as regressors

[22]We estimate loadings on the principal components using 3-months of daily observations from months $t-3$ through $t-1$. Specifically, we first extract principal components, then run a stock-by-stock time series regression to estimate loadings. In this procedure, we use data from prior to t only to avoid mechanical reductions in the variance of $\eta_{i,t}$ related to in-sample fit of the residual in month t. We use a high frequency approach to allow for time-varying loadings while minimizing issues related to missing data from stocks which recently entered the sample. Since the goal of this

in the cross-sectional regression in (9), which allows us to isolate the residual. Our estimate of the granular residual is the weighted average of the stock-level error terms from this regression. In our baseline specification, we use the first three principal components of the cross-section of stock returns as factors, though we demonstrate robustness to alternative statistical representations of the common factors in Section 3.4 below.

In our empirical exercises, *unadjusted index returns* refer to excess returns of the value-weighted stock market index. To address effects from the granular residuals, we construct *adjusted index returns* (r_t^{adj}) by subtracting our empirical measure for the granular residuals from the unadjusted index returns:

$$r_t^{adj} = r_t - \eta_t, \tag{10}$$

where r_t is the *unadjusted index return* during a month t and η_t is our empirical measure for the granular residuals during a month t. Note that r_t^{adj} remains an excess return so long as η_t has zero risk price (i.e., is orthogonal to the SDF).

Table 1 reports summary statistics for the monthly excess returns – both the unadjusted and adjusted index – and the granular residuals. Panel A reports mean, standard deviation, skewness, kurtosis, median and interquartile range (IQR) for those of the value-weighted index. Our estimates indicate that the granular residuals in the CRSP value-weighted index are negatively skewed and have a large kurtosis compared to the normal distribution.[23] The median – a robust measure of central tendency for skewed/leptokurtotic distributions – of the realized granular residuals is very close to zero. More importantly, the standard deviation of the granular residual

regression is to isolate the error term, we note that the same estimates would obtain if we instead were to control for coefficients on rotated versions of the same factors (f_t, g_t). Thus, for convenience, we use estimated loadings on unrotated principal components rather than those estimated on rotated factors when estimating (9) in practice.

[23]The astute reader will observe that the standard deviation of the adjusted index is actually higher relative to the unadjusted one. This happens for two reasons. First, there are a few very large realizations of our estimated granular residual (results are robust to winsorizing these more extreme observations), which inflates the standard deviation estimates. Second, individual stocks' factor loadings are estimated imperfectly, which implies that our estimated of the adjusted index equals $f_t + \sum_{i=1}^{N_t} w_{i,t-1} [\eta_{i,t} - \hat{\eta}_{i,t}]$. This additional term reflects estimation errors in the portion of each stock's return that is explained by common factors and can add to the realized volatility of the adjusted index. Note however that much of variance of this additional term is driven by systematic volatility factors, so its contribution to our volatility estimates may be more likely to reflect systematic sources of uncertainty than the granular residual.

Table 1: Summary statistics of excess returns and granular residuals

Sample	Mean (×100)	Stdev. (×100)	Skew.	Kurt.	Median (×100)	IQR (×100)
Panel A: CRSP value-weighted index						
Unadjusted index returns	0.626	5.407	0.168	10.761	0.985	5.756
Granular residuals	−0.313	3.110	−3.061	25.996	0.068	2.557
Adjusted index returns	0.939	7.281	1.289	16.128	1.147	6.718
Panel B: CRSP equal-weighted index						
Unadjusted index returns	1.715	7.564	2.612	27.061	1.595	6.754
Granular residuals	0.003	0.022	3.768	52.715	0.000	0.009
Adjusted index returns	1.712	7.548	2.614	27.099	1.590	6.731

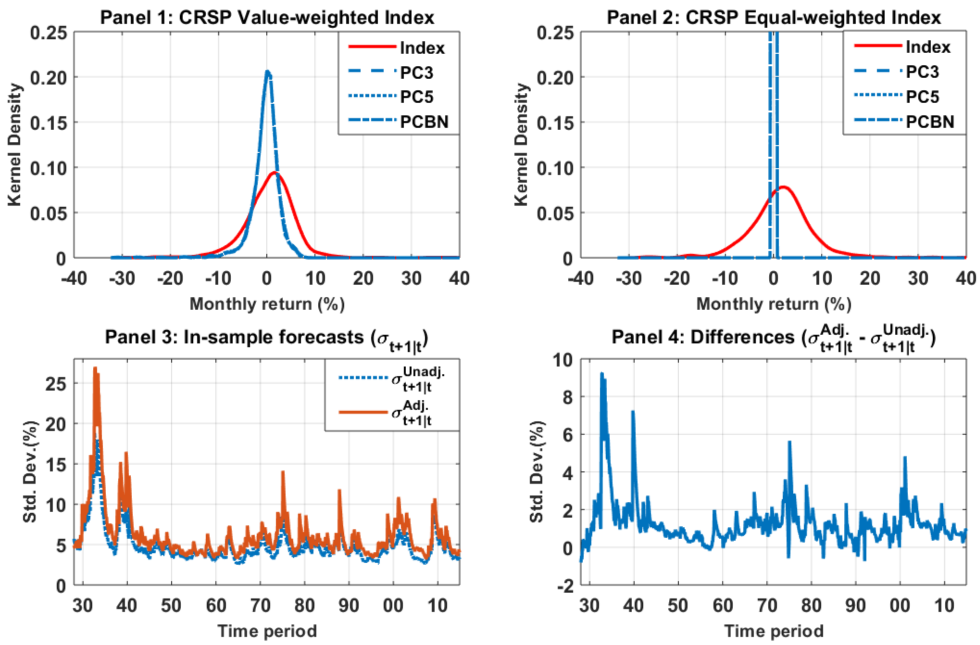
Note: This table provides summary statistics of monthly excess returns and granular residuals in the Center for Research in Security Prices (CRSP) index. The reported excess returns are both unadjusted index returns and adjusted index returns, where the latter returns are constructed by subtracting the granular residuals from the unadjusted index returns. This table reports mean, standard deviation, skewness, kurtosis for those of value-weighted index (Panel A) and of equal-weighted index (Panel B). Given skewness and excess kurtosis appeared in our sample, we also report median and interquartile range (IQR) as robust statistics for mean and standard deviation.

is sizable, equal to 3.1 percent, which is about 3/5 of the standard deviation of the unadjusted index returns. Likewise, the IQR, a robust measure for a variability for leptokurtic distributions is large relative to the unadjusted index.

Panel B reports the same summary statistics for the the equal-weighted index. When all individual stocks have same weights, the granular residuals are small; the standard deviation of the granular residuals is less than 1/100 of that of the unadjusted index returns. Consequently, the summary statistics for the adjusted index returns are almost identical to those for the unadjusted index returns. Small granular contributions arise because idiosyncratic returns are almost totally diversified in the equal-weighted index.

The top two panels in Figure 2 provides a visualization of the findings in Table 1. We present the density of the index return, as well as a kernel estimate of the density of the granular residual as estimated from our baseline specification with three principal components, as well as a variety of other specifications which are considered in our later robustness exercise (see the discussion in Section 3.4 for more details). We overlay these additional specifications to note that our estimates of the magnitude of the granular residual are fairly insensitive to the factor model we use to estimate it. These estimates suggest granular residuals add sizable variation to the value-weighted index (Panel 1), whereas the contribution to the the equal-weighted index is negligible (Panel 2).

Figure 2: Time-series behavior of the granular residual and index returns



Note. This figure highlights the relevance of granular measurement errors for empirical exercises. The top two panels plot kernel density estimates of monthly returns on CRSP indices together with those of granular measurement errors estimated in a number of ways. In each panel, the red solid line plots the density for CRSP value- (Panel 1) and equal-weighted (Panel 2) indices while various blue dotted lines indicate the corresponding densities of estimated granular measurement errors. Since, in Panel 2, the kernel density of all granular measurement errors is extremely concentrated around 0, we truncate the vertical axis. Then, the left bottom panel (Panel 3) plots in-sample forecasts for the conditional volatility in the CRSP value-weighted index, obtained by estimating the GARCH(1,1) model. The red solid line represents the conditional volatility forecasts obtained from the adjusted index returns, whereas blue dotted line represents those from the unadjusted index returns. Panel 4 plots differences in those volatility forecasts. Here a positive difference indicates a smaller conditional volatility forecast from the unadjusted index returns than that from the adjusted index returns.

3.3 Baseline risk-return tradeoff estimates

In this subsection, we investigate the effects of the granular residuals on the estimates for the risk-return tradeoff relationship using GARCH volatility forecasting models. To demonstrate the effects of the granular residuals, we estimate coefficients in the GARCH volatility forecasting model using excess returns of the CRSP value-weighted index ('Unadjusted') and also using excess returns that are adjusted by the granular residuals ('Adjusted'). Specifically, our empirical tests will estimate the following specification

$$r_{t+1} = \phi_0 + \phi_1 \cdot r_t + \gamma \cdot \sigma^2_{m,t+1} + u_{t+1}, \tag{11}$$

$$u_{t+1} = \sigma_{m,t+1}\varepsilon_{t+1}, \tag{12}$$

$$\sigma^2_{t+1} = \alpha_0 + \alpha_1 u^2_t + \alpha_2 \sigma^2_t, \tag{13}$$

where u_t is a forecasting error from the previous period. Here the conditional variance forecasts (σ^2_{t+1}) following the GARCH(1,1) process can be viewed as a weighted average of the squared forecasting error and the conditional variance forecast from the previous period, with corresponding weights being α_1 and α_2 respectively. We provide unreported coefficient estimates in Appendix B. Throughout this section, we estimate parameters in the GARCH volatility forecasting models by numerical maximum likelihood estimation (MLE). After estimating parameters, we calculate their asymptotic standard errors by approximating the first and second derivatives of the log-likelihood function at maximum likelihood estimates.[24]

Panels 3 and 4 in Figure 2 plot the levels and differences between our baseline volatility measures for the adjusted and unadjusted indices, respectively. As is clear from Panel 4, the difference between the two exhibits substantial variation at both high and low frequencies. This nontrivial difference lends empricial support to our conjecture that granular residuals add noise to the lagged forecasting errors used to generate the conditional variance forecasts within the GARCH forecasting models.[25]

While we adopt the GARCH(1,1) as our baseline volatility forecasting model, there are numerous extensions of GARCH models highlighting characteristics such as volatility persistence, asymmetry and a leptokurtic distribution of financial asset returns. Accordingly, we complement our baseline GARCH(1,1) with estimates from two nonlinear GARCH models such as a threshold GARCH (GJR-GARCH(1,1)) of Glosten et al. (1993) and an exponential GARCH (EGARCH(1,1)) in Nelson (1991). These two models, whose specifications for volatility are detailed in Appendix B, are widely adopted in the empirical literature for the risk-return trade-off given their ability to capture "leverage effects", a known feature of financial asset volatility.[26]

[24]See details for numerical MLE and calculation of asymptotic standard errors in Hamilton (Time Series Analysis, 1994, pp. 133-148).

[25]For example, Byun (2016) shows that cross-sectional dispersion in the returns of different stocks, a measure of aggregate idiosyncratic risk similar to our notion of granular residuals, does not help forecast volatility of the S&P 500 index when using the cross-sectional dispersion as an additional explanatory variable in the GARCH forecasting models. See Byun (2016) for detailed explanation and also for alternative channel that the cross-sectional dispersion could improve volatility forecasts indirectly.

[26]See, e.g., French et al. (1987), Schwert (1990), Franses and Van Dijk (1996) , and also Poon and Granger (2003) and Brownlees et al. (2012) for reviews of the literature.

Table 2: Risk-return tradeoff coefficient (γ) across GARCH models

	Conditional Variance ($\sigma^2_{t+1\|t}$)		Conditional Volatility ($\sigma_{t+1\|t}$)	
Models	Value	Equal	Value	Equal
Panel A: GARCH				
Unadjusted	1.319	2.160***	0.140	0.427***
	(0.920)	(0.532)	(0.103)	(0.104)
Adjusted	1.550**	2.165***	0.286***	0.427***
	(0.618)	(0.532)	(0.103)	(0.104)
$\gamma^{Unadj.} - \gamma^{Adj.}$	−0.231	−0.005	−0.146	−0.001
Panel B: GJR-GARCH				
Unadjusted	1.150	2.815***	0.109	0.411***
	(1.131)	(0.911)	(0.124)	(0.116)
Adjusted	1.986***	2.819***	0.288**	0.411***
	(0.705)	(0.908)	(0.118)	(0.116)
$\gamma^{Unadj.} - \gamma^{Adj.}$	−0.837	−0.004	−0.179	−0.001
Panel C: EGARCH				
Unadjusted	1.604	2.983***	0.123	0.346***
	(1.058)	[0.616]	[0.102]	[0.090]
Adjusted	1.989***	2.989***	0.211**	0.346***
	[0.598]	[0.617]	[0.095]	[0.090]
$\gamma^{Unadj.} - \gamma^{Adj.}$	−0.385	−0.006	−0.088	−0.001

Note: This table compares γ estimates and standard errors (in parentheses) from the unadjusted index returns with those from the adjusted index returns. The first two columns report γ estimates obtained from our baseline specification, which adopts conditional variance ($\sigma^2_{t+1|t}$) as a proxy for risk. The last two columns report those from alternative specification, which adopts conditional volatility ($\sigma_{t+1|t}$) as a proxy for risk. We estimate parameter estimates using excess returns of CRSP value- and equal-weighted indices. We consider three popular GARCH volatility forecasting models such as a GARCH (1,1) (Panel A), a Threshold GARCH model of Glosten et al. (1993) (Panel B) and an exponential GARCH model of Nelson (1991) (Panel C). In each panel, the first four rows report maximum likelihood estimates and robust standard errors of γ except for a few parameter estimates from EGARCH models. Their standard errors, reported in the square brackets, are obtained by approximating the information matrix by evaluating outer products at maximum likelihood estimates. The fifth row ($\gamma^{Unadj.} - \gamma^{Adj.}$) reports differences in γ estimates. */**/*** represent the statistical significance at 90%, 95% and 99% respectively.

The first column in Table 2 compares γ estimates from excess returns on the unadjusted value-weighted index ('Unadjusted') with those from the adjusted index ('Adjusted'). In Panel A, we find a positive and statistically significant estimate for γ from the adjusted index returns, which contrasts with a postive but statistically insignificant estimate for γ from the unadjusted index returns. Furthermore, consistent with the attenuation bias discussed earlier, our estimate for γ from the unadjusted index returns is smaller than that from the adjusted index returns. These results are robust across all three alternative volatility forecasting models (Panels B and C), where the attenuation bias is even larger with models featuring leverage effects.

In the second column of Table 2, we estimate risk-return tradeoff coefficients for excess returns of the CRSP equal-weighted index. In contrast to our results for the value-weighted index, unadjusted and adjusted index returns to provide similar estimates for the risk-return tradeoff for the equal-weighted index given that the number of stocks in the equal weighted index is sufficiently large for the law of large numbers to hold almost exactly. Accordingly, our corrections matter little in this case: estimates of γ are nearly invariant to index adjustment across all volatility forecasting models. In a result that, to our knowledge, is new to the literature, we find strong evidence of a risk-return tradeoff. Our γ estimates are large and highly statistically significant for both adjusted and unadjusted equal-weighted indices.

Before moving on to the next section, we further investigate the risk-return tradeoff relationship under an alternative model specification. One possibility is to adopt conditional volatility rather than conditional variance as a proxy for a time-varying risk, as studied by Baillie and DeGennaro (1990) and French et al. (1987):

$$r_{t+1} = \phi_1 + \phi_2 \cdot r_t + \gamma \cdot \sigma_{t+1} + u_{t+1} \tag{14}$$

The third and fourth columns in Table 2 report estimates for the risk-return relationship under (14). The estimation results are consistent with those in the first two columns. For the value-weighted index, the adjusted index returns provide positive and statistically significant γ estimates, which are larger in magnitude than those from the unadjusted index returns. For the equal-weighted index, we confirm that both the unadjusted and adjusted index returns provide similar-sized, statistically significant, γ estimates across all volatility forecasting models considered.

3.4 Robustness exercises

Next, we discuss further robustness checks for the effects of the granular residuals in the evaluation of the risk-return relationship. Specifically, we summarize further empirical evidence obtained from alternative market definitions, alternative econometric specifications for estimating the granular residual and volatility of the true factor, alternative sample periods, and alternative approach. We place unreported results from various robustness checks at Appendix B.

The sample composition of the CRSP universe exhibits non-trivial variation over time. As observed above, the number of stocks increases dramatically as new markets are added. Here, Table 3 considers the robustness of our main findings to alternative

Table 3: Risk-return tradeoff coefficient (γ) with alternative market indices

	Conditional Variance ($\sigma^2_{t+1\|t}$)		Conditional Volatility ($\sigma_{t+1\|t}$)	
Models	Value	Equal	Value	Equal
Panel A: All NYSE stocks				
Unadjusted	1.351	2.142***	0.145	0.406***
	(0.898)	(0.468)	(0.102)	(0.095)
Adjusted	1.493**	2.147***	0.270***	0.406***
	(0.585)	(0.470)	(0.095)	(0.095)
$\gamma^{Unadj.} - \gamma^{Adj.}$	−0.142	−0.005	−0.125	−0.001
Panel B: All NYSE-AMEX-NASDAQ stocks, excluding bottom decile by size				
Unadjusted	1.311	2.207***	0.139	0.340***
	(0.919)	(0.738)	(0.104)	(0.109)
Adjusted	1.930**	2.309***	0.294***	0.355***
	(0.792)	(0.751)	(0.111)	(0.109)
$\gamma^{Unadj.} - \gamma^{Adj.}$	−0.619	−0.102	−0.155	−0.015
Panel C: S&P 500 (1964.7 - 2014.12)				
Unadjusted	2.230	4.071**	0.147	0.378*
	(1.719)	(1.953)	(0.146)	(0.207)
Adjusted	2.895*	4.080**	0.236	0.378*
	(1.753)	(1.954)	(0.172)	(0.210)
$\gamma^{Unadj.} - \gamma^{Adj.}$	−0.665	−0.009	−0.089	−0.000

Note: This table compares γ estimates and standard errors (in parentheses) from the unadjusted index returns with those from the adjusted index returns. The first two columns report γ estimates obtained from our baseline specification, which adopts conditional variance ($\sigma^2_{t+1|t}$) as a proxy for risk. The last two columns report those from alternative specification, which adopts conditional volatility ($\sigma_{t+1|t}$) as a proxy for risk. We estimate parameter estimates using excess returns of the CRSP value- and equal-weighted indices of equities in New York Stock Exchange (Panel A), and the NYSE-AMEX-NASDAQ universe excluding the bottom decile of stocks using NYSE size breakpoints (Panel B), and value- and equal-weighted portfolios of stocks in the S&P 500 index (Panel C). In each panel, the first four rows report maximum likelihood estimates and robust standard errors of γ. The fifth row ($\gamma^{Unadj.} - \gamma^{Adj.}$) reports differences in γ estimates. Panel C uses a shorter sample period from July 1964 to December 2014, whereas other panels use the total sample period from January 1928 to December 2014. */**/*** represent the statistical significance at 90%, 95% and 99% respectively.

market definitions. First, we use the CRSP value-weighted index consisting only of equities in the New York Stock Exchange, for comparison with the results of earlier research.[27] The results from alternative sample are consistent with those reported in Table 2 – insignificant, smaller, and less precisely-estimated coefficients for the unadjusted index. Also, following Lewellen (2015), we consider an index of "all but

[27]For example, Chou (1988) finds a positive and statistically significant relationship using GARCH-M over the weekly sample from July 1962 to December 1985. Although finding a positive estimate using the same model, French et al. (1987) document a statistically insignificant relation from the monthly sample spanning February 1928 to December 1984. In contrast, Glosten et al. (1993) document a negative relationship from April 1951 to December 1989 using modified GARCH-M models.

small" stocks. This restriction has little impact on the value-weighted index, but it could in principle have had a large effect on the equal-weighted index because a large number of tiny stocks (below the 10th NYSE size percentile breakpoint) are excluded from the analysis.[28] However, our results for the strong risk-return tradeoff for the equal-weighted portfolio are quite unaffected by this narrower sample restriction. Finally, we repeat the analysis for stocks in the S&P 500, for which we can get a list of index constitutents starting in 1964. Estimates for the adjusted value-weighted index are slightly noisier though quite consistent with our findings above.

Our baseline specification uses three principal components to estimate idiosyncratic residuals and the entire CRSP sample period. In Appendix Table B3, we construct alternative measures of granular residuals using estimates for idiosyncratic returns relative to various other specifications, including other principal components specifications and the Fama-French 3 factors. Results are quite similar across factor specifications and, while we only report GARCH(1,1) estimates for brevity, across volatility forecasting models. In Appendix Table B4, we also evaluate the risk-return relationship during two subsample periods split in July 1962, motivated by earlier researchers' finding of a weaker risk-return relationship over the early sample period (Turner et al. 1989, Chou et al. 1992, Whitelaw 1994).[29] Results are similar to those presented above and, consistent with findings in the literature, stronger in the later sample.

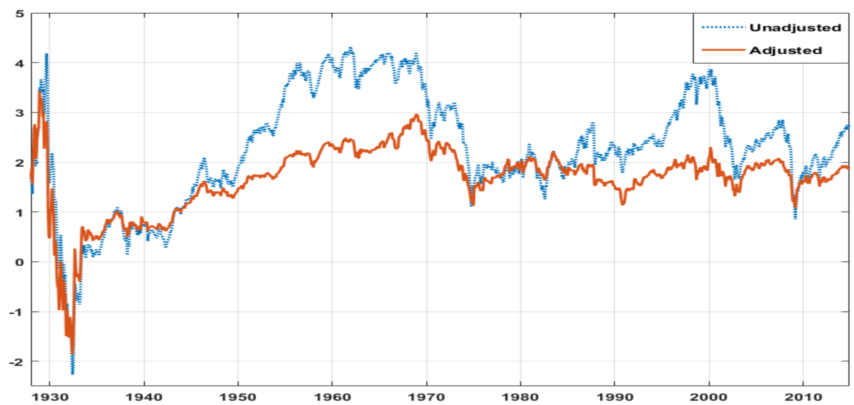
Finally, we also consider a bivariate model approach that is motivated from insignificant contribution of granular residuals to the volatility of the CRSP equal-weighted index.[30] Specifically, we use the the equal weighted index return to indirectly infer the volatility of the value-weighted index, which allows us to obtain an estimate of the risk-return tradeoff for the value-weighted index by maximizing a

[28]Note that estimated principal components changes with the set of stocks considered in the market portfolio, so the value-weighted estimates could change even if the fraction of market value associated with the small stocks is very small.

[29]We split our sample in July 1962 since most new entrants into the CRSP index occur after this point. The second subsample encompasses sample periods in which the risk-return trade-off has been studied by earlier researchers (Chou 1988, Baillie and DeGennaro 1990, Nelson 1991, Goyal and Santa-Clara 2003 and Guo and Whitelaw 2006 to name a few).

[30]The alternative approach shares the intuition of Goyal and Santa-Clara (2003), who find a positive risk-return relationship by regressing excess returns of the value-weighted index on the realized variance of the equal-weighted index returns ('average stock variance'). The robustness of the earlier result was questioned by a few researchers due to a lack of for an extended sample period (Wei and Zhang, 2005), because of a selection bias (Bali et al., 2005), and because of a regime-dependent relationship (Angelidis and Tessaromatis, 2009).

Figure 3: Risk return tradeoff coefficients from time-varying parameter models



Note. This figure plots estimates for time-varying risk-return tradeoff coefficient (i.e., price for risk), obtained by estimating the time-varying parameter (TVP) model of Chou et al. (1992). The TVP model allow the risk-return tradeoff coefficient (γ) to be time-varying while maintaining the same conditional variance forecasts following the GARCH(1,1) model. The (red) solid line represents risk-price estimates obtained from the unadjusted index returns, whereas the (blue) dotted line represents those from the unadjusted index returns.

joint likelihood for the observed value-weighted and equal-weighted returns. Results, which are reported in Appendix Table B7, are similar to the more simple adjustment proposed in the main text.

3.5 Inferences about time-varying risk prices vs quantities

As discussed above, granular measurement errors can generate a time varying attenuation bias which can resemble time-varying risk prices. With this in mind, we revisit the GARCH forecasting exercise from above and allow for a time-varying risk-return tradeoff, following Chou et al. (1992), for Unadjusted and Adjusted index returns, respectively.[31] The model thus decomposes the risk premium into a time-varying price of risk (γ_t) and a time-varying quantity for risk ($\sigma^2_{t+1|t}$).

Figure 3 compares γ_t estimates from Unadjusted and Adjusted index returns. Other than a large decline in estimated risk prices – likely driven by the outsized impact of the Great Depression – at the start of our sample, estimated time-varying risk prices are much less volatile for the the Adjusted index relative to its Unadjusted counterpart.[32] From 1960 onwards, Adjusted estimates are quite stable while

[31]In Appendix B.4, we provide detailed explanations of the estimation procedure and report conditional volatility forecasts (quantity of risk), which are quite similar to Figure 2.

[32]Related to this finding in the time series, in unreported additional analysis, we find similar

unadjusted estimates feature substantial low frequency fluctuations. Consistent with our formal argument above, Figure 3 suggests that granular measurement errors may indeed lead to estimates that overstate the importance of time variation in risk prices.

4 Cross-section: Granular residuals, market beta, and a reassessment of the size anomaly

In this section, we consider the implications of granular residuals for cross-sectional asset pricing tests. As discussed above, the presence of a non-trivial granular residual can generate substantial biases in estimates of stock or portfolio-level exposures to market risk (β_i). In section 4.1, we provide some analytical results on how unpriced granular residuals also influence tests of the cross-sectional predictions of the model outlined in Section 2 above, then present our empirical results.

4.1 Econometric biases

Suppose that a researcher has an estimate of f_{t+1} as well as some other factors g_{t+1} which, so as to be conformable with the theory as written above, have been orthogonalized with respect to the market factor f_{t+1}. When the factors f_{t+1} and g_{t+1} are tradable portfolios, the law of iterated expectations and (3) imply:

$$E[r_{i,t+1}] = E[E_t\left[r_{i,t+1}\right]] = \beta_i \cdot \gamma \cdot E[\sigma^2_{t+1}] + E[\Lambda' \Omega_t \omega_i] = \beta_i E[r_{t+1}] + \omega_i' E[g_{t+1}]. \quad (15)$$

where, in the multifactor model above, the second term captures a stock's exposure to other priced factors which are orthogonal to f_{t+1}. Thus, the main empirical prediction of the theory is that, after controlling for exposures to other orthogonal priced factors, expected returns line up with β_i, where the slope of the relationship is equal to the equity premium.

Since it is common to conduct cross-sectional asset pricing tests using portfolios, define analogous measures (β_p, ω_p') for the return on a portfolio $r_{p,t+1}$. Usually, researchers estimate factor loadings by running the following portfolio-level regression:

$$r_{p,t+1} = \alpha_p + \beta_p f_{t+1} + \omega_p' g_{t+1} + \eta_{p,t+1}. \quad (16)$$

effects on estimates of the market risk premium obtained from the cross-section using Fama-MacBeth regressions. These results are available upon request.

Given data for a single portfolio or a small number of portfolios, one would often directly test the theoretical prediction that the intercept $\alpha_p = 0$.[33]

A standard approach for estimating (16) is to run an OLS regression of a portfolio's excess return on r_{t+1} and other factors g_{t+1}, where g_{t+1} is assumed to be tradable and rotated so as to be orthogonal to f_{t+1} and η_{t+1}. Due to this orthogonality assumption, the multivariate OLS estimate of β_p and its univariate (i.e., excluding g_{t+1}) counterpart both converge to

$$\underset{T\to\infty}{\text{plim}}\widehat{\beta}_p^{OLS} = \beta_p \cdot \left\{ \frac{Var\left[f_{t+1}\right]}{Var\left[f_{t+1}+\eta_{t+1}\right]} \right\} + \frac{Cov\left[\eta_{t+1}, \eta_{p,t+1}\right]}{Var\left[f_{t+1}+\eta_{t+1}\right]} \equiv \beta_p + \kappa_p, \tag{17}$$

where κ_p denotes the bias in β_p, which will be useful later. Hence $\widehat{\beta}_p^{OLS}$ is affected by the granular residuals in two ways; a downward bias represented by the term in the curly brackets and an upward bias represented by the direct correlation between the granular residual η_{t+1} and the comparable idiosyncratic error term of the portfolio $\eta_{p,t+1}$ (itself a weighted average of stock-level η terms). While the upward bias in the second term is generally small except for a portfolio comprised of very large stocks, the term in the curly brackets is substantially less than 1, indicating that $\widehat{\beta}_p^{OLS}$ is likely to be attenuated downwards for most individual stocks. However, if the researcher uses value-weighted portfolios as test assets, the direction of the bias is not immediately clear: note that these two forces exactly offset for the market portfolio, which has a true loading on the true factor of 1 (by assumption/normalization) as well as a slope coefficient of 1 when regressed on itself.

Biases in β_p generate corresponding biases in estimates portfolio α_p's, which should equal 0 when the theory holds. Under our assumptions on g_{t+1}, α_p converges to

$$\underset{T\to\infty}{\text{plim}}\ \widehat{\alpha}_p^{OLS} - \alpha_p = \underset{T\to\infty}{\text{plim}}\ \widehat{\alpha}_p^{OLS} = \kappa_p \cdot E\left[r_{t+1}\right], \tag{18}$$

so, given a positive market risk-premium, $\widehat{\alpha}_p^{OLS}$ does not line up with the theoretical prediction of zero unless $\widehat{\beta}_p^{OLS}$ is unbiased, and downward-biased betas lead to positive

[33] Note that β_p is the plim of the OLS regression coefficient which would obtain from regressing $r_{p,t+1}$ on the true factor f_{t+1} and orthogonal factors g_{t+1}, where $\eta_{p,t+1}$ is the residual from that regression. Time variation in the weights will imply that the portfolio-level loadings are time-varying. If this is the case, $\eta_{p,t+1}$ includes a term which equals $(\beta_{pt} - \beta_p)f_{t+1} + (\omega'_{pt} - \omega'_p)g_{t+1}$. However, it is more standard to assume that portfolio-level loadings are constant in empirical work.

alphas.[34] Therefore, biases in $\widehat{\beta}_p^{OLS}$ introduced by granular measurement errors would lead a researcher to reject the theory model even if it holds.

Given data for multiple portfolios and estimated factor loadings, one can also test equation (15) using time series averages of returns for multiple portfolios by

$$\frac{1}{T}\sum_{t=1}^{T} r_{p,t} = \lambda_0 + \lambda_1 \hat{\beta}_p + \lambda_2' \hat{\omega}_p + u_p. \tag{19}$$

Note that, under our orthogonalization assumptions, $\hat{\omega}_p$ is consistently estimated even if the presence of the granular residual generates biased estimates of $\hat{\beta}_p$. Thus, while it is more common to run the multiple regression and freely estimate λ_2, the algebra is simpler – and estimates would be more precise under the null that portfolios were priced properly – if we instead impose the theoretical restriction that $\lambda_2 = E[g_{t+1}]$. Thus, we can estimate λ_1 by a univariate regression of average returns of portfolios which are modified to be "neutral" with respect to the other factors, $\frac{1}{T}\sum_{t=1}^{T} r_{p,t} - \hat{\omega}_p' \frac{1}{T}\sum_{t=1}^{T} g_t$, on a constant and $\hat{\beta}_p$.

If one estimates λ_1 from this second-pass cross-sectional regression, the implied market risk premium – the slope of the securities market line – converges to

$$\underset{T,N\to\infty}{\text{plim}}\ \widehat{\lambda}_1 = E[r_{t+1}] \cdot \left\{ \frac{Var\left[\beta_p\right] + Cov\left[\kappa_p, \beta_p\right]}{Var(\beta_p) + Var\left[\kappa_p\right] + 2Cov\left[\kappa_p, \beta_p\right]} \right\}. \tag{20}$$

This expression is similar to one which appears in equation (7). Here, the use of a mis-measured proxy for the priced factor is associated with a multiplicative bias in $\widehat{\lambda}_1$, of which direction is not immediately clear. $\widehat{\lambda}_1$ is downward biased if and only if $Var\left[\kappa_p\right] \geq -Cov\left[\kappa_p, \beta_p\right]$.[35] In the applications we consider, we find that the covariance term is strongly negative (small stocks tend to have very high β_p, accordingly large negative bias), suggesting that $\widehat{\lambda}_1$ is upward-biased empirically. Thus, a researcher could reject the prediction that the slope of the securities market line λ_1 equals $E[r_{t+1}]$ even when it holds with the true β_p.

[34] To see this, note that $\widehat{\alpha}_p^{OLS} = \frac{1}{T}\sum_{t=1}^{T} r_{p,t} - \hat{\beta}_p \frac{1}{T}\sum_{t=1}^{T} r_t - \hat{\omega}_p' \frac{1}{T}\sum_{t=1}^{T} g_t$. When $Cov_t[f_{t+1}, g_{t+1}] = Cov_t[\eta_{t+1}, g_{t+t}] = 0$, $\underset{T\to\infty}{\text{plim}}\ \hat{\omega}_p = \omega_p$ regardless of whether the true factor f_{t+1} or the mis-measured proxy $f_{t+1} + \eta_{t+1}$ is used in the regression to estimate the factor loadings. Therefore, biases in $\hat{\alpha}_p$ will be driven solely by biases in $\hat{\beta}_p$ under the null that the theory holds.

[35] For instance, this would happen if $Corr(\kappa_p, \beta_p) \approx -1$ (as would happen if all portfolios had a common multiplicative bias in betas) and $Var(\beta_p) > Var(\kappa_p)$.

4.2 Methodology

4.2.1 Estimating β_p consistently

In order to examine whether these biases are empirically meaningful in the cross-section, we require an estimate of β_p which remains consistent in the presence of granular measurement errors. We propose a simple instrumental variables-like (IV) approach to address this potential concern, where we use principal components (PCs) extracted from the panel of CRSP stocks as instruments for the true market factor. PCs are a popular in the empirical finance literature as a way to span the dimensions of risk in a linear setting without taking a stand on the economic identity of the risk factors.[36] Also, these PCs will satisfy a necessary condition for being valid (i.e., $\eta_{p,t+1}$ will be asymptotically uncorrelated with the PCs for all portfolios) provided that the number of stocks becomes arbitrarily large relative to the number of factors.[37]

Because the set of stocks we use to construct the PCs and individual stocks' factor loadings may vary over time, we estimate PCs and portfolio loadings over rolling windows. Our baseline specification uses three PCs over 60 month rolling windows in the post-1963 sample of returns, though results are robust to different numbers of PCs, window sizes, and start dates, as described below. Over each window, we extract the PCs from the set of CRSP stocks which have a return reported for every month in the window. Our IV estimate is obtained by a procedure similar to 2-stage least squares (2SLS). Specifically, within each window, we project the market excess return on the PCs and a constant. Then, in the second stage we recover β_p^{IV} by regressing the portfolio excess return $r_{p,t+1}$ onto the projection of the market onto the PCs and a constant. In some cases, other pricing factors will be orthogonalized with respect to this linear projection, which will imply that IV betas from univariate and multivariate regressions will be numerically identical.[38] We contrast our IV results with standard OLS estimates obtained by regressing each portfolio return (possibly

[36] See, e.g., Giglio and Xiu (2017) for a recent use and discussion of PCs spanning properties.

[37] Having a large number of stocks relative to the number of factors is necessary but not sufficient. If f_{t+1} is not spanned by the PCs and exposures to the unspanned component of f_{t+1} vary across portfolios, we will not obtain consistent estimates of β_p. Our finding that results are stable as we vary fitting window lengths and number of PCs is consistent with omitted factors not being a large issue, but, as with any exclusion restriction in an IV procedure, it is ultimately untestable.

[38] This exact numerical result will not hold for OLS estimates, though it turns out approximately to be the case empirically. Because of the orthogonalization we impose on g_{t+1} so as to be conformable with the theory, ours is slightly more complicated than a standard 2SLS procedure. However, we will nonetheless use the shorthand "IV" when describing it.

after subtracting estimated exposures to orthogonal factors) on a constant and the CRSP value-weighted excess return.

4.2.2 Controlling for other orthogonal factors and time aggregation

Even if we estimate β_p consistently, our estimates of λ_1 in equation (19) are not necessarily correct unless we control for exposure to other common, priced factors g_{t+1}, or if it happens to be the case that β_p and ω_p are uncorrelated in the cross-section or $\lambda_2 = 0$. Since the seminal papers of Fama and French (1993, 1996), the size and value premiums have featured heavily in the empirical finance literature. When one sorts stocks based on their market capitalization or book-to-market, one is able to generate large α_p relative to the CAPM. Below, we will argue that the size premium shrinks considerably once we control for the effects of the granular residual, a result which is in line with the predictions of the analysis in the prior section. However, a sizable value premium remains, consistent with exposures to a common value factor helping to explain the cross section of expected returns.

Recall that our theoretical framework above assumed that additional factors g_{t+1} are orthogonal to the true market factor f_{t+1}. Accordingly, for each rolling window, we propose a simple procedure to construct a version of the Fama-French High Minus Low (HML) factor which is orthogonal to the projection of the market return onto the principal components.[39] Specifically, within each window, we project the HML factor on the linear projection of the market return onto the PCs. From this regression, we compute the orthogonal HML (OHML) factor as the constant plus the regression residual.[40] We then estimate $\omega_{p,t}$ (where the time subscript indicates the last period of the rolling window used in the estimation) on a constant and OHML. The "second stage" regression is run with $r_{p,t-j} - \omega_{p,t} OHML_{t-j,t}$, $j \in \{t-H+1, t\}$ as the dependent variable.[41] This keeps portfolio means the same across IV and OLS specifications and makes the graphs easier to interpret, as only beta differs between

[39] See Appendix C for a more thorough description.

[40] In other words, OHML is constructed by taking the original HML factor and subtracting off the portion of its return explained by its estimated loading on the projection of the market return onto the PCs. The first order condition of the OLS regression guarantees that the covariance between OHML and our IV is 0 by construction over the rolling window, though it makes the definition of OHML dependent on the specific fitting window in question.

[41] As mentioned above, for the IV estimates, these results are numerically identical to those obtained by running a multivariate regression with a constant, the projection of the market onto the PCs, and OHML as regressors for each rolling window.

methods. Controlling for exposures to the orthogonal value factor has nontrivial effects on our estimates of λ_1 below.

For each rolling fitting window, we obtain an estimate of $\beta_{p,t}$ and an estimate of the expected portfolio return, which is either a rolling average of portfolio excess returns ($\bar{r}_{p,t} \equiv \frac{1}{H}\sum_{j=0}^{H-1} r_{p,t-j}$) or an average excess return after subtracting exposures to orthogonal pricing factors – ($\bar{r}_{p,t}^{ORTH} \equiv \frac{1}{H}\sum_{j=0}^{H-1}[r_{p,t-j} - \omega_{p,t} OHML_{t-j,t}]$), where H is the length of the fitting window and t is the time subscript associated with associated with the last observation in each fitting window.[42] To get a single scalar estimate for each portfolio excess return and beta, we report the average of these estimates over all fitting windows. All statistics are corrected for the overlapping structure of these estimates using a weighted bootstrap procedure described in Appendix C.

4.2.3 Choice of sample period and test assets

All of our test portfolios are sourced from Kenneth French's data library.[43] Consistent with most cross-sectional asset pricing papers, we use monthly data for US stocks. In addition to being a standard source, these portfolios are constructed in a standardized way and, conveniently, the tables report value-weighted returns as well as the average market cap of stocks in each portfolio. These statistics allows us to test predictions above which relate biases in β_p (i.e. κ_p) with firm size.

We include in our sample (1) all portfolios constructed by forming deciles based on a single characteristic, which yields 15 portfolios, and (2) all portfolios constructed by double sorting into 5 size quintiles and 5 (or 7, in the case of net income) quintiles based on additional characteristics, which yields another 9 portfolios. These portfolios are listed in Appendix Table C1. These double-sorts which condition on size are particularly useful for illustrating the relationship between average firm sizes and κ_p. Since some portfolio returns only available post 1963, and also because the CAPM is known to perform better over the 1927-1962 period, we focus on the post-1963 sample here. Similar results obtain over the full sample, which we report in the Appendix.

[42] Conclusions from models with additional pricing factors included are quite similar, so we do not report these results for brevity. However, they are available upon request from the authors.

[43] `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`

4.3 Empirical results

4.3.1 Representative portfolios

We begin by studying the effects of correcting for the granular residual for some well-known portfolio sorts from the empirical finance literature, before showing that our findings are robust across many different sorts and specifications. In Figure 4, we display our OLS and IV estimates for the 10 size sorted portfolios, the 25 size by book-to-market (BM) sorted portfolios, the 10 market beta sorted portfolio, and the 25 size by beta sorted portfolios. The left panel plots raw univariate betas, while the right panel orthogonalizes the portfolios with respect to OHML. Blue numbers plot $(\beta_{IV}, \bar{r}_i)$, where $\bar{r}_i$ is the mean return of portfolio i.[44] Similarly, red numbers indicate $(\beta_{OLS}, \bar{r}_i)$. Analogously colored lines display the fitted least-absolute deviation (LAD) line for corresponding to each set of portfolios.[45] The black line plots the expected return-beta relationship (securities market line, "SML") implied by the theory.

Panel 1, at the top left, plots size-sorted portfolios. Other than the largest size portfolio, almost all the IV betas are larger then the corresponding OLS ones, where the rightward horizontal movements (indicating a downward bias in OLS betas) are decreasing in size, which is inversely associated with both OLS and IV betas. Since mean returns on the vertical axis are the same, this generates an IV-SML ($\lambda_1\beta_p$) which is considerably flatter and closer to the CAPM-implied value relative to its OLS counterpart. Correspondingly, pricing errors α_p – vertical distances between expected returns and the black line – are almost uniformly smaller for all 10 size portfolios, consistent with the size premium being essentially eliminated by addressing the attenuation bias associated with the granular residual. Panel 2 shows that results are unchanged by controlling for effects of OHML exposure.

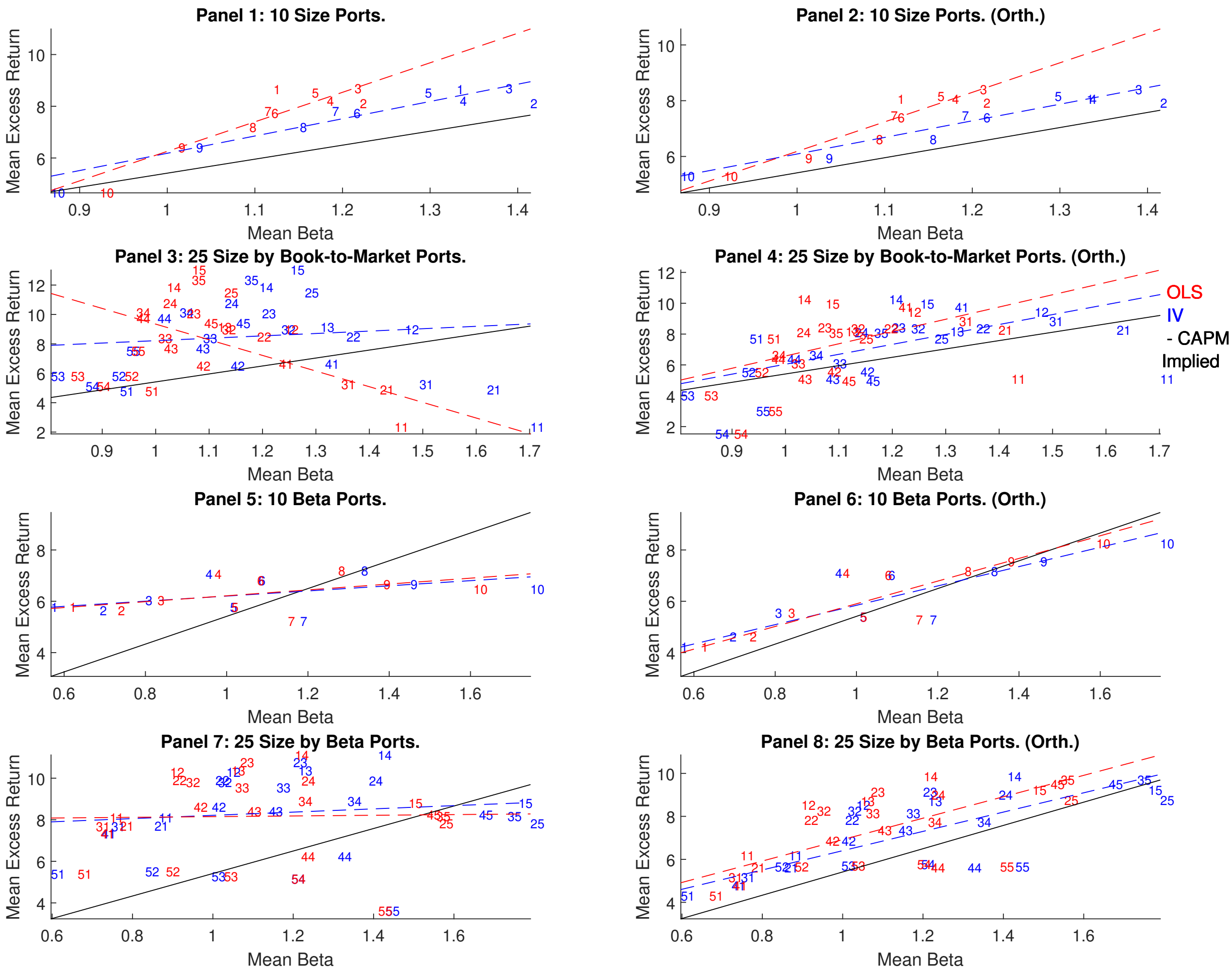
Panel 3 displays the same variables for the 25 size and book-to-market sorted portfolios.[46] The OLS-SML is negatively sloped indicating a negative risk-return relationship. The IV-SML is positively sloped, on the other hand, though the line is too flat relative to the prediction of the theory. With the exception of the small-growth

[44]Since betas are estimated on a 60 month rolling window, the mean returns are also calculated over each window. Thus, $\bar{r}_i$ is the average of the $\bar{r}_{p,t}$ over rolling windows. Similarly, the betas are the mean betas over windows.

[45]We use LAD instead of OLS since estimates are less sensitive to several "corner portfolios" in the 25 portfolio sort (e.g., small growth stocks) which are often difficult to explain via factor models.

[46]The numbering scheme is as follows: the first and second numbers indicate quintile of the first and sorting variable, respectively. For example, portfolio 55 means "large-value."

Figure 4: Expected return-beta relationship for size, size × book-to-market, market beta, and size × market beta sorted portfolios



Note. This figure displays SML estimates for various portfolio sorts, with and without the orthogonalization procedure explained in the main text. The left column is unorthogonalized, while the right is orthogonalized. The title of each plot is the name of the portfolio sort used to construct the colored SMLs. The red line estimates betas on portfolios using OLS over 60 month rolling windows, the blue line uses 3 PCs to instrument for the market before calculating the SML, and the black line is the CAPM implied SML (slope equal to market risk premium). All quantities are calculated over the 60 month window and then averaged across windows. Each number refers to the portfolio number in the sort. For the 25 portfolio sorts, the first number refers to the first sorting variable (size), and the second number refers to the second sorting variable (either book-to-market or beta).

portfolio (whose expected returns are notoriously hard to explain), most portfolios' expected returns move closer to the CAPM-implied SML. Once again, this happens because most expected returns lie above the theoretical SML and most IV betas are noticeably larger than their OLS counterparts. Other than portfolios in the largest size category (whose IV modestly decrease), all IV betas shift to the right.

Panel 4 of Figure 4 repeats the analysis of Panel 3 with the orthogonalized portfolios. In this case, portfolios' OHML exposures differ substantially and our orthogonalization procedure significantly steepens the estimated OLS beta risk-return relationship. However, the IV risk-return relationship is also changed so that the IV-SML is even closer to the CAPM-SML. In addition, despite the fact that the dispersion in IV betas is usually higher than OLS, expected returns more closely align with IV betas than OLS ones.

The third and fourth row of Figure 4 plots results for 10 portfolios sorted by market beta and 25 portfolios sorted by size and beta, respectively. Beginning with Panel 5, again we observe substantial differences between OLS and IV betas. In this case, high OLS beta portfolios tend to have higher IV betas, and vice versa for low beta portfolios. Prior to orthogonalization, the well-established "flatness" of the SML is exacerbated by our corrections: estimated betas are even further apart, making the limited variation in expected returns even more puzzling. However, after orthogonalization, for both OLS and IV, the slope of the SML is closer to the CAPM implied SML, though perhaps a bit too flat nonetheless. We see more clear evidence consistent with OHML being priced, suggesting the econometrician may miss the effects of the granular residual without correcting for omitted variable bias coming from a negative correlation between market betas and OHML loadings.

Finally, Table 4 offers perhaps the most direct test of our predictions about the effects of the granular residual in the cross section by studying size and beta-sorted portfolios. Specifically, Table 4 provides our estimate of the bias in β_p for our unorthognalized specification, though results with the orthogonalized ones are virtually identical. Recall from equation (17) that true betas are attenuated downward by a multiplicative factor – implying that, holding the direct covariance between r_{t+1} and $\eta_{p,t+1}$ (which depends mostly on size) constant, the attenuation bias is larger for portfolios with higher exposures to the true factor. Within size quintiles, our estimates suggest that OLS betas accurately reflect the ordering of IV betas, though magnitudes are incorrect due to effects of the granular residual. Thus, we closely approximate this comparative static by moving down columns of Table 4: i.e., increasing beta, holding size fixed. Indeed, biases are more strongly negative as we move down every in *every* column of the table. Likewise, holding true factor exposures constant, the OLS beta increases with size. As we move across columns within the same row, we are changing size holding beta roughly constant. Consistent with equation (17),

Table 4: Difference between OLS and IV betas for size × beta-sorted portfolios

Market beta quintile	Size quintile Small	2	3	4	Large	Small - Large
Low	-0.12	-0.09	-0.04	0	0.07	-0.19
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)
2	-0.14	-0.11	-0.08	-0.05	0.05	-0.19
	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)
3	-0.17	-0.13	-0.1	-0.05	0.03	-0.2
	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.02)
4	-0.21	-0.17	-0.12	-0.09	0	-0.21
	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.02)
High	-0.27	-0.23	-0.18	-0.13	-0.02	-0.26
	(0.03)	(0.02)	(0.02)	(0.02)	(0.01)	(0.03)
High - Low	-0.15	-0.14	-0.14	-0.13	-0.09	
	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	

Note: This table reports estimates of the bias induced in market betas associated with the presence of the granular residual in cross-sectional asset pricing tests, which tests the predictions of equation (17). Our model predicts that the difference between the IV and OLS coefficients is increasing in beta and decreasing in size, consistent with the results in the table. Estimates are from monthly returns of 25 size and beta-sorted portfolios from Kenneth French's data library from 1963 to the present. This table reports the time series average of rolling regression estimates of beta which are formed by regressing raw excess returns on either the market excess return (OLS) or a linear combination of 3 principal components (IV) constructed using the cross-section of stocks. Standard errors are in parentheses.

$\beta_{IV} - \beta_{OLS}$ always increases as we move across columns within a row. These patterns are uniformly monotonic, holding for every cell in the table.

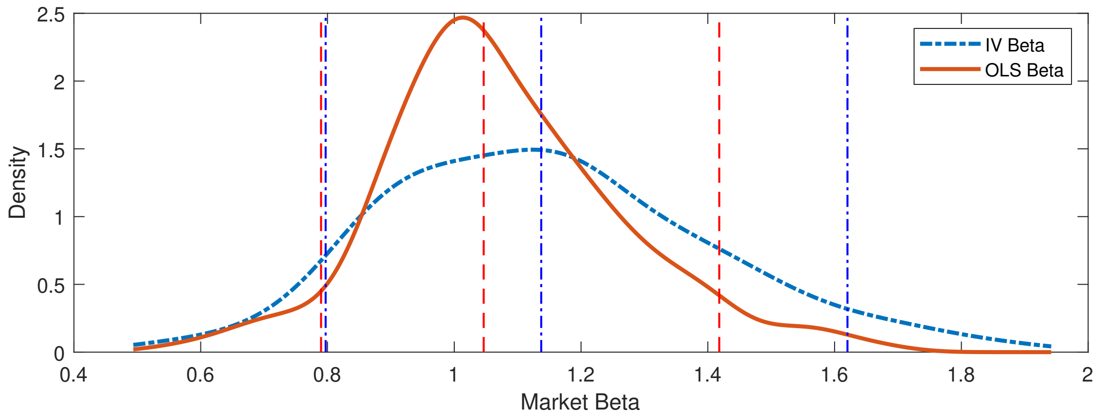
Panels 7 and 8 of Figure 4 characterize the relationship between expected returns and beta for the size × beta-sorted portfolios. In the non-orthogonalized specification, both estimates of the SML are too flat relative to the theory. However, consistent with the results in panels 5-6, we find that correcting for OHML exposures changes this result substantially. In the vast majority of cases, portfolios move closer to the CAPM-SML in the IV specification relative to OLS, consistent with our correction for the granular residual moving the data closer to the theoretical benchmark.

4.3.2 Large sample of portfolios

The previous section demonstrated effects of our IV corrections for a few selected sets of portfolios. In this section, we repeat our analysis for a larger set of portfolios to confirm the robustness of our main results.

We begin by confirming that our common finding that most portfolios' betas are attenuated downward holds more broadly. Figure 5 plots a kernel density estimate

Figure 5: Distribution of OLS and IV market betas for a large number of univariate characteristic and bivariate characteristic × size sorted portfolios



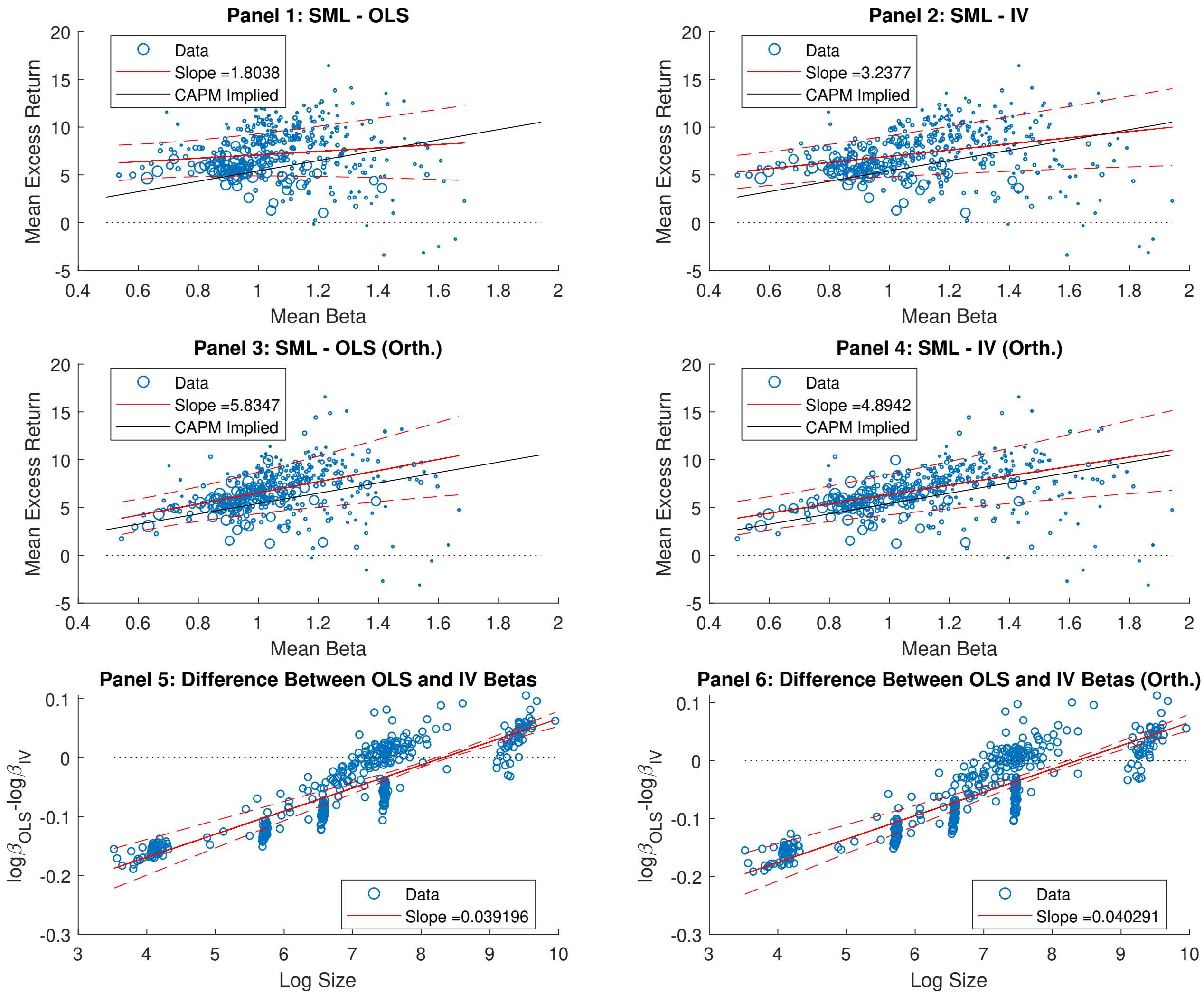
			Market beta percentile								
Estimate	Mean	Std	1	5	10	25	50	75	90	95	99
IV	1.147	0.257	0.587	0.797	0.853	0.952	1.137	1.308	1.491	1.620	1.820
OLS	1.075	0.189	0.638	0.790	0.877	0.954	1.046	1.180	1.324	1.418	1.603
IV - OLS	0.072	0.086	-0.057	-0.045	-0.031	-0.006	0.061	0.140	0.193	0.222	0.263

Note. Figure plots kernel density estimates of OLS and IV estimates of market betas of 14 different sets of decile portfolios formed on various characteristics and 11 5 × 5 double sorted portfolios formed on these characteristics and size (from Kenneth French's website) along with their β estimates obtained by ordinary least squares (OLS) estimation versus our proposed instrumental variables procedure. Horizontal axis reports the time series average of a sequence of betas obtained from rolling 60-month regressions. Color-coded vertical lines indicate the 5th, 50th, and 95th percentiles of the two distributions. We also tabulate a number of univariate summary statistics of the two distributions.

of average betas across the full 25 set of sorted portfolios described above. The blue curve is the density for average IV betas, and the red corresponds with average OLS betas. Analogously-colored vertical lines indicate the 5th, 50th, and 95th percentiles for each distribution. The table beneath the figure displays more percentiles of the distributions in levels, as well as some summary statistics for the distribution of differences between IV and OLS betas (i.e., $-\kappa_p$). The IV density has fatter tails, especially in the right, and is in general shifted rightwards relative to the red. In the table, we see that IV percentiles are mostly larger than the OLS percentiles, where differences are quite large in the right tail. While the average bias, at -.07, might not sound large relative to the average OLS beta (1.07), it is quite substantial relative to the cross-sectional standard deviation of OLS betas (0.19). Likewise, the standard deviation of the bias in beta, at 0.086, is quite large relative to dispersion in OLS betas, and the standard deviation of IV betas is 36% larger relative to OLS estimates.

Next, we turn to the expected return-beta relationship. The top two rows of

Figure 6: Expected return-beta relationship and beta bias for a large number of univariate characteristic and bivariate characteristic × size sorted portfolios



Note. This figure plots SMLs for many different portfolios (listed in table (B7) using IV and OLS, with and without orthogonalization (first two rows). The final row compares the difference between OLS and IV betas as a function of the average size of the firms in a given portfolio, with and without orthogonalization. In the top two rows, circles denote portfolios, and the size of the circle is related to the size of the firms in that portfolio. The red line in each plot is an LAD best fit line, and the dashed lines are 95% confidence intervals. In the top two rows, the black line is the CAPM implied SML.

Figure 6 provide exact analogs to the scatter plots in Figure 4 for our universe of sorts. Given the larger number of portfolios, we plot OLS and IV estimates in separate panels, where sizes of dots correspond with average firm size. Panels 1-2 in top row presents estimates of the relationship between expected returns and betas estimated by OLS and IV, respectively, without orthogonalization. As in Figure 4, we plot a linear estimate of the line of best fit using a LAD regression and associated 95% confidence intervals estimated using a bootstrap inference procedure. Despite the fact that our IV correction increases dispersion in betas, the slope of the IV-SML at 3.2% annualized is actually about 80% larger than the OLS-SML (1.8%), consistent with the IV betas better reflecting cross-sectional differences in priced risk exposures relative to OLS, though still flatter than magnitudes implied by the historical equity premium (5.4% for the post-1963 sample).

While differences between OLS and IV betas are large regardless of whether we control for exposures to other factors, similar to many of the portfolios in Figure 4, portfolios' heterogeneous exposures to the value factor appear to be an important source of omitted variable bias in Panels 1-2. Accordingly, Panels 3-4 of Figure 6 subtract off estimated exposures to the orthogonal value factor. In both cases, the slope of the SML is steeper – slopes of the OLS-SML and IV-SML are 5.8% and 4.9% respectively – bracketing our 5.4% estimate of the equity premium.[47] Further, given that many of the OLS portfolios lie above the CAPM-SML, the rightward shifts induced by our IV procedure for most portfolios' betas often push expected returns closer to the CAPM-SML. As we report in Table 5 below, our estimate of the slope of IV-SML is about 1 standard error below the historical average equity premium and highly statistically significantly different from zero. While plenty of unexplained variation remains, results in panel 4 point to a robust link between our unbiased estimate of beta and expected returns in line with the theory.

Next, panels 5-6 display the relationship between market capitalization and estimated biases in beta. The y-axis shows the log difference between OLS and IV betas for our universe of sorts. The x-axis the log market capitalization of each sort. If the primary difference for a deviation between OLS and IV betas is related to the granular residual, larger stocks should systematically tend to have higher estimated OLS betas. This is exactly what these panels show: The relationship between log market capitalization and log difference in betas is positive. Results are quite similar regardless of orthogonalization. This is roughly the visual analog of moving along the final row of Table 4. Average firm size has a lot of explanatory power for biases in betas, consistent with a nontrivial role played by the granular residual.

[47] Note that we face a higher hurdle with IV betas, since, all else constant, a downward multiplicative bias in most OLS betas would lead to an upward-biased estimate of the OLS-SML slope.

Table 5: Robustness of cross-sectional results to alternative specifications

	Fitting		3 PCs		4 PCs		5 PCs	
	Window	Method	Orth	Non-Orth	Orth	Non-Orth	Orth	Non-Orth
α : Lo-Hi Size Portfolio	60	IV	-0.57	0.68	-0.26	0.85	-0.02	0.998
			(0.93)	(0.95)	(0.93)	(0.95)	(0.92)	(0.94)
		OLS	2.29	3.52	2.45	3.52	2.55	3.52
			(0.95)	(0.96)	(0.95)	(0.96)	(0.95)	(0.96)
	48	IV	-0.71	0.56	-0.26	0.84	0.002	1.01
			(0.88)	(0.92)	(0.9)	(0.92)	(0.9)	(0.91)
		OLS	2.14	3.44	2.36	3.44	2.46	3.44
			(0.92)	(0.93)	(0.93)	(0.93)	(0.93)	(0.93)
	36	IV	-0.997	0.38	-0.14	0.86	0.2	1.12
			(0.89)	(0.92)	(0.88)	(0.9)	(0.87)	(0.89)
		OLS	1.95	3.39	2.37	3.39	2.48	3.39
			(0.92)	(0.94)	(0.93)	(0.94)	(0.93)	(0.94)
Size SML Slope	60	IV	5.95	6.68	6.39	6.83	6.67	6.94
			(0.77)	(0.53)	(0.8)	(0.55)	(0.83)	(0.56)
		OLS	10.59	11.42	10.93	11.42	11.14	11.42
			(1.54)	(1.13)	(1.56)	(1.13)	(1.58)	(1.13)
	48	IV	5.9	6.72	6.41	6.94	6.7	7.04
			(0.73)	(0.5)	(0.79)	(0.53)	(0.81)	(0.54)
		OLS	10.52	11.07	10.96	11.07	11.18	11.07
			(1.42)	(0.97)	(1.46)	(0.97)	(1.48)	(0.97)
	36	IV	5.49	6.04	6.37	6.67	6.79	6.85
			(0.7)	(0.46)	(0.8)	(0.49)	(0.84)	(0.5)
		OLS	10.11	10.66	10.87	10.66	11.14	10.66
			(1.38)	(0.85)	(1.41)	(0.85)	(1.43)	(0.85)
All Portfolio SML Slope	60	IV	4.89	3.24	5.11	3.23	5.25	3.32
			(0.58)	(0.33)	(0.6)	(0.32)	(0.59)	(0.3)
		OLS	5.83	1.8	5.92	1.8	5.99	1.8
			(0.73)	(0.45)	(0.73)	(0.45)	(0.73)	(0.45)
	48	IV	4.79	3.76	5.23	3.91	5.32	3.49
			(0.55)	(0.34)	(0.59)	(0.35)	(0.59)	(0.35)
		OLS	5.87	2.18	6.1	2.18	6.12	2.18
			(0.67)	(0.45)	(0.69)	(0.45)	(0.7)	(0.45)
	36	IV	4.51	3.76	5.02	3.94	5.31	4.06
			(0.52)	(0.27)	(0.58)	(0.3)	(0.6)	(0.32)
		OLS	5.57	2.51	5.95	2.51	6.14	2.51
			(0.66)	(0.54)	(0.71)	(0.54)	(0.71)	(0.54)
Beta Bias Slope	60		0.04	0.04	0.04	0.04	0.03	0.03
			(0.003)	(0.003)	(0.003)	(0.002)	(0.003)	(0.002)
	48		0.04	0.04	0.04	0.04	0.04	0.04
			(0.003)	(0.003)	(0.003)	(0.002)	(0.003)	(0.002)
	36		0.05	0.05	0.04	0.04	0.04	0.04
			(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)

Note: This table demonstrates the robustness of many of the results from Figures 4-6 to alternative model specifications over the post-1963 sample. We allow for different numbers of principal components and 3, 4, and 5-year fitting windows, respectively, and with and without subtracting off estimated orthogonal HML exposure. First we present estimates of the alpha from a portfolio which is long the smallest decile portfolio by size and short the largest size decile portfolio, which can be graphically determined by summing up the distances between the vertical lines in 4. Next, we plot the slope of the securities market line (the red and blue dashed lines in Panels 1 and 3 of Figure 4). Next, we plot the slope of the securities market line from the larger set of portfolios (this is analogous to Panels 1-4 of Figure 6). Finally, we plot the slope of the line of best fit from panels 5-6 of Figure 6.

Table 5 reports point estimates and standard errors for some of the quantities discussed above and shows that our results are robust to various fitting window lengths (60, 48, and 36 months) and number of PCs used to form our instrument (3, 4, and 5) for the post-1963 sample.[48] We feature four summary statistics. For each statistic, we show the OLS and IV estimates, with and without orthogonalization.

First, we show the alpha of the long-short portfolio of the smallest stocks minus the largest stocks ($1-10$ of the 10 size sorted portfolios). Consistent with our baseline estimates in Figure 4, the OLS estimates of the annualized pricing error for this long-short portfolio is 3.5% (t = 3.6) and 2.3% (t = 2.4) using the non-orthogonalized and orthogonalized specifications, respectively. In contrast, corresponding IV estimates are 0.7% (t = 0.7) and -0.6% (t = -0.6), respectively. Whereas OLS alphas for the size premium are always statistically significant, IV estimates are always insignificant despite being estimated with similar levels of precision. Second, related to Panels 1-2 of Figure 4, we next report estimates of the slope of the SML for 10 size-sorted portfolios implied by the OLS and IV estimates. This measure is substantially upward biased for OLS beta and in line with the historical equity premium when IV betas are used instead, a result which holds across a wide variety of specifications.

Third, we report SML slopes computed using all portfolios, corresponding with slopes of the fitted lines in the top four panels of Figure 6. Here, it is clear that orthogonalization is important. When we do not orthogonalize, the IV slope is larger, though still flatter relative to the prediction of the theory. However, once we orthogonalize with respect to the value factor, both OLS and IV estimates are generally quite close to the 5.4% average equity premium across a wide variety of specifications.

Finally, we display the slope of the bottom two panels in Figure 6 across our different specifications. Again, biases in portfolio level betas are robustly linked with average firm size throughout.

5 Correlations: Market Returns and Real Activity

In this section, we examine whether the granular residual has any predictive power for macroeconomic aggregates. In most macro-finance models, returns which are uncorrelated with changes in aggregate state variables do not command risk premia. We address this question indirectly, by answering two closely related questions. Suppose

[48] Table C2 in the Appendix shows similar results for the full sample since 1927.

that the market return increases by 1%. By how much would an econometrician revise her forecast about future macroeconomic conditions on the basis of that information? Suppose instead that the true market factor increases by 1%, by how much, if at all, would these forecasts change?

The difference between these two questions depends on two factors. First, the magnitude of the granular residual affects the extent to which one can obtain different answers in the first place. Second, the difference between the two estimates will reflect the difference in informativeness of the true factor and granular residual about future macro variables. If the granular residual and the true factor are equally informative about the macroeconomy, both coefficients should be the same. When the granular residual is non-negligible and completely uninformative about the macroeconomy, we would expect the latter coefficient to be larger than the former.

To answer these questions, we follow the method in Schmidt (2016), which is itself related to Jordà (2005).[49] While we refer the reader to that paper for further details, the main result is that if we have a variable Y_t, we can estimate a quantity analogous to an impulse response in a VAR to an unexpected stock return at horizon over horizons $h = 1, ..., H$ by running the following regression:

$$Y_{t+h} = \alpha_h + \beta_h r_{t+1} + \gamma_h Y_t + \phi_h(L)Y_t + \psi_h' X_t + \epsilon_{t+h} \tag{21}$$

where r is the stock return, $\phi(L)$ is a lag polynomial, X_t is a vector of controls, and ϵ_{t+h} is the forecast error. Then, the IRF is simply $\{\beta_h\}_{h=1}^{H}$. Our choice of Y variables are log industrial production (IP) growth, log employment growth, initial claims divided by lagged employment, log per capita consumption growth, log compensation growth, the ADS business conditions index (Aruoba, et al (2009)), the unemployment rate, log GDP growth, and the Chicago Fed National Activity Index.

We consider two specifications. First, we estimate equation (21) by OLS, where importantly, r_t is the value weighted market index return. Second, we use our results from the previous section to estimate equation (21) by IV. That is, let F_t denote the

[49]The identifying assumptions proposed in Schmidt (2016) rely on correct specification of the mean of returns and allow for the conditional expectation of the macro variable Y_{t+h} to potentially be misspecified. This is advantageous given that returns are much closer to random walks than many of the macro variables in the consideration set.

fitted value of the first-stage regression:

$$r_t = a + b' PC_t + u_t$$

where PC_t are our three principal components used in the rolling baseline specification in the previous section.[50] Next, we perform 2SLS with F_t as an instrument for r_t. Let $\beta_{IV,h}$ be the coefficient in equation (21) when r_{t+1} is replaced by the fitted value from the first stage.

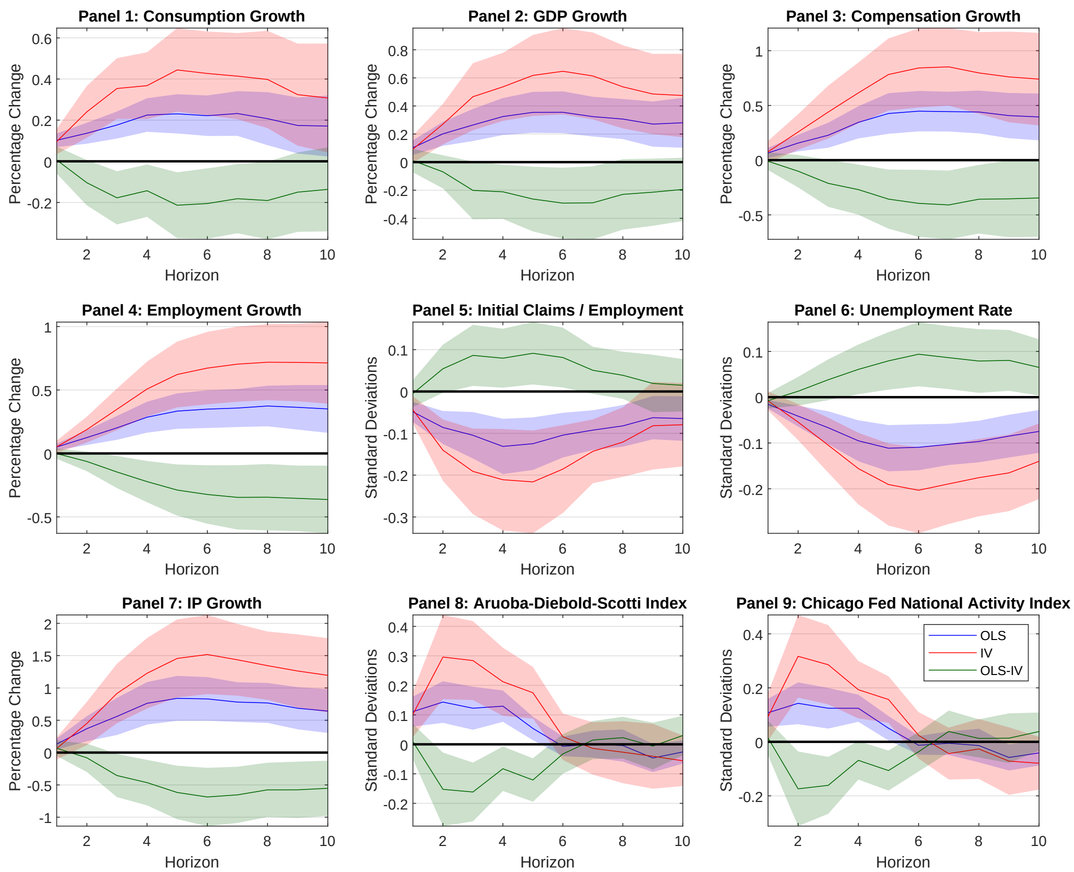
If the granular residual is uninformative (or less informative) about contemporaneous or future real activity, we expect $|\beta_h| \leq |\beta_{h,IV}|$. The reason is simple: in that case, the return measure is contaminated by classical measurement error which creates an attenuation bias. Our IV estimate corrects for this attenuation bias, which is only present if indeed the granular residual is indeed uninformative (or less informative) about Y_{t+h} than the true factor.

Figure 7 shows the impulse response functions for the macroeconomic aggregates listed above. The blue line and blue confidence bands (computed from Newey-West standard errors) correspond to OLS estimation of equation (21). The red ones correspond to our IV estimation. In every panel, we clearly see the blue line is closer to the zero line at each horizon. This means that $\beta_h < \beta_{IV,h}$ for each $h = 1, ..., H$. The green line in Figure 7 displays the difference $\beta_h - \beta_{IV,h}$ for each h. Thus, it is the difference between the OLS and IV impulse response functions. In essentially all cases, both the IV and OLS impulse responses are have the same signs, but the OLS one is attenuated towards 0. The confidence bands are created by estimating the system of OLS and IV regressions jointly using Newey-West standard errors and exactly identified GMM. These estimates suggest that the substantial attenuation effects documented above are statistically significant.

Taking stock, differences in magnitudes are often quite large, implying that the true factor is considerably more informative about future conditions. While this evidence is only suggestive, this indepenent result provides one potential economic rationale for the true factor likely commanding a higher risk price relative to the granular residual, an assumption we used to motivate our empirical tests and derivations of econometric biases above.

[50]Note that, since we are using rolling windows, there will be multiple estimates of F_t for a given t (aside from the first and last date). We take the average across windows for a given date.

Figure 7: Impulse response functions to value weighted market returns and instrumented returns



Note. This figure displays impulse response functions of selected macroeconomic aggregates to the value weighted market return (blue) and an instrumented version of the market return (red), as well as the difference in impulse responses (green). The impulse response functions are calculated from regressions at different horizons of equation (21). The point estimates correspond to $\{\beta_h\}_{h=1}^{H}$. The blue line corresponds to an OLS based impulse response function from equation (21). The red line corresponds to using the fitted value from our principal components regressions from Section 6 as an instrument for r_t. The confidence intervals are computed from Newey-West standard instruments with 10 lags. The descriptions of the macroeconomic variables can be found in the main text. The green lines is the difference between the OLS and IV IRFs.

6 Conclusion

This paper presents evidence from the US stock market consistent with the presence of a substantial, nonpriced component of the market index return which is driven by idiosyncratic shocks to very large stocks which are not diversified away. As we argue through analytical calculations, empirical corrections, and simulation exercises, the presence of a nontrivial granular residual can lead a researcher to be more likely to reject predictions of standard asset pricing theory. Quantitatively, after correcting for the presence of these factors, we find stronger evidence of a link between the conditional mean and volatility of the market portfolio. In addition, we demonstrate that estimated market exposures can often be substantially biased in the cross-section and provide evidence that correcting for the presence of unpriced granular residuals allows a single factor CAPM model to explain the size anomaly. Finally, we find evidence consistent with a stronger correlation between stock returns and future real economic activity after making simple corrections.

While we focus on empirical asset pricing tests related to the market return, our argument applies much more broadly given that value-weighted portfolios are ubiquitous among standard test assets and pricing factors. While value-weighting has many advantages, we uncover one potential disadvantage. Further, our mechanism also helps to make sense of potentially "puzzling" results: such as the observation that the CAPM works better on announcement days (Savor and Wilson, 2014) or that the CAPM works well for jumps in market returns but not well for diffusive shocks (Bollerslev et al., 2016) within a fairly standard model. While ours is not the only explanation, the basic intuition for these types of results within our framework is that the relative variance of the true priced factor relative to the granular measurement error could be larger in these cases, causing standard measures to more accurately reflect the quantity of risk.

References

Acemoglu, D., V. M. Carvalho, A. Ozdaglar, and A. Tahbaz-Salehi (2012): "The network origins of aggregate fluctuations," *Econometrica*, 80, 1977–2016.

Angelidis, T. and N. Tessaromatis (2009): "Idiosyncratic risk matters! A regime switching approach," *International Review of Economics & Finance*, 18, 132–141.

Bai, J. and S. Ng (2002): "Determining the number of factors in approximate factor models," *Econometrica*, 70, 191–221.

——— (2006): "Confidence intervals for diffusion index forecasts and inference for factor-augmented regressions," *Econometrica*, 74, 1133–1150.

Baillie, R. T. and R. P. DeGennaro (1990): "Stock returns and volatility," *Journal of Financial and Quantitative Analysis*, 25, 203–214.

Bali, T. G., N. Cakici, X. S. Yan, and Z. Zhang (2005): "Does idiosyncratic risk really matter?" *Journal of Finance*, 60, 905–929.

Black, F. (1976): "Studies in stock price volatility changes, Proceedings of the 1976 Business Meeting of the Business and Economic Statistics Section," *American Statistical Association*, 177–181.

Bollerslev, T. (1986): "Generalized autoregressive conditional heteroskedasticity," *Journal of Econometrics*, 31, 307–327.

Bollerslev, T., S. Z. Li, and V. Todorov (2016): "Roughing up beta: Continuous versus discontinuous betas and the cross section of expected stock returns," *Journal of Financial Economics*, 120, 464–490.

Bollerslev, T. and J. M. Wooldridge (1992): "Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances," *Econometric reviews*, 11, 143–172.

Brownlees, C., R. Engle, and B. Kelly (2012): "A practical guide to volatility forecasting through calm and storm," *Journal of Risk*, 14, 3.

Byun, S. (2016): "The usefulness of cross-sectional dispersion for forecasting aggregate stock price volatility," *Journal of Empirical Finance*, 36, 162–180.

Campbell, J. Y. (1996): "Understanding risk and return," *Journal of Political economy*, 104, 298–345.

Campbell, J. Y., M. Lettau, B. G. Malkiel, and Y. Xu (2001): "Have individual stocks become more volatile? An empirical exploration of idiosyncratic risk," *Journal of Finance*, 56, 1–43.

Campbell, J. Y. and T. Vuolteenaho (2004): "Bad beta, good beta," *American Economic Review*, 94, 1249–1275.

Campbell, S. D., S. Delikouras, D. Jiang, and G. M. Korniotis (2016): "The human capital that matters: Expected returns and high-income households," *The Review of Financial Studies*, 29, 2523–2563.

Chou, R. Y. (1988): "Volatility persistence and stock valuations: Some empirical evidence using GARCH," *Journal of Applied Econometrics*, 3, 279–294.

Chou, R. Y., R. F. Engle, and A. Kane (1992): "Measuring risk aversion from excess returns on a stock index," *Journal of Econometrics*, 52, 201–224.

Christie, A. A. (1982): "The stochastic behavior of common stock variances: Value, leverage and interest rate effects," *Journal of Financial Economics*, 10, 407–432.

Connor, G. and R. A. Korajczyk (1986): "Performance measurement with the arbitrage pricing theory: A new framework for analysis," *Journal of Financial Economics*, 15, 373–394.

——— (1988): "Risk and return in an equilibrium APT: Application of a new test methodology," *Journal of Financial Economics*, 21, 255–289.

Davis, S. J., J. Haltiwanger, R. Jarmin, J. Miranda, C. Foote, and E. Nagypal (2006): "Volatility and Dispersion in Business Growth Rates: Publicly Traded versus Privately Held Firms [with Comments and Discussion]," *NBER macroeconomics annual*, 21, 107–179.

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

Fama, E. F. and K. R. French (1993): "Common risk factors in the returns on stocks and bonds," *Journal of financial economics*, 33, 3–56.

——— (1996): "Multifactor explanations of asset pricing anomalies," *The journal of finance*, 51, 55–84.

Fama, E. F. and J. D. MacBeth (1973): "Risk, return, and equilibrium: Empirical tests," *Journal of political economy*, 81, 607–636.

Franses, P. H. and D. Van Dijk (1996): "Forecasting stock market volatility using (nonlinear) GARCH models," *Journal of Forecasting*, 229–235.

French, K. R., G. W. Schwert, and R. F. Stambaugh (1987): "Expected stock returns and volatility," *Journal of Financial Economics*, 19, 3–29.

Gabaix, X. (2011): "The granular origins of aggregate fluctuations," *Econometrica*, 79, 733–772.

Giglio, S. and D. Xiu (2017): "Inference on risk premia in the presence of omitted factors," *Working paper*.

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.

Goyal, A. and P. Santa-Clara (2003): "Idiosyncratic risk matters!" *Journal of Finance*, 58, 975–1007.

Guo, H. and R. F. Whitelaw (2006): "Uncovering the risk–return relation in the stock market," *Journal of Finance*, 61, 1433–1463.

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

Heaton, J. and D. Lucas (2000): "Portfolio choice and asset prices: The importance of entrepreneurial risk," *The Journal of Finance*, 55, 1163–1198.

Jordà, Ò. (2005): "Estimation and inference of impulse responses by local projections," *American Economic Review*, 95, 161–182.

Koopman, S. J. and E. Hol Uspensky (2002): "The stochastic volatility in mean model: empirical evidence from international stock markets," *Journal of Applied Econometrics*, 17, 667–689.

Lewellen, J. (2015): "The Cross-section of Expected Stock Returns," *Critical Finance Review*, 4, 1–44.

Lundblad, C. (2007): "The risk return tradeoff in the long run: 1836–2003," *Journal of Financial Economics*, 85, 123–150.

Malloy, C. J., T. J. Moskowitz, and A. Vissing-Jørgensen (2009): "Long-run stockholder consumption risk and asset returns," *The Journal of Finance*, 64, 2427–2479.

Mayers, D. (1973): "Nonmarketable assets and the determination of capital asset prices in the absence of a riskless asset," *The Journal of Business*, 46, 258–267.

Merton, R. C. (1973): "An intertemporal capital asset pricing model," *Econometrica*, 867–887.

NELSON, D. B. (1991): "Conditional heteroskedasticity in asset returns: A new approach," *Econometrica*, 347–370.

POON, S. H. AND C. W. GRANGER (2003): "Forecasting volatility in financial markets: A review," *Journal of Economic Literature*, 41, 478–539.

ROLL, R. (1977): "A critique of the asset pricing theory's tests Part I: On past and potential testability of the theory," *Journal of financial economics*, 4, 129–176.

ROSS, S. A. (1976): "The arbitrage theory of capital asset pricing," *Journal of Economic Theory*, 13, 341–360.

SAVOR, P. AND M. WILSON (2014): "Asset pricing: A tale of two days," *Journal of Financial Economics*, 113, 171–201.

SCHMIDT, L. (2016): "Climbing and Falling Off the Ladder: Asset Pricing Implications of Labor Market Event Risk," Working paper.

SCHWERT, G. W. (1990): "Stock returns and real activity: A century of evidence," *Journal of Finance*, 45, 1237–1257.

STOCK, J. H. AND M. W. WATSON (2002): "Forecasting using principal components from a large number of predictors," *Journal of the American Statistical Association*, 97, 1167–1179.

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.

WEI, S. X. AND C. ZHANG (2005): "Idiosyncratic risk does not matter: A re-examination of the relationship between average returns and average volatilities," *Journal of Banking & Finance*, 29, 603–621.

WHITELAW, R. F. (1994): "Time variations and covariations in the expectation and volatility of stock market returns," *Journal of Finance*, 49, 515–541.

Appendix

A Motivating evidence on granularity in the value-weighted index

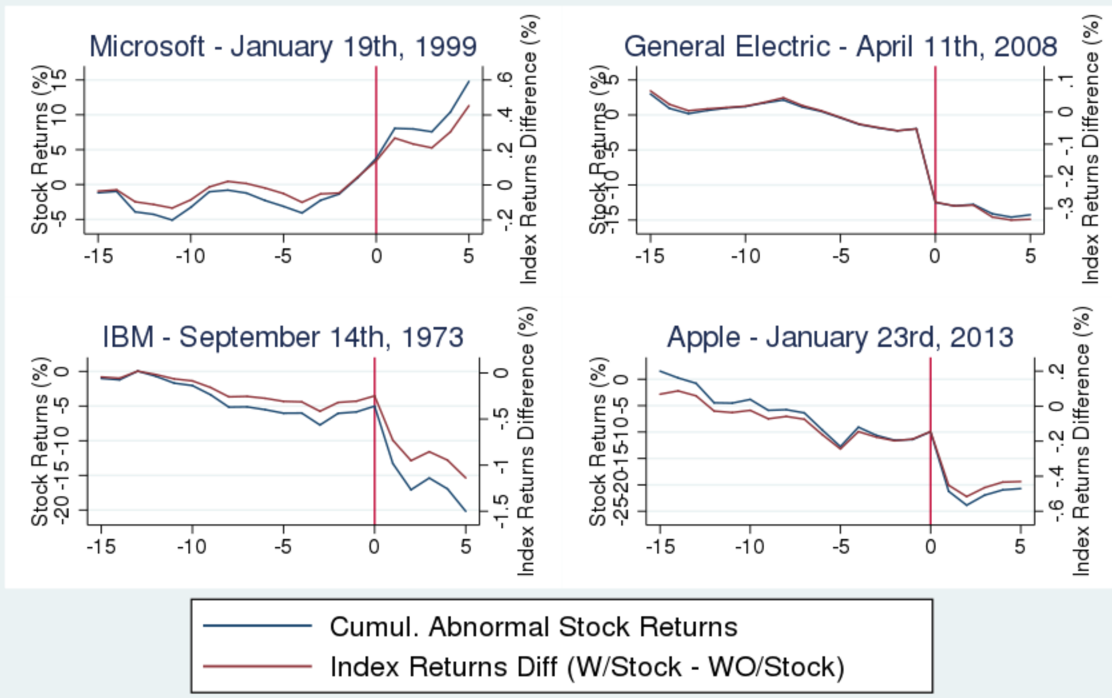
The extent to which the granular residual η_{t+1} is quantitatively large is an empirical question. To support our discussion and empirical tests in the main text, we provide some empirical evidence from the U.S. stock market, suggesting this is likely the case.

Given the significant market concentration among largest stocks, drastic changes in a large company's stock price can singlehandedly pull down entire market returns by nontrivial amounts. Figure 8 plots cumulative abnormal stock returns for four megacap companies around the events of interest (blue lines) along with their counterfactual impacts on the value-weighted market index return (red lines). The cumulative abnormal returns are constructed by aggregating residuals obtained from a single market factor model. To do so, we estimate a stock's exposure on the market factor (i.e., beta) from an ordinary least squares regression of excess individual stock returns on the excess CRSP value-weighted index returns, using daily observations from 5 months to 1 month prior to the event. To demonstrate a megacap company's nontrivial contribution to the value-weighted market index return, we estimate its counterfactual impacts as differences in returns of the value-weighted market index with the stock of interest and without the stock. This latter calculation captures the direct impact of each company-specific market return on the index relative to a counterfactual in which it had had identical returns as the value-weighted average of other firms over the same period. The red vertical lines indicate the event date of interest.

Starting from the top-left quadrant and moving clockwise, we have the following events. The first three examples are associated with earnings announcements of Microsoft (positive news), GE (negative news), and Apple (positive earnings, but below expectations, triggering a selloff).[51] Prior to these announcements, their stocks accounted for sizable shares in the CRSP value-weighted index; Microsoft, GE, and Apple had weights of 2.8%, 2.5%, and 3.2% of the index. Each of these announcements was associated with a change in market value of around 15-20%; thus, the difference between the index with and without each stock in each month is 0.4%, -0.3%, and -0.4%, respectively. The last example comes from IBM, against which a Federal

[51]For more details, see Fisher, "Microsoft's Profit Up 75% in Quarter." New York Times, January 20, 1999; Associated Press, "Shares Sink as G.E. Disappoints." New York Times, April 12, 2008; and Wingfield, "Heady Returns", New York Times, January 24, 2014.

Figure 8: Examples of aggregate effects of news-driven returns for megacap companies



Notes: All stock price data comes from CRSP whereas the market factors and risk-free rates to calculate abnormal returns comes from Fama-French. The market index used is value-weighted where the difference represents the value-weighted index with the individual stock contribution minus the value-weighted index without the stock contribution. Individual company betas are calculated using the three months prior to the event.

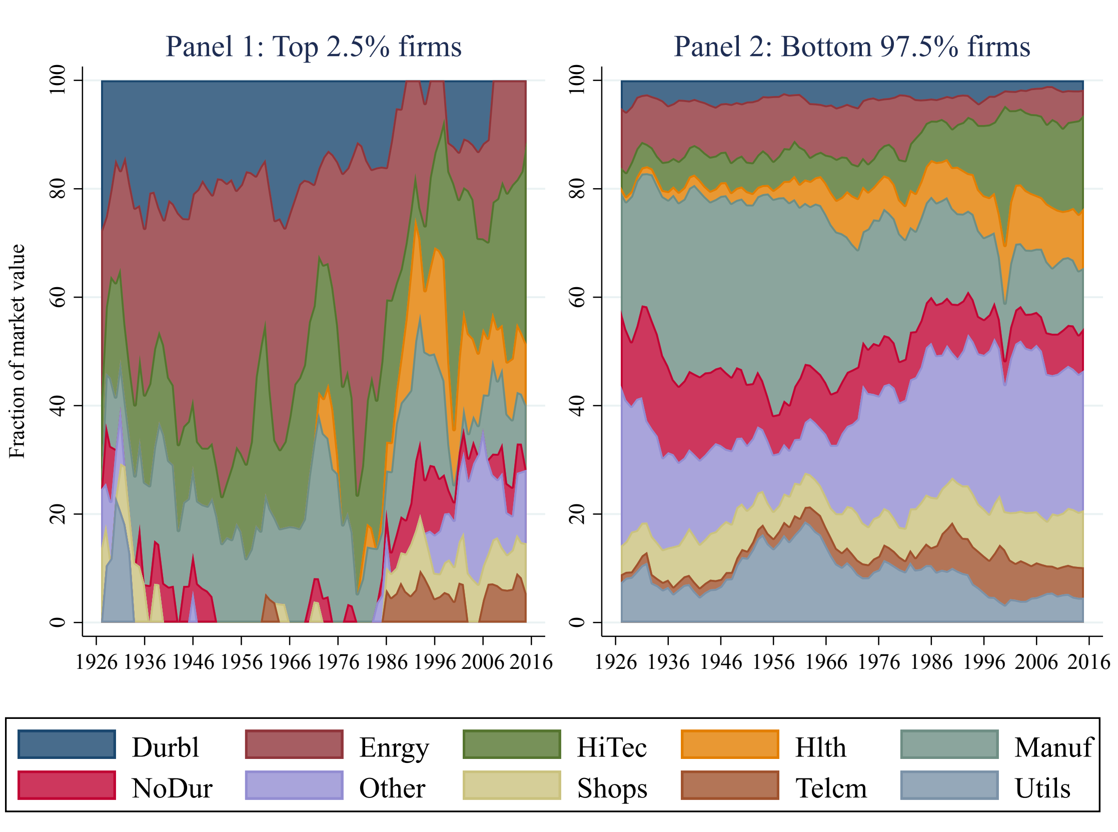
court, in favor of Telex, handed down an antitrust ruling. [52] Given the IBM's weight of 5.3% prior to the announcement, its 20% abnormal returns associated with the announcement triggered an additional 1% decline in the index in September 1973.

Next, we explore to what extent firms at the very top of the market cap distribution are representative of the broader economy. To this end, we compare historical changes in industrial compositions for a small number of large-sized firms with those for all other firms. For each of the Fama-French 10 industries, Figure 9 plots a fraction of combined market capitalizations of the top 2.5% stocks (Panel 1) and those of remaining stocks in the CRSP universe (Panel 2).[53] Comparing Panel 1 with Panel 2, we note that the industry composition of this "megacap" group is quite different from the rest of the market as a whole, and it exhibits much more pronounced time series

[52]See, e.g., Vartan, "Stocks Rise Again in Heavy Trading." New York Times, September 26, 1973 and Vartan "Prices of Stocks Climb Despite Plunge by I.B.M." New York Times, September 18, 1973.

[53]This threshold rules implies that there are usually between 20-35 firms included in this sample, though counts get somewhat higher in the 1990s (peaking at 45) and are closer to 15 prior to 1940, and we note that results are quite similar for other cutoffs.

Figure 9: Industry composition of megacaps vs other stocks in CRSP value-weighted index



Note: This figure plots the share of total market value within each industry, by year, of stocks in the top 2.5% of the market value distribution (left panel) versus all other stocks (right panel) in the CRSP universe. Industries are defined using the Fama-French 10 industry classification using firm SIC codes, see Kenneth French's website for detailed definitions. We take the SIC code from Compustat when available; otherwise, we use the SIC code from CRSP. We use the number of stocks in the NYSE to classify stocks as belonging to the top 2.5% of the market value distribution to avoid generating mechanical increases in the composition of firms in the top category as many small AMEX and NASDAQ firms are added to the database. The Other category includes mining, construction, building materials, transportation, hotels, business services, entertainment, and finance.

fluctuations relative to the set of smaller firms. For instance, firms in the energy and durables categories are associated with almost half of the market value in the top group over the first two thirds of the sample (see Panel 1), whereas the two categories combined make up less than 20% of the sample of smaller firms throughout the entire sample period (see Panel 2). Likewise, tech firms are somewhat overrepresented, especially later in the sample.

As discussed above, the distinction between aggregate and idiosyncratic shocks becomes quite blurry when one considers a large shock experienced by one of the largest firms in the economy. Large changes in large firms' valuations (e.g., earnings announcements) may act as bellwethers for other firms. In our examples above, GE's disappointing earnings announcement, which was driven by losses in its financial services division in the spring of 2008, may indeed have provided a negative signal about the fortunes of other companies in the financial services sector prior to the burst of the financial crisis. Likewise, Microsoft's surprisingly positive earnings announcement in 1999 was likely to be perceived as conveying positive news for other tech stocks as well, and perhaps the Telex ruling conveyed some useful information about future antitrust policy which might have aggregate implications.[54] Second, large firms are connected to many other firms in the economy through customer-supplier relationships, suggesting that network effects will further amplify aggregate effects of these shocks (Acemoglu et al., 2012).

However, even if the word "idiosyncratic" may not describe these shocks perfectly, it is plausible that movements in the market index driven by idiosyncratic news about very large firms (coming from industries that change over time and may not necessarily be representative of the broader economy) are less likely to cause concern for investors relative to price movements which are common across a large number of individual stocks. We provide additional direct and indirect evidence consistent with this indeed being the case above.

B Further details about time-series analysis

This section provide empirical results that are omitted from our main manuscript. This section also provide detailed explanations for empirical exercises, which are used to support our story about the granular and also used to check robustness of our empirical results.

B.1 Alternative GARCH models and coefficient estimates

In a study of the relationship between expected excess stock returns and volatility, Glosten et al. (1993) introduced the GJR-GARCH model, which allows for a larger feedback from prior squared negative returns relative to positive returns. Denoting by I_t an indicator variable for a positive forecasting error in period t, the conditional variance is given by

$$\sigma_{t+1}^2 = \alpha_0 + \alpha_1 u_t^2 + \alpha_2 \sigma_t^2 + \alpha_3 I_t u_t^2, \tag{22}$$

where $I_t = 1$ for $u_t > 0$ and 0 otherwise.

Next, Nelson (1991) proposed the EGARCH model, which specifies the conditional

[54]Note that, to the extent that the news conveyed by these announcements had a large impact on valuations of other stocks, the differences between the index return with and without the stocks in question would be pushed towards zero.

variance in logarithmic form as

$$\ln(\sigma_{t+1}^2) = \alpha_0 + \alpha_1 \left(|\widetilde{u}_t| - E\,|\widetilde{u}_t|\right) + \alpha_2 \ln(\sigma_t^2) + \alpha_3 \widetilde{u}_t, \tag{23}$$

where $\widetilde{u}_t = u_t/\sigma_t$ is a standardized forecast error in period t and $E\,|\widetilde{u}_t| = 2/\pi$ for a normally distributed $\widetilde{u}_t$. Note that the above asymmetric GARCH models capture the leverage effect through α_3: with $\alpha_3 < 0$, the conditional variance (σ_{t+1}^2) is higher for a negative than for a positive forecasting error (u_t).[55] To complement our baseline GARCH(1,1) model, we also estimate the above alternative models using (11), (12) and (22) for GJR-GARCH, and (11), (12) and (23) for EGARCH.

Table B1 displays quasi-MLE estimates and asymptotic standard errors (in parentheses) obtained from the GARCH (Panel A), GJR-GARCH (Panel B) and EGARCH (Panel C) models, which correspond to the results reported at the first two columns in Table 2. That is, estimates are obtained by using both the Unadjusted and Adjusted index returns of the CRSP value- and equal-weighted index returns, while conditional variance forecasts are used to proxy for agents' time-varying stock market risk. In general, parameter estimates except for γ are qualitatively similar to those provided by earlier researchers. For example, parameter estimates for the GARCH forecasting model such as α_0, α_1 and α_2 are statistically significant at any conventional size, which are commonly observed when using both the Unadjusted and Adjusted index returns of the CRSP value- and equal-weighted index. Furthermore, we confirm the leverage effects: α_3 estimates for GJR-GARCH model are positive and statistically significant (Panel B) while those for EGARCH model are negative and statistically significant (Panel C).

Table B2 reports those estimates and asymptotic standard errors when conditional volatility forecasts, rather than conditional variance forecasts, are used to proxy for time-varying stock market risk. In general, parameter estimates are similar to those in Table B1. One exception is smaller γ estimates than those from the conditional variance specification. Such results are expected given the different scale of the adopted risk proxy (i.e., by considering square roots of the conditional variance forecasts), resulting in smaller magnitudes in response to smaller-sized risk proxy.

B.2 Alternative measures for granular residuals

In this subsection, we construct alternative measures of granular residuals using estimates for idiosyncratic returns relative to other principal components specifications. The primary motivation for considering these statistical factors is to soak up more variation in returns that might be associated with risk. Accordingly, we consider

[55] The recognition of the leverage effect goes back to Black (1976), noting a negative correlation between current returns and future volatility, and is further investigated by Christie (1982). According to the leverage effect, a reduction in the equity value would raise the debt-to-equity ratio, hence raising the riskiness of the firm's equity as manifested by an increase in future volatility.

five principal components of the cross-section of stock returns as well as optimally-chosen number of factors following Bai and Ng (2002). Furthermore, we consider arbitrarily large numbers of factors whose combined explanatory power exceeds at least 50% of daily variations of individual returns at each month. Finally, we also consider the Fama-French 3 factors.

Table B3 reports results from our alternative measures for the granular residuals. Panel A shows that the statistical properties of the alternative granular residuals are similar to those reported in Table 1, in terms of medians and interquartile ranges. Panel B confirms the positive and statistically significant estimates for the risk-return tradeoff relationship across all forecasting models when using the Adjusted index returns that net out granular residuals. Within each volatility forecasting model, the γ estimates across alternative granular residuals are at similar magnitudes and statistical significance to those reported in Table 2. Panel C shows that attenuation biases in γ estimates are present in all constructions of the granular residual, although the sizes of biases differ slightly across alternative residuals.

Before concluding this section, several points stand out. First, we may need a large numbers of factors, here about 20 factors on average in our sample, to achieve at least 50% of explanatory power for individual returns; the maximum and minimum numbers of factors are 26 and 2, respectively. Second, we find that the average optimal number of factors is about 3.[56] Third, when the granular residuals are obtained by using the Fama-French 3 factors, our results are similar to those with the optimally-chosen number of factors.

B.3 Alternative sub-samples

Table B4 reports estimates for the risk-return relationship coefficients, obtained from estimating three GARCH forecating models across two subsample periods. The first two columns report γ estimates from the CRSP value-weighted index during our two subsample periods. For comparison, we report those from the CRSP equal-weighted index in the last two columns. As before, we compare γ estimates obtained from the Unadjusted index returns with those from the Adjusted index returns for all forecasting models, while the last row in each panel reports the difference between the two γ estimates.

Three findings are worth noting. First, we confirm attenuation biases in the γ estimates again. We find, from all forecasting models and also from both subsample periods, that the Unadjusted (value-weighted) index returns produce smaller γ estimates than the Adjusted index returns (Columns 1-2).[57] In contrast, such differ-

[56]We consider two criteria introduced in Bai and Ng (2002). In principle, the two criteria may not provide the same numbers of factors. In our sample, the two criteria select the same numbers of factors for most periods. The maximum (minimum) numbers of factors is 18 (1).

[57]We find a negative but insignificant γ estimate (-0.068) from EGARCH when using the unad-

ences in γ estimates disappear when using the CRSP equal-weighted index returns (Columns 3-4). Second, we find smaller γ estimates from the first subsample than those from the second subsample. Even for a case addressing the granular residuals ('Adjusted'), we find a small (and accordingly statistically insignificant) γ estimate from the first subsample, consistent with earlier empirical results. To address empirical issues caused by using highly volatile stock market returns, Lundblad (2007) suggests evaluating the risk-return relationship from a sufficiently long dataset.[58] Third, we find larger biases in γ estimates when using the CRSP value-weighted index returns for the second subsample than those for the first subsample. These results can be explained by the relative sizes of the granular residuals. The relative size – measured as a ratio of sample standard deviations of the granular residuals to the unadjusted index returns – for the second subsample is 0.648, which is larger than those for the total (0.578) and for the second subsample (0.526).

B.4 Time varying parameter model

In this subsection, we provide detailed explanations of the time-varying parameter in the mean model (henceforth, TVP model) of Chou et al. (1992), used to investigate the effects of the granular residuals on the potentially time-varying relation between risk and return.

Let ξ_{t+1} be a state variable representing time-varying risk-return tradeoff coefficient (i.e., price of risk). Then our problem with the state-space representation becomes,

$$r_{t+1} = \phi_1 \cdot r_t + \gamma_{t+1} \cdot \sigma_{t+1}^2 + u_{t+1},$$

$$\gamma_{t+1} = \gamma_t + v_{t+1},$$

where u_{t+1} and v_{t+1} are assumed to be uncorrelated Gaussians with zero means and with variances σ_{t+1}^2 and σ_v^2. The state variable γ_{t+1} measures the risk-premium per unit of variance because σ_{t+1}^2 measures (conditional) variance of stock returns. The conditional variance of stock returns follows the GARCH process

$$\sigma_{t+1}^2 = \alpha_0 + \alpha_1 \cdot \eta_t^2 + \alpha_2 \cdot \sigma_t^2,$$

where the innovation process is defined as

$$\begin{aligned}\eta_{t+1} &\equiv r_{t+1} - E_t\left[r_{t+1}\right] \\ &= u_{t+1} + \sigma_{t+1}^2 \cdot \left(\xi_{t+1} - E_t\left[\xi_{t+1}\right]\right)\end{aligned}$$

justed index returns during the first subsample period.

[58] From this perspective, magnitudes (or sizes) of γ estimates are expected to be inversely related to volatilities of stock market index returns. The sample standard deviation of the CRSP value-weighted index returns for the first subsample is 6.641, which is larger than both for the total (5.407) and the second subsample (4.415).

where $\xi_{t+1} - E_t[\xi_{t+1}] = v_{t+1}$.

To estimate the TVP model, we construct a sequence of log likelihood function by using the Kalman filter and the state space representation of the TVP model. To investigate the effects of the granular residuals, we estimate the TVP model by using the Unadjusted index returns, and also by using the Adjusted index returns. Table B5 reports quasi-MLE estimates from the TVP model, which are comparable to the corresponding estimates in Table B1. For comparison, Table B6 provides summary statistics for risk prices (γ_{t+1}, Panel A) and risk quantities (σ_{t+1}, Panel B) obtained from the TVP model along with those from the fixed parameter model (FP), which is our baseline GARCH model reported in the first two columns of Panel A in Table B1.

Regarding the risk quantities in Panel B, we find quantitively similar summary statistics for both the Unadjusted and Adjusted index returns. That is, the Unadjusted index returns (the Adjusted index returns) produce quantatively similar summary statistics from both the TVP model and the FP model. However, when comparing risk quantities from the Unadjusted index returns with those from the Adjusted index returns, we find that the Unadjusted index returns provide smaller conditional volatility forecasts than the Adjusted index returns. Figure B1 confirms these observations, which are consistent with those from the FP model displayed in Figure 2.

Turning to the risk prices in Panel A, we find that both the mean and the variance of the risk prices obtained from the Unadjusted index returns are larger than those from the Adjusted index returns. In particular, the mean from the Unadjusted index returns is 2.274, which is far above the estimate (1.319) from our baseline exercise in Table 2. In contrast, the Adjusted index returns provide a correctly-sized average risk prices (1.660), which is close to our baseline estimate (1.653) for the risk-return relationship. These results indicate that the granular residuals influence estimates for the time-varying risk prices at sizable magnitudes. As we noted in explanations for Figure 3, the granular residuals in the Unadjusted index returns cause both the level and the variance of risk prices to be overestimated by a sizable magnitude. Hence, analysts are likely to *overemphasize* the importance of time variations in risk prices if the granular residuals have sizable contribution to time variations in stock market index returns.

B.5 Bivariate model

In this subsection, we provide detailed explanations for our alternative bivariate approach, introduced in Section 3.4, which is motivated by the insignificant contribution of the granular residuals to the excess returns of the CRSP equal-weighted index. To begin with, let us introduce notation that will be used for a bivariate model below. Let variables (and associated parameters) having subscripts 'v' represent those for the value-weighted index. For example, $r_{v,t+1}$ is an excess return on the value-weighted index in period $t+1$, whose conditional variance forecast is denoted by $\sigma^2_{v,t+1}$. Similarly, let variables (and associated parameters) having subscripts 'e' represent those for the equal-weighted index.

Consider the following bivariate forecasting model for excess returns on the value- and equal-weighted indices;

$$\begin{bmatrix} r_{v,t+1} \\ r_{e,t+1} \end{bmatrix} = \begin{bmatrix} \phi_{v,0} \\ \phi_{e,0} \end{bmatrix} + \begin{bmatrix} \phi_{v,1} & 0 \\ 0 & \phi_{e,1} \end{bmatrix} \cdot \begin{bmatrix} r_{v,t} \\ r_{e,t} \end{bmatrix} + \begin{bmatrix} \gamma_v & 0 \\ 0 & \gamma_e \end{bmatrix} \cdot \begin{bmatrix} \sigma^2_{v,t+1} \\ \sigma^2_{e,t+1} \end{bmatrix} + \begin{bmatrix} u_{v,t+1} \\ u_{e,t+1} \end{bmatrix}, \tag{24}$$

where $u_{v,t+1}$ and $u_{e,t+1}$ are forecasting errors in excess returns of the value- and equal-weighted indices, respectively. When estimating the above bivariate model separately for excess returns of each index, estimation results coincide with those reported in Table B1.

Denoting by η_{t+1} the granular residual – an aggregate measure of idiosyncratic returns in period $t+1$ – we model the forecasting errors as follows:

$$\begin{bmatrix} u_{v,t+1} \\ u_{e,t+1} \end{bmatrix} = \begin{bmatrix} \sigma_{v,t+1} \cdot \varepsilon_{v,t+1} + \eta_{t+1} \\ \sigma_{e,t+1} \cdot \varepsilon_{e,t+1}, \end{bmatrix}, \tag{25}$$

where ε_{t+1} is a mean-zero martingale difference sequence with $E\left[\varepsilon^2_{t+1}\right] = 1$. And η_{t+1} is an another mean-zero martingale difference sequence, which may or may not exhibit a well-defined behavior (i.e., non-stationary). For simplicity, we calculate the likelihood assuming $E\left[\eta^2_{t+1}\right] = \sigma^2_\eta$, as we are not interested in time-varying properties of the granular residuals. Notice that η_{t+1} appears only in the forecasting error in the value-weighted index return. This is because idiosyncratic returns in the equal-weighted index are expected to be sufficiently diversified to apply the law of large numbers. Furthermore, our empirical results in Section 3 confirm such a conjecture.

To calculate the quasi-likelihood the bivariate model, we make two simplifying assumptions on the relationship between volatilities of the true factor in the equal-weighted index and its contribution to the value-weighted index. The first assumption is on the innovation process (i.e., the source of the true market shock) and the second assumption is for the relative size of two volatilities (i.e., scale). For simplicity, we assume a common aggregate market shock drives both unobservable true factor processes. Denoting by ε_{t+1} the common aggregate market shock – an innovation process for the true factor – we assume $\varepsilon_{t+1} = \varepsilon_{v,t+1} = \varepsilon_{e,t+1}$. Then we impose a simple relationship between the sizes of the two volatilities: $\sigma_{v,t+1} = \rho \cdot \sigma_{e,t+1}$ for some constant ρ. Thus the innovation in the equal-weighted index is always a scaled version of the common innovation in the value-weighted index. Then, the conditional volatility in this component of value-weighted index return differs from the equal-weighted index return by a scalar multiple ρ. Note that these two assumptions help to separate the granular residuals from the forecasting error. The forecasting error is then used to generate conditional variance forecasts for the true market factor.

To estimate the bivariate model, we rewrite the model using the two assumptions. After replacing $\sigma^2_{v,t+1}$ by $\rho^2 \cdot \sigma^2_{e,t+1}$ in (24) and also replacing $\sigma_{v,t+1}$ by $\rho \cdot \sigma_{e,t+1}$ in (25), the bivariate model becomes

$$\begin{bmatrix} r_{v,t+1} \\ r_{e,t+1} \end{bmatrix} = \begin{bmatrix} \phi_{v,0} \\ \phi_{e,0} \end{bmatrix} + \begin{bmatrix} \phi_{v,1} & 0 \\ 0 & \phi_{e,1} \end{bmatrix} \cdot \begin{bmatrix} r_{v,t} \\ r_{e,t} \end{bmatrix} + \begin{bmatrix} \gamma_v \cdot \rho^2 \\ \gamma_e \end{bmatrix} \cdot \sigma^2_{e,t+1} + \begin{bmatrix} u_{v,t+1} \\ u_{e,t+1} \end{bmatrix}, \quad (26)$$

and

$$\begin{bmatrix} u_{v,t+1} \\ u_{e,t+1} \end{bmatrix} = \begin{bmatrix} \rho & 1 \\ 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} \sigma_{e,t+1}\varepsilon_{t+1} \\ \eta_{t+1} \end{bmatrix}, \quad (27)$$

where the process for the conditional variance follows the GARCH(1,1) process,

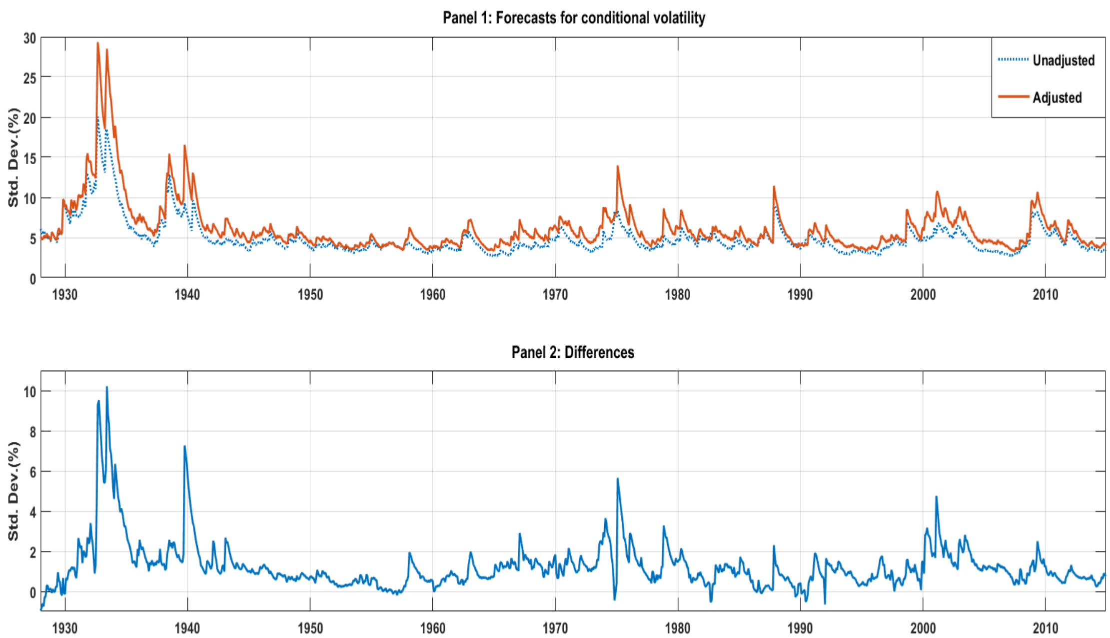
$$\sigma^2_{e,t+1} = \alpha_0 + \alpha_1 \cdot u^2_{e,t} + \alpha_2 \cdot \sigma^2_{e,t}, \quad (28)$$

where $\alpha_0 > 0$, and $0 < \alpha_1, \alpha_2 < 1$ with $\alpha_1 + \alpha_2 < 1$.

Note that σ^2_η represents the average variance of the granular residuals under this model specification. While the granular residuals exhibit a volatility clustering behavior, we do not consider conditional variances for the granular residuals for two reasons. First, we are mainly interested in conditional variances of the stock market index return, which are used for studying the risk-return relationship. Second, time-variation in the granular residuals is driven by realizations of idiosyncratic returns, which are assumed to be unpriced and not to enter the mean equation. Accordingly, omitting a model for these dynamics yields a much more parsimonious specification.

Table B7 reports quasi-MLE estimates for the bivariate model, where we again estimate the bivariate model of equations (26) and (27) along with the three GARCH forecasting models. In general, we find that estimates are qualitatively consistent with those reported in Table B1. In particular, we find positive and statistically significant estimates (γ_v) for the risk-return relationship in the value-weighted index by using conditional variance forecasts from the equal-weighted index returns. Here the γ_v estimates are larger than the γ estimates obtained from both the Unadjusted and the Adjusted index returns, reported in Table B1. However, we find large robust standard errors due to large differences in the two information matrix estimates, respectively obtained from the first- ('outer product method') and the second-order approximations ('second derivative method'). While asymptotic standard errors can be calculated from either the outer product method or the second derivative method, it is widely acknowledged that large robust standard errors indicate a possibility for misspecification. In fact, the bivariate model is written under two strong assumptions, which are necessary for the estimation of parameters numerically yet also impose strong restrictions on the behaviors of parameters (and accordingly conditional variance forecasts).

Figure B1: Quantity of risk



Note: This figure plots forecasts for the conditional volatility of the CRSP value-weighted index, where forecasts are obtained by estimating the time-varying parameter in the mean (TVP) model of Chou et al. (1992). The TVP model allows the risk-return tradeoff coefficient (γ) to be time-varying (γ_t) while maintaining the same conditional variance forecasts following the GARCH(1,1) model. In Panel A, the (red) solid line represents the conditional volatility forecasts obtained from the adjusted index returns, whereas the (blue) dotted line represents those from the unadjusted index returns. Panel B plots differences in volatility forecasts, where a positive difference indicates a smaller conditional volatility forecast from the unadjusted index returns than that from the adjusted index returns.

Table B1: Parameter estimates: Conditional Variance

	Value				Equal			
Models	Unadjusted		Adjusted		Unadjusted		Adjusted	
Panel A: GARCH								
ϕ_0	0.005**	(0.002)	0.004	(0.002)	0.006**	(0.002)	0.006**	(0.002)
ϕ_1	0.074**	(0.036)	0.199***	(0.034)	0.197***	(0.034)	0.196***	(0.034)
γ	1.319	(0.920)	1.550**	(0.618)	2.160***	(0.532)	2.165***	(0.532)
α_0 $\left(\times 10^3\right)$	0.069***	(0.027)	0.131***	(0.042)	0.143***	(0.046)	0.142***	(0.045)
α_1	0.134***	(0.024)	0.138***	(0.030)	0.150***	(0.032)	0.150***	(0.032)
α_2	0.845***	(0.025)	0.837***	(0.031)	0.821***	(0.032)	0.821***	(0.032)
Panel B: GJR-GARCH								
ϕ_0	0.004	(0.002)	0.000	(0.003)	0.002	(0.004)	0.002	(0.003)
ϕ_1	0.100***	(0.039)	0.240***	(0.039)	0.236***	(0.038)	0.236***	(0.038)
γ	1.150	(1.131)	1.986***	(0.705)	2.815***	(0.911)	2.819***	(0.908)
α_0 $\left(\times 10^3\right)$	0.102**	(0.050)	0.173**	(0.070)	0.178**	(0.080)	0.177**	(0.080)
α_1	0.042	(0.028)	0.009	(0.028)	0.049	(0.038)	0.049	(0.039)
α_2	0.840***	(0.037)	0.849***	(0.040)	0.817***	(0.044)	0.817***	(0.044)
α_3	0.145***	(0.064)	0.191***	(0.054)	0.180***	(0.067)	0.179***	(0.067)
Panel C: EGARCH								
ϕ_0	0.003	(0.002)	−0.000	[0.002]	0.002	[0.002]	0.002	[0.002]
ϕ_1	0.085**	(0.037)	0.234***	[0.032]	0.241***	[0.034]	0.241***	[0.034]
γ	1.604	(1.058)	1.989***	[0.598]	2.983***	[0.616]	2.989***	[0.617]
α_0	−0.269***	(0.098)	−0.218***	[0.050]	−0.219***	[0.054]	−0.219***	[0.054]
α_1	0.223***	(0.032)	0.187***	[0.024]	0.249***	[0.026]	0.249***	[0.026]
α_2	0.956***	(0.016)	0.961***	[0.009]	0.961***	[0.009]	0.961***	[0.009]
α_3	−0.098***	(0.027)	−0.113***	[0.017]	−0.089***	[0.019]	−0.089***	[0.019]

Note: This table reports maximum likelihood estimation (MLE) estimates of model parameters in the univariate GARCH (Panel A), GJR-GARCH (Panel B) and EGARCH models (Panel C) without (Unadjusted) and with adjusting granular residuals (Adjusted). The parameter estimates and asymptotic standard errors (in parentheses) are shown for excess returns of the value- and equal-weighted indices. The table reports maximum likelihood estimates and robust standard errors except for a few parameter estimates from EGARCH models. Their standard errors, reported in the square brackets, are obtained by approximating the information matrix by evaluating outer products at maximum likelihood estimates. */**/*** represent the statistical significance at 90%, 95% and 99%, respectively.

Table B2: Parameter estimates: Conditional Volatility

	Value				Equal			
Models	Unadjusted		Adjusted		Unadjusted		Adjusted	
Panel A: GARCH								
ϕ_0	0.001	(0.004)	-0.006	(0.005)	-0.010^{*}	(0.006)	-0.010^{*}	(0.006)
ϕ_1	0.075^{**}	(0.036)	0.201^{***}	(0.034)	0.198^{***}	(0.034)	0.197^{***}	(0.034)
γ	0.140	(0.103)	0.286^{***}	(0.103)	0.427^{***}	(0.104)	0.427^{***}	(0.104)
α_0 $(\times 10^3)$	0.068^{***}	(0.025)	0.128^{***}	(0.042)	0.136^{***}	(0.042)	0.136^{***}	(0.044)
α_1	0.133^{***}	(0.024)	0.137^{***}	(0.029)	0.149^{***}	(0.030)	0.149^{***}	(0.030)
α_2	0.846^{***}	(0.024)	0.838^{***}	(0.027)	0.824^{***}	(0.028)	0.824^{***}	(0.028)
Panel B: GJR-GARCH								
ϕ_0	0.001	(0.005)	-0.009	(0.006)	-0.011^{*}	(0.006)	-0.011^{*}	(0.006)
ϕ_1	0.100^{***}	(0.038)	0.238^{***}	(0.037)	0.233^{***}	(0.035)	0.233^{***}	(0.035)
γ	0.109	(0.124)	0.288^{*}	(0.118)	0.411^{***}	(0.116)	0.411^{***}	(0.116)
α_0 $(\times 10^3)$	0.098^{**}	(0.045)	0.158^{***}	(0.059)	0.155^{***}	(0.056)	0.155^{***}	(0.056)
α_1	0.045^{*}	(0.026)	0.024	(0.025)	0.072^{**}	(0.030)	0.072^{**}	(0.030)
α_2	0.842^{***}	(0.036)	0.847^{***}	(0.038)	0.822^{***}	(0.036)	0.822^{***}	(0.036)
α_3	0.143^{***}	(0.062)	0.176^{***}	(0.051)	0.140^{***}	(0.047)	0.139^{***}	(0.047)
Panel C: EGARCH								
ϕ_0	0.001	[0.004]	-0.006	[0.005]	-0.007	[0.005]	-0.007	[0.005]
ϕ_1	0.081^{**}	[0.033]	0.232^{***}	[0.033]	0.236^{***}	[0.035]	0.235^{***}	[0.035]
γ	0.123	[0.102]	0.211^{*}	[0.095]	0.346^{***}	[0.090]	0.346^{***}	[0.090]
α_0 $(\times 10^3)$	-0.249^{***}	[0.056]	-0.187^{***}	[0.045]	-0.179^{***}	[0.042]	-0.180^{***}	[0.042]
α_1	0.224^{***}	[0.033]	0.193^{***}	[0.024]	0.252^{***}	[0.026]	0.252^{***}	[0.026]
α_2	0.959^{***}	[0.009]	0.966^{***}	[0.008]	0.967^{***}	[0.007]	0.968^{***}	[0.007]
α_3	-0.095^{***}	[0.022]	-0.105^{***}	[0.016]	-0.073^{***}	[0.018]	-0.073^{***}	[0.018]

Note: See notes to Table B1.

Table B3: Risk-return tradeoff coefficient (γ) with alternative modeling assumptions

Models	Baseline	PC5	PC50%	PCBN	FF3
Panel A: Summary statistics					
No. of factors	3	5	20	3	3
Median ($\times$100)	0.068	0.063	0.014	0.056	−0.023
IQR ($\times$100)	2.577	2.544	2.389	2.568	2.035
Panel B: Estimates for risk-return tradeoff coefficient (γ_2)					
GARCH	1.550**	1.590**	1.630**	1.684***	1.689**
	(0.618)	(0.637)	(0.666)	(0.646)	(0.770)
GJR-GARCH	1.986***	2.066***	2.159***	2.162***	2.321**
	(0.705)	(0.727)	(0.798)	(0.759)	(1.046)
EGARCH	1.989***	2.055***	2.101**	2.106***	2.195***
	[0.598]	[0.614]	(0.824)	(0.814)	(0.855)
Panel C: Differences ($\gamma_2^{Unadj.} - \gamma_2^{Adj.}$)					
$GARCH$	−0.231	−0.271	−0.311	−0.365	−0.370
$GJR-GARCH$	−0.837	−0.917	−1.010	−1.012	−1.171
$EGARCH$	−0.385	−0.451	−0.498	−0.503	−0.591

Note: This table reports results using our alternative granular residuals for the total sample period (Jan 1928 - Dec 2014). Panel A shows the median number of empirical/statistical factors used (No. of factors), median (Median) and interquartile range (IQR) of alternative measures for the granular residuals. Panel B reports γ estimates and standard errors (in parentheses) when using the Adjusted index returns, obtained by subtracting alternative granular residuals from the Unadjusted index returns, across three GARCH forecasting models. Panel C reports differences in in γ estimates, which are comparable to those reported at each panel's last row in Table 2. The alternative measures for the granular residuals are constructed from asymptotic principal component analysis introduced in Connor and Korajczyk (1986, 1988). To do so, we obtain a certain number of eigenvectors from the previous 3-months of daily individual returns. Then we estimate monthly idiosyncratic returns relative to those eigenvectors, which become our statistical factors explaining individual returns. Lastly, we construct our alternative measures for the granular residuals using alternative estimates for idiosyncratic returns relative to the eigenvectors. While the baseline considers 3 statistical factors, here we consider 5 statistical factors (PC5), arbitrary number of factors explaining at least 50% of daily variations in individual returns (PC50%), and an optimal number of factors (PCBN) suggested by Bai and Ng (2002). We also consider widely adopted Fama-French 3 factors (FF3). */**/*** represent the statistical significance at 90%, 95% and 99%, respectively.

Table B4: Risk-return tradeoff coefficient (γ) over different sub-samples

	Value		Equal	
Models	1928.1-1962.6	1962.7-2014.12	1928.1-1962.6	1962.7-2014.12
Panel A: GARCH				
Unadjusted	0.556	2.800	1.585***	6.092**
	(0.993)	(1.893)	(0.492)	(2.758)
Adjusted	0.906	5.247**	1.586***	6.139**
	(0.598)	(2.586)	(0.491)	(2.791)
$\gamma^{Unadj.} - \gamma^{Adj.}$	−0.350	−2.447	−0.000	−0.047
Panel B: GJR-GARCH				
Unadjusted	0.364	1.411	1.662***	4.209**
	(1.304)	(1.831)	(0.522)	(1.936)
Adjusted	0.879	3.726*	1.661***	4.234**
	(0.692)	(1.963)	(0.520)	(1.947)
$\gamma^{Unadj.} - \gamma^{Adj.}$	−0.515	−2.315	0.001	−0.025
Panel C: EGARCH				
Unadjusted	−0.068	2.083	1.907***	3.987
	[0.936]	[2.172]	[0.677]	(2.560)
Adjusted	0.663	4.860***	1.911***	4.005
	(0.770)	[1.874]	[0.678]	(2.566)
$\gamma^{Unadj.} - \gamma^{Adj.}$	−0.731	−2.777	−0.004	−0.017

Note: This table reports estimates of the risk-return tradeoff coefficients for the first and second half of our sample period. We refer the reader to the text and notes for Table 2 for additional details.

Table B5: Parameter estimates: Time-varying parameter model

Models	Unadjusted		Adjusted	
ϕ_0	0.083**	(0.036)	0.213***	(0.036)
ϕ_1	0.119*	(0.067)	0.073*	(0.041)
α_0 ($\times 10^3$)	0.071***	(0.025)	0.133***	(0.045)
α_1	0.124***	(0.022)	0.137***	(0.028)
α_2	0.851***	(0.023)	0.836***	(0.027)

Note: This table reports maximum likelihood estimation (MLE) estimates of model parameters in the time-varying parameter model of Chou et al. (1992). For volatility forecasting model, the univariate GARCH-M model is estimated without (unadjusted) and with adjusting granular residuals (adjusted). Robust standard errors are calculated following Bollerslev and Wooldridge (1992). */**/*** represent the statistical significance at 90%, 95% and 99%, respectively.

Table B6: Summary statistics for prices and quantities of risk

	TVP		FP	
	Uadjusted	Adjusted	Unadjusted	Adjusted
Panel A: Price of risk				
Mean	2.275	1.826		
Std	1.153	0.737		
Min	-2.268	-1.953		
Q1	1.686	1.566		
Median	2.270	1.945		
Q3	3.079	2.280		
Max	4.318	3.398		
Panel B: Quanity of risk				
Mean	4.945	5.863	4.944	5.856
Std	2.222	2.821	2.254	2.815
Min	2.635	3.279	2.550	3.211
Q1	3.714	4.323	3.672	4.296
Median	4.323	5.032	4.307	5.012
Q3	5.292	6.245	5.279	6.222
Max	20.052	27.245	18.830	24.541

Note: Quantity of risk is the square root of conditional variance forecasts. TVP represents the time-varying parameter in the mean model introduce in Chou et al. and FP represents the fixed parameter model that is our baseline GARCH-M model. In our baseline model, estimates for prices of risk are 1.319 from Unadjusted index returns and 1.653 from Adjusted index returns.

Table B7: Parameter estimates: Alternative approach

Parameters	GARCH		GJR-GARCH		EGARCH	
$\phi_{v,0}$	0.003*	(0.002)	0.001	(0.002)	0.002	(0.002)
$\phi_{v,1}$	0.097***	(0.023)	0.051**	(0.023)	0.076***	(0.024)
γ_v	2.488***	(0.640)	3.981***	(0.566)	3.041***	(0.845)
$\phi_{e,0}$	0.005***	(0.002)	−0.003	(0.002)	−0.002	(0.002)
$\phi_{e,1}$	0.177***	(0.024)	0.175***	(0.023)	0.189***	(0.023)
γ_e	2.564***	(0.418)	4.727***	(0.537)	4.195***	(0.593)
ρ	0.666***	(0.010)	0.686***	(0.010)	0.676***	(0.010)
σ_v	0.024***	(0.001)	0.023***	(0.001)	0.023***	(0.001)
$\alpha_0\,(\times 1000)$	0.148***	(0.041)	0.400***	(0.077)	−292.161***	(53.300)
α_1	0.812***	(0.023)	0.744***	(0.035)	0.237***	(0.025)
α_2	0.159***	(0.023)	0.000	(0.009)	0.948***	(0.009)
α_3			0.273***	(0.042)	−0.122***	(0.017)

Note: This table reports parameter estimates and asymptotic standard errors (in parentheses) from the alternative forecasting equation, which is closely related to the insight provided by Goyal and Santa-Clara (2006). Given granular measurement errors' insignificant contribution on the CRSP equal-weighted index returns, we include conditional variances of the equal-weighted index in the forecasting equation of the CRSP value-weighted index returns. To do so, we assume that two excess returns are highly correlated, the degree of correlation being represented by ρ. We further assume that excess returns of the CRSP value-weighted index are decomposed into excess returns explained by the CRSP equal-weighted index and the granular measurement errors, of which unconditional variance is denoted by σ_v. Subscripts 'v' and 'e' represent parameters in the forecasting equation for the excess value- and equal-weighted index returns . In general, maximum likelihood estimates of model parameters provide large robust standard errors, indicating a potential misspecification issue for this approach. Hence, the reported standard errors are from the second derivative method, which is known to provide more robust standard errors than alternative methods, such as the outer product method.

C Further results on cross-sectional analysis

In this section we provide a more detailed explanation of our orthogonalization procedure used in Section 4.

We extract principal components from the CRSP universe of stocks. In order to accommodate possible time-variation in betas and PC loadings, we extract these components in a rolling manner. Our main specification uses a 60 month overlapping window, and our sample starts in 1963-01 and goes until 2014-12, inclusive.

For each window, we only keep stocks that appear in the CRSP database for the full window length. Then, we extract the first three principal components from this cleaned sample. For the value weighted market return, we use the excess market return series from Kenneth French's web site.[59]

Over each window, we run the following regression:

$$r_t = \alpha_{m,\tau} + \sum_{k=1}^{K} \beta_{m,k,\tau} f_{k,t} + \epsilon_{m,t} \tag{29}$$

where r_t is the excess return of the market over the risk-free rate, $f_{k,t}$ is the k th principal component extracted from the CRSP sample, and $\epsilon_{m,t}$ is an error term. Note that the parameters, $\alpha_{m,\tau}$ and $\beta_{m,k,\tau}$ are time-subscripted with τ , which indicates the window over which they are estimated. In our baseline, $K = 3$. Define the fitted value of this regression as $F_{t,\tau}$, where once again τ indexes windows, and for each window $t = 1, ..., 60$.

The empirical finance literature has demonstrated that a number of factors can help explain the cross-section of expected returns. Our goal is to focus only on the effects of the granular residual on estimation of the market beta, and so we do not want our estimates to be confounded by omitted factors. Thus, we will want to capture these influences without affecting our estimates of the IV beta. One of the best documented factors which seems to help explain the cross-section of expected returns and return predictability is the book-to-market factor. Thus, we use the HML factor from Kenneth French's website to control for this source of risk, though we have found similar results in alternative models which allow for additional priced factors.

Our goal is to construct a version of the HML factor which is orthogonal to our projection of the market factor onto the PCs. Let H_t be HML. We run the following regression over each window:

$$H_t = \alpha_{H,\tau} + \beta_{H,\tau} F_{t,\tau} + \epsilon_{H,t}$$

[59]The correlation between this series and a series constructed it from CRSP stocks directly is above 0.99.

We collect the following long-short portfolio:

$$H_t - \hat{\beta}_{H,\tau} F_{t,\tau} = \alpha_{H,\tau} + \epsilon_{H,t}$$

Note that all elements of the left-hand side are linear combinations of stocks, and so is indeed a long-short portfolio, which makes the right-hand side a long-short portfolio as well. Importantly:

$$\text{Cov}\left(\alpha_{H,\tau} + \epsilon_{H,t}, F_{t,\tau}\right) = 0$$

Call this long-short portfolio $g_{t,\tau}$. Finally, define:

$$\hat{F}_\tau = \begin{bmatrix} 1_T & g_\tau & F_\tau \end{bmatrix}$$

where T is the length of a window (60 months), and 1_T is vector of ones of length T . This is our regressor matrix for calculating betas over each window.

For our cross-sectional exercises, we use returns of characteristic-sorted portfolios, of which list is provided at Table C1. Let $r_{p,t}$ denote the return of one of these portfolios in excess of the risk-free rate. The time-series of available dates varies portfolio to portfolio. We calculate betas for portfolios only when the window has a full time-series of portfolio returns.

For each window, τ , we run the following time-series regression:[60]

$$r_{p,t} = \alpha_p + \omega_p' g_t + \beta_p F_t + \epsilon_{p,t}, \quad p = 1, ..., P_\tau$$

where P_τ is the number of characteristic sorted portfolios with full data over window τ. In order to facilitate the graphing and visualization of a security market line (SML), we will need to be able to remove the influence of our non-IV factors. Define:

$$\tilde{r}_{p,t} = r_{p,t} - \hat{\omega}_p' g_t = \alpha_p + \beta_p F_t + \epsilon_{p,t}$$

Thus, we will plot the SML using $E[\tilde{r}_{p,t}]$ as our measure of mean returns.

We contrast this SML with the one estimated by the usual, OLS, procedure. That is, for each characteristic portfolio we also run:

$$\tilde{r}_{p,t} = \alpha_{OLS} + \beta_{p,OLS} r_{m,t} + \epsilon_{p,t,OLS}$$

where $r_{m,t}$ is the excess market return over the appropriate window. Notice that we are using $\tilde{r}_{p,t}$ on the left hand side. Thus, for each portfolio, when we plot points in beta-mean return space, the only difference between OLS estimated and IV estimated points will be in betas.

Table C2 provides shows our procedure is robust to various different assumptions.

[60]We drop the τ subscript for convenience, but as explained above, each regressor matrix is constructed window-by-window.

The text in the main part of the paper describes each statistic. Here we discuss the calculation of the standard errors.

These are computed using a weighted bootstrap procedure. This works as follows. Let S be the number of simulations we wish to use in our bootstrap. We draw S vectors of exponential random variables, $X_s,\ s = 1, ..., S$ with mean 1 of length T, where T is the length of our data series. One weight is associated with each calendar time observation. Then, within each simulation s and for each rolling window, we estimate factor loadings using weighted OLS regressions using the corresponding subset X_s as weights. This analysis treats the following inputs as raw data: the market excess return, the PCs, the non-orthogonalized HML factor, and the characteristic sorted portfolios. Using this re-weighted data, we re-estimate all portfolio means and factor loadings and reproduce all the statistics of interest. (Note that second stage regressions, such as SML estimates, do not need to be weighted, because they simply aggregate information from the original set of weighted statistics.) The standard errors in Table (C2) are all then computed as standard deviations of the S bootstrapped statistics.

Table C1: Portfolios Used in Cross-Sectional Tests

5-by-5 Size and...	10 Portfolio Sorts
Book-to-Market	Size
Operating Profitability	Book-to-Market
Investment	Investment
Momentum	Operating Profitability
Short-Term Reversal	Momentum
Variance	Short-Term Reversal
Net Issuance (5-by-7)	Earnings - Price Ratio
Beta	Cash Flows - Price Ratio
Accruals	Dividend - Price Ratio
	Accruals
	Net Issuance
	Variance
	Residual Variance
	Beta
	Industry

Note: This table lists the portfolio sorts from Kenneth French's website used in Section 6 of the paper for cross-sectional tests. The left panel displays the names of the five-by-five sorts and the one seven-by-five sort. The right panel shows the 10 portfolio sorts. All of the left column sorts are "size and..." so we simply list the second sorting variable.

Table C2: Robustness of cross-sectional results to alternative specifications for the full sample

	Fitting		3 PCs		4 PCs		5 PCs	
	Window	Method	Orth	Non-Orth	Orth	Non-Orth	Orth	Non-Orth
α : Lo-Hi Size Portfolio	60	IV	0.5	2.23	0.71	2.38	0.87	2.53
			(1.06)	(1.01)	(1.05)	(1.01)	(1.05)	(1.02)
		OLS	3.16	5.03	3.24	5.03	3.29	5.03
			(1.1)	(1.09)	(1.09)	(1.09)	(1.09)	(1.09)
	48	IV	0.31	1.89	0.63	2.15	0.8	2.29
			(1.07)	(1.01)	(1.07)	(1.01)	(1.07)	(1.01)
		OLS	2.95	4.72	3.07	4.72	3.14	4.72
			(1.11)	(1.08)	(1.11)	(1.08)	(1.11)	(1.08)
	36	IV	-0.06	1.31	0.48	1.67	0.72	1.93
			(1.09)	(1)	(1.08)	(0.99)	(1.07)	(0.99)
		OLS	2.62	4.21	2.86	4.21	2.89	4.21
			(1.14)	(1.06)	(1.14)	(1.06)	(1.13)	(1.06)
Size SML Slope	60	IV	9.3	10.44	9.71	10.76	10.09	11.05
			(1.07)	(0.58)	(1.12)	(0.62)	(1.18)	(0.64)
		OLS	13.33	16.3	13.56	16.3	13.69	16.3
			(1.63)	(1.23)	(1.65)	(1.23)	(1.66)	(1.23)
	48	IV	8.22	9.78	8.88	10.19	9.27	10.51
			(0.97)	(0.54)	(1.04)	(0.57)	(1.11)	(0.59)
		OLS	12.53	15.81	12.76	15.81	12.9	15.81
			(1.51)	(1.12)	(1.53)	(1.12)	(1.55)	(1.12)
	36	IV	7.53	8.9	8.46	9.43	8.91	9.83
			(0.87)	(0.47)	(0.97)	(0.52)	(1.04)	(0.54)
		OLS	12.01	15.31	12.47	15.31	12.58	15.31
			(1.44)	(1)	(1.47)	(1)	(1.49)	(1)
All Portfolio SML Slope	60	IV	5.12	3.96	5.36	4.14	5.37	4.24
			(0.81)	(0.65)	(0.84)	(0.71)	(0.85)	(0.76)
		OLS	6.19	1.92	6.21	1.92	6.33	1.92
			(1.15)	(0.84)	(1.14)	(0.84)	(1.14)	(0.84)
	48	IV	4.91	4.11	5.35	4.18	5.49	4.23
			(0.7)	(0.5)	(0.74)	(0.56)	(0.77)	(0.63)
		OLS	6.51	2.49	6.51	2.49	6.26	2.49
			(1.09)	(0.88)	(1.1)	(0.88)	(1.12)	(0.88)
	36	IV	4.57	3.88	5.07	4.02	5.33	4.09
			(0.69)	(0.46)	(0.73)	(0.5)	(0.75)	(0.55)
		OLS	6.34	2.96	6.6	2.96	6.66	2.96
			(1.03)	(0.88)	(1.08)	(0.88)	(1.08)	(0.88)
Beta Bias Slope	60		0.04	0.04	0.03	0.03	0.03	0.03
			(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
	48		0.04	0.04	0.03	0.04	0.03	0.03
			(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
	36		0.04	0.04	0.04	0.04	0.03	0.03
			(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)

Note: This table demonstrates the robustness of many of the results from Figures 4-6 to alternative model specifications. We allow for different numbers of principal components and 3, 4, and 5-year fitting windows, respectively, and with and without subtracting off estimated orthogonal HML exposure. First we present estimates of the alpha from a portfolio which is long the smallest decile portfolio by size and short the largest size decile portfolio, which can be graphically determined by summing up the distances between the vertical lines in 4. Next, we plot the slope of the securities market line (the red and blue dashed lines in Panels 1 and 3 of Figure 4). Next, we plot the slope of the securities market line from the larger set of portfolios (this is analogous to Panels 1-4 of Figure 6). Finally, we plot the slope of the line of best fit from panels 5-6 of Figure 6.

D Simulation exercises

In this section, we demonstrate the quantitative relevance of the biases associated with the granular residuals discusses in Sections 3 and 4. After providing detailed explanations for our Monte Carlo simulation exercises, we report our simulation results under various conditions, in which sizes of the granular residuals are determined by characteristics of idiosyncratic returns and also by those of the stock market index. We show that the granular residuals cause attenuation biases in estimates for the risk-return relationship as well as biases in estimates for market risk exposures. Lastly, we investigate potential effects from model misspecification, for which the risk-return relationship is estimated from the GARCH model while the true process for the factor variance follows the stochastic volatility model. As mentioned previously, it is worth noting that the granular residuals in our simulation exercise are endogenous, which are arise naturally in excess returns of the value-weighted index.

D.1 Overview of simulation results

We perform a Monte Carlo analysis to demonstrate the quantitative relevance of the biases associated with the granular residuals discussed above. The granular residuals are endogenously generated from the fat-tailed size distribution of individual stocks, which itself is generated from realizations of the true market factor and idiosyncratic returns. For simplicity, we assume a single factor data-generating process (i.e., $g_{t+1} = 0$). In our simulation exercises, all stocks have the same characteristics (initial prices, idiosyncratic return distribution, and the number of shares outstanding), though we allow for some heterogeneity in individual stocks' market-risk exposures (β_i). We provide more detailed descriptions of our simulation procedure in the following subsection, but our basic objective is to generate a concentrated distribution of portfolio weights similar to the one observed in the data and plotted in section 2. In Section 3, we propose methods for correcting for the presence of granular residuals, and we confirm that they work well in these simulation exercises. In these simulations, results using our corrections are very similar to those which work directly with the true factors. Thus, here we simply compare the effects of running standard empirical asset pricing tests with the true factor and the market portfolio (which is also "contaminated" by the presence of the granular residual η_{t+1}).

One common empirical approach to test the implication (5) in the time series is to use a GARCH-in-mean model with a variance forecast in the excess return mean equation.[61] Following this convention in the literature, our empirical tests will

[61] To deal with potential misspecification, it is a common practice to include an intercept and a lagged excess return in the forecasting equation (e.g., Koopman and Hol Uspensky, 2002)

estimate the following specification

$$\sigma_{m,t+1}^2 = h\left(u_s, \sigma_{m,s}^2 \text{ for } s \leq t\right), \tag{30}$$

where u_{t+1} is a forecast error, the time-varying volatility of which is modeled in (12) with ε_{t+1} being a martingale difference sequence with unit variance. The conditional variance of u_{t+1} follows the GARCH process $h(\cdot)$ as a function of forecasting errors and variance forecasts available in period t. For example, $h\left(u_t, \sigma_{m,t}^2; \alpha_0, \alpha_1, \alpha_2\right) = \alpha_0 + \alpha_1 u_t^2 + \alpha_2 \sigma_{m,t}^2$ is a GARCH(1,1) model whose persistence is determined by $\alpha_1 + \alpha_2$. When the law of large numbers holds and the variance forecasting model is correctly-specified, such a test is well-grounded in theory (which would predict that $\phi_0 = \phi_1 = 0$).

Assuming that processes for the true factor and those for individual excess returns follow this simple model, we generate an exogenous process for the true market factor with a conditional variance following the GARCH(1,1) model with parameter estimates in Table (D1).[62] Then we generate processes for individual returns following (1) by using individual betas and idiosyncratic returns that are drawn from the same distribution. Individual betas and idiosyncratic returns $\eta_{i,t}$ are drawn from a Student-t distribution. For tractability, we scale down the standard deviation so that the distribution has a smaller standard deviation than its empirical counterpart.

Figure D1 exhibits Monte Carlo densities of the estimated risk-return tradeoff coefficients (γ) and biases in estimated betas of size portfolios (β_p). Panel 1 confirms the attenuation bias discussed above. The distribution of γ estimates obtained from the noisy proxy for the true market factor ('Market portfolio') are downward-biased, whereas the distribution of γ estimates from the true market factor ('True factor'). These results are consistent with (7) and our empirical results below, confirming non-trivial effects from the granular residuals by influencing estimates of the conditional factor variance (i.e., conditional variance of the true market factor). Panel 2 confirms the biases in OLS-estimated betas. For most portfolios, we find severely downward biased beta estimates obtained by regressing portfolio returns on the value-weighted market index returns, which is a noisy proxy due to the granular residuals. (Note that biases would be monotonically decreasing in size if we built in an empirically realistic negative correlation between size and market exposures.) One exception is the beta estimate of the largest portfolio, where positive biases stem from the mechanical positive association between large-sized stock's idiosyncratic shocks and the granular residuals in the market factor proxy. These results are consistent with (17)

[62] For our simulation purpose, we begin by estimating the GARCH(1,1) model using adjusted returns of the CRSP value-weighted index index. To do so, we estimate individual betas from the two-stage least squares estimation using daily CRSP equal-weighted index returns as an instrumental variable to value-weighted index returns. Then we estimate granular residuals as weighted averages of unexplained individual returns (i.e., residuals) from a cross-sectional regression of monthly individual returns on estimated betas. Table (D1) report parameter estimates from both unadjusted and adjusted index returns across three sample periods.

and generate patterns in $\hat{\alpha}$ with the opposite sign by equation (18).

So far, we have demonstrated that the granular residuals – undiversified idiosyncratic returns in a value-weighted index return – provide biased estimates for both the risk-return relationship and the portfolios' market risk exposures. The biased risk-return relationship stems from a fundamental source of misspecification and applies when the relevant variances are known. Thus, our conclusion is not specific to a particular method for estimating conditional variances, which was a popular topic in earlier research. In what follows, we propose two approaches to correct for the effects of the granular residuals in empirical asset pricing tests. Specifically, we address the attenuation bias in γ estimates directly, by subtracting empirical estimates for the granular residuals from market returns. We address the bias in β estimates indirectly, by using common factors in the cross section of stocks as instrumental variables for the value-weighted return.

D.2 Detailed description of simulation exercise

For our simulation exercises, we assume that processes for the true factor and those for individual returns follow the simple model of individual returns, introduced in Section 2. We further assume that the conditional variance of the true factor follows the GARCH(1,1) model in equation (13). When generating a process for the true factor, we use coefficient estimates for α_1 and α_2 that are obtained from the adjusted excess returns over the total sample period, reported in Table D1, which are going to be introduced in the following section. Following Lundblad (2007), we calibrate the value for α_0 to match the unconditional standard deviation of the simulated true factor process with that of the adjusted index returns (7.116 percent). Throughout our simulation exercises, the true risk-return tradeoff coefficient (γ) is set to 1.653.

Regarding characteristics for individual stock returns, we assume that all individual stocks have the same prices in the initial period and that their idiosyncratic returns each period are drawn from the same distribution. The individual stocks differ only by their initial exposures on the true market factor, which can be viewed as the market betas of individual stocks (henceforth, stock market betas). We generate a highly leptokurtic distribution of stock market betas by drawing candidates from a Student-t distribution with 10 degrees of freedom. Then we rescale the drawn candidates adequately – dividing by b – and add 1 to the rescaled candidates to make the distribution of the initial stock market betas centered around 1. Lastly, we winsorize extreme observations from both tails (2.5% observations from each tail) by corresponding cutoff values. These steps are intended to reproduce a typical distribution of the stock market betas observed in the data. We use the same initial stock market betas for all our simulation exercises. For simplicity, we normalize the number of stocks outstanding for each individual stock to be 1. Such normalization indicates that individual stocks' weights are determined by their price shares in the value-weighted index.

For each simulation, we draw a shock for the true market factor and idiosyncratic returns for 'N' individual stocks for '$T+s$' periods. We discard the first 's' periods of simulated processes to ensure the granular residuals endogenously appeared in excess returns of the value-weighted index. For the first simulation exercises in the following section, we set $N = 1,000$, $T = 1,000$, $s = 500$ and $b = 50$ to generate a sufficiently long and stationary sample, used for empirical evaluation of the risk-return relationship. For the second simulation exercises, we set $N = 1,000$, $T = 500$, $s = 60$ and $b = 4$ to generate a sample with individual stocks whose betas exhibit a sufficient degree of cross-sectional variation. The latter sample is used to evaluate market-risk exposures of individual stocks as well as portfolios constructed from individual stocks. While we draw the true market shock from a standard normal distribution, we draw idiosyncratic returns from a mean-zero normal distribution whose variance is four times larger than that of the true market factor. This follows from observations in Campbell et al. (2001), documenting that the market factor accounts for about 19% of monthly variation in individual stock returns. Given simulated processes for the true market factor and for the excess returns of the individual stocks, we generate the excess returns of the value-weighted index and the granular residuals. In Section D.3, we estimate the risk-return tradeoff coefficient (γ) from both the Unadjusted and Adjusted index returns using the GARCH(1,1) model, which is used to generate the true market factor. In Section D.4, we estimate the market risk exposure coefficients (βs) of individual stocks and portfolios in three ways, which are going to be discussed later in detail.

D.3 Biases in risk-return tradeoff relationships

Table D2 reports the distribution of generated processes for the true factor, granular residuals and excess returns obtained from 5,000 simulated sample paths. For each case, we report averages of mean, standard deviation, skewness, kurtosis of the generated processes. Panel A.1 report those from our baseline simulation exercise, where idiosyncratic returns are drawn from a mean-zero normal distribution with a variance two times larger (scale=2) than that of the true factor. Since it is matched with the standard deviation of the adjusted index returns, the average standard deviation of the simulated true factor is 0.07123 by construction. The averages of the mean, skewness, kurtosis of the simulated true factor are smaller than those of the adjusted index returns reported in Table 1 yet these differences do not affect our simulation results. On average, the distribution of the simulated granular residuals is zero, negatively skewed and highly leptokurtic. More importantly, the average standard deviation of the simulated granular residuals is 0.03844, which corresponds to about half of the true factor's standard deviation. The sizable standard deviation of the simulated granular residuals confirm our findings in Table 1 and adds further variability to the simulated excess returns. The average standard deviation of the excess returns is larger than that of the true factor, indicating that the excess return is a mis-measured proxy for the true factor. Given the sizable granular residuals, the average skewness and kurtosis of the excess returns slightly differ from those of the true factor while the average mean is close to that of the true factor.

Then we estimate the risk-return tradeoff coefficient (γ) from the GARCH(1,1) model using the simulated true factor and the excess returns. As mentioned earlier, the true risk-return tradeoff coefficient is set to 1.653 for all simulated paths. While the simulated true factor processes correspond to the adjusted index returns, the simulated excess returns are used in place of the unadjusted index returns, which are mis-measured due to the simulated granular residuals. Panel A.1 in Table D3 reports results from our baseline simulation. The first two columns report the mean and standard deviation of γ estimates and the following two columns report the bias and mean squared errors (MSE) of γ estimates relative to the true risk-return coefficient ($\gamma = 1.653$).[63] The last column indicates which parameter for each simulation case which is changed relative to the baseline simulation. From our baseline simulation, we find negatively biased coefficient estimates from both the adjusted and the unadjusted index returns. A potential explanation for the negatively biased estimate from the adjusted index returns is due to model misspecification. When simulating processes for the true market factor, we use the values for α_1, α_2 and γ, which are obtained by estimating the equation (11) – an empirical specification including an intercept and an additional regressor such as the lagged excess return. When estimating the risk-return tradeoff coefficient (γ) from the generated processes, we use the true model specified in equation (1), which does not include an intercept as well as the lagged true market factor.

More importantly, we find that the unadjusted index returns possess a more negative bias than the adjusted index returns. The unadjusted index returns, because of the large negative bias, have a larger MSE than the adjusted index returns, even though the latter estimates have a larger standard deviation relative to when the unadjusted index returns are used. Note that variation in the coefficient estimates in our simulation is driven by both variation in the true market factor process and variation in idiosyncratic returns, for which we randomly draw numbers from the normal distribution. To isolate effects of the granular residuals from those of uncertainty in the true market factor, we control for the variation in the true market factor process, fixing the realizations of true market factor process to one which provides an estimate which lines up with the population value, but varies realizations of idiosyncratic returns (i.e., the granular residual). The third row ('Unadjusted*') reports estimation results from the unadjusted index returns, where the variation across simulations is

[63] Let $\widetilde{\gamma}$ be the true risk-return tradeoff coefficient assumed in our simulation. Denoting by γ_s an estimate obtained during a simulation s for $s = 1, \cdots, S$, the bias is calculated as $Bias\left(\gamma\right) = \sum_{s=1}^{S}\left(r_s - \widetilde{\gamma}\right)/S$. And the mean squared errors is calculated as $MSE = Bias\left(\gamma\right)^2 + Var\left(\gamma\right)$, where $Var\left(r\right) = \sum_{s=1}^{S}\left(r_s - \widetilde{\gamma}\right)^2/S$.

purely driven by random realizations of idiosyncratic returns. Holding the true market factor process fixed, we find that the unadjusted index returns still give negatively biased estimates.

Given sizable biases from the granular residuals, we further investigate the role of idiosyncratic returns by drawing those from a mean-zero normal distribution with twice larger variance than the baseline. Panel A.2 in Table D2 reports the distribution of generated processes again. We do not report the distribution of the true market factor; since the true market factor process is not influenced by idiosyncratic returns, its characteristics are generally similar across different sets of simulation exercises. When idiosyncratic returns have larger variances than before, we find that the granular residuals have a larger standard deviation, which increases the standard deviation for the excess returns. The consequence of the larger variability in the granular residuals and, accordingly in the excess returns, is a larger negative bias and MSE, which is shown in Panel A.2 in Table D3. The average of the risk-return coefficient estimates is 0.601, which is far smaller than the true coefficient (1.689) as well as the mean of the true market factor (1.646).

Figure D2 exhibits Monte Carlo densities of the estimated risk return tradeoff coefficient (γ) obtained from our baseline simulation. Panel 1 confirms our findings in Panel A.1 in Table D3. The excess return ('Unadjusted'), which is a noisy proxy for the true market factor, gives negatively biased estimates for γ. The negative bias is mainly due to the granular residuals whose variation is driven by the realization of idiosyncratic returns. This conjecture is confirmed from our controlled experiment, where variation in the granular residuals (accordingly the excess returns) is purely driven by realizations of idiosyncratic returns. Although they have a small variance, the excess return ('Unadjusted*') still give negatively biased estimates for γ. Next, Panel 2 shows that the size of negative bias is associated with an idiosyncratic return volatility. When idiosyncratic returns have larger variance, the excess return ('Unadjusted (scale=4)') gives a larger negative bias than the baseline. The enlarged negative bias dominates a reduced variance, leading to a larger MSE than the baseline.

Next, we investigate the effects of other characteristics that potentially affect the granular residuals as well as the stock index returns. To do so, we consider two variations on our baseline simulation, where both variations are closely related to properties of the granular residuals. First, we consider a Student-t distribution, in which idiosyncratic returns exhibit higher kurtosis than those from the normal distribution. Specifically, we draw idiosyncratic returns from a Student-t distribution with 10 degrees of freedom ($\nu = 10$) and 7 degrees of freedom ($\nu = 7$), respectively. In these simulation exercises, the kurtoses for idiosyncratic returns are 4 ($\nu = 10$) and 5 ($\nu = 7$), which are larger than 3 obtained from the standard normal distribution. Second, we consider a larger number of individual stocks (N) than in the baseline simulation. From our exercises for $2,000$ and $5,000$ individuals in the stock index, we expect to answer the question for the role of idiosyncratic returns, providing a plausible guide for investments in the stock market index. Put another way, our main question in these simulation exercises lies on whether the idiosyncratic returns could be sufficiently diversified in the stock market index, which is a typical assumption in standard asset pricing theory.

Panels B and C in Table D2 report distribution of generated processes for the granular residuals and the excess returns from 5,000 simulated sample paths again. Here we generate the processes by changing our simulation setups as described above. When idiosyncratic returns are drawn from a highly leptokurtic distribution such as the Student-t distribution, the granular residuals have a larger standard deviation and kurtosis and also have a more negatively skewed distribution than the baseline (B.1). Furthermore, we find larger effects from a higher kurtosis for the idiosyncratic-return distribution, where all deviations in the standard deviation, skewness, kurtosis from the baseline model become larger when idiosyncratic returns have larger kurtoses (B.2). The highly leptokurtic idiosyncratic-return distribution, through the granular residuals, affect distributions of the excess returns, providing a higher standard deviation and a smaller skewness than the baseline. In contrast, effects for the kurtosis of the excess returns are ambiguous. Panel C confirms our intuition that the effects of the granular residuals are smaller for a large number of individuals in the stock index. The standard deviation of the granular residuals using a larger number of assets in the stock market index is smaller than that from the baseline ($N = 1,000$). On the other hand, the skewness and kurtosis of the granular residuals using a large number of stocks are larger than those from the baseline. More importantly, the granular residuals' variability, represented by the standard deviation, is the same magnitude as that of the excess returns. Even with $5,000$ individual stocks, the standard deviation is 0.02911, which is slightly smaller than 1/2 of that of the excess returns. The sizable standard deviation indicates that the excess returns are still noisy despite diversification effects.

Table D3 confirm our conjectures. When idiosyncratic returns exhibit excess kurtosis, the negative bias in estimates for the risk-return tradeoff coefficient become larger than in the baseline when using the unadjusted index returns (B.1). Furthermore, the estimates become more negatively biased if idiosyncratic returns exhibit a higher excess kurtosis (B.2). Although leptokurtic idiosyncratic returns result in smaller standard deviations, the negative biases dominate the reduced variances in the calculation of the MSE, giving large MSEs than the baseline. In Panel C, we confirm the limited diversification effects in the presence of the granular residuals. We find smaller negative biases from larger numbers of individuals stocks (C.1 and C.2) than the baseline (A.1). We find that the effects from the granular residuals, which can be viewed as noise in the stock market index, are reduced when using a large number of individual stocks. However, as our results in Panel B indicate, the effectiveness of the diversification could be heavily influenced by the characteristics of idiosyncratic returns. That is, depending on the idiosyncratic return distribution, one may need a larger number of individual stocks than that considered to achieve a sufficient diversification in the stock market index. Figure D3 exhibits Monte Carlo densities of the estimated risk return tradeoff coefficient (γ), confirming our findings in this subsection.

D.4 Biases in market-risk exposures

For reference, Table D4 reports the distribution of generated processes for the true factor, granular residuals and excess returns obtained from 5,000 simulated sample paths, which are used for our simulation exercises in this subsection.

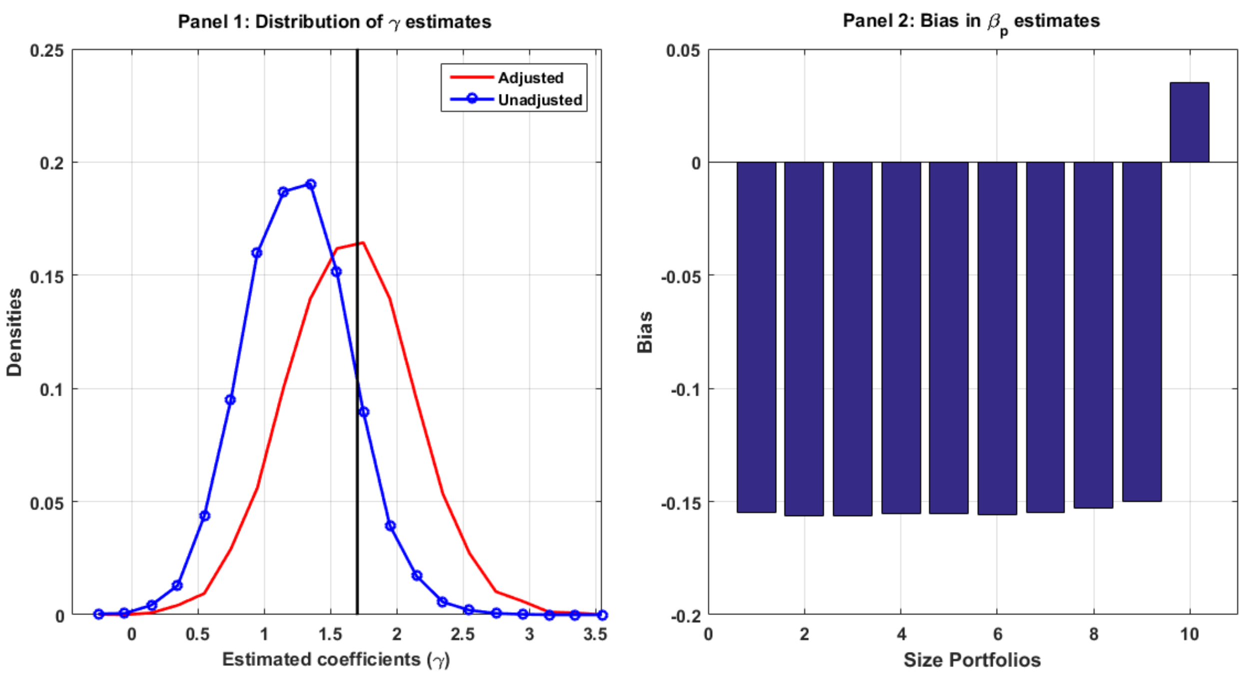
Table D5 reports average of asymptotic bias, variance and mean squared error (MSE) of individual stocks' beta estimates. Again, simulation results confirm our intuition, previously noted in Section 2. When the effects from the granular residuals are adequately corrected, we find that OLS-β estimates of individual stocks, on average, are downward biased. And the downward biases due to the granular residuals become more severe for larger granular residuals, whose sizes are largely influenced by the distribution of idiosyncratic returns. That is, we find larger biases from larger idiosyncratic variances (Panel A.2), and also from highly leptokurtic idiosyncratic returns (Panel B). Lastly, we find reduced biases in OLS-β estimates due to improved diversification in the stock market index. However, our simulation results indicate that the diversification effect itself may not be sufficient enough to eliminate the downward bias. In Panel C, the OLS-β estimates are smaller than the true individual betas by 0.053 on average. In contrast, we find small-sized biases from 2SLS-β estimates, for which we run the two-stage least squares estimation (2SLS) using returns on the equal-weighted index as an instrumental variable. And our findings from 2SLS-β estimates are robust regardless of characteristics of the stock market index as well as those of idiosyncratic returns.

Next, we investigate the effects of the granular residuals on portfolios' beta estimates. Given two characteristics of individual stocks such as their market capitalizations and betas, we form ten Beta portfolios based on betas of individual stocks, and also form ten Size portfolios based on market capitalizations (sizes) of individual stocks, proxied by their prices. Specifically, we form the Beta portfolios after controlling for sizes of individual stocks and also form the Size portfolios after controlling for individual betas. We control these characteristics since a stock's market capitalization in our simulation exercises tends to be correlated with its beta. To see this, recall the simple model introduced in Section 2. A stock's price at each period is endogenously determined by a realization of the true market factor and that of its idiosyncratic return, of which *ex-ante* distribution is identical. In such a case, a stock having a higher beta tends to be larger in size, provided by a positive market risk-premium assumed during our simulation exercise. The portfolios are rebalanced/updated at the end of each period. By construction, our Beta portfolios have similar market capitalizations (i.e. similarly sized) while having different portfolio-level betas. Similarly, our Size portfolios have betas of 1 while having different market capitalizations at the portfolio level.

Table D6 compares biases in OLS-β estimates ('Uncorrected) with those in 2SLS-β estimates ('Corrected') for the Beta (Panel A) and Size portfolios (Panel B). In both panels, we also report average market capitalizations (sizes) and betas of portfolios, where one can confirm the adopted portfolio formation mechanism. In Panel A, we find that the granular residuals cause attenuation biases in OLS-β estimates, which contrasts to small biases observed from 2SLS-β estimates. In particular, magnitudes of the attenuation biases are proportional to portfolios' betas, confirming our conjecture provided in equation 17. Our findings in Panel B are consistent with those in Panel A. One exception is that the OLS-β estimate for the largest Size portfolio is positively biased, which is also expected from the last term in equation 17. Here OLS-β estimates for the remaining Size portfolios produce similarly-sized attenuation biases since all Size portfolios' betas are set to 1 during the formation of the Size portfolios.

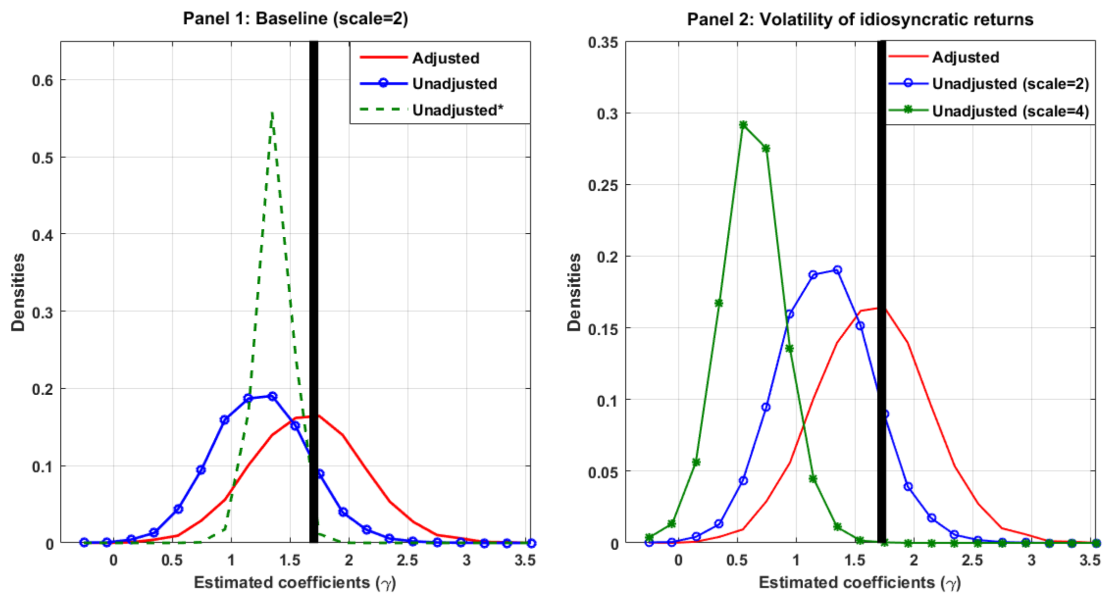
Then, we repeat the simulation exercises by considering various conditions that are relevant to the sizes of the granular residuals. Table D7 reports asymptotic biases (Panel A) in OLS-β estimates and mean squared errors for estimates (Panel B), obtained from ten Size portfolios. Again, our simulation results from the Size portfolios confirm those from individual stocks in Table D5. Lastly, Figures D4 and D5 confirm our findings. That is, OLS-β estimates for the smallest Size portfolio are downward biased, whereas those for the largest Size portfolio are upward biased.

Figure D1: Simulated biases in time-series intertemporal risk premium (γ) and cross sectional market factor exposure (β) estimates



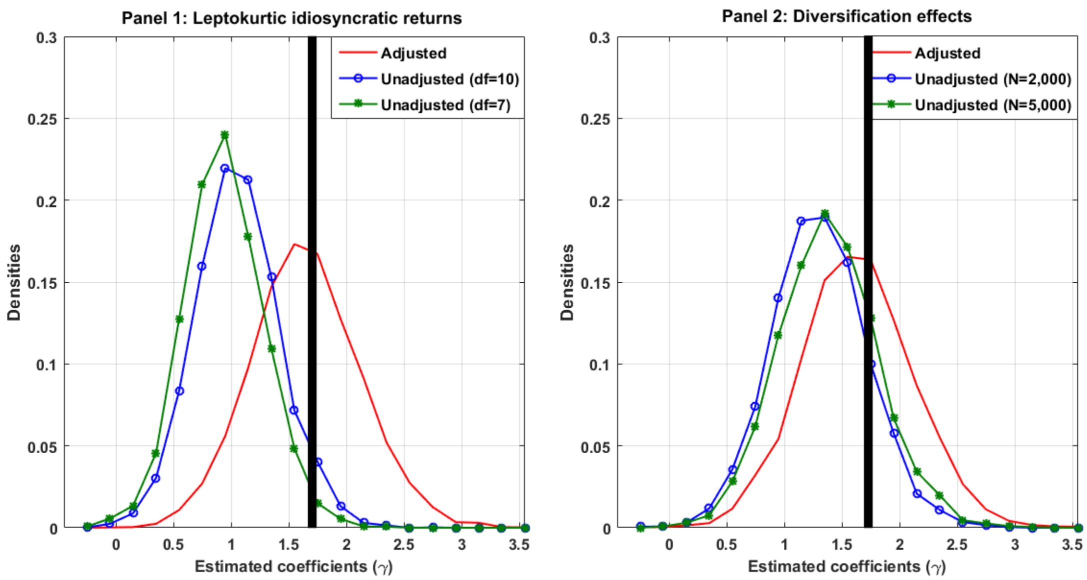
Note. This figure presents Monte Carlo densities of the estimated risk-return tradeoff coefficient (γ, Panel 1) and biases in the market-risk exposures of Size portfolios (β_p, Panel 2). In Panel 1, the solid line denotes densities of the risk-return tradeoff coefficient estimated by using simulated processes for the true market factor ('Adjusted'), whereas the line marked by circles denotes those estimated by using simulated processes for the market factor proxy ('Unadjusted'). The vertical solid line indicates the true risk-return coefficient assumed during our simulation exercise ($\gamma = 1.653$). Panel 2 presents biases in portfolio-level beta estimates obtained by regressing simulated returns of Size portfolios on simulated processes for the market factor proxy. At each period during our simulation exercise, the Size portfolios are formed based on market capitalizations of individual stocks after controlling for individual betas. Market-risk exposures (i.e., betas) of all Size portfolios are set to be 1 by construction. The true market factor and its proxy are generated for 5,000 times by simulating conditional variance of the true factor process and idiosyncratic returns of 1,000 individual stocks. While generating the conditional variance of the true factor process following GARCH(1,1) model using parameter estimates in Table (2), idiosyncratic returns are generated from mean-zero normal distribution whose variance is assumed to be two times larger than variance of true factor process. During our simulation exercise, we assume that all stocks have same initial prices, yet have heterogenous degrees in responding to the true market factor (i.e., different betas). Individual stocks' returns (accordingly, their capitalizations and weights in the market index) are determined by realizations of the true market factor – through different betas – and also by realizations of the idiosyncratic returns, which are drawn from the same distribution. Given an endogenously generated weight distribution, the granular residuals are also generated endogenously during our simulation exercise. For each simulation in Panel 1, samples are drawn for 1,500 periods whereas the first 500 samples are discarded and the remaining 1,000 samples are used to estimate the risk-return coefficient. For each simulation in Panel 2, samples are drawn for 560 periods where the last 60 samples are used to estimate the market-risk coefficient.

Figure D2: The risk return tradeoff and variance of idiosyncratic returns



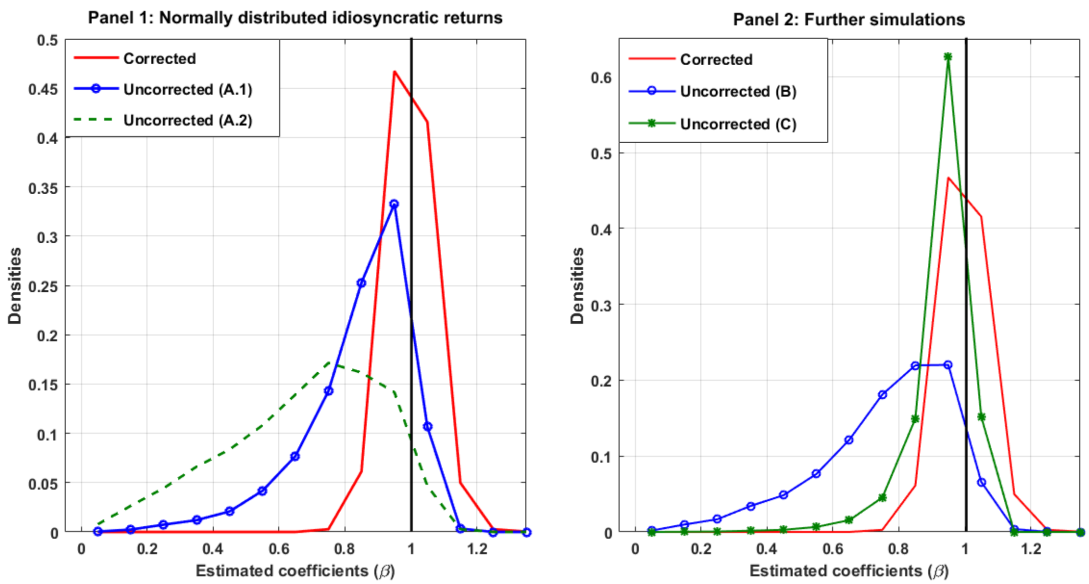
Note: This figures presents Monte Carlo densities of the estimated risk return tradeoff coefficient, γ. The solid line denotes densities of the risk-return tradeoff coefficient estimated by using adjusted index returns ('Adjusted'). The solid line marked by circles (o) denotes those estimated by using unadjusted index returns ('Unadjusted (scale=2)'), where idiosyncratic returns have two times larger variances than the true market factor. In Panel 1, the dashed line denotes those estimated by using unadjusted index returns ('Unadjusted*'), where the excess index returns are generated by holding the process for the true market factor. In Panel 2, the solid line marked by stars (*) denote those estimated by using unadjusted index returns ('Unadjusted (scale=4)') again, where idiosyncratic returns become to have four times larger variances than the true market factor. The excess returns and granular residuals are generated for 5,000 times by simulating conditional variance of the true factor process and idiosyncratic returns of 1,000 individual stocks (N=1,000). While generating the conditional variance of the true factor process following GARCH-M model by using parameter estimates in Table B1, idiosyncratic returns are generated from mean-zero normal distribution whose variance is assumed to be two (Panel 1) and four times ('Unadjusted (scale=4)' in Panel 2) larger than variance of the true factor process. Assuming all stocks have same number of outstanding shares, initial prices are also assumed to be identical (normalized to 1). Despite such identical distribution of initial weights, we generate granular residuals endogenously. To do so, samples are drawn for 1,500 periods whereas the first 500 samples are discarded and the remaining 1,000 samples are used to estimate risk-return coefficients (T=1,000). Across all Monte Carlo simulations, the risk-return coefficient is set to be 1.653 and the unconditional variance of the true factor process is matched with the unconditional variance of the observed excess return net of granular residuals. The vertical line indicates the true risk-return coefficient assumed during our simulation exercise.

Figure D3: The risk return tradeoff and the granular residuals



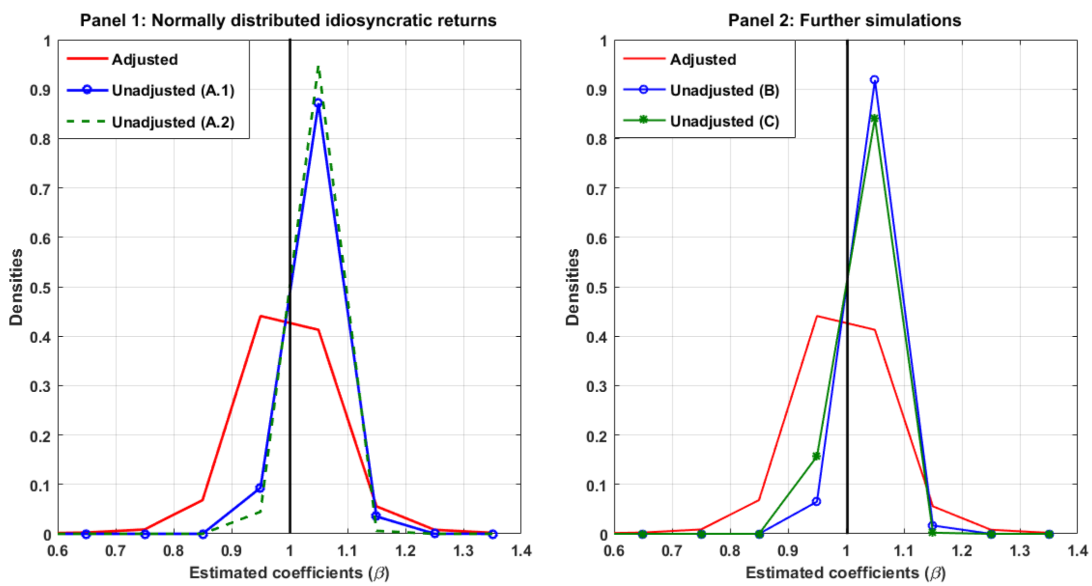
Note: This figures presents Monte Carlo densities of the estimated risk return tradeoff coefficient, γ. The excess returns and granular residuals are generated for 5,000 times by simulating conditional variance of the true factor process following GARCH-M model by using parameter estimates in Table B1, and also by simulating individual returns of 1,000 individual stocks (Panel 1) and larger number of individual stocks (Panel 2). For Panel 1, idiosyncratic returns are generated from mean-zero Student-t distribution whose degrees of freedom parameter is assumed to be ten and seven. Excess kurtoses of the Student-t distributed idiosyncratic returns are 1 and 2, respectively. Here the solid line denotes densities of the risk-return tradeoff coefficient estimated by using adjusted index returns ('Adjusted'). And the solid line marked by circles (o) denote those estimated by using unadjusted index returns when idiosyncratic returns are drawn from a Student-t distribution with 10 degrees of freedom ('Unadjusted (df=10)'). Similarly, the solid line marked by stars (*) denotes those when idiosyncratic returns are drawn from a Student-t distribution with 7 degrees of freedom ('Unadjusted (df=7)'). For Panel 2, idiosyncratic returns are generated from mean-zero normal distribution whose variance is assumed to be two times larger than the true factor variance. Here the solid line denotes densities of the risk-return tradeoff coefficient estimated by using adjusted index returns ('Adjusted'). The solid line marked by circles (o) denote those estimated by using unadjusted index returns of 2,000 individuals ('Unadjusted (N=2,000)'). And the solid line marked by stars (*) denote those estimated by using unadjusted index returns of 5,000 individuals ('Unadjusted (N=5,000)'). Remaining conditions are identical to the baseline simulation. Assuming all stocks have same number of outstanding shares, initial prices are also assumed to be identical (normalized to 1). Despite such identical distribution of initial weights, we generate granular residuals endogenously. To do so, samples are drawn for 1,500 periods whereas the first 500 samples are discarded and the remaining 1,000 samples are used to estimate risk-return coefficients (T=1,000). Across all Monte Carlo simulations, the risk-return coefficient is set to be 1.653 and the unconditional variance of the true factor process is matched with the unconditional variance of the observed excess return net of granular residuals. The vertical line indicates the true risk-return coefficient assumed during our simulation exercise.

Figure D4: The granular residuals and the smallest portfolio's beta estimate



Note: This figures presents Monte Carlo densities of the estimated risk measure coefficient, β of the smallest Size portfolio. The excess returns and granular residuals are generated for 5,000 times by simulating conditional variance of the true factor process following GARCH-M model by using parameter estimates in Table B1. For Panel 1, idiosyncratic returns are generated from normal distribution whose variance are two times ('Case A.1') and four times ('Case A.2') larger than that of the true market factor. The (red) solid line denotes Monte Carlo densities of 2SLS-β estimates ('Corrected'), for which the effects from the granular residuals are corrected. The (blue) solid line with circles (o) denotes those of OLS-β estimates ('Uncorrected (A.1)'), for which simulated processes for Case A.1 are used. The dotted line denotes those of OLS-β estimates ('Uncorrected (A.2)'), for which simulated processes for Case A.2 are used. In Panel 2, the (red) solid line again denotes the same Monte Carlo densities of 2SLS-β estimates ('Corrected'), plotted in Panel 1. The (blue) solid line with circles (o) denotes those of OLS-β estimates ('Uncorrected (B)') obtained from simulated processes, in which idiosyncratic returns are generated from mean-zero Student-t distribution with degrees of freedom parameter being ten. Lastly, the (green) solid line with stars (*) denotes Monte Carlo densities of OLS-β estimates ('Uncorrected (C)'), for which simulated processes from 5,000 individual stocks are used. Remaining conditions are identical to the baseline simulation (A.1). Assuming all stocks have same number of outstanding shares, initial prices are also assumed to be identical (normalized to 1). Despite such identical distribution of initial weights, we generate granular residuals endogenously. To do so, samples are drawn for 1,500 periods whereas the first 500 samples are discarded and the remaining 1,000 samples are used to estimate risk-return coefficients (T=1,000). Across all Monte Carlo simulations, the risk-return coefficient is set to be 1.653 and the unconditional variance of the true factor process is matched with the unconditional variance of the observed excess return net of granular residuals. The vertical line indicates the true risk-return coefficient assumed during our simulation exercise.

Figure D5: The granular residuals and the largest portfolio's beta estimate



Note: This figures presents Monte Carlo densities of the estimated risk measure coefficient, β of the largest Size portfolio. The excess returns and granular residuals are generated for 5,000 times by simulating conditional variance of the true factor process following GARCH-M model by using parameter estimates in Table B1. For Panel 1, idiosyncratic returns are generated from normal distribution whose variance are two times ('Case A.1') and four times ('Case A.2') larger than that of the true market factor. The (red) solid line denotes Monte Carlo densities of 2SLS-β estimates ('Corrected'), for which the effects from the granular residuals are corrected. The (blue) solid line with circles (o) denotes those of OLS-β estimates ('Uncorrected (A.1)'), for which simulated processes for Case A.1 are used. The dotted line denotes those of OLS-β estimates ('Uncorrected (A.2)'), for which simulated processes for Case A.2 are used. In Panel 2, the (red) solid line again denotes the same Monte Carlo densities of 2SLS-β estimates ('Corrected'), plotted in Panel 1. The (blue) solid line with circles (o) denotes those of OLS-β estimates ('Uncorrected (B)') obtained from simulated processes, in which idiosyncratic returns are generated from mean-zero Student-t distribution with degrees of freedom parameter being ten. Lastly, the (green) solid line with stars (*) denotes Monte Carlo densities of OLS-β estimates ('Uncorrected (C)'), for which simulated processes from 5,000 individual stocks are used. Remaining conditions are identical to the baseline simulation (A.1). Assuming all stocks have same number of outstanding shares, initial prices are also assumed to be identical (normalized to 1). Despite such identical distribution of initial weights, we generate granular residuals endogenously. To do so, samples are drawn for 1,500 periods whereas the first 500 samples are discarded and the remaining 1,000 samples are used to estimate risk-return coefficients (T=1,000). Across all Monte Carlo simulations, the risk-return coefficient is set to be 1.653 and the unconditional variance of the true factor process is matched with the unconditional variance of the observed excess return net of granular residuals. The vertical line indicates the true risk-return coefficient assumed during our simulation exercise.

Table D1: Parameter estimates: One-factor model

Models	Unadjusted		Adjusted	
Panel A: 1928.1 - 2014.12				
ϕ_0	0.005**	(0.002)	0.004	(0.002)
ϕ_1	0.074**	(0.035)	0.199***	(0.034)
γ	1.319	(0.920)	1.689**	(0.771)
α_0 $\left(\times 10^3\right)$	0.069***	(0.027)	0.137***	(0.041)
α_1	0.134***	(0.024)	0.142***	(0.030)
α_2	0.845***	(0.025)	0.825***	(0.032)
Panel B: 1928.1 - 1964.6				
ϕ_0	0.008***	(0.003)	0.007**	(0.003)
ϕ_1	0.102*	(0.060)	0.173***	(0.058)
γ	0.556	(0.993)	0.940	(0.773)
α_0 $\left(\times 10^3\right)$	0.059	(0.038)	0.085**	(0.042)
α_1	0.144***	(0.038)	0.177***	(0.042)
α_2	0.848***	(0.035)	0.817***	(0.036)
Panel C: 1964.7 - 2014.12				
ϕ_0	0.002	(0.004)	-0.006	(0.008)
ϕ_1	0.049	(0.043)	0.214***	(0.042)
γ	2.800	(1.914)	5.269*	(2.753)
α_0	0.092**	(0.042)	0.242**	(0.099)
α_1	0.117***	(0.033)	0.098**	(0.040)
α_2	0.841***	(0.036)	0.817***	(0.052)

Note: This table reports maximum likelihood estimation (MLE) estimates of model parameters in the univariate GARCH model across three sample periods: Total sample (Panel A), the first sample (Panel B), the second sample (Panel C). In each panel, we report parameter estimates and asymptotic standard errors (in parentheses) for excess returns of the CRSP value-weighted index returns without (Unadjusted) and with adjusting granular residuals (Adjusted). */**/*** represent the statistical significance at 90%, 95% and 99%, respectively.

Table D2: Monte Carlo distribution of simulated processes (1)

	Mean (×100)	Stdev. (×100)	Skewness	Kurtosis
Panel A: Normal distributed idiosyncratic returns				
A.1: Baseline (scale=2)				
True factor	0.818	7.123	0.447	6.895
Granular residuals	0.001	3.844	−0.061	4.671
Excess returns	0.819	8.162	0.304	5.507
A.2: Higher volatility of idiosyncratic returns (scale=4)				
Granular residuals	0.000	7.956	−0.046	4.172
Excess returns	0.818	10.755	0.129	4.230
Panel B: Student-t distributed idiosyncratic returns				
B.1: Leptokurtic idiosyncratic returns (d.f.=10)				
Granular residuals	−0.007	4.966	−0.069	5.492
Excess returns	0.795	8.766	0.245	5.210
B.2: Higher kurtosis of idiosyncratic returns (d.f.=7)				
Granular residuals	−0.006	6.066	−0.111	7.884
Excess returns	0.803	9.514	0.173	6.162
Panel C: Diversification effects				
C.1: Larger number of individual stocks (N=2,000)				
Granular residuals	0.002	3.506	−0.083	4.861
Excess returns	0.804	8.012	0.323	5.712
C.2: Even larger number of individual stocks (N=5,000)				
Granular residuals	−0.005	2.911	−0.083	5.267
Excess returns	0.795	7.729	0.333	5.841

Note: This table provides average of summary statistics such as mean, standard deviation, skewness, kurtosis of simulated processes for the true factor, excess returns and granular residuals in the value-weighted index. Using parameter estimates in Table B1, those processes are obtained from 5,000 Monte Carlo simulations. Panel A reports the average of summary statistics when idiosyncratic returns are generated from a normal distribution, where variance of idiosyncratic returns are assumed to be two times (A.1) and four times (A.2) larger than variance of the true factor process. Panel B reports those when idiosyncratic returns are generated from a Student-t distribution. Degrees of freedom parameters are 10 and 7, which represent leptokurtic idiosyncratic returns having excess kurtosis of 1 (B.1) and 2 (B.2) respectively. Lastly, Panel C reports those when considering larger number of individual stocks compared to the baseline having 1,000 individual stocks. Number of individual stocks are 2,000 (C.1) and 5,000 (C.2). For each simulation, conditional variance of the true factor process is generated from a GARCH(1,1) model using parameter estimates in Table B1. Given a generated process for the conditional variance, the true factor process is generated as a sum of risk-premium and innovation process of the true factor process. The risk-premium is defined as a product of a risk-return tradeoff coefficient (price for risk) and a conditional variance (quantity for risk), and the risk-return tradeoff coefficient, γ, is set to 1.653. The innovation process follows mean-zero normal distribution having time-varying conditional variance of the true factor process. Lastly, we match unconditional variance of the true factor process with unconditional variance of the 'adjusted' excess value-weighted index return used for the empirical analysis.

Table D3: Monte Carlo distribution of estimated risk-return tradeoff coefficient (γ)

	Mean	Stdev.	Bias	MSE	Note
Panel A: Normal distributed idiosyncratic returns					
A.1: Baseline					scale=2
Adjusted	1.646	0.458	−0.007	0.210	
Unadjusted	1.195	0.370	−0.458	0.347	
Unadjusted*	1.278	0.125	−0.375	0.156	
A.2: More volatile idiosyncratic returns					scale=4
Unadjusted	0.601	0.256	−1.052	1.172	
Panel B: Student-t distributed idiosyncratic returns					
B.1: Leptokurtic distribution	0.986	0.324	−0.667	0.549	d.f.=10
B.2: More leptokurtic distribution	0.887	0.309	−0.766	0.681	d.f.=7
Panel C: Diversification effects					
C.1: Larger number of stocks	1.225	0.381	−0.428	0.328	N=2,000
C.2: Even larger number of stocks	1.330	0.384	−0.323	0.252	N=5,000

Note: This table provides summary statistics of γ distribution obtained from GARCH(1,1) model when using unadjusted and adjusted index returns. To isolate effects of endogenously generated granular residuals from those of uncertainty in conditional variance process, we also estimate risk-return tradeoff coefficients from the simulated processes in which the conditional variance of the true factor process is fixed, providing exactly same estimate for the risk-return tradeoff coefficient (γ=1.653). The third row in A.1 (under 'Unadjusted*') reports estimation results of which the variation across simulations is purely driven by randomness of idiosyncratic returns. See footnote in the previous table for detailed explanations regarding Monte Carlo simulations.

Table D4: Monte Carlo distribution of simulated processes (2)

	Mean (×100)	Stdev. (×100)	Skewness	Kurtosis
Panel A: Normal distributed idiosyncratic returns				
A.1: Baseline (scale=2)				
True factor	0.727	6.665	0.932	9.198
Granular residuals	0.001	1.145	−0.076	5.654
Excess returns	0.728	6.772	0.889	8.827
A.2: Higher volatility of idiosyncratic returns (scale=4)				
Granular residuals	0.001	2.063	−0.099	6.095
Excess returns	0.728	7.006	0.805	8.173
Panel B: Student-t distributed idiosyncratic returns (d.f.=10)				
Granular residuals	0.002	1.552	−0.105	6.463
Excess returns	0.730	6.862	0.870	8.604
Panel C: Diversification effects (N=5,000)				
Granular residuals	−0.001	0.596	−0.088	6.333
Excess returns	0.727	6.695	0.932	9.147

Note: This table provides average of summary statistics such as mean, standard deviation, skewness, kurtosis of simulated processes for the true factor, excess returns and granular residuals in the value-weighted index. Using parameter estimates in Table B1, those processes are obtained from 5,000 Monte Carlo simulations. Panel A reports the average of summary statistics when idiosyncratic returns are generated from a normal distribution, where variance of idiosyncratic returns are assumed to be two times (A.1) and four times (A.2) larger than variance of the true factor process. Panel B reports those when idiosyncratic returns are generated from a Student-t distribution with a degrees of freedom parameter being 10, which represents leptokurtic idiosyncratic returns having excess kurtosis of 1. Lastly, Panel C reports those when considering larger number of individual stocks (N=5,000) compared to the baseline having 1,000 individual stocks. For each simulation, conditional variance of the true factor process is generated from a GARCH(1,1) model using parameter estimates in Table B1. Given a generated process for the conditional variance, the true factor process is generated as a sum of risk-premium and innovation process of the true factor process. The risk-premium is defined as a product of a risk-return tradeoff coefficient (price for risk) and a conditional variance (quantity for risk), and the risk-return tradeoff coefficient, γ, is set to 1.653. The innovation process follows mean-zero normal distribution having time-varying conditional variance of the true factor process. Lastly, we match unconditional variance of the true factor process with unconditional variance of the 'adjusted' excess value-weighted index return used for the empirical analysis.

Table D5: Monte Carlo distribution of estimated beta (β)

	Bias	Variance	MSE
Panel A: Normal distributed idiosyncratic returns			
A.1: Baseline (scale = 2)			
Corrected	0.005	0.046	0.046
Uncorrected	−0.128	0.052	0.070
A.2: More volatile idiosyncratic returns (scale =4)			
Corrected	0.010	0.076	0.077
Uncorrected	−0.277	0.077	0.159
Panel B: Student-t distributed idiosyncratic returns (d.f.=10)			
Corrected	0.064	1.097	1.101
Uncorrected	−0.609	0.041	0.436
Panel C: Diversification effects (N=5,000)			
Corrected	0.002	0.063	0.063
Uncorrected	−0.053	0.047	0.050

Note: This table reports asymptotic bias, variance and mean squared error (MSE) of individual beta estimates. The reported are average of those statistics among 1,000 individual beta estimates in Panels A and B, and average among 5,000 individual betas in Panel C.

Table D6: Monte Carlo distribution of estimated portfolio-level market risk exposures (β_p)

	1	2	3	4	5	6	7	8	9	10
Panel A: Beta portfolios										
Bias										
Uncorrected	−0.037	−0.051	−0.060	−0.063	−0.065	−0.071	−0.075	−0.077	−0.080	−0.095
Corrected	0.003	0.005	0.000	0.004	0.003	0.005	0.004	0.002	0.004	0.003
Characteristics										
Size	7.002	7.080	7.128	7.157	7.198	7.218	7.240	7.281	7.325	7.354
Beta	0.599	0.756	0.834	0.894	0.948	0.999	1.055	1.122	1.217	1.356
Panel B: Size portfolios										
Bias										
Uncorrected	−0.154	−0.156	−0.156	−0.155	−0.155	−0.155	−0.155	−0.153	−0.149	0.035
Corrected	0.006	0.005	0.005	0.007	0.007	0.005	0.006	0.006	0.006	−0.001
Characteristics										
Size	1.562	2.812	3.582	4.204	4.765	5.311	5.880	6.518	7.340	9.368
Beta	0.997	0.998	0.999	0.999	1.000	1.000	1.001	1.001	1.002	0.999

Note: This table reports results from our simulation exercises for the evaluation of the portfolio-level market risk exposure. In each panel, the first two rows report portfolios' characteristics such as sizes and betas of portfolios, where the portfolio size represents a logarithm of portfolio-level capitalization. And the following two rows report asymptotic bias of portfolio-level beta estimates with (Corrected) and without correcting for the effects from the granular residuals (Uncorrected). Panel A reports those of ten Beta portfolios constructed from individual stocks after controlling for individual stocks' sizes, and Panel B reports those of ten Size portfolios after controlling for individual stocks' betas.

Table D7: Monte Carlo distribution of estimated market risk exposures (β_p) for Size portfolios

Cases	1	2	3	4	5	6	7	8	9	10
Panel A: Bias										
Case A.1	−0.154	−0.156	−0.156	−0.155	−0.155	−0.155	−0.155	−0.153	−0.149	0.035
Case A.2	−0.323	−0.326	−0.324	−0.324	−0.324	−0.325	−0.325	−0.323	−0.321	0.034
Case B	−0.233	−0.232	−0.234	−0.233	−0.234	−0.234	−0.232	−0.232	−0.228	0.036
Case C	−0.062	−0.062	−0.062	−0.062	−0.062	−0.062	−0.062	−0.062	−0.061	0.013
Panel B: Mean squared error										
Case A.1	0.050	0.050	0.049	0.049	0.049	0.049	0.049	0.048	0.047	0.002
Case A.2	0.159	0.159	0.159	0.158	0.158	0.159	0.159	0.158	0.156	0.002
Case B	0.094	0.093	0.093	0.093	0.094	0.094	0.093	0.093	0.091	0.002
Case C	0.012	0.012	0.012	0.012	0.012	0.012	0.011	0.011	0.011	0.000

Note: This table reports asymptotic biases (Panel A) and mean squared errors (Panel B) of estimates for risk exposures of Size portfolios, constructed from individual stocks. The Size portfolios are formed based on market capitalizations of individual stocks after controlling for individual stocks' betas. At the end of each period, we update and rebalance the Size portfolios. At both panels, we report results from our simulation exercises executed under four cases. While 'Case A.1' represents our baseline simulation, 'Case A.2' considers the idiosyncratic variance, which are four times larger than the true factor variance. 'Case B' considers Student-t distributed idiosyncratic returns, of which excess kurtosis is set to 1 (i.e., degrees of freedom parameter being 10). 'Case C' considers large number of individual stocks (N=5,000).