The unintended consequences of grouping in tests of asset pricing models

Robert R. Grauer a,*, Johannus A. Janmaat b

a Faculty of Business Administration, Simon Fraser University, 8888 University Drive, Burnaby, BC, Canada V5A 1S6
b Department of Economics, Acadia University, Wolfville, Nova Scotia, Canada B0P 1X0

Received 19 July 2001; accepted 13 May 2003
Available online 9 January 2004

Abstract

We identify a number of unintended consequences of grouping when the capital asset pricing model is true and when it is false. When the model is true, grouping may cause fundamental problems with the most basic capital asset pricing and cross-sectional regression relationships. For example, with traditional grouping, the market portfolio is super-efficient—unless securities in each group are value weighted. Yet, when the model is grossly false, grouping may cause the model to appear to be absolutely correct. Ironically, the only way this can occur is when securities in each group are value weighted. To make matters worse, when the model is false, the slope of a cross-sectional regression of expected returns on betas fitted to grouped data may be either steeper or flatter than when the regression is fitted to ungrouped data. In other words, grouping may exacerbate the very problem it was meant to alleviate.

The analysis in this paper is based on three ideas.

First, early tests of the mean-variance capital asset pricing model (MV CAPM) are conducted on individual securities. When means are regressed on betas, the slope is too flat and the intercept is too high. A possible explanation follows from the observation that, in any univariate
regression, measurement error in the independent variable will cause the estimated slope to be flatter than the true slope.

In order to alleviate the measurement-error problem most subsequent tests of asset pricing models employ grouped data. Securities are sorted by some characteristic such as beta, market value of equity, book-to-market equity, or earnings-to-price. Then, each security is assigned to one (and only one) group.

Second, securities plot on the Security Market Line (SML) if and only if the market portfolio is MV efficient. S1

Arguably S1 is the most fundamental theoretical statement one can make about the MV CAPM.

Third, in a cross-sectional regression of means on betas, the intercept is the return on a zero-beta portfolio and the slope is the excess return on the market portfolio. In a Generalized Least Squares (GLS) regression, the slope and intercept sum to the expected return on the market. S2

Arguably S2 is the most fundamental set of statements that one can make about the cross-sectional regression methodology used to test the MV CAPM.

The econometric benefits of grouping in reducing measurement error in cross-sectional regressions of means on betas are beyond question. Moreover, there are clear benefits in reducing the dimension of the covariance matrix needed to trace out the minimum-variance frontier in tests that focus on the MV efficiency of the market portfolio. However, nothing is without cost. In this paper, we examine the costs or unintended consequences of two forms of grouping in a world where there is no measurement error. In traditional grouping, securities are combined into portfolios and the grouped data are employed in the tests. In Fama–French grouping, securities are combined into portfolios, each security in a portfolio is assigned its portfolio beta, and individual security data are employed in the tests. The analysis builds on, and extends, the body of literature that questions tests of the CAPM based on logical rather than statistical considerations.

Let \( x_m \) and \( w_q \) be \( n \)-vectors containing the weights in the market portfolio and portfolio \( q \), where \( x_{jm} \) and \( w_{jq} \) are security \( j \)'s weight in the respective portfolios. The securities in portfolio \( q \) are defined to be value weighted when \( w_{jq} = x_{jm} / \sum_k x_{km} \) for the securities included in portfolio \( q \), and zero for the securities not included in portfolio \( q \). (For notational purposes, vectors and matrices are set in bold type, vectors are column vectors, a prime indicates transposition e.g. \( x' \) is the row vector corresponding to the column vector \( x \), and scalars are italicized.) Then, the three primary costs of traditional grouping may be stated as follows. First, suppose that the CAPM is true i.e. statement S1 holds for the \( n \)-primitive securities in the economy. The only way that statement S1 holds for grouped data is if we value weight the securities in each group. Otherwise, the market portfolio is super-efficient.
relative to (plots outside) the grouped frontier. Second, grouping may cause the CAPM to appear to be absolutely correct when it is patently false. In a truly ironic twist, the only way this can occur is if we value weight the securities in each group. Third, when the CAPM is false, the slope of a cross-sectional regression of expected returns on betas fitted to grouped data may be either steeper or flatter than when the regression is fitted to the ungrouped data. The primary costs of Fama–French grouping occur when the CAPM is true—statements S1 and S2 do not hold.

In the first direct test of the Sharpe (1964)–Lintner (1965) MV CAPM, Douglas (1969) finds the estimated slope of the SML is too flat and the intercept is too large. In an influential paper, Miller and Scholes (1972) replicate the Douglas study and provide a detailed analysis of the possible econometric difficulties involved in estimating the relationship. They conclude that measurement error in the betas seems to contribute to the Douglas result. That is, with any univariate regression measurement error in the independent variable can cause the estimated slope to be flatter than the true slope. Shortly thereafter, Black et al. (1972), Blume and Friend (1973), and Fama and MacBeth (1973) produce the first extensive tests of the model, grouping all NYSE stocks into ten equal-weighted beta-ranked portfolios in order to reduce the measurement error in the betas. With the reduction in measurement error, the relationship between the portfolios’ expected returns and betas is near linear. Nonetheless, the slope is still too flat and the intercept is too large.

More recently, Fama and French (1992) find that the slope is much too flat and (statistically) insignificantly different from zero in the post-1963 period. They report: “In a shot straight at the heart of the SLB model, the average slope from the regressions of returns on $\beta$ alone . . . is 0.15% per month and only 0.46 standard errors from 0”. Fama and French (1992, p. 438). In fact, size and book-to-market equity combine to capture most of the cross-sectional variation in average stock returns associated with beta, size, leverage, book-to-market equity, and earnings–price ratios. Since the last four variables are measured precisely for individual stocks, there is no reason to employ grouped data in regressions containing these variables. Therefore, Fama and French group securities to estimate betas, assign the group’s beta to each stock in the group, and employ individual stock data in the regressions.

Roll (1977) is the first to appreciate the importance of statement S1. He argues that the theory asserts that a particular portfolio, the market portfolio, is MV efficient. Thus, the theory is not testable (at least in terms of the SML) unless that portfolio is observable and used in the tests. The subsequent papers in this stream of the literature sidestep this fundamental problem by implicitly assuming that all the assets

---

2 This result assumes that each security is assigned to one and only one group. However, if contrary to the grouping method employed in any paper we are aware of, securities were assigned to more than one group, it is possible that the grouped and ungrouped frontier could coincide. In particular, if we assigned the securities to two portfolios or groups that plot on the ungrouped frontier, we could trace out the whole ungrouped frontier with the grouped data. But, an analysis of this problem is beyond the scope of this paper.

3 See, for example, Sections 9–4 in Johnston (1972) or Jensen (1972, p. 366).

4 SLB is shorthand for Sharpe, Lintner, and Black.
are observed. However, like Roll (1977), the papers make their points in a single-period framework employing population parameters and illustrate the points with numerical examples.

Roll and Ross (1994) and Kandel and Stambaugh (1995), reacting to Fama and French’s (1992) finding that there is virtually no relationship between expected returns and betas, highlight the danger of focusing exclusively on mean-beta space. Roll and Ross (1994) demonstrate that a market proxy can be almost MV efficient even though the slope from an ordinary least squares (OLS) regression of population expected returns on population betas is zero. Conversely, Kandel and Stambaugh (1995) show that there can be a near perfect OLS fit between means and betas calculated relative to a portfolio that is grossly inefficient. More importantly, they show that in a GLS regression of mean returns on betas, the slope and $R^2$ are determined uniquely by the mean-variance location of the market index relative to the minimum-variance boundary. In contrast to OLS, GLS gives a zero slope only if the mean return on the market index equals that of the global minimum-variance portfolio. This latter result is derived earlier by Roll (1985).

However, neither Roll and Ross nor Kandel and Stambaugh verify whether the minimum-variance frontier contains a portfolio with all positive weights. Thus, we cannot be sure whether their results hold if the CAPM is true i.e. if the positively weighted market portfolio is MV efficient. Furthermore, any reasonable proxy portfolio should also contain all positive weights. While Roll and Ross are able to construct one example where a proxy portfolio contains all positive weights, the proxies in Kandel and Stambaugh’s paper do not.

Grauer (1999) overcomes these shortcomings by examining scenarios where the CAPM is true and where it is false. When the MV CAPM is true, each of the following four statements holds: (1) the market portfolio is MV efficient, (2) there is at least one positively weighted efficient portfolio, (3) the market portfolio is the tangency portfolio, and, (4) securities plot on the SML. In the examples where the CAPM is false, none of the statements holds. When the CAPM is true, the findings are not as dramatic as those reported by Roll and Ross and Kandel and Stambaugh. For example, the slopes of OLS and GLS regressions that employ betas calculated from almost efficient positively weighted proxy portfolios are too flat—but not equal to zero. On the other hand, when the CAPM is false and the market portfolio is almost efficient, the slope of a regression of means on market betas may be zero. Perhaps more importantly, the coefficients of OLS, GLS, or both OLS and GLS regressions that employ market betas may take on exactly the same values as when the model is true, even though the market is grossly inefficient.

In this paper, we extend this line of analysis to identify the costs of traditional grouping, where securities are combined into portfolios and grouped data are employed in the tests, and Fama–French grouping, where securities are combined into...
portfolios, each security in a portfolio is assigned its portfolio beta, and individual security data are employed in the tests.

Of course, we are not the first to point out problems associated with grouping. But for the most part the problems associated with grouping studied by others are quite different from those examined here. The exceptions are Roll (1977, 1979). Roll (1977, pp. 131–132) foreshadows one of the problems we raise when he warns “the widely-used portfolio grouping procedure can support the theory even when it is false. This is because individual asset deviations from exact linearity can cancel out in the formation of portfolios.” Roll (1979) examines the problem in detail. Litzenberger and Ramaswamy (1979) are the first to correct for estimation error using individual assets and Huang and Litzenberger (1988) provide an excellent discussion of the econometric problems associated with grouping.

More recently, on the theoretical side, Berk (2000) analyzes the implications of sorting data into groups and then running the tests within each group as Daniel and Titman (1997) do. He shows that by picking enough groups to sort into a researcher can destroy the within-group explanatory power of a correct asset pricing model. Lo and MacKinlay (1990) focus on data snooping biases arguing that tests of asset pricing models may yield misleading inferences when data are grouped according to some empirically motivated characteristic such as market value of equity.

On the empirical side, Kim (1995, 1997) employs individual assets and a weighted least squares version of maximum likelihood estimation to deal with the measurement error problem. Liang (2000) argues that the portfolio approach could suffer when the sorting variables contain the true values plus measurement errors because the grouped measurement errors would be embedded in the data. He develops a random sampling approach to overcome the problem and applies it to investigate beta shifts, showing that previous results about beta shifts are driven by measurement errors. It is also disconcerting to see that the results of empirical tests differ depending on how securities are grouped. See, for example, the differences between the results reported in Fama and French (1992) and Kothari et al. (1995). Moreover, Fama and French (1996, p. 74) assert that Fama and French’s (1993) three-factor model outperforms the CAPM because: “The average absolute pricing errors (intercepts) of the CAPM are large (25–30 basis points per month), and they are three to five times those of the three-factor model (5–10 basis points per month)” when portfolios are formed on the basis of firm-specific attributes, e.g., size, book-to-market equity, earnings to price, sales rank, or cash flow to price. Yet, Fama and French (1997) find that the pricing errors of the two models are of similar magnitude when the models are confronted with industry data, while Grauer (2002) reports that the pricing errors for the three-factor model are larger in a different industry data set. Finally, it is unsettling that different ways of grouping by size itself can lead to different inferences. See Kan and Zhang (1995).

The paper proceeds as follows. Section 1 analyzes the costs of the traditional grouping method. Section 2 describes the data used to create examples. Section 3 contains the examples that illustrate the costs of both traditional and Fama–French grouping. Section 4 contains a summary. In the body of the paper, we transform the ungrouped $n$-asset MV problem into a grouped MV problem defined over $q$...
portfolios. In Appendix A, we formulate the MV problem as an $n$-asset problem subject to a set of linear constraints. This formulation explicitly captures the restrictions on individual assets that are imposed by grouping. In turn, it provides the intuitive reason why, when the CAPM is true, the grouped frontier, which is subject to the budget constraint and the grouping restrictions, must lie inside the ungrouped frontier, which is only subject to the budget constraint—unless the grouping restrictions are not active at the single point where the two frontiers are tangent. We then establish the equivalence of the two ways of analyzing grouping illustrating the ideas with a numerical example.

1. Analysis of the traditional grouping method

Consider the MV problem

\[
\text{Max } L = \{\eta'x - (1/2)x'\Sigma x\} + \lambda (1 - t'x),
\]

where $\mu$, $x$, and $t$ are $n$-vectors containing unity plus expected rates of return, portfolio weights, and ones, respectively, $\Sigma$ is an $n \times n$ positive-definite covariance matrix of asset returns, $\lambda$ is a Lagrange multiplier associated with the budget constraint $t'x = 1$, and $t$ is a scalar parameter. The optimality conditions are

\[
\mu = (\lambda/t)1 + (1/t)\Sigma x \quad \text{and} \quad t'x = 1,
\]

where $\Sigma x$ is an $n$-vector whose $j$th element is the covariance of the return on asset $j$ with the return on the optimal portfolio. The optimal portfolio and its expected return can be written as linear functions of $t$, while the variance of the return on the optimal portfolio is a quadratic function of $t$. The minimum-variance frontier is traced out as $t$ varies from $-\infty$ to $+\infty$. (See Appendix A for details.) Eq. (2) shows that there is a linear relationship between each asset’s mean and its covariance with the optimal portfolio.

These results are simply mathematical relationships that hold for any MV problem. But, if we make the traditional CAPM homogeneity assumptions in an economy where there are only $n$ risky assets, the same efficient set mathematics can be used to trace out the economy’s minimum-variance frontier. However, although the equilibrating process underlying the CAPM guarantees that the weights in the market portfolio are strictly positive, it takes a little ingenuity to construct examples where this holds. Best and Grauer (1985) showed that the “trick” is to work backwards through the efficient set mathematics. Given a covariance matrix and a pre-specified positively weighted market portfolio, we construct a set of $(\Sigma, x_m)$-compatible means (SML means) and use these means in the examples, where the CAPM is true, that are presented in Section 3. More specifically, we construct a set of $(\Sigma, x_m)$-compatible means

\[
\mu = \theta_1 t + \theta_2 \Sigma x_m,
\]

where $\Sigma$ is given, $x_m$ is the pre-specified positively weighted market portfolio, and $\theta_1$ and $\theta_2$ are (positive) scalar constants. Then, when we substitute the $(\Sigma, x_m)$-compatible means from Eq. (3) into Eq. (1) and solve for the representative investor with risk tolerance $t = t_m$, the market portfolio $x_m$ is MV efficient, securities plot on
the SML, where from Eqs. (2) and (3), $\theta_1 = \lambda(t_m)/t_m = r_z$, $\theta_2 = (\mu_m - r_z)/\sigma_m^2 = 1/t_m$, and $r_z$ is the zero-beta rate. 6 We say the CAPM is true if these relations hold for the $n$ underlying assets.

In order to study the effects of grouping, we assume that the $n$ assets are combined into $q$ portfolios, where consistent with the empirical literature each security is assigned to one and only one group or portfolio, and $n$ is greater than $q$. Let $W = [w_1, \ldots, w_q]$ be an $n \times q$ weighting (grouping) matrix, whose columns contain the weights in the $q$ portfolios. The $q$th column $w_q^0 = [w_{1q}, \ldots, w_{jq}, \ldots, w_{nq}]$ contains the weights in portfolio $q$, where $w_{jq}$ is zero if security $j$ is not included in portfolio $q$. Naturally, each of the columns must sum to one. Using $W$, we transform the ungrouped $n$-asset MV problem into a grouped MV problem defined over $q$ portfolios.

The mean vector and covariance matrix for the grouped data are $\mu_q = W^0\mu_n$ and $\Sigma_q = W^0 \Sigma_n W$, where (for emphasis when needed) a subscript indicates the dimension of a vector or matrix. The MV problem for the $q$ portfolios is

$$\text{Max } L = \{\mu_q^0 x_q - (1/2)x_q^0 \Sigma_q x_q\} + \lambda(1 - x_q^0 x_q).$$

Note that the form of the efficient set mathematics for the grouped problem defined over $q$ portfolios is identical in form to the ungrouped problem defined over $n$ assets. Moreover, the solution defined in terms of the $q$ portfolios can be transformed back into the original $n$ assets using the relation

$$x_n = W x_q.$$  

Suppose then that the CAPM is true, and we work with grouped data. That would not seem to cause any problems. It is well known, for example, that if individual assets plot on the SML, then any portfolio, whose weights sum to one, must plot on the SML. So grouping is not going to affect the SML relationship. However, grouping may still cause fundamental problems. When we group, we constrain the way we hold the individual assets in each group. It follows that the grouped (constrained) frontier must lie inside the ungrouped frontier. At best, the grouped and ungrouped frontiers can have a single point in common. Thus, when the CAPM is true: (1) the market portfolio will plot outside the grouped frontier, (2) statement S1 will not hold for the grouped data, (3) Kandel and Stambaugh’s (1995) measure of relative efficiency will be undefined, and (4) Kandel and Stambaugh’s (1987) frontier correlation will exceed one—unless the grouping constraints are inactive at the point where the grouped frontier is tangent to the ungrouped frontier. 7

---

6 Alternatively, one could explicitly add the riskless asset to the objective function (1) and obtain the same result.

7 Presumably neither the ungrouped or grouped frontiers are exactly the same as the frontier employed in multivariate tests of the CAPM. Under one formulation of the multivariate test, the market portfolio $x_m$ is efficient relative to a set of $n$ assets if it is impossible to construct a portfolio combining the market and the $n$ assets such that the resulting potential performance exceeds the market’s performance $(\mu_m - r)/\sigma_m$. However, there is a dimensionality problem with the resulting covariance matrix because the market portfolio is a linear combination of the $n$ assets. In order to overcome the problem, Jobson and Korkie (1982) suggest that the test combine $n - 1$ assets with the market, while Gibbons et al. (1989) and MacKinlay (1987) assume that the market portfolio is not a linear combination of the $n$ assets.
constraints will be inactive with the tangency point at the market portfolio is if we value weight the securities in each portfolio.

Suppose instead that the CAPM is false, i.e., that the underlying assets do not plot on the SML. We develop scenarios where the CAPM is false employing the method used by Grauer (1999). This allows us to perturb the SML means to create a variety of (sometimes bizarre) scenarios where the CAPM is false in the $n$-asset universe. For example, an OLS and/or GLS regression of expected returns on betas can have the same intercept and slope as when the CAPM is true even though the tangency portfolio takes on seemingly unrealistic characteristics. Alternatively, the slope of a regression of means on market betas can be zero even though the market portfolio is almost MV efficient.

Then, we extend the analysis to examine the effects of grouping. We show that value-weighted grouping—the only type of grouping that avoids problems when the CAPM is true—may cause the CAPM to appear to be completely correct when it is clearly false. Conversely, the slope of a cross-sectional regression of expected returns on betas fitted to value-weighted grouped data may be flatter than when the regression is fitted to the ungrouped data.

We begin by reviewing the method Grauer (1999) used to make the CAPM false in an $n$-asset universe. Consider the general linear regression model $y = Xb + e$, where $y$ is an $n$-vector containing $n$ observations on the dependent variable, $X$ is an $n \times k$ matrix containing $n$ observations on the $k$ independent variables, $b$ is a $k$-vector of parameters, and $e$ is an $n$-vector of random errors. In an OLS regression, it is assumed that the error terms are independently and identically distributed (iid) with zero means and covariance matrix $\Sigma = \sigma^2 I$, where $I$ is an $n \times n$ identity matrix. The OLS coefficients are $b_{\text{ols}} = (X'X)^{-1}X'y$. In a GLS regression, the assumption that the error terms are iid is relaxed. In this case, the error terms are assumed to have zero means and covariance matrix $\Sigma$. The GLS coefficients are $b_{\text{gl}} = (X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}y$.

When the CAPM is false in an $n$-asset universe, an OLS and/or a GLS regression of expected returns on betas can yield exactly the same intercepts and slopes as in the case when the model is true. To show this, let $y_n = \mu_n + e_n$, where $e_n$ is a vector containing perturbations from the SML means $\mu_n$. In addition, let $X$ contain a vector of ones and market betas, e.g., $X = [\iota, \beta]$.

The OLS coefficients are $b_{\text{ols}} = (X'X)^{-1}X'y + (X'X)^{-1}X'e_n = b_0 + \Delta b_{\text{ols}}$. If we choose $e_n$ so that $X'e_n = 0$ and/or $X'\Sigma^{-1}e_n = 0$, where $0$ is a 2-vector of zeros, we have exactly the same OLS and/or GLS coefficients as when the CAPM is true. Likewise, by choosing $e_n$ so that $(X'X)^{-1}X'e_n$ and/or $(X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}e_n$ have specific values we can make the OLS and/or GLS coefficients take on almost any desired value.

---

8 We could also consider the perturbations $e_n$ to be returns on a second factor, which together with the betas completely describe the cross-section of expected returns. Under this interpretation, the regression of means on betas is misspecified. However, in a regression of means on betas and second-factor returns, the intercept and beta-slope coefficient will be unchanged from the case where the CAPM is true when $e_n$ is orthogonal to $\iota$ and $\beta$. 

---
When we pool the data into \( q \) portfolios, we apply the matrix \( W \) so that
\[
W' y_n = W' \mu_n + W' e_n,
\]
which may be written as
\[
y_q = \mu_q + e_q,
\]
where for example the vector of deviations in the means after grouping is
\[
e_q = W e_n.
\]
When the CAPM is false, if we value weight the securities in each portfolio and choose the perturbations \( e_n \) so that
\[
e_q = 0_q,
\]
then after pooling, the model will be absolutely correct. Ironically, value-weighted grouping—the only type of grouping that avoids problems when the CAPM is true—is the only form of grouping that can cause this problem when the CAPM is false. Conversely, when the model is false, the slope of a cross-sectional regression of expected returns on betas fitted to value-weighted grouped data may be either steeper or flatter than when the regression is fitted to the ungrouped data.

However, we do not have to restrict the analysis to value-weighted grouping. With the right deviations, satisfying
\[
X' W W' e_n = 0 \quad \text{and/or} \quad X' W (W \Sigma W)^{-1} W' e_n = 0,
\]
the OLS and/or GLS regression coefficients from the grouped data can incorrectly indicate that the model is correct. This can occur even when \( X' e_n \neq 0 \) and/or \( X' \Sigma^{-1} e_n \neq 0 \), so that the regression coefficients on the ungrouped data would correctly indicate that the model is wrong. Unfortunately, there is no direction to the bias. Thus, with ungrouped data, \( X e_n \) may be near zero indicating that an OLS regression of expected returns on betas will have almost the same slope as when the model is true. At the same time with grouped data, \( X W W' e_n \) can be far from zero. Moreover, there is no reason to believe that \( X' W_1 W'_1 e_n = X' W_2 W'_2 e_n \) when the data are grouped in different ways. Thus, it is quite possible that grouping with \( W_1 \) might make the slope of the SML estimated from grouped data appear too steep and grouping with \( W_2 \) might make it look too flat.

The coefficient of determination, \( R^2 \), is not immune to the effect of grouping either. First, for the ungrouped data,
\[
R^2 = (\mu_n' X (X' X)^{-1} X' \mu_n)/(\mu_n' \mu_n).
\]
After we group, even when we have not introduced any perturbation into the means, the \( R^2 \) for the grouped data is given by
\[
R^2_q = (\mu_q' W W' X (X' W W X)^{-1} X' W W \mu_n)/(\mu_q' W \mu_n).
\]
When the CAPM is true, these two values are equal to one. When we introduce a perturbation to the means, i.e., when the CAPM is false, it is easy to demonstrate that \( R^2_q \) will decrease. However, the effect on \( R^2_q \) is not so clear. If \( e_n \) and \( W \) are orthogonal, then \( R^2_q \) does not change, no matter how great the change in \( R^2_n \). In contrast, another \( W \) can magnify the impact of \( e_n \), and result in a larger change in \( R^2_q \) than in \( R^2_n \).

2. The data

Two data sets are employed in the examples. The first set contains 20 individual assets. The second set contains 100 size portfolios of all New York Stock Exchange (NYSE) stocks. The data sets are selected to provide realistic mean vectors and covariance matrices proceeding as follows. For the 20-asset universe, 120 months of data are selected at random from the Center for Research in Security Prices (CRSP) data. For the 100 size portfolio universe, returns from the 1926 to 1995 period are employed. We use the historic return data to construct covariance matrices
for the two universes. Given a covariance matrix and a pre-specified positively weighted market portfolio we construct a set of \((\Sigma, x_m)\)-compatible means (or equivalently SML means) from Eq. (3) and use these means in the examples.

The CAPM makes no predictions about the magnitude of the weights in the market portfolio. But, in order to be consistent with the “size” effect, i.e., small firms have higher returns than large firms, we assign high (low) market portfolio weights to low (high) expected return assets. In the 20 asset universe, we employ a sum-of-digits weighting scheme of the form \(n/N, (n-1)/N, \ldots, 1/N\), where the sum of the \(n\) integers from 1 to \(n\) \((n = 20)\) is \(N = n(n+1)/2\), to generate the weights in the market portfolio. In general, we assign high (low) market portfolio weights to the low (high) mean assets when the CAPM is true. Although the relationship between the portfolio weights and means is not exactly one-to-one, the weight of the low-mean asset is 20 times the weight of the high-mean asset. In the 100 size portfolio universe, we take the average capitalization weights of the 100 size portfolios over the 1926–1995 period to be the weights in the market portfolio.9

While there is an infinite number of mean vectors from the two parameter family of \((\Sigma, x_m)\)-compatible means, \(\theta_1\) is chosen so that the zero-beta rate is equal to the riskfree rate of 0.5% per month (6% per annum), and \(\theta_2\) is chosen so that the excess return on the MV-efficient market portfolio is 0.75% per month. This latter value corresponds to the excess return of the Standard & Poor’s 500 Index over Treasury bills during the last half century. Thus, in mean–beta space, the population means in percent per month are \(\mu = 0.5\bar{i} + 0.75\beta_m\), when the hypothesis “the market portfolio is MV efficient” is true.

3. Examples

3.1. The traditional grouping method

The results for the traditional grouping method are shown in Table 1 and Figs. 1–5. The 20 primitive securities are grouped into four size portfolios. In each portfolio, securities are value weighted, except in example 2. The CAPM is true in the 20-asset universe shown in Table 1, examples 1 and 2, and in Figs. 1 and 2. The model is false in Table 1, examples 3–5, and in Figs. 3–5. Recall two points. First, when the MV CAPM is true, each of the following four statements holds: (1) the market portfolio is MV efficient, (2) there is at least one positively-weighted efficient portfolio, (3) the

---

9 Given a covariance matrix, it is easy to make any portfolio MV efficient. However, in general, the \((\Sigma, x_m)\)-compatible means will not be consistent with the scenario of the low-mean asset having the largest weight in the market portfolio. In the individual-asset universe, we simply had to guess which assets to weight in which ways to produce the low-mean high-weight scenario. On the other hand, in the size-portfolio universe, the average capitalization weights and the \((\Sigma, x_m)\)-compatible means are consistent with the low-mean high-weight scenario.
The capital asset pricing model is true in a 20 asset universe shown in examples 1 and 2. It is false in examples 3–5. The riskless interest rate is 0.5% and the expected return on the market portfolio is 1.25%. Hence, $b_0 = 0$ and $b_1 = \mu_m - r = 0.75%$. (Regressions employing the 20 individual assets are labeled ungrouped data.) The 20 securities are grouped into 4 size portfolios. Under the traditional grouping method securities in each portfolio are value weighted, except in example 2 where the weights are (0.01, 0.01, 0.96, 0.01, 0.01). (Regressions employing the portfolio data are labeled grouped data.) Under the Fama–French grouping method securities in each portfolio are equally weighted and portfolio betas are assigned to individual securities. (Regressions employing these data are labeled Fama–French data.) The ungrouped and grouped data in examples 1–5 are plotted in Figs. 1–5. Kandel and Stambaugh (1995) define $w_m = (\mu_m - \mu_g)/(\mu_v - \mu_g)$ to be a measure of the relative efficiency of the market portfolio, where $\mu_v$ and $\mu_g$ are the expected return of the minimum-variance portfolio with the same variance as the market portfolio and the global minimum-variance portfolio. The $R^2$ of the generalized least squares regression is equal to the square of the relative efficiency measure. Kandel and Stambaugh (1987) define the frontier correlation as $\sigma_v/\sigma_g$, where $\sigma_v$ is the standard deviation of the minimum-variance portfolio with the same expected value of $\mu$ as the market portfolio.

<table>
<thead>
<tr>
<th>Example Data</th>
<th>Ordinary least squares</th>
<th>Generalized least squares</th>
<th>Efficiency measures</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Intercept</td>
<td>Slope</td>
<td>$R^2$</td>
</tr>
<tr>
<td>1. Ungrouped</td>
<td>0.00</td>
<td>0.75</td>
<td>1.00</td>
</tr>
<tr>
<td>Grouped</td>
<td>0.00</td>
<td>0.75</td>
<td>1.00</td>
</tr>
<tr>
<td>Fama–French</td>
<td>0.00</td>
<td>0.75</td>
<td>0.84</td>
</tr>
<tr>
<td>2. Grouped</td>
<td>0.00</td>
<td>0.75</td>
<td>1.00</td>
</tr>
<tr>
<td>3. Ungrouped</td>
<td>-0.30</td>
<td>1.00</td>
<td>0.09</td>
</tr>
<tr>
<td>Grouped</td>
<td>0.00</td>
<td>0.75</td>
<td>1.00</td>
</tr>
<tr>
<td>Fama–French</td>
<td>-0.07</td>
<td>0.81</td>
<td>0.05</td>
</tr>
<tr>
<td>4. Ungrouped</td>
<td>0.90</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Grouped</td>
<td>0.00</td>
<td>0.75</td>
<td>1.00</td>
</tr>
<tr>
<td>Fama–French</td>
<td>-0.01</td>
<td>0.76</td>
<td>0.05</td>
</tr>
<tr>
<td>5. Ungrouped</td>
<td>0.00</td>
<td>0.75</td>
<td>0.46</td>
</tr>
<tr>
<td>Grouped</td>
<td>0.12</td>
<td>0.64</td>
<td>0.45</td>
</tr>
<tr>
<td>Fama–French</td>
<td>0.06</td>
<td>0.70</td>
<td>0.33</td>
</tr>
</tbody>
</table>

Table 1
Results for the cross-sectional regressions $\mu_j - r = b_0 + b_1 \beta_j + \epsilon_j$

The market portfolio is the tangency portfolio, and (4) securities plot on the SML. Moreover, the examples are constructed so that the riskless interest rate is 0.5% and the expected return on the market portfolio is 0.5% + 0.75% = 1.25%. Hence, in the regression $\mu_j - r = b_0 + b_1 \beta_j + \epsilon_j$, the intercept $b_0 = 0$ and the slope $b_1 = \mu_m - r = 0.75%$. Second, when the CAPM is false in the examples in this section, each of the four statements is false. And, the legend in each figure identifies which of the statements is most obviously false.

In Table 1, example 1, and in Fig. 1, the fundamental theoretical relationship S1 holds for both ungrouped and grouped data. The ungrouped frontier, generated from the 20 securities, and the grouped frontier, generated from the four portfolios, are tangent at the market portfolio. But, with grouped data, the relationship only holds if securities in each group or portfolio are value-weighted. In Table 1, example 2, and in Fig. 2, the fundamental CAPM relationship S1 does not hold for grouped
Fig. 1. The CAPM is true in a 20 asset universe. The 20 securities (dark circles) are grouped into four size portfolios (open squares). In each portfolio, securities are value weighted using market portfolio weights. In mean–standard deviation space, the market portfolio is mean-variance efficient. The ungrouped frontier (derived from the securities) and grouped frontier (derived from the portfolios) are tangent at the market portfolio. In mean–beta space, securities and portfolios plot on the security market line. The fundamental CAPM result—securities plot on the security market line if and only if the market portfolio is mean-variance efficient—holds for both grouped and ungrouped data in only one case: when securities in each group or portfolio are value weighted. See Table 1 example 1.

Fig. 2. The CAPM is true in a 20 asset universe. The 20 securities are grouped into four size portfolios. In each portfolio, securities are weighted (0.01, 0.01, 0.96, 0.01, 0.01). In this case, the grouped data plot on the security market line. But the market portfolio is super-efficient relative to the grouped frontier. Furthermore, grouping causes problems with Kandel and Stambaugh’s (1987, 1995) efficiency measures. With grouped data, the measure of relative efficiency is undefined and the frontier correlation exceeds one. See Table 1 example 2.
data when securities are weighted (0.01, 0.01, 0.96, 0.01, 0.01) in each size portfolio. Although the four portfolios plot on the SML, the market portfolio plots outside the grouped frontier. This, in turn, causes problems with Kandel and Stambaugh’s (1987, 1995) efficiency measures. In this example, the grouped frontier lies entirely inside the ungrouped frontier. In fact, the standard deviation of the grouped frontier’s global minimum-variance portfolio is larger than the market portfolio’s standard deviation. Thus, with grouped data, Kandel and Stambaugh’s (1995) measure of relative efficiency is undefined and Kandel and Stambaugh’s (1987) frontier correlation exceeds one.

The CAPM is false in Table 1, examples 3–5. In Table 1, example 3, and in Fig. 3, the CAPM is clearly incorrect in mean–standard deviation space. The market portfolio’s mean is 1.25%. On the other hand, the ungrouped frontier’s tangency portfolio has a mean of minus 9026%. Furthermore, the ungrouped data do not plot on the security market line. With grouped data, however, the model is absolutely correct. See Table 1 example 3.

![Graph showing the CAPM in a 20 asset universe.](image)

Fig. 3. The CAPM is false in a 20 asset universe. The 20 securities are grouped into four size portfolios. In each portfolio, securities are value weighted. The CAPM is clearly false in mean–standard deviation space. The market portfolio’s mean is 1.25%. On the other hand, the ungrouped frontier’s tangency portfolio has a mean of minus 9026%. Furthermore, the ungrouped data do not plot on the security market line. With grouped data, however, the model is absolutely correct. See Table 1 example 3.

10 This non-standard weighting scheme is simply used to illustrate the result that the market portfolio is super-efficient relative to the grouped frontier when the CAPM is true unless the securities in each group are value weighted. Appendix A contains an example where the market portfolio is super-efficient relative to a grouped frontier when the securities in each group are equal weighted.
Fig. 4. The CAPM is false in a 20 asset universe. The 20 securities are grouped into four size portfolios. In each portfolio, securities are value weighted. In this case, the CAPM is clearly false in mean–beta space. The ordinary least squares regression of expected excess returns on betas has a zero slope. With generalized least squares, the slope is negative. With grouped data, however, the model is absolutely correct. (Although the ungrouped frontier looks somewhat similar to the ungrouped frontier in Fig. 3, the tangency portfolio’s mean is only 5.5% in this example.) See Table 1 example 4.

Fig. 5. The CAPM is false in a 20 asset universe. The 20 securities are grouped into four size portfolios. In each portfolio, securities are value weighted. The market portfolio plots inside both frontiers. Securities and portfolios do not plot on the security market line. With ungrouped data, the ordinary least squares and generalized least squares regressions of expected excess returns on betas have intercepts of 0.0 and slopes of 0.75. After grouping, the ordinary least squares slope (0.64) and the generalized least squares slope (0.33) are too flat. Ironically, grouping exacerbates the very problem it was meant to alleviate. See Table 1 example 5.
returns and betas according to an OLS regression. In addition, the slope and intercept do not sum to the expected excess return on the market. With GLS, the relation is negative. Furthermore, Kandel and Stambaugh’s (1995) relative efficiency measure is negative. But, with value-weighted grouped data, the model appears to be absolutely correct.

In Table 1, example 5, and in Fig. 5, the slope of a cross-sectional regression of expected returns on betas fitted to value-weighted grouped data is flatter than when the regression is fitted to the ungrouped data. With ungrouped data, the OLS and GLS intercepts and slopes of zero and 0.75 are identical to what they are when the CAPM is true. With grouped data, however, the OLS and GLS intercepts are too large and the OLS slope (0.64) and the GLS slope (0.33) are too flat. Ironically, value-weighted grouping exacerbates the very problem it was meant to alleviate. Moreover, the grouped $R$-squares are smaller than the ungrouped $R$-squares.

### 3.2. The Fama–French grouping method

We illustrate the problems with Fama–French grouping employing examples. We focus on four examples where the CAPM is true in a universe consisting of 100 size portfolios of all NYSE stocks. The 100 size portfolios, which are assumed to be the primitive assets in the examples, are assigned into portfolios sorted either by size, or first by size and then by beta, or first by beta and then by size. Portfolios are value-weighted or equal weighted, each asset in a portfolio is assigned its portfolio beta, and all 100 assets are employed in the regressions. In addition, we examine four examples from the 20-asset universe—one where the CAPM is true and three where it is false. In both data sets the riskless interest rate is 0.5% and the expected return on the market portfolio is 1.25%.

The first problem with Fama–French grouping is that the fundamental capital asset pricing relationship $S1$ does not hold when the CAPM is true. This is shown in Fig. 6. The market portfolio is MV efficient, but securities do not plot on the SML because they have been assigned portfolio betas. This contrasts with traditional grouping where grouped data plot on the SML when the CAPM is true, but the market portfolio plots outside the grouped efficient frontier—unless securities in each group are value weighted.

Fama–French grouping creates five additional problems for GLS regressions of means on betas. First, in a GLS regression of expected excess returns on betas, the fundamental methodological premise is that the slope and intercept will sum to the expected excess return on the market regression independent of whether or not the CAPM is true, see Roll (1985). Unfortunately, Fama–French grouping

---

11 The relationship holds for an OLS regression when the CAPM is true and for an equal-weighted market portfolio when the CAPM is false. In other cases it does not hold. For example, Roll and Ross (1994) derive a minimum-variance frontier of portfolios that produce betas that have no OLS relation to expected returns. Each of the minimum-variance portfolios has a different expected rate of return. Yet, the sum of the slope and the intercept of an OLS regression of means on betas (calculated from any of the portfolios) is the same: the arithmetic average of the individual asset means.
also compromises statement S2—when the model is true and when market portfolio betas are employed in the regression. 12 This is shown in the four examples in Table 2 and in Table 1, example 1. 13 Second, the GLS slope is too flat—in some cases much too flat. When the CAPM is true in the 20-asset universe, the slope is 0.61. In the 100 asset universe, the slope drops to as low as 0.12 when the groups are equal weighted. Third, the $R^2$ from the GLS regression is not equal to the square of Kandel and Stambaugh’s (1995) relative efficiency measure.

Fourth, the results depend on the number of groups and how the securities are sorted into the groups. The four examples in Table 2 show that the smaller the number of portfolios, the flatter the GLS slope, and the further the sum of the slope plus the intercept is from the correct value of 0.75. For example, with securities assigned betas from 5, 10, and 25 equally weighted size portfolios, the GLS slopes are only 0.12, 0.27, and 0.34 respectively. The corresponding sums of the slope plus the intercept are 0.44, 0.48, and 0.52—far from the correct value of 0.75. Sorting first by size

---

12 Note that Fama and French run OLS regressions only and that the betas from equal-weighted portfolios are assigned to individual securities. While their method yields the correct slope and intercept in this case, it is perhaps serendipitous for two reasons. First, as noted, Roll and Ross (1994) and Kandel and Stambaugh (1995) demonstrate the shortcomings of OLS regressions in the absence of grouping. Second, when the CAPM is true, the only way traditionally grouped data conform to the most basic theoretical relationships is when securities in each portfolio are value weighted.

13 Table 1 examples 3–5 show that the same result holds when the model is false. In each of these examples, the excess return on the market is 0.75, but the sum of the GLS slope plus the GLS intercept is 0.41, 1.72 and 0.44, respectively.
Table 2
Results for the cross-sectional regressions \( \mu_j - r = b_0 + b_1 \beta_j + \epsilon_j \) when individual assets are assigned portfolio betas

<table>
<thead>
<tr>
<th>Example</th>
<th>Ordinary least squares</th>
<th>Generalized least squares</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Intercept</td>
<td>Slope</td>
</tr>
<tr>
<td>1. Value weighted</td>
<td>0.08</td>
<td>0.70</td>
</tr>
<tr>
<td>Equally weighted</td>
<td>0.00</td>
<td>0.75</td>
</tr>
<tr>
<td>2. Value weighted</td>
<td>0.03</td>
<td>0.73</td>
</tr>
<tr>
<td>Equally weighted</td>
<td>0.00</td>
<td>0.75</td>
</tr>
<tr>
<td>3. Value weighted</td>
<td>0.00</td>
<td>0.75</td>
</tr>
<tr>
<td>Equally weighted</td>
<td>0.00</td>
<td>0.75</td>
</tr>
<tr>
<td>4. Value weighted</td>
<td>0.00</td>
<td>0.75</td>
</tr>
<tr>
<td>Equally weighted</td>
<td>0.00</td>
<td>0.75</td>
</tr>
</tbody>
</table>

The capital asset pricing model is true in a universe consisting of 100 size portfolios of all NYSE stocks. The riskless interest rate is 0.5% and the expected return on the market portfolio is 1.25%. Hence, \( b_0 = 0 \) and \( b_1 = \frac{\mu_m - r}{\sigma^2_m} = 0.75 \). The 100 assets are assigned to portfolios sorted by size or first by size and then by beta. Portfolios are either value weighted or equally weighted as in Fama and French (1992), each asset in a portfolio is assigned its portfolio beta, and individual asset data are employed in the regressions. In example 1, the 100 assets are sorted into 5 size portfolios with 20 assets in each portfolio. See Fig. 6. In example 2, the 100 assets are sorted into 10 size portfolios. There are 10 assets in each portfolio. In example 3, the 100 assets are sorted into 25 size portfolios with 4 assets in each portfolio. In example 4, the 100 assets are sorted into 5 size portfolios. Each size portfolio is then sorted into 5 beta portfolios. There are 4 assets in each portfolio. Some of the figures in the table have been rounded.

and then by beta mitigates, but does not remove, these effects. In Table 2, example 3, where there are 25 size portfolios, the GLS slope is 0.34. On the other hand, in Table 2, example 4, where the 100 assets are sorted into 5 size portfolios and then each size portfolio is sorted into 5 beta portfolios again producing 25 portfolios, the GLS slope is 0.61.

Fifth, the results differ according to the order of the sorts. Early tests of the CAPM form groups based on a beta sort. More recent tests form groups based on size or book-to-market equity sorts. In addition, the recent tests tend to sort on two variables, say size and beta. In order to check whether the sort, or the order of the sort, affects the results reported in Table 2, where the 100 assets are sorted either by size or first by size and then by beta, the 100 assets are assigned into portfolios again—this time sorted either by beta or first by beta and then by size. The results show that the GLS slopes and \( R \)-squares are larger when the assets are sorted by beta rather than by size. For instance, Table 2, example 2 shows that with 10 equal-weighted size portfolios, the GLS slope is 0.27 and the \( R \)-square is 0.18. With 10 equal-weighted beta portfolios (not shown in the table), the GLS slope is 0.42 and the \( R \)-square is 0.51. On the other hand, the GLS slopes and \( R \)-squares are smaller when the data are sorted first by beta and then by size rather than first by size and then by beta. For instance, in Table 2, example 4 when the 100 assets are sorted into five equal-weighted size portfolios and then each size portfolio is sorted into five beta portfolios, the GLS slope is 0.61 and the \( R \)-square is 0.74. When the 100 assets
are sorted first by beta and then by size (not shown in the table), the GLS slope is 0.51 and the R-square is 0.60.

4. Summary and concluding comments

Empirical analysis is an art that balances costs against benefits. The econometric benefits of grouping in reducing measurement error in cross-sectional regressions of means on betas are beyond question e.g. grouping alleviates measurement error in the betas that cause the estimated slope to be flatter than the true slope. Moreover, grouping provides significant benefits in reducing the dimension of the covariance matrix needed to trace out the minimum-variance frontier in tests that focus on the MV efficiency of the market portfolio. However, nothing is without cost. In this paper, we examine the unintended consequences or costs of grouping in a world where there is no measurement error. Two scenarios and two grouping methods are examined. In the first scenario, the CAPM holds exactly in an \( n \)-asset economy. In the second, it is false. With traditional grouping, the underlying securities are combined into portfolios and grouped data are employed in the tests. With Fama–French grouping securities are combined into portfolios, each security in a portfolio is assigned its portfolio beta, and individual security data are employed in the tests.

We identify two unintended consequences associated with traditional grouping when the CAPM is true. First, the fundamental theoretical relationship—securities plot on the SML if and only if the market portfolio is MV efficient—will not hold for grouped data, unless securities in each group are value weighted. Grouped data plot on the SML when the CAPM is true. The reason is straightforward. If individual securities plot on the SML, then any portfolio, whose weights sum to one, must also plot on the SML. But, the market portfolio will plot outside the grouped minimum-variance frontier, unless securities in each group are value weighted. Second, the relations between Kandel and Stambaugh’s (1987, 1995) correlation coefficient, the GLS R-square, and measures of relative efficiency do not hold, unless securities in each group are value weighted. (See Figs. 1 and 2.)

Likewise, there are two unintended consequences of traditional grouping when the CAPM is false. First, grouping may cause the model to appear to be absolutely correct when it is patently false. We present examples where, with ungrouped data, the tangency portfolio is driven to near minus infinity in expected return—standard deviation space in one case and where there is no relation between expected return and beta in another. Yet, with grouped data, the model appears to be absolutely correct. (See Figs. 3 and 4.) Ironically, the only way this can occur is when securities in each group are value-weighted, the only form of grouping that avoids problems when the CAPM is true. To make matters worse, the slope of a cross-sectional regression of expected returns on betas fitted to value weighted grouped data may be either steeper or flatter than when the regression is fitted to ungrouped data. In other words, grouping may exacerbate the very problem it was meant to alleviate. (See Fig. 5.)
When the CAPM is true, Fama–French grouping creates problems with the fundamental theoretical and cross-sectional regression relationships captured in statements S1 and S2. The market portfolio is MV efficient, but securities do not plot on the SML. In a GLS regression of expected excess returns on betas, the slope and intercept do not sum to the expected excess return on the market portfolio, and the slope can be much too flat (see Fig. 6). In addition, the sum of the (OLS and GLS) slope and intercept differs depending on the number of groups, whether or not the groups are sorted, and on the order of the sort.

Acknowledgements

Financial support from the Social Sciences and Humanities Research Council of Canada is greatly appreciated, as is the assistance of William Ting and Christopher Fong. The paper was presented at the Northern Finance Association meetings in Winnipeg, Simon Fraser University, and the University of California at Berkeley. We thank Reo Audette, Nils Hakansson, John Heaney, John Herzog, Raymond Kan, Peter Klein, Geoff Poitras, and Simon Wheatley for valuable comments. In addition, we thank the referees whose insightful comments greatly improved the paper. Of course, we are responsible for any remaining errors.

Appendix A

In the body of the paper, we use the weighting matrix $W$ to transform the ungrouped $n$-asset MV problem into a grouped MV problem defined over $q$ portfolios. Alternatively, we formulate the MV problem as an $n$-asset problem subject to a set of linear constraints. This formulation explicitly captures the restrictions on individual assets that are imposed by grouping. More importantly, it provides the intuitive reason why, when the CAPM is true, the grouped frontier that is subject to a budget constraint and the grouping restrictions must lie inside the ungrouped frontier that is only subject to a budget constraint—unless the grouping restrictions are inactive at the single point where the two frontiers are tangent. Finally, we establish the equivalence of the two ways of analyzing grouping illustrating the ideas with a numerical example.

The restrictions on the underlying $n$ assets implicit in $W$ can be captured explicitly in terms of a set of linear constraints of the form $Rx_n = 0$, where $R$ is an $(n - q) \times n$ restriction matrix and $0$ an $n - q$ vector. Recall from equation (5) that $x_n = Wx_q$, which reproduces the weights on each of the primitive assets, given only the weights on the $q$ portfolios. Since we are reducing $n$ primitives to $q$ portfolios, there must be $n - q$ restrictions, the rank of $R$. Further, since $Rx_n = 0$, we must also have that $Rx_q = RWx_q = 0$. For nonzero weights this requires that $RW = 0$, where in this case $0$ is an $(n - q) \times q$ matrix. In fact, any matrix with rank $n - q$, and dimensions $(n - q) \times n$, satisfying $RW = 0$ is a restrictions matrix that can generate the portfolios in $W$. 
Table 3
A numerical example illustrating the effects of grouping

Panel A: Inputs to the problem

<table>
<thead>
<tr>
<th>Mean vector</th>
<th>Covariance matrix</th>
<th>Market weights</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \mu = \begin{bmatrix} 1.19 \ 1.30 \ 1.31 \ 1.35 \ 1.40 \ 1.42 \end{bmatrix} )</td>
<td>( \Sigma = \begin{bmatrix} 0.2 &amp; 0 &amp; 0 &amp; 0 &amp; 0 &amp; 0 \ 0 &amp; 0.8 &amp; 0 &amp; 0 &amp; 0 &amp; 0 \ 0 &amp; 0 &amp; 1.4 &amp; 0 &amp; 0 &amp; 0 \ 0 &amp; 0 &amp; 0 &amp; 2.5 &amp; 0 &amp; 0 \ 0 &amp; 0 &amp; 0 &amp; 0 &amp; 10.0 &amp; 0 \ 0 &amp; 0 &amp; 0 &amp; 0 &amp; 0 &amp; 16.0 \end{bmatrix} )</td>
<td>( x_m = \begin{bmatrix} 0.45 \ 0.25 \ 0.15 \ 0.10 \ 0.03 \ 0.02 \end{bmatrix} )</td>
</tr>
</tbody>
</table>

Panel B: R-matrices that keep assets 1 and 6, 2 and 5, 3 and 4 in fixed proportions

<table>
<thead>
<tr>
<th>Value-weighted</th>
<th>Equal-weighted</th>
</tr>
</thead>
<tbody>
<tr>
<td>( R = \begin{bmatrix} -1.0 &amp; 0 &amp; 0 &amp; 0 &amp; 0 &amp; 0.45/0.02 \ 0 &amp; -1.0 &amp; 0 &amp; 0 &amp; 0.25/0.03 &amp; 0 \ 0 &amp; 0 &amp; -1.0 &amp; 0.15/0.10 &amp; 0 &amp; 0 \end{bmatrix} )</td>
<td>( R = \begin{bmatrix} -1.0 &amp; 0 &amp; 0 &amp; 0 &amp; 0 &amp; 1.0 \ 0 &amp; -1.0 &amp; 0 &amp; 0 &amp; 1.0 &amp; 0 \ 0 &amp; 0 &amp; -1.0 &amp; 1.0 &amp; 0 &amp; 0 \end{bmatrix} )</td>
</tr>
</tbody>
</table>
The data for the mean-variance problem given by Eq. (A.1) is recorded in Panel A. A value-weighted and an equal-weighted restriction matrix are recorded in Panel B. The left-hand columns of Panel C contain the output from the efficient-set mathematics when the budget constraint is the only constraint. The middle (right-hand) columns contain the output when the value-weighted (equal-weighted) restriction matrix is added to the budget constraint.
Clearly \( R \) is not unique. But assuming the portfolios in \( W \) are distinct, e.g. any single asset is represented in only one portfolio, it is easy to construct a restriction matrix. Recall that \( W = [w_1, \ldots, w_q] \), where the elements of portfolio (column) \( q \) are \( w'_q = [w_{1q}, \ldots, w_{jq}, \ldots, w_{mq}] \), and \( w_{jq} \) is zero if the security \( j \) is not included in portfolio \( q \). We can arbitrarily choose any nonzero element of \( w_q \) to be the reference element, so we choose the first one. The relationship between this asset and any other asset in portfolio \( q \) that must be maintained is \( x_j/w_{jq} = x_k/w_{kq} \), where \( x_j \) and \( x_k \) are the \( j \)th and \( k \)th elements of \( x_n \). This can be represented as

\[
[0, \ldots, 0, -1, 0, \ldots, w_{jq}/w_{kq}, 0, \ldots, 0]x_n = 0.
\]

Constructing a row like this for each nonzero element in \( w_q \), and repeating the process for each column in \( W \) gives us the rows of a restriction matrix that will reproduce \( W \). Assembling them together into one matrix gives us an \( R \) that satisfies \( RW = 0 \).

We capture the grouping problem, as well as the single-constraint problem given in Eq. (1), as special cases of the MV problem subject to general linear constraints

\[
\text{Max} L = \{\eta'x - (1/2)x'\Sigma x\} + \lambda'(B - Ax),
\]

where \( A \) is a \( k \times n \) constraint matrix, \( B \) is a \( k \)-vector containing the right-hand side of the constraints and \( \lambda \) is a \( k \)-vector containing the Lagrange multipliers. See Markowitz (1959), Sharpe (1970), and Best and Grauer (1990). When \( \tau'x = 1 \) is the only constraint, \( A = \tau' \), \( B = 1 \), and \( \lambda \) is a scalar. The constraints \( \tau'x = 1 \) and \( Rx = 0 \) are easily accommodated with \( A' = [1' R'] \) and \( B' = [1' 0'] \). The optimality conditions are

\[
\eta - \Sigma x - A'\lambda = 0, \quad \text{and} \quad Ax = B. \tag{A.2}
\]

Solving Eq. (A.2) gives the solution and multiplier vectors

\[
x(t) = h_0 + th_1, \quad \text{and} \quad \lambda(t) = g_0 + tg_1 \tag{A.3}
\]

with \( h_0 = \Sigma^{-1}A'c^{-1}B \), \( h_1 = [\Sigma^{-1}\mu - \Sigma^{-1}A'c^{-1}a^*] \), \( g_0 = -c^{-1}B \), and \( g_1 = c^{-1}a^* \), where \( a^* = A \Sigma^{-1} \mu \) is a \( k \)-vector and \( c^* = A \Sigma^{-1} A' \) is a \( k \times k \) matrix. (In the special case where \( \tau'x = 1 \) is the only constraint, \( a^* \) and \( c^* \) reduce to the well-known efficient set constants \( a = \tau'\Sigma^{-1}\mu \) and \( c = \tau'\Sigma^{-1}1 \).) The solution is such that \( \tau'h_0 = 1 \) and \( \tau'h_1 = 0 \). The parametric equation for the minimum-variance frontier is

\[
\mu_p = \alpha_0 + \alpha_1 t, \quad \text{and} \quad \sigma_p^2 = \gamma_0 + \gamma_1 t + \gamma_2 t^2, \tag{A.4}
\]

where \( \alpha_0 = \mu'h_0 = c^{-1}a^* \) is the mean of the global minimum-variance portfolio on the frontier, \( \alpha_1 = \mu'h_1 = b - a^*c^{-1}a^* \) is the mean of the \( h_1 \)-vector and \( b = \mu'\Sigma^{-1}\mu \), \( \gamma_0 = h_0'\Sigma h_0 \) is the variance of the global minimum-variance portfolio, \( \gamma_1 = 2h_0'\Sigma h_1 = 0 \), and \( \gamma_2 = h_1'\Sigma h_1 = \alpha_1 \). From Eq. (A.4) and the chain rule of calculus, the slope of the minimum-variance frontier in mean–standard deviation space is \( d\mu_p/d\sigma_p = \sigma_p/t \).

At this point, we employ a numerical example: (1) to illustrate the efficient set mathematics for the constrained \( n \)-asset problem; (2) to show that when the CAPM is true the market portfolio will plot outside the grouped frontier unless the groups are value weighted using the weights in the market portfolio; and (3) to show that the
analysis with portfolios based on the grouping matrix $W$ is equivalent to the constrained analysis with individual assets using the restriction matrix $R$.

Consider an economy where large (small) firms have low (high) means. The data and output from the efficient-set mathematics are given in Table 3 and the corresponding frontiers are graphed in Fig. 7. Table 3 Panel A gives the inputs to the MV problem: the mean vector, the covariance matrix and the weights in the market portfolio. We have chosen the data so that many of the calculations may be verified by inspection, and all of them may be verified with any software package that permits matrix manipulation. For example, it is easily verified that $\mu$ is $(\Sigma, x_m)$-compatible with $\theta_1 = \lambda(t_m)/t_m = r_z = 1.1$, $\theta_2 = (\mu_m - r_z)/\sigma_m^2 = 1/t_m = 1$, and $t_m = 1$.

While many empirical studies group by size in order to maximize the spread in the means and betas of the portfolios, we present a nonstandard grouping procedure to illustrate that it does not matter how we group as long as we value weight the securities using the weights in the market portfolio. Thus, instead of forming three portfolios consisting of assets 1 and 2, 3 and 4, and 5 and 6, respectively; we group asset 1 with 6, 2 with 5, and 3 with 4. Panel B shows two $R$-matrices that keep assets 1 and 6,
2 and 5, and 3 and 4 in fixed proportions. It is easily verified that $\mathbf{R}x_m = \mathbf{0}$, when we value weight the portfolios, and that $\mathbf{R}x_m \neq \mathbf{0}$, when we equal weight them. Panel C contains the output from the efficient set mathematics for three cases, when the MV problem is given by Eq. (A.1), the solutions and multipliers by Eq. (A.3), and the portfolio means and variances by Eq. (A.4). The left-hand columns contain the results when the budget constraint is the only constraint, i.e., when $A = \mathbf{1}'$, $B = 1$, and $\lambda$ is a scalar. The remaining columns contain the results when the $\mathbf{R}$-matrices are added to the constraints, i.e., when $A' = [\mathbf{1} \mathbf{R}]$ and $B' = [\mathbf{1} \mathbf{0}]$. The middle columns employ the value-weighted $\mathbf{R}$-matrix from Panel B and the right-hand columns employ the corresponding equal-weighted $\mathbf{R}$-matrix.

Turning to the solutions in Panel C, note that each of the three $\mathbf{h}_0$-vectors contains the weights in the global minimum-variance portfolio on each of the frontiers given by the three sets of constraints. The three $\mathbf{h}_1$-vectors show how the solutions change as a function of the parameter $t$. When $t = t_m = 1$, the market portfolio is the optimal solution to the single-constraint case and to the value-weighted restrictions case as is shown in the left-hand and middle columns. But the underlying $n$-securities obey the value-weighted restrictions at one point on the ungrouped frontier, where the budget constraint is the only constraint, and at every point on the value-weighted grouped frontier, when the value-weighted restrictions are imposed! Furthermore, the equal-weighted restrictions prevent the market portfolio from ever being on the equal-weighted grouped frontier. That is, the $\mathbf{h}_0$- and $\mathbf{h}_1$-vectors in the right-hand columns do not span the market portfolio $x_m$.

Turning next to the Lagrange multipliers reported in Panel C, we see that there is one multiplier associated with the budget constraint in the left-hand columns, and four multipliers associated with the budget constraint and three value-weighted and equal-weighted grouping constraints in the middle and right-hand columns. When $t = t_m = 1$ and the multiplier associated with the budget constraint is divided by $t$, $\lambda(t_m)/t_m$ equals 1.1, which may be interpreted as either the riskless or zero-beta return. Furthermore, when $t = t_m = 1$, the three multipliers associated with the value-weighted restrictions matrix are equal to zero. This is another way of seeing that the value-weighted constraints relax at one point, where the value-weighted grouped frontier is tangent to the ungrouped frontier (see Fig. 7). It is apparent from the multipliers associated with the equal-weighted grouping constraints that there is no value of $t$ at which the equal-weighted constraints will relax. Therefore, the market will be super-efficient relative to the equal-weighted grouped frontier. (See Fig. 7.)

Suppose instead that we work with grouped data, when the securities are value-weighted according to the formula $w_j = x_{jm}/\sum x_{km}$, where as before $x_{jm}$ is the weight of security $j$ in the market portfolio and the sum in the denominator is over all $k$ securities in the portfolio. Then, the weighting matrix is given by

$$
\mathbf{W}' = \begin{bmatrix}
0.45/0.47 & 0 & 0 & 0 & 0 & 0.02/0.47 \\
0 & 0.25/0.28 & 0 & 0 & 0.03/0.28 & 0 \\
0 & 0 & 0.15/0.25 & 0.10/0.25 & 0 & 0
\end{bmatrix}.
$$

It is easily verified that the mean vector is $\mu_q' = (\mathbf{W}'\mu_n)' = [1.1998, 1.3107, 1.3260]$ and the diagonal elements of the covariance matrix $\Sigma_q = \mathbf{W}'\Sigma_n\mathbf{W}$ are equal to $(0.2123,$
0.7526, 0.9040), where the elements of $\mathbf{p}_q$ and $\mathbf{S}_q$ have been rounded to four decimal places. From Eq. (A.3) evaluated at $t_m = 1$, $\mathbf{x}_q = (0.47, 0.28, 0.25)$. Finally, given that $\mathbf{x}_n = \mathbf{Wx}_q$, it may be verified by inspection that $\mathbf{x}_m$ is MV efficient, and that $r_z = \lambda(t_m)/t_m = 1.1$.

References