

Industrial Effects and the CAPM: From the Views of Robustness and Longitudinal Data Analysis

Tsung-Chi Cheng¹, Hung-Neng Lai², and Chien-Ju Lu¹

¹*National Chengchi University* and ²*National Central University*

Abstract: The traditional approach by Fama and Macbeth (1973) to the validity of an asset pricing model suffers from two drawbacks. Firstly, it uses the ordinary least squares (OLS) method, which is sensitive to outliers, to estimate the time-series beta. Secondly, it takes averages of the slope coefficients from cross-sectional regressions which ignore the importance of time-series properties. In this article, robust estimators and a longitudinal approach are applied to avoid the problems of these two kinds. We use data on the electronics industry in Taiwan's stock market during the period from September 1998 to December 2001 in order to examine whether betas from the Capital Asset Pricing Model (CAPM) are a valid measure of risk and whether industries to which the firms belong explain excess returns. The methods we propose lead to more explanatory power than the traditional OLS results.

Key words: Capital Asset Pricing Model, industrial effects, linear mixed-effect model, longitudinal data analysis, robust estimation.

1. Introduction and Background

The Capital Asset Pricing Model (CAPM) of Sharpe (1964) and Lintner (1965) has been widely used in measuring the relationship between the expected return on a security and its risk in financial markets. However, it has been under attack from all corners since its inception. The most serious assault in recent years, arguably, was initiated by Fama and French (1992, 1993, hereafter FF), who advocate an APT-like three-factor model to replace the CAPM.

The evidence presented by Fama and French does not convince everybody. Competing theories have been offered to explain the cross-sectional variation of stock returns (Lakonishok *et al.*, 1994; Deniel and Titman, 1997) those who support the CAPM also provide evidence to prove that it is still alive and well (Kothari *et al.*, 1995). In this paper we would like to offer some opinions from two aspects to contribute to the literature.

Our main contributions firstly come from the estimation technique. Traditionally, the estimation of the relationship between risk and expected return consists

of two stages following Fama and Macbeth (1973, hereafter FM). The first is that betas are estimated from time series regressions by ordinary least squares (OLS) estimators, and the second stage tests the relationship between mean returns and betas across firms or portfolios by taking time series averages of the slope coefficients from monthly cross-sectional regressions (Shalit and Yitzhaki, 2001).

The above approach has been modified by the ways in which the variables are defined and how the betas are estimated; see the reviews respectively by Kothari *et al.* (1995) and Chou and Liu (2000). However, much of the research maintains the structure of a two-stage estimation, which is not desirable from statistical points of view in many ways. Specifically, Shanken (1992) proves that the estimates in two-stage regressions are consistent if the sample length tends to infinity and if the error terms are independently and identically distributed over time, and the assumptions are only approximations to real-world data. In this paper we propose another way to estimate the relationship between returns and betas by using robust estimates to prevent potential outliers at the first stage and by using the approach of the longitudinal research to deal with the estimation problem at the second stage.

Although the OLS produces the optimal estimate of linear model coefficients and provides a convenient distribution theory for inference when the errors are Gaussian, it is well known that the OLS objective is particularly sensitive to outliers and leverage points. The problem may partly be solved by altering some seemingly outliers or leverage points. For example, FF (p. 439) reset the values of the largest and smallest 0.05% of observations for some of the explanatory variables, but this kind of approach is *ad hoc* to say the least. It essentially needs robust estimators to cope with multiple outliers. We believe that by using an appropriate technique the estimation of the beta or other variables can be greatly improved.

We are not the first ones to apply robust statistics to conduct tests of asset pricing. Knez and Ready (1997) extend FF's monthly cross-sectional regressions by applying a least trimmed squares (LTS) regression and show that the significance of size and book-to-market ratio obtained by OLS regressions in FF (1992) is due to positive skewness in the return distributions. However, their robust tests on CAPM focus only on the cross-sectional averages across firms in each portfolio and across months, whereas we apply robust estimates and employ the longitudinal approach. In addition, the two-level linear mixed-effects (LME) model allows us to formulate the industrial and company effects elegantly by treating them as random effects. By contrast, we have to use dummy variables to represent the effects in the OLS regressions.

FM's approach uses independent *t*-tests to examine whether time series averages of betas from the monthly cross-sectional regressions are a valid measure

of risk. Nevertheless, it essentially creates the so-called repeated measurement problem in the statistical literature: some important information might be discarded when only using a cross-sectional analysis. The panel data analysis is one of the approaches to deal with the problem of this kind, for it takes the time-series properties of the observations into account (Hsiao, 1986). The common trait of longitudinal research is that the entity under investigation or unit of study is observed or measured at more than one point in time, possibly repeatedly, and that it develops over time.

If the beta cannot fully explain the variation of excess returns, then what can? Our second main contribution from this paper is to provide some evidence of a possible candidate: industrial effects. It is widely believed among academics and practitioners that the industries to which firms belong are related to their expected returns, presumably because they are proxies for different risks. However, several studies report that industry factors are weak or imprecise (Roll, 1992; Fama and French, 1997; Alexander and Thistle, 1999). We examine this issue by going a step further. Those papers use the U.S. Government's Standard Industry Classification (SIC) codes or Goldman Sachs Classification to classify stocks into several industries. Because they try to make their sample comprehensive, their classification may be too crude. In contrast, we only focus on a small number of stocks to improve the precision of classification.

The sample we choose is a group of stocks representing companies in Taiwan's electronics industry. This sample is selected for two reasons. Firstly, the sample period is between 1998 and 2001, during which electronics stocks attracted most of the attention from investors. Trading activities concentrated on those stocks and they sometimes accounted for nearly 90% of the daily trading volumes. A study of the behavior of electronics stocks is indeed a study of the trading of general Taiwanese investors.

Secondly, the electronics industry in Taiwan is complete. The listed companies include those ranging from designing and producing integrated circuits to computer retailers and distributors. We can make the classification more precise in the industry, and if there is no industrial effect at all in the industry, then the effect may not be found elsewhere. Specifically, we classify the companies into upper, middle, and lower streams to examine whether their stock returns are related to the level in which they are located. By focusing on the industrial effects of the various levels of the electronics industry, we hope to detect a strong relationship between returns and the industry, which is difficult to uncover under a broad classification.

Our findings can be summarized as follows. Firstly, the betas estimated by the LTS regressions are better able to capture the time-series variations of the excess returns than the OLS. The LTS betas also perform well in both the cross-

sectional regressions analysis of the FM procedure and the LME model. Secondly, under the FM procedure, the LTS beta is significant in explaining the returns only after interacting with the industrial dummies. Under the LME model, both the OLS betas and the LTS betas are significant in all models. Thirdly, the industrial effects are visible under the FM procedure as well as the LME model. Overall, there is evidence for and against the CAPM. The robust estimate of the beta successfully explains part of the returns under the LME model, while it leaves a large part of the returns unexplained in which the industrial effects may play roles.

The rest of this article is organized as follows. Section 2 gives a brief introduction about robust estimators. Apart from the LTS mentioned above, we also discuss the RDL_1 estimator, which is particularly used in the linear regression model with both continuous and discrete regressors. Section 3 provides a preliminary analysis of the data. Section 4 modifies the second stage of the FM procedure to present a cross-sectional analysis of stock returns in the presence of industrial effects with both OLS and robust estimations. The robust estimates yield more reasonable conclusions to the aspect of CAPM. Section 5 considers the longitudinal data analysis to treat the estimation problem during the second stage. Some conclusions and comments are drawn in Section 6.

2. Robust Statistics

In this section we briefly review the application of robust statistics to regression models. A linear regression model can be written as

$$y_i = \beta_0 + \sum_{j=1}^p \beta_j x_{ij} + \epsilon_i, \quad i = 1, \dots, n, \quad (2.1)$$

where there are p explanatory variables and they are usually quantitative, β_j is the j th regression coefficient, and β_0 denotes the intercept term. We often make certain idealized assumptions about the error term, ϵ_i , to be independent and identically distributed with a normal distribution, $N(0, \sigma^2)$, for the purpose of statistical inferences. A least squares analysis weights each observation equally in achieving the parameter estimates. Therefore, it is sensitive to outlying cases. The robust methods, however, enable the observations to be weighted unequally, so that observations with large residuals are down-weighted.

Many robust methods have been discussed and proposed since the middle of the 1960s (Huber, 1981). One of the desirable properties for a robust estimator is one with a higher breakdown point. The (finite sample) breakdown point of an estimator refers to the smallest proportion of observations which when altered can cause the value of the estimator to be arbitrarily large or small (see

Hampel *et al.*, (1986)). We focus on two estimators with a high breakdown point here. The first one is the least trimmed squares (LTS) estimator proposed by Rousseeuw (see Rousseeuw and Leory (1987)), which Knez and Ready (1997) apply to the CAPM. The second one is RDL_1 proposed by Huber and Rousseeuw (1997), which is particularly used in the case when the regression model consists of binary regressors.

2.1 Least trimmed squares estimator

For the regression model (2.1), OLS finds any estimate of $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_p)$ to minimize the sum of the residuals, that is

$$\min_{\boldsymbol{\beta}} \sum_{i=1}^n e_i^2, \quad (2.2)$$

where $e_i = e_i(\hat{\boldsymbol{\beta}}) = y_i - x_i^T \hat{\boldsymbol{\beta}}$ is the residual and $\hat{\boldsymbol{\beta}} = (\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_p)$ is the estimator. To obtain the LTS, we first let

$$e_{(1),n}^2 \leq e_{(2),n}^2 \leq \dots \leq e_{(n),n}^2$$

be the ordering of the residuals. The objective function of LTS is to

$$\min_{\hat{\boldsymbol{\beta}}} \sum_{i=1}^q e_{(i),n}^2, \quad (2.3)$$

where q is $[n/2] + [(p+1)/2]$ and $[\cdot]$ denotes the integer part. It has the largest breakdown point for LTS (see Rousseeuw and Leory (1987, p.132)). Atkinson and Cheng (1999) discuss the choice of q . They show that one can get more stable results for the detection of outliers as well as highly efficient estimates when more data are fitted, provided q is small enough to exclude outlying cases. Zeman *et al.* (2001) suggest that $[0.75n]$ is a reasonable value for q in most empirical studies.

In contrast to OLS, the LTS estimator has a high breakdown point which ensures it is robust enough to resist multiple outliers. The LTS fits only those q observations with the smallest residuals and trims the rest where $q \leq n$. Note that since the LTS depends only on the residuals, it generally will not trim the same number of observations from the upper and lower tails of the distribution. A more detailed concept about LTS can be referred to that in Rousseeuw and Leory (1987).

2.2 RDL₁ estimator

If there are qualitative explanatory variables, then it is conventional in practice to encode such regressors by binary dummy variables. Extending model (2.1) to a model with continuous and discrete regressors, the model is expressed in the following form

$$y_i = \beta_0 + \sum_{j=1}^p \beta_j x_{ij} + \sum_{l=1}^k \gamma_l I_{il} + \epsilon_i, \quad (2.4)$$

where I_l denotes the dummy variable with elements either zero or one. Huber and Rousseeuw (1997) point out that we cannot simply run the robust regression, such as LTS, by treating the dummy variable in the same way as the continuous regressors as normally done when using OLS in model (2.1), since this may lead to a problem of singular matrices. Therefore, they propose the RDL₁ estimator which can withstand contaminations in the data for model (2.4). Here, RD stands for robust distance, and L₁ indicates the least absolute values method in the estimates of the regression coefficients.

The RDL₁ consists of three stages: identifying leverage points, downweighting the leverage points when estimating the parameters, and estimating the residual scale. In the first stage we look for leverage points, i.e., outliers in set $\mathbf{X} = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ where the components of $\mathbf{x}_i = (x_{i1}, \dots, x_{ip})$ are the continuous regressors and \mathbf{X} has the dimension of p . The minimum volume ellipsoid estimator (MVE) of Rosseeuw (1985) is used to obtain the robust Mahalanobis distances. The MVE consists of a robust location estimator $T(\mathbf{X})$ defined as the center of the small ellipsoid containing half the points of \mathbf{X} , as well as a scatter matrix $C(\mathbf{X})$ given by the shape of that ellipsoid. The robust distances are defined as

$$RD(\mathbf{x}_i) = \sqrt{(\mathbf{x}_i - T(\mathbf{X}))C(\mathbf{X})^{-1}(\mathbf{x}_i - T(\mathbf{X}))^T}, \quad i = 1, 2, \dots, n.$$

If the regressors \mathbf{x}_i are observational (rather than designed) with a multivariate normal distribution, then the $T(\mathbf{X})$ and $C(\mathbf{X})$ are consistent for the underlying parameters (Davies, 1992). For large n , the $(RD(\mathbf{x}_i))^2$ would thus approximate χ^2 distributed with degrees of freedom p (Rousseeuw and Leroy, 1987). Consequently, observations with larger values of $RD(\mathbf{x}_i)$ can be identified as leverage points.

Based on the robust distances, the second step is to compute weights w_i by

$$w_i = \min \left\{ 1, \frac{p}{(RD(\mathbf{x}_i))^2} \right\}, \quad i = 1, \dots, n.$$

The parameters (β, γ) of model (2.4) are then estimated by a weighted L_1 procedure

$$\min \sum_{i=1}^n w_i |e_i|,$$

where e_i denotes the i th residual in the fitting. In the final step the scale of the residuals is estimated by

$$\hat{\sigma} = 1.4826 \text{median}_i |e_i|,$$

where the constant 1.4826 makes the estimator consistent at normal errors.

The robust estimate $(\hat{\beta}, \hat{\gamma}, \hat{\sigma})$ can now be used to detect regression outliers, by flagging the observations whose absolute standardized residual $|e_i/\hat{\sigma}|$ exceeds 2.5. The entire three-stage procedure is called the RDL_1 estimation, and its statistical properties and its S-PLUS code can be found in Huber and Rousseeuw (1997).

3. Preliminary Analysis

3.1 Data description

We begin our investigations by applying the FM procedure to data related to electronics stocks in Taiwan. The data cover the period from September 1998 to December 2001. The choice of the sample period is a compromise between the length of the period and the availability of the data, for many of these stocks were not listed on the Taiwan Stock Exchange until 1998. Monthly stock returns on individual securities and the market index are obtained from the database of the Taiwan Economic Journal, in which stock prices have been adjusted for any ex dividend. The sample consists of 48 stocks in which there are 23 companies in the upper stream (semiconductor and printed circuit boards), 14 in the middle stream (motherboards and components), and 11 in the lower stream (PCs and monitors).

We have implemented some graphical methods for the preliminary analysis. Because the graphs are too tedious to be shown here, we only report a part of them and summary descriptions about the data are as follows.

1. The box plots of excess returns for each month presented in Figure 1 show that there exist outlying observations.
2. The time series plot of excess returns shows that these levels have different, but not obvious patterns.
3. The histogram of excess returns of all companies indicates that the excess returns are skewed to the right.

4. The scatter plots of returns versus risk for each company show that there exists a linear relationship between them and several obvious and potential outliers appear.

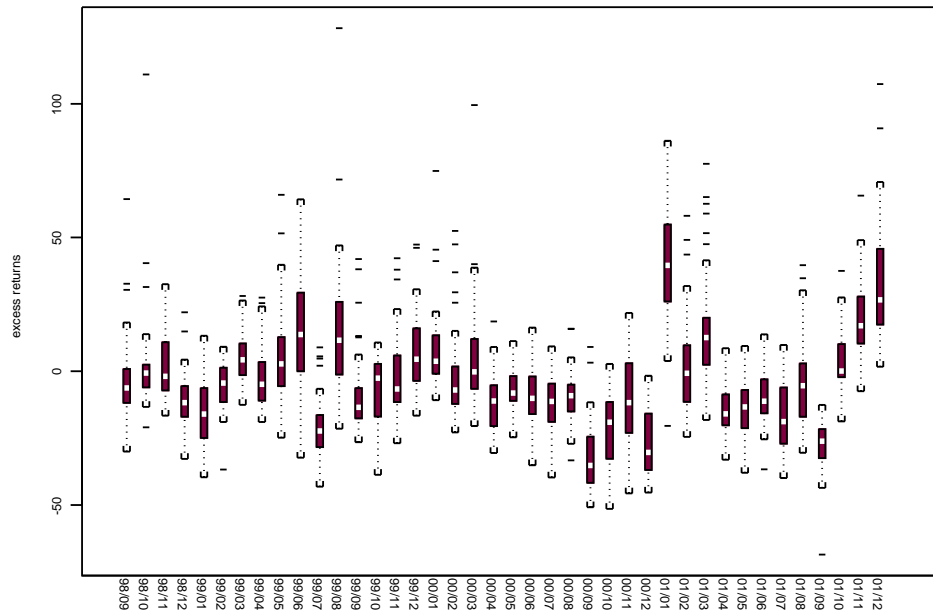


Figure 1: Box plots of excess returns corresponding to each company.

In sum, using cross-sectional regression analysis by OLS may be affected by those potential outliers. Furthermore, there exist variations among the monthly excess returns. The assumption of constant variance when applying OLS might be violated.

3.2 Estimating betas

Fama and Macbeth (1973) conduct two stages of regression analysis to examine the CAPM. The theoretical form of CAPM can be expressed as

$$E(R_i) = R_f + \beta_i(E(R_m) - R_f),$$

where R_i is the rate of return of security i , R_f is the risk-free rate, and R_m is the market return. Because β_i is not observable, it has to be estimated from a time-series analysis in the first stage:

$$(R_{jt} - R_{ft}) = \alpha_j + \beta_j R_{ct} + \epsilon_{jt}, \quad j = 1, 2, \dots, 48, \quad t = 1, 2, \dots, 40, \quad (3.1)$$

where $(R_{jt} - R_{ft})$ is j security's excess returns at time t and the risk-free interest rate, R_{ft} , is the secondary market interest rate of the 10-year Treasury bond at time t . Term R_{ct} is the market excess returns at time t , $R_{mt} - R_{ft}$, and the rate of return on the Taiwan Stock Exchange Capitalization Weighted Price Index (TAIEX) of the Exchange is used as the proxy for R_{mt} .

To compare the difference between the OLS and the LTS in estimating the betas, we conduct 48 time series regressions for each company. We then take averages of the estimated coefficients for the 48 models, which are denoted as $\bar{\alpha} = \sum_{j=1}^{48} \hat{\alpha}_j / 48$ and $\bar{\beta} = \sum_{j=1}^{48} \hat{\beta}_j / 48$. Finally, t -statistics are used to test the hypotheses of $\alpha = 0$ and $\beta = 0$.

The results are presented in Table 1. Note that this is not the standard procedure by Fama and Macbeth (1973), who use the "rolling beta" approach, as illustrated in the next section. The FM procedure yields multiple betas for each security, which are difficult to summarize. Therefore, at this stage we simply follow Fama and French (1992) who implement the full sample to estimate betas. Terms TS1 and TS2 indicate the OLS and LTS used to estimate the coefficients, respectively. The coefficients presented in the table are the averages of the estimated time series regression coefficients of 48 companies. Both $[0.7n]$ and $[0.8n]$ for the value of q used in LTS objective function (2.3) yield similar results. Here, we report the results when $q = [0.8n]$.

Table 1: Time series regression analysis for each company

		TS1	TS2
$\bar{\alpha}$		3.4137**	-0.6827
	$t(\bar{\alpha})$	(6.3877)	(-1.2090)
$\bar{\beta}$		1.3262**	1.1068**
	$t(\bar{\beta})$	(26.1410)	(20.7775)
Average R^2		0.4810	0.4785

"*" denotes significance at the 10 percent level, and "***" denotes significance at the 5 percent level.

When taking the average of the coefficients (as presented in Table 1) the betas of both models have similar values, which are higher than one, meaning that the returns of electronics stocks are very sensitive to the returns of the market. The t -values of the betas indicate that they are significant. The estimated intercept is not significant by the LTS, whereas it is somewhat large by the OLS. In such a regression model, the intercept of a good pricing model should be close to zero (Merton, 1973; Ross, 1976), and Table 1 indicates that the mispricing of

the CAPM in TS1 is due to outliers. Without the effect of those outliers when applying the LTS, TS2 is more in line with the CAPM.

We also examine the plots of 95% confidence interval estimates for the coefficients of each company from TS1 and TS2. To save space, we only summarize the results. It appears that the coefficients may differ among companies and some of them show a considerable difference from others. However, for the majority of the stocks, zero is well in the confidence intervals of intercepts and one is in the interval of betas. This suggests that the CAPM works well for the majority of the stocks, and the poor performance of TS1 in Table 1 results from a few outliers.

4. Allowing Industrial Effects

4.1 Two-stage regressions

In this section we conduct the FM procedure to include betas and the variables of industrial effects (“stream” dummies). In the first stage the beta of model (3.1) is estimated using the rolling method in Fama and Macbeth (1973). For example, the first twelve months (from September 1998 until September 1999, $t = 1$ to 12) are used to obtain the slope estimate which leads to be the beta of the 13th month, denoted as $\hat{\beta}_{j,13}$ for security j . We next use the second 12 months (from October 1998 until October 1999, $t = 2$ to 13) to obtain the beta of the 14th month. Iteratively, we then will get 28 months (from October 1999 until December 2001) of estimated betas to be the explanatory variables of the cross-sectional regression analysis as follows:

$$\begin{bmatrix} \hat{\beta}_{1,13} & \hat{\beta}_{1,14} & \cdots & \hat{\beta}_{1,40} \\ \hat{\beta}_{2,13} & \hat{\beta}_{2,14} & \cdots & \hat{\beta}_{2,40} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\beta}_{48,13} & \hat{\beta}_{48,14} & \cdots & \hat{\beta}_{48,40} \end{bmatrix}. \quad (4.1)$$

The 12-month period to estimate betas is called the formation period, and the 28-month period is called the testing period. The coefficient of beta is the risk premium. Therefore, there are 28 months used in the cross-sectional regression analysis.

For the cross-sectional regression in the second stage, we consider three different models. The first one includes beta as the only one explanatory variable; that is,

$$R_{jt} - R_{ft} = \lambda_{0j} + \lambda_{1j}\hat{\beta}_{jt} + \epsilon_{jt}, \quad j = 1, 2, \dots, 48. \quad (4.2)$$

The second model uses dummy variables to capture the stream effects:

$$R_{jt} - R_{ft} = \lambda_{0j} + \lambda_{1j}\hat{\beta}_{jt} + \lambda_{2j}I_1 + \lambda_{3j}I_2 + \epsilon_{jt}, \quad (4.3)$$

$$j = 1, 2, \dots, 48,$$

where

$$I_1 = \begin{cases} 1 & \text{if the company belongs to the lower stream} \\ 0 & \text{otherwise,} \end{cases}$$

and

$$I_2 = \begin{cases} 1 & \text{if the company belongs to the middle stream} \\ 0 & \text{otherwise.} \end{cases}$$

The third model takes account of the interaction of streams and betas:

$$R_{jt} - R_{ft} = \lambda_{0j} + \lambda_{1j}\hat{\beta}_{jt} + \lambda_{2j}I_1 + \lambda_{3j}I_2 + \lambda_{4j}I_1\hat{\beta}_{jt} + \lambda_{5j}I_2\hat{\beta}_{jt} + \epsilon_{jt}, \quad (4.4)$$

$$j = 1, 2, \dots, 48.$$

Models (4.2), (4.3), and (4.4) yield 28 regression equations for $t = 13, \dots, 40$.

Parts of our graphical results for the estimated betas and comments on those figures are summarized as follows.

1. Plots (1) and (2) of Figure 2 show the box plots of the estimated betas for each month computed by OLS and LTS approaches, respectively. There appear outliers for several months. Similar plots for each company are not shown here. Both OLS and LTS results show that the estimated betas for each company are more fluctuating than those for each month.
2. The time series plots of the estimated beta by OLS and by LTS reveals that the trends have three different types in the three streams with the OLS beta and that the beta with OLS is smoother than the LTS method.
3. The scatter plots of security excess returns and betas show that the betas estimated by OLS are more concentrated in the X -direction than those estimated by LTS.

4.2 OLS

Following FM, we conduct month-by-month cross-sectional regressions of stock excess returns on the beta and calculate the time series averages of each regression's intercept and slope. The t -statistics are used to test whether the averages are different from zero.

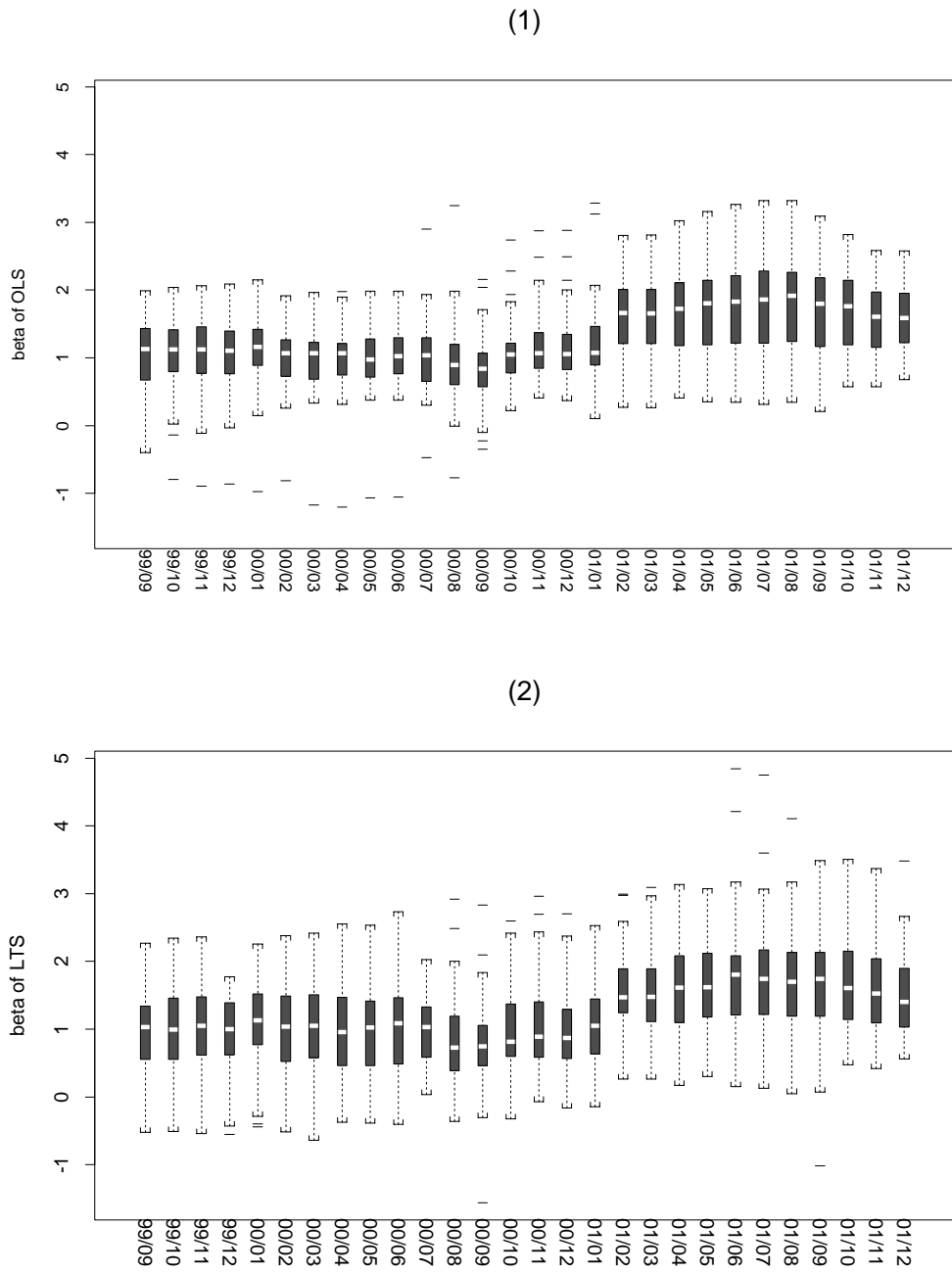


Figure 2: Box plots of extended betas for each month: (1) by OLS and (2) by LTS.

We first apply the OLS to FM's procedure. In Table 2, OLS1 to OLS3 are the averages of the estimated month-by-month cross-sectional regressions coefficients for models (4.2) to (4.4), respectively. The only significant variable in the three models is the intercept, indicating neither the beta nor the stream dummy is able to explain the cross-sectional variations of the stocks.

Table 2: The average results of cross-sectional regressions using OLS

	OLS1	OLS2	OLS3
Intercept	-6.1423**(-2.4638)	-5.5934**(-2.2100)	-4.0033(-1.4841)
$\beta_{j,t}$	1.1789 (0.8880)	0.7738 (0.5710)	-0.3621(-0.2827)
I_1		-0.8922 (-0.6023)	-2.8814(-0.8492)
I_2		0.2153 (0.1886)	-2.5964(-0.8974)
$\beta_{j,t} \times I_1$			1.8075 (0.9385)
$\beta_{j,t} \times I_2$			1.4313 (0.6744)
Average R^2	0.0564	0.1227	0.1690

“**” denotes significance at the 10 percent level, and “***” denotes significance at the 5 percent level.

Examining the coefficients associated with the variables still offers some insights of the excess returns. The average risk premium of beta in OLS2 is smaller than in OLS1, indicating that stream dummies may explain part of the risk premium. Furthermore, when the interaction of the dummy variables and betas is also concluded in OLS3, the average risk premium of beta is negative for the upper stream while that for the middle and lower streams it is positive. Although the estimated coefficients are not statistically significant, it suggests that risk premia vary greatly across different streams of stocks.

4.3 Robust results

We now apply a robust method to examine the effect of outliers and report the results in Table 3. LTS1 denotes the results of model (4.2) obtained by LTS, whereas RDL_12 and RDL_13 are those respectively of (4.3) and (4.4) used by the RDL_1 estimator. As explained in Section 2.2, LTS cannot be applied here, because it may produce singular matrices in a regression model with dummy variables.

Similar to Table 2, the intercept is significant in the result of LTS1 and RDL_12 , but not in RDL_13 . Moreover, the estimated coefficients of the other variables in LTS1 and RDL_12 are not significant, but RDL_13 is quite interesting. The coefficient of $\hat{\beta}_{j,t}$ is small and insignificant, and those of $\hat{\beta}_{j,t} \times I_1$ and $\hat{\beta}_{j,t} \times I_2$ are large and statistically significant. This implies that the risk premium is

small for the upper stream companies and is big for the middle and lower stream companies. The lower stream companies have the biggest risk premium. In addition, the coefficient of I_1 is negative and significant, which means that a stream effect exists in the electronics industry that cannot be explained by the beta. Negative I_1 and I_2 imply that the returns of the middle and the lower stream companies are smaller than those of the upper stream companies.

In terms of the signs of coefficients, the results of Table 3 are not very different from those of Table 2, but in terms of statistical significance, those of Table 3, especially of RDL₁3, are much stronger. Therefore, the outlying observations weaken the relationship between return, risk, and industrial effects, and the relationship holds for the majority of stocks.

Table 3: The average results of cross-sectional regressions using robust estimators

	LTS1	RDL ₁ 2	RDL ₁ 3
Intercept	-7.8958**(-2.7925)	-6.6353**(-2.4414)	-3.9678 (-1.2907)
$\beta_{j,t}$	1.2304 (1.1082)	1.2205 (0.9902)	-1.3729 (-1.1022)
I_1		-0.3558 (-0.2493)	-7.9242**(-3.2196)
I_2		0.4342 (0.2793)	-4.8489 (-1.5155)
$\beta_{j,t} \times I_1$			6.3859** (3.1999)
$\beta_{j,t} \times I_2$			4.8679** (2.2651)
Average R^2	0.1012		

“**” denotes significance at the 10 percent level, and “***” denotes significance at the 5 percent level.

5. Longitudinal Data Analysis

In this section the longitudinal data analysis is applied to deal with the problem of CAPM. Two strong reasons motivate us to employ the longitudinal data technique in the investigation of CAPM here. Firstly, the FM approach of averaging coefficients of cross-sectional regressions ignores time-series properties of the returns within the same company or the same industry, whereas the longitudinal data analysis takes both time-series and cross-sectional properties into account. Secondly, the linear mixed-effect model for longitudinal data provides a convenient way to include the stream effect by treating it as a random effect. Moreover, the variation among different companies is included in the model. Other advantages of the longitudinal data analysis can be referred to that of Pinheiro and Bates (2000), in which applications and theoretical discussions about this approach are given.

The two-level linear mixed-effects model is used to analyze these data with the form of

$$\begin{aligned}
 Y_{ij} &= \beta_{ij}\boldsymbol{\lambda} + \mathbf{Z}_{i,j}b_i + \mathbf{Z}_{ij}b_{ij} + \epsilon_{ij} \quad i = 1, 2, 3 \quad j = 1, 2, \dots, n_i \\
 b_i &\sim N(0, \sigma_1^2), \quad b_{ij} \sim N(0, \sigma_2^2), \quad \epsilon_{ij} \sim MVN(\mathbf{0}, \sigma_0^2 \mathbf{I}),
 \end{aligned}
 \tag{5.1}$$

where $\boldsymbol{\lambda}$ is a 2×1 parameter vector of fixed effects; β_{ij} is a 28×2 matrix of betas (we omit the “hat” notation of β hereafter); b_i is the first-level (stream) random effects; and b_{ij} is the second-level (company) random effects, which refers to company j within stream i . Therefore, $n_1 = 23$, $n_2 = 14$, and $n_3 = 11$ in model (5.1).

Note that the variance functions can be used to model the variance structure of the within-group errors to allow heteroscedasticity in model (5.1). For example, the variance function model for a two-level model is

$$\begin{aligned}
 \text{Var}(\epsilon_{ijt}|b_i, b_{ij}) &= \sigma^2 g^2(\mu_{ijt}, \nu_{ijt}, \delta), \\
 i &= 1, 2, 3, \quad j = 1, 2, \dots, n_i, \quad t = 13, 14, \dots, 40,
 \end{aligned}$$

where $\mu_{ijt} = E[Y_{ijt}|b_i, b_{ij}]$, ν_{ijt} is a vector of variance covariates, δ is a vector of variance parameters, and $g(\cdot)$ is the variance function, assumed to be continuous in δ . Furthermore, in the context of mixed-effect models, correlation structures are used to model dependence among the within-group errors.

To establish a general framework for the correlation structure of model (5.1), the within-group errors within the same innermost level of group are

$$\begin{aligned}
 \text{cor}(\epsilon_{ijt}, \epsilon_{ijt'}) &= h[d(p_{ijt}, p_{ijt'}), \phi], \\
 i &= 1, 2, 3, \quad j = 1, 2, \dots, n_i, \quad t, t' = 13, 14, \dots, 40,
 \end{aligned}$$

where $d(p_{ijt}, p_{ijt'})$ denotes the distance between the corresponding *positions*, p_{ijt} and $p_{ijt'}$. The details of the variance (eg., the choice of $g(\cdot)$) and correlation structures of model (5.1) can be referred to Pinheiro and Bates (2000, Chapter 5) and Diggle *et al.* (1996, Chapter 5).

To determine the random effects, we presume that the different streams and companies will affect the excess returns of each security in different ways. Therefore, the first-level random effect indicates the stream and the second-level random effect denotes the company. The random effect matrix $\mathbf{Z}_{i,j}$ is a vector of 1s, which means that the first-level random effect has an equal effect on the security’s excess returns under those companies in the same stream. The random effect matrix \mathbf{Z}_{ij} is a vector of 1s which means that the second-level random effect affects the security’s excess returns on an equal weight, but changes with different streams and companies.

5.1 LME with OLS betas

Table 4 shows the results of model (5.1) where the values of β are obtained by the OLS approach. The four models all assume that the random effects are “stream” and “company within stream”. The difference is in the assumption of the within-group error ϵ_{ij} . The error term of OLS4 is identically, independently, and normally distributed as in (5.1). OLS5 assumes that within-group errors are heteroscedastic; that is, the variances of ϵ_{ij} are different for each stream and company. OLS6 assumes that the variances are correlated with AR(1), and OLS7 assumes that the variances are both heteroscedastic and correlated with AR(1).

Table 4: The results of LME using OLS betas

	OLS4	OLS5	OLS6	OLS7
Intercept	-7.5753** (-5.7104)	-8.6983** (-6.9143)	-7.4185** (-4.9375)	-8.5712** (-6.1385)
β_{jt}	2.7165** (2.9697)	3.5623** (3.7398)	2.7327** (2.6425)	3.5523** (3.3645)
σ_0	21.937	15.766	21.959	15.717
σ_1	0.00214	0.00281	0.00214	0.00153
σ_2	0.00391	0.00389	0.00390	0.00279
ϕ			0.1346	0.1171
AIC	12120.68	12031.38	12101.14	12019.23
BIC	12146.69	12301.88	12132.36	12294.93

“*” denotes significance at the 10 percent level, and “**” denotes significance at the 5 percent level.

In comparison with OLS1 in Table 2, of which the average risk premium 1.1789 is insignificant, those risk premiums of the panel data analysis, OLS4, OLS5, OLS6, and OLS7 in Table 4, are all significant. The risk premiums estimated by the LME model are all more than twice those in Table 2, which means that the cross-sectional analysis may ignore important information in the time series.

For all the LME models from OLS4 to OLS7, the standard deviations of random effects σ_1 and σ_2 are both smaller relative to the within-group standard deviation σ_0 . It seems that the random effects of stream and company can only explain a small fraction of variations in excess returns. Finally, ϕ is the coefficient of the AR(1) model and is not very large, which suggests that the self correlation of variance is not strong.

5.2 LME with LTS betas

We now use the betas estimated by LTS to carry out a similar analysis as in the previous subsection. The numbers 4 to 7 affixed after “LTS” in Table 5 are the same assumptions on model (5.1) as discussed in the results of Table 4.

Table 5 reveals that the standard deviations of random effects σ_1 and σ_2 are both smaller relative to the within-group standard deviation σ_0 . It seems that the random effects of stream and company are small relative to the within-group variations.

In comparison with the average risk premium of the cross-sectional analysis without the stream effect, which is LTS1 in Table 3, the average risk premium is insignificant, while those in Table 5 are all significant. Again, it means that the cross-sectional analysis may ignore important information in the time series for a given company.

Table 5: The results of LME using LTS betas

	LTS4	LTS5	LTS6	LTS7
Intercept	-6.1223** (-5.3375)	-6.8944** (-6.5311)	-5.7270** (-4.4825)	-6.6124** (-5.7326)
β_{jt}	1.6856** (2.1094)	2.1675** (2.7064)	1.5098* (1.7123)	2.0170** (2.3208)
σ_0	21.973	16.073	21.995	15.935
σ_1	0.01319	0.00963	0.01276	0.00923
σ_2	0.01273	0.00930	0.01223	0.00885
ϕ			0.1357	0.1150
AIC	12125.30	12038.18	12105.49	12025.43
BIC	12151.31	12308.68	12136.70	12301.13

“*” denotes significance at the 10 percent level, and “**” denotes significance at the 5 percent level.

In comparison with the results in Table 4, interesting patterns finally emerge. All the intercepts and the coefficients of betas of the four models of Table 5 are smaller than their counterparts in Table 4, while the standard deviations of the random effects greatly multiply. Similar to the findings in Section 4, the contrast of Table 4 and Table 5 indicates that outliers weaken the explanatory power of beta as well as the industrial effects.

5.3 Model comparison and diagnostics

We next consider the panel data analysis of linear mixed-effects models with a different assumption about the within-group errors, which leads to several models at hand. A model selection about these LME results is needed. Here, we use the Akaike information criterion (AIC) and Bayesian information criterion (BIC); see Pinheiro and Bates (2000). They are defined as follows:

$$\begin{aligned}\text{AIC} &= -2 \times \log(\text{likelihood}) + 2 \times p \\ \text{BIC} &= -2 \times \log(\text{likelihood}) + p \times \log n,\end{aligned}$$

where p represents the number of parameters in the fitted model, and n represents the number of observations. Both criteria show that the smaller the values are, the better the fit.

The results are shown in the bottom rows of Table 4 and Table 5. For the AIC, model OLS7 is the best since it has the smallest values, but we can see that the AIC of model LTS7 is larger than that of OLS7 by just 0.052%. Therefore, for the criterion of AIC, we conclude that linear mixed-effects models OLS7 and LTS7 are both reasonable modeling for the data.

6. Conclusions

We have applied robust estimators and panel data analysis to re-examine the CAPM. There are more variations in the LTS betas than the OLS betas, and the former are more successful in capturing the time-series variations of excess returns. This implies that the inclusion of outliers undermines the relationship between individual and market excess returns. Furthermore, the use of the LTS betas in the second-stage analysis, no matter if the traditional FM procedure or the LME model is employed, increases the t -values of the coefficients associated with the betas.

The advantage of the longitudinal analysis is to take account of both the time-series and cross-sectional properties of the observations. However, our LME model does not give further support to the CAPM. While the betas are significant in all of the LME models, the intercepts are bigger in those models than in the FM procedure, which suggests that the beta fails to fully explain the returns. Allowing for heteroscedasticity in the error terms to increase the precision of the estimation, it also increases the t -values of both the intercept and the coefficient of the beta. Allowing for autocorrelation in the error terms does not change the result very much.

Does the industrial/stream effect exist in the market? Our findings suggest that it does, especially when the LTS betas are used. In the FM procedure, the

stream effects are not only visible in the dummies, but they are also related to the betas. The significance of the interaction terms between betas and the stream dummies suggests that the returns of the companies from different streams behave differently with the betas. In the LME model, the stream and the random effects are somewhat small, but they are larger when the LTS betas are used.

Taking all the evidence together, CAPM cannot fully explain the excess returns. The industrial effects may help explain part of the returns, but not all. However, we only consider here the electronics industry in Taiwan's stock market. For future studies, first of all, including the other industries in the whole market is needed. Secondly, more tests are required by considering a wide range of explanatory variables, including macroeconomic variables such as those in Chen *et al.* (1986), or risk factor variables such as in Fama and French (1993), which have been shown to successfully explain cross-sectional stock average returns of stocks in the U.S.

References

- Alexander, D. L. and Thistle, P. D. (1999). Market power, efficiency and the dispersion of systematic risk. *Review of Industrial Organization* **14**, 377-390.
- Atkinson, A. C. and T.-C. Cheng. T.-C. (1999). Computing the least trimmed squares regression with the forward search. *Statistics and Computing* **9**, 251-263.
- Chen, N.-F. and Roll, R. and Ross, S. A. (1986). Economic forces and the stock market. *Journal of Business* **59**, 383-403.
- Chou, P.-H. and Liu, Y.-F. (2000). The cross section of expected returns in Taiwan: Characteristics, single factor, or multi factors? *Review of Securities and Futures Markets* **12**, 1-32.
- Daniel, K. Titman, S. (1997). Evidence on the characteristics of cross sectional variation in stock returns. *Journal of Finance*, 1-33.
- Davies, L. (1992). The asymptotic of Rousseeuw's minimum volume ellipsoid estimator. *Annals of Statistics* **20**, 1828-1843.
- Diggle, P. J. and Liang, K. Y. and Zeger, S. L. (1996). *Analysis of Longitudinal Data*. Clarendon Press.
- Fama, E. F. and French, K. R. (1992). The cross-section of expected stock returns. *Journal of Finance* **47**, 427-465.
- Fama, E. F. and French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* **33**, 3-56.
- Fama, E. F. and French, K. R. (1997). Industry costs of equity. *Journal of Financial Economics* **43**, 153-193.

- Fama, E. F. and Macbeth, J. D. (1973). Risk, return and equilibrium: Empirical tests. *Journal of Political Economy* **81**, 607-636.
- Hampel, F. R., Ronchetti, E. M., Rousseeuw, P. J. and Stahel, W. A. (1986). *Robust Statistics: The Approach Based on Influence Functions*. John Wiley.
- Hsiao, C. (1986). *Analysis of Panel Data*. Cambridge University Press.
- Huber, P. J. (1981). *Robust Statistics*. John Wiley.
- Hubert, M. and Rousseeuw, P. J. (1997). Robust regression with both continuous and binary regressors. *Journal of Statistical Planning and Inference* **57**, 153-163.
- Knez, P. J. and Ready, M. J. (1997). On the robustness of size and book-to-market on cross-sectional regressions. *Journal of Finance* **52**, 1355-1382.
- Kothari, S. P. Shanken, J. and Sloan, R. G. Another look at the cross-section of expected returns. *Journal of Finance* **50**, 185-224.
- Lakonishok, J., Shleifer, A. and Vishny, R. W. (1994). Contrarian investment, extrapolation, and risk. *Journal of Finance* **49**, 1541-1578.
- Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics* **47**, 13-37.
- Merton, R. C., (1973). An intertemporal capital asset pricing model. *Econometrica* **41**, 867-887.
- Pinheiro, J. C. and D. M. B. Bates, D. M. B. (2000). *Mixed-Effects Models in S and S-PLUS*. Springer.
- Ross, S. A. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory* **13**, 341-360.
- Rousseeuw, P. J. (1985). Multivariate estimation with high breakdown point. In *Mathematical Statistics and Applications* (Edited by Grossmann, W. and G. Pflug, G. Vincze, I. and Wertz, W.) 283-297. Reidel.
- Rousseeuw, P. J. and A. M. Leroy, A. M. (1987). *Robust Regression and Outlier Detection*. Wiley.
- Shalit, H. and S. Yitzhaki, S. (2001). Estimating beta (Working Paper), Hebrew University of Jerusalem.
- Shanken, J. (1992). On the estimation of beta-pricing models *Review of Financial Studies* **5**, 1-33.
- Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance* **19**, 425-442.
- Zaman, A. and Rousseeuw, P. J. and Orhan, M. (2001). Econometric applications of high-breakdown robust regression techniques. *Econometrics Letters* **71**, 1-8.

Received April 11, 2004; accepted September 27, 2004.

Tsung-Chi Cheng
Department of Statistics
National Chengchi University
Taipei, Taiwan
chengt@nccu.edu.tw

Hung-Neng Lai
Department of Finance
National Central University
Taoyuan, Taiwan
hnlai@cc.ncu.edu.tw

Chien-Ju Lu
Department of Accounting
National Chengchi University
Taipei, Taiwan
chienru@alumni.nccu.edu.tw