บทความวิจัย

# Using Industry Earnings to Predict Market Earnings

**Dr.Sukanya Rukphanichmanee**
*Lecturer of School of Accountancy,*
*University of the Thai Chamber of Commerce*

**ABSTRACT**

In this study, I aim to improve a predictive ability of aggregate earnings by using industry earnings changes to predict market earnings changes. Using industry earnings to predict, I find some evidences that relaxing the prediction model to allow for more than one coefficients helps increase the predictive ability of aggregate earnings. However, this improvement is sensitive to the industry classification and aggregation scheme. To be specific, industrial prediction (IP) model is more precise than simple regression model in predicting future market earnings changes when industry earnings are classified based on SIC one-digit and value weighted schemes but if I classified earnings based on Fama French (either 10 or 17 industries) and equally weight them, the IP model shows no improvement.

**Keywords:** Aggregate Earnings, Earnings Prediction, Industry Earnings, Market Earnings

บทความวิจัย

# การใช้กำไรของอุตสาหกรรมในการคาดการณ์กำไรของตลาด

**ดร.สุกัญญา รักพานิชมณี**

*อาจารย์ประจำคณะบัญชี*

*มหาวิทยาลัยหอการค้าไทย*

**บทคัดย่อ**

วัตถุประสงค์ของการวิจัยนี้ เพื่อเพิ่มความสามารถในการพยากรณ์ของรายได้รวมของทุกบริษัทในตลาดหลักทรัพย์ของสหรัฐอเมริกา โดยอาศัยรายได้ของทุกบริษัทในแต่ละกลุ่มอุตสาหกรรมในการพยากรณ์ ผลการวิจัยพบว่าการใช้วิธีการแยกรายได้รวมของตลาดเป็นแต่ละอุตสาหกรรมช่วยเพิ่มความสามารถในการพยากรณ์ของรายได้รวมของทุกบริษัทในตลาดหลักทรัพย์ อย่างไรก็ดี ความสามารถในการพยากรณ์ที่เพิ่มขึ้นนี้ ขึ้นอยู่กับวิธีการจัดประเภทกลุ่มอุตสาหกรรมและการถ่วงน้ำหนัก โดยหากจัดกลุ่มอุตสาหกรรมโดยใช้วิธีการของ Fama French การใช้รายได้ของแต่ละอุตสาหกรรมในการพยากรณ์ไม่ช่วยเพิ่มความสามารถในการพยากรณ์ของรายได้รวมของทุกบริษัทในตลาดหลักทรัพย์ แต่ความสามารถในการพยากรณ์ของรายได้รวมนี้จะเพิ่มขึ้นอย่างมีนัยสำคัญหากคำนวณโดยใช้รายได้ของกลุ่มอุตสาหกรรมที่จัดตามหลัก SIC one-digit และถ่วงน้ำหนักด้วยมูลค่าตลาด

**คำสำคัญ:** รายได้รวมของทุกบริษัทในตลาดหลักทรัพย์ การพยากรณ์รายได้ รายได้รวมของอุตสาหกรรม รายได้รวมของตลาด

## 1. Introduction

Earnings prediction has always been in the interest of investors, managers and accounting professionals. The earnings prediction is used for fundamental analysis such as forecasting a company's future performance, determining a firms' stock price, evaluating managers' ability, and making business decisions. Overall, earnings prediction is important to the company and investors in making investment decision. Because of its importance, many studies such as Finger (1994), Foster (1977) and Dichev and Tang (2009) are conducted to test the earnings predictability. While predicting a company's earnings at the firm level to some extent has received attention from accounting researchers, the market level earnings forecast has not been extensively explored. My study is motivated by the fact that investors nowadays become more diversify on their investment portfolio and tend to pay more attention on how an overall market is doing. Constructing a more efficient model to predict future aggregate earnings should help them make better investing decision. Therefore, the objective of this study is to establish a model that is more precise in predicting future aggregate earnings.

## 2. Literature review and hypothesis development

Consistent with Kothari at al. (2006), Cready and Gurun (2010) document that current month aggregate earnings changes which idiosyncratic firms risk is averaged away have a predictive ability to predict future aggregate earnings changes. They use time series regression to regress seasonal monthly aggregate earnings changes on lagged aggregate earnings changes. Their results show that aggregate earnings changes are statistically significant predict future aggregate earnings changes up to the next 4 quarters. This confirms that earnings possess a predictive ability at the aggregate level. Even though Cready and Gurun's results are impressive, they only use a simple regression model to regress aggregate earnings changes on lag earnings changes. This method assumes that all observations across different industries associate with the market earnings at the same level. To improve an ability to predict future aggregate earnings, I relax this assumption by allowing for more than one coefficient.

King (1966) is one among the first who talks about the industry effect. He uses a factor analysis to test his argument and shows that stock prices for firms in the same industry exhibit a common movement that goes beyond the market effect. To be specific, his results show that 50% of stock price movements are explained by movements in the market index and the additional 10–20% of the residual variances is then explained by industry index. Based on King (1966) results and a connection between the stock price and firms earnings, Brown and Ball (1967) conduct a study that test a relationship among earnings of an individual firm, earnings of the other firms in its industry and the earnings of all firms in the market. They use the following earnings variables in their analysis: 1) net income 2) operating income 3) net income + after tax interest expenses 4) adjusted EPS 5) operating income deflated by total assets and 6) net income + after tax interest expense deflated by total assets. Besides the high correlation between earnings from various industries and market earnings, their results also show that earnings from different industries associate with market earnings in the different levels.

Based on King (1966) and Brown and Ball (1967) findings, it may not be reasonable to assign only one coefficient to all observations in the aggregate earning prediction model. A procedure that treats all of the data equally would give less precisely measured points more influence than they should have and would give high precisely measured points less influence than they actually have. In this paper, I will attempt to assign each group of the observations the proper amount of influence over the parameter estimates. I assume that a model that classifies observations based on their industries and assigning them a proper weight based on the predictive power will give me a more precise estimation. This model thereafter is called industrial prediction (IP) model. A more detail about this model will be elaborated below in models section. I hypothesize that IP model will be a better prediction method and my hypothesis can be rephrased below.

*$H_0$: Squared of prediction error from simple prediction model is equal to squared of prediction error from industrial prediction (IP) model.*

*$H_A$: Squared of prediction error from industrial prediction (IP) model is lower than squared of prediction error from simple prediction model.*

## 3. Research Design

Quarterly earnings data from Q1 1985 to Q4 2004 are obtained from Compustat fundamentals and later are used to predict seasonal quarterly earnings changes during 1995 to 2004. Data are formed in the rolling basis and have 40 quarters each length. For example, to predict market earnings changes for Q1 1995, I use earnings data during Q1 1985 to Q4 1994 and then the earnings data from Q2 1985 to Q1 1995 are required to predict Q2 1995 earnings. Seasonal differenced quarterly earnings ($\Delta$E) are defined as a difference between earnings per share this quarter (t) and earnings per share four quarters prior (t–4). Earnings are measured before extraordinary items such as gain (loss) on disposal of a discontinued division, and, to ensure that fiscal quarters are aligned, the sample is restricted to firms with a March, June, September or December fiscal year ends. I scale seasonal differenced quarterly earnings by previous quarter price ($\Delta E_t/P_{t-1}$). This ratio is firm-level ratio which later will be equally weight or value weight to form market earnings changes. The equal-weighted aggregate market changes ($\Delta$E/P-ew) is simply the average of firm-level ratio based on its market values at the beginning of the quarter or lagged market value (MV). These same procedures are also used to form industry seasonal earnings changes. The only difference is that industry earnings changes are grouped based on Standard Industrial Classification (SIC) classification schemes[1]. I decided to define industry based on this classification scheme because among available schemes, SIC is the oldest and most widely used by researchers and economists. Following this criterion, industry earnings changes of 10 industries which are agriculture, mining, constructions, manufacturing, transportation, wholesales trade, retail trade, finance, services and public admin, and other are formed. To avoid the impact of the outliers, I exclude stocks with price below \$1 or above \$10,000 and the top and bottom 1% of firms ranked by $\Delta E_t / P_{t-1}$ before calculating earnings changes.

I use simple regression (equation 1) to determine predictive ability of market earnings. The equation represents prediction model that does not allow for the variation in coefficient among observations from different industries. The regression process is repeatedly done to estimate changes in earnings for Q1 1995 to Q4 2004.

$$\Delta\hat{X}_{mkt,k} = \hat{a}_{0,t} + \hat{a}_{1,t}(\Delta X_{mkt,k-1}) + \varepsilon t \quad \text{for } k = t-1 \text{ to } t-40 \qquad \text{...(1)}$$

Where

$\Delta\hat{X}_{mkt,k}$ = $\Delta E_{mkt,k} / P_{k-1}$

t = period from Q1 1995 to Q4 2004

k = quarter period starting from t–1 to t–40. This indicates that $\hat{a}_{0,t}$ and $\hat{a}_{1,t}$ are an intercept and a coefficient estimated from observations during period t–1 to t–40.

I examine the predictive ability of each industry's earnings by using equation 2. This equation takes into account the different effects that each industry contributes to the market. Regression following this model is also repeatedly done for each industry to get the coefficient estimate for Q1 1995 to Q4 2004.

For each industry,

$$\Delta\hat{X}_{ind\ i,k} = \hat{a}_{0,t} + \hat{a}_{1,t}(\Delta X_{ind\ i,k-1}) + \varepsilon t \quad \text{for } k = t-1 \text{ to } t-40 \qquad \text{...(2)}$$

Where

$\Delta\hat{X}_{ind\ i,k}$ = $\Delta E_{ind\ i,k} / P_{k-1}$

i = industry 1,2,3,...,n

t = period from Q1 1995 to Q4 2004

k = quarter period starting from t–1 to t–40. This indicates that $\hat{a}_{0,t}$, $\hat{a}_{1,t}$ are an intercept and a coefficient estimated from observations during period t–1 to t–40.

The estimated industry seasonal earnings changes deflated by price ($\Delta\hat{X}_{ind\ i,k}$) calculated from equation 2 are then equally weighted or value weighted to form market earnings changes. By using equal-weighted method (EW), sum of estimated industry earnings changes deflated by price ($\Delta\hat{X}_{ind\ i}$) is divided by number of industries.

[1] Industry earnings changes are grouped based on one-digit SIC code but because many studies such as Clarke (1989) and Bhojraj et al (2003) document a limitation of the SIC industry classification scheme, later in the robustness test I will aggregated earnings changes based on Fama French 10 and 17 industries respectively.

Equation 3 below shows how to equal weight the estimated industry seasonal earnings changes.

$$\frac{\sum_{i=1}^{n} \Delta\hat{X}_{ind\ i}}{n} \quad ...(3)$$

Where

$\Delta\hat{X}_{ind\,i}$ = estimated seasonal differenced quarterly earnings deflated by lagged price

i = industry 1,2,3,...,n

n = number of industries

For value-weighted method (VW), estimated earnings changes deflated by price for each industry ($\Delta\hat{X}_{ind\,i}$) calculated from equation 2 are multiplied by the industry market value ($MV_{ind}$) at the beginning quarter which is defined as stock price of all firms in the industry multiply by outstanding shares of all outstanding firms in the industry at the beginning of the quarter. Then they are summarized and divided by total market value ($MV_{mkt,}$) of all firms in the market at the beginning of the quarter. This will yield estimated average VW-industry changes in earnings. Below equation 4 shows how to valued weight estimated industry seasonal earnings change.

$$\sum_{i=1}^{n} \frac{\Delta\hat{X}_{ind\,i} \times MV_{ind\,i}}{MV_{mkt}} \quad ...(4)$$

Where

$\Delta\hat{X}_{ind\,i}$ = estimated changes in earnings of industry i deflated by lagged price

i = industry 1,2,3,...,n

$MV_{ind\,i}$ = total market value (stock price * outstanding shares) of all firms in industry i at the beginning of the quarter

$MV_{mkt}$ = total market value (stock price * outstanding shares) of all firms in the market at the beginning of the quarter

Actual seasonal aggregate earnings changes deflated by lagged price ($\Delta E_k / P_{k-1}$) during Q1 1995 to Q4 2004 are obtained from Computstat and are used to compare with the estimated seasonal earnings changes deflated by lagged price ($\Delta\hat{E}_k / P_{k-1}$) calculated from equation 1, 3 and 4. The differences are prediction errors which later are squared to eliminate the sign effect. The squared prediction error yield from IP model will be compared with squared prediction error from simple regression model (equation 1) by using t-test. If the IP model yields the lower squared prediction error, I will reject the null and conclude that the multi-equation earning prediction model is better than the simple regression model in predicting the future market earnings.

## 4. Results

### 4.1 Summary Statistics

Table 1 reports means, standard deviation, minimum and maximum for market and industry earnings changes. Panel A shows that during Q1 1985 to Q4 2004 period, the mean value of equal-weighted (valued weighted) quarterly market earnings changes is 0.0002750 (–0.0002619) with standard deviation of 0.0043658 (0.0017875) and the mean value of equally-weighted (valued weighted) quarterly industry earnings changes is 0.0007918 (–0.0002261) with standard deviation of 0.0047000 (0.0026795). Table 1 Panel B also provides summary statistics on the same variable

**Table 1** Descriptive Statistics for market earnings changes and industry earnings changes

Panel A and panel B show the descriptive statistics for quarterly market earnings changes and quarterly industry earnings changes during Q1 1985 to Q4 2004 and during Q1 1995 to Q4 2004 respectively. N is a number of quarters. $\Delta E_k$ or quarterly earnings changes are earnings this quarter minus earning four quarters prior. $\Delta E_k / P_{k-1}$ is quarterly earnings changes deflated by lagged price. This ratio is calculated for each firm and then either equally weight or value weight to aggregate. The sample consists of firms with a March, June, September or December fiscal year end. This sample excludes stock with price below \$1 and above \$10,000 and the top and bottom 1% of firms ranked by $\Delta E_k / P_{k-1}$.

| | N | Mean | Std Dev | Min | Max |
|---|---|---|---|---|---|
| **Panel A (1985 to 2004 or 80 quarters)** | | | | | |
| Equal-weighted (EW) market earnings changes deflated by lagged price ($\Delta E_{mkt,k}/P_{k-1}$) | 80 | 0.0002750 | 0.0043658 | –0.0083970 | 0.0084107 |
| Equal-weighted (EW) industry earnings changes deflated by lagged price ($\Delta E_{ind,k}/P_{k-1}$) | 80 | 0.0007918 | 0.0047000 | –0.0086300 | 0.0118736 |
| Value-weighted (VW) market earnings changes deflated by lagged price ($\Delta E_{mkt,k}/P_{k-1}$) | 80 | –0.0002619 | 0.0017875 | –0.0053430 | 0.0074368 |
| Value-weighted (VW) industry earnings changes deflated by lagged price ($\Delta E_{ind,k}/P_{k-1}$) | 80 | –0.0002261 | 0.0026795 | –0.0099540 | 0.0082535 |
| **Panel B (1995 to 2004 or 40 quarters)** | | | | | |
| Equally-weighted (EW) market earnings changes deflated by lagged price ($\Delta E_{mkt,k}/P_{k-1}$) | 40 | 0.0013128 | 0.0045110 | –0.0083970 | 0.0084107 |
| Equal-weighted (EW) industry earnings changes deflated by lagged price ($\Delta E_{ind,k}/P_{k-1}$) | 40 | 0.0019243 | 0.0050004 | –0.0073788 | 0.0118736 |
| Value-weighted (VW) market earnings changes deflated by lagged price ($\Delta E_{mkt,k}/P_{k-1}$) | 40 | –0.0001523 | 0.0024689 | –0.0053432 | 0.0074368 |
| Value-weighted (VW) industry earnings changes deflated by lagged price ($\Delta E_{ind,k}/P_{k-1}$) | 40 | –0.0001279 | 0.0033765 | –0.0099540 | 0.0082535 |

for the shorter time period, Q1 1995 to Q4 2004. It shows that the mean value of equal-weighted (valued weighted) quarterly market earnings changes is 0.0013128 (–0.0001523) with standard deviation of 0.0045110 (0.0024689) and the mean of equal-weighted (valued-weighted) quarterly industry earnings changes is 0.0019243 (–0.0001279) with standard deviation of 0.0050004 (0.0033765). Compared to KLW (2006) results, the mean value of my data is smaller because I leave out the high inflation period (1970s).

### 4.2 Equal-weighted aggregate earnings changes and one-digit SIC code

In this section industry seasonal earnings changes are aggregated based on one-digit SIC code. Table 2 reports the squared prediction error calculated from the simple regression (equation 1) and industrial prediction (IP) model (equation 2), A prediction error is the difference between an actual seasonal market earnings changes deflated by lagged price ($\Delta E_k/P_{k-1}$) and an estimated seasonal market earnings changes deflated by

**Table 2** Summary of squared prediction errors from equal-weighted simple regression and industrial prediction (IP) model and the differences of squared prediction error from these two models. Sample used in industrial prediction (IP) model is aggregated based on one-digit SIC code.

| Prediction Quarter | Squared prediction error yield from simple regression | Squared prediction error yield from EW industrial prediction model | Difference in squared prediction error |
|---|---|---|---|
| 1995Q1 | 0.0000464 | 0.0000080 | 0.0000384 |
| 1995Q2 | 0.0000015 | 0.0000093 | (0.0000078) |
| 1995Q3 | 0.0000004 | 0.0000013 | (0.0000010) |
| 1995Q4 | 0.0000019 | 0.0000017 | 0.0000001 |
| 1996Q1 | 0.0000018 | 0.0000142 | (0.0000124) |
| 1996Q2 | 0.0000009 | 0.0000034 | (0.0000025) |
| 1996Q3 | 0.0000035 | 0.0000007 | 0.0000029 |
| 1996Q4 | 0.0000005 | 0.0000053 | (0.0000048) |
| 1997Q1 | 0.0000001 | 0.0000091 | (0.0000090) |
| 1997Q2 | 0.0000005 | 0.0000008 | (0.0000003) |
| 1997Q3 | 0.0000001 | 0.0000016 | (0.0000015) |
| 1997Q4 | 0.0000049 | 0.0000009 | 0.0000040 |
| 1998Q1 | 0.0000002 | 0.0000005 | (0.0000003) |
| 1998Q2 | 0.0000087 | 0.0000000 | 0.0000087 |
| 1998Q3 | 0.0000119 | 0.0000049 | 0.0000070 |
| 1998Q4 | 0.0000211 | 0.0000053 | 0.0000159 |

**Table 2** Summary of squared prediction errors from equal-weighted simple regression and industrial prediction (IP) model and the differences of squared prediction error from these two models. Sample used in industrial prediction (IP) model is aggregated based on one-digit SIC code. (Cont.)

| Prediction Quarter | Squared prediction error yield from simple regression | Squared prediction error yield from EW industrial prediction model | Difference in squared prediction error |
|---|---|---|---|
| 1999Q1 | 0.0000035 | 0.0000045 | (0.0000010) |
| 1999Q2 | 0.0000023 | 0.0000043 | (0.0000019) |
| 1999Q3 | 0.0000047 | 0.0000239 | (0.0000192) |
| 1999Q4 | 0.0000000 | 0.0000015 | (0.0000015) |
| 2000Q1 | 0.0000096 | 0.0000121 | (0.0000025) |
| 2000Q2 | 0.0000032 | 0.0000020 | 0.0000012 |
| 2000Q3 | 0.0000001 | 0.0000012 | (0.0000011) |
| 2000Q4 | 0.0000261 | 0.0000572 | (0.0000311) |
| 2001Q1 | 0.0000046 | 0.0000287 | (0.0000241) |
| 2001Q2 | 0.0000014 | 0.0000009 | 0.0000005 |
| 2001Q3 | 0.0000230 | 0.0000490 | (0.0000260) |
| 2001Q4 | 0.0000178 | 0.0000018 | 0.0000160 |
| 2002Q1 | 0.0000612 | 0.0001248 | (0.0000636) |
| 2002Q2 | 0.0000037 | 0.0000668 | (0.0000631) |
| 2002Q3 | 0.0000219 | 0.0000501 | (0.0000281) |
| 2002Q4 | 0.0000230 | 0.0000656 | (0.0000426) |
| 2003Q1 | 0.0000004 | 0.0000118 | (0.0000113) |
| 2003Q2 | 0.0000121 | 0.0000320 | (0.0000199) |
| 2003Q3 | 0.0000085 | 0.0000002 | 0.0000083 |
| 2003Q4 | 0.0000083 | 0.0000092 | (0.0000009) |
| 2004Q1 | 0.0000010 | 0.0000042 | (0.0000032) |
| 2004Q2 | 0.0000009 | 0.0000037 | (0.0000028) |
| 2004Q3 | 0.0000036 | 0.0000380 | (0.0000343) |
| 2004Q4 | 0.0000147 | 0.0000330 | (0.0000183) |

T-Stat[2]: 1.81

[2] The t-test assesses whether the means of two groups are statistically different from each other.

lagged price calculated ($\Delta\hat{E}_k/P_{k-1}$). The results show that only 11 out 40 sample period quarters that IP model is more accurate than simple regression model in predicting the future market earnings. Moreover, using T-test to compare squared prediction error from these models, I find that the squared prediction error from simple regression model is statistically significantly less than the squared prediction error yield from IP model with the t-stat = 1.81. This result suggests that giving a similar weight to each industry earnings does not improve the predictive ability of the aggregate earnings. I suspect that the earnings of larger industry may have more predictive power than the earnings of the small industry. Next section will investigate this assumption.

### 4.3 Value-weighted aggregate earnings changes and one-digit SIC code

Similar to section 4.1, the sample used in this section is aggregated based on one-digit SIC code. In this section, however, the sample is value-weighted. The main difference between equally-weighted and value-weighted method is that value-weighted method assigns more weight to the larger industry. The results show that for 23 out of 40 quarters, the squared prediction error calculated from the industrial prediction

**Table 3** Summary of squared prediction errors from value-weighted simple regression and industrial prediction (IP) model and the differences of squared prediction error from these two models. Sample used in industrial prediction (IP) model is aggregated based on one-digit SIC code (10 industries).

| Prediction Quarter | Squared prediction error yield from simple regression | Squared prediction error yield from VW industrial prediction model | Difference in squared prediction error |
|---|---|---|---|
| 1995Q1 | 0.0000002 | 0.0000007 | (0.0000006) |
| 1995Q2 | 0.0000002 | 0.0000018 | (0.0000016) |
| 1995Q3 | 0.0000002 | 0.0000001 | 0.0000001 |
| 1995Q4 | 0.0000070 | 0.0000000 | 0.0000070 |
| 1996Q1 | 0.0000007 | 0.0000005 | 0.0000002 |
| 1996Q2 | 0.0000001 | 0.0000009 | (0.0000007) |
| 1996Q3 | 0.0000006 | 0.0000001 | 0.0000005 |
| 1996Q4 | 0.0000131 | 0.0000024 | 0.0000107 |
| 1997Q1 | 0.0000087 | 0.0000000 | 0.0000087 |
| 1997Q2 | 0.0000102 | 0.0000000 | 0.0000102 |
| 1997Q3 | 0.0000022 | 0.0000000 | 0.0000022 |
| 1997Q4 | 0.0000014 | 0.0000002 | 0.0000013 |

**Table 3** Summary of squared prediction errors from value-weighted simple regression and industrial prediction (IP) model and the differences of squared prediction error from these two models. Sample used in industrial prediction (IP) model is aggregated based on one-digit SIC code (10 industries). (Cont.)

| Prediction Quarter | Squared prediction error yield from simple regression | Squared prediction error yield from VW industrial prediction model | Difference in squared prediction error |
|---|---|---|---|
| 1998Q1 | 0.0000014 | 0.0000015 | (0.0000001) |
| 1998Q2 | 0.0000025 | 0.0000008 | 0.0000017 |
| 1998Q3 | 0.0000009 | 0.0000009 | (0.0000001) |
| 1998Q4 | 0.0000016 | 0.0000030 | (0.0000014) |
| 1999Q1 | 0.0000087 | 0.0000006 | 0.0000081 |
| 1999Q2 | 0.0000016 | 0.0000006 | 0.0000010 |
| 1999Q3 | 0.0000005 | 0.0000016 | (0.0000011) |
| 1999Q4 | 0.0000008 | 0.0000018 | (0.0000010) |
| 2000Q1 | 0.0000012 | 0.0000002 | 0.0000009 |
| 2000Q2 | 0.0000003 | 0.0000001 | 0.0000002 |
| 2000Q3 | 0.0000000 | 0.0000001 | (0.0000001) |
| 2000Q4 | 0.0000037 | 0.0000011 | 0.0000026 |
| 2001Q1 | 0.0000001 | 0.0000016 | (0.0000015) |
| 2001Q2 | 0.0000031 | 0.0000109 | (0.0000078) |
| 2001Q3 | 0.0000003 | 0.0000031 | (0.0000028) |
| 2001Q4 | 0.0000113 | 0.0000009 | 0.0000103 |
| 2002Q1 | 0.0000458 | 0.0000072 | 0.0000386 |
| 2002Q2 | 0.0000001 | 0.0000007 | (0.0000006) |
| 2002Q3 | 0.0000107 | 0.0000073 | 0.0000034 |
| 2002Q4 | 0.0000072 | 0.0000001 | 0.0000071 |
| 2003Q1 | 0.0000090 | 0.0000012 | 0.0000078 |
| 2003Q2 | 0.0000008 | 0.0000038 | (0.0000031) |
| 2003Q3 | 0.0000000 | 0.0000071 | (0.0000071) |
| 2003Q4 | 0.0000344 | 0.0000365 | (0.0000022) |
| 2004Q1 | 0.0000508 | 0.0000031 | 0.0000478 |
| 2004Q2 | 0.0000022 | 0.0000141 | (0.0000119) |
| 2004Q3 | 0.0000041 | 0.0000006 | 0.0000035 |
| 2004Q4 | 0.0000152 | 0.0000038 | 0.0000114 |

T-Stat: –[illegible].69

(IP) model is less than that is calculated from simple regression model. With T-statistic equals to –1.69, the results supports my hypothesis that IP model is superior to simple regression model as the squared prediction error from this model is statistically significantly less than the squared prediction error yield from simple regression model.

## 5. Robustness test

My assumption for the robustness test is that the industry classification schemes and a number of industries retained in the aggregation process influence a predictive ability of earnings changes. In this section, therefore, I use an alternative classification scheme to aggregate industry earnings changes. I decided to use Fama French industry classification scheme (both 10 and 17 industries) because Bhojraj et al (2003) document a high degree of correspondence among SIC, NAICS, and Fama French classifications. To be specific, 84% of the firms grouping by Fama French algorithms are agreed with two-digit SIC grouping. Moreover, in this section I will also use the alternative regression model to test whether my results depend on the regression method. In summary, in this section I will test the sensitivity of my results to an industry classification scheme, number of industries and regression model.

### 5.1 Equal-weighted aggregate earnings changes and 10 industries Fama French industry classification scheme

The industry earnings changes in this section are grouped based on Fama French 10 industries and then later is equally weighted to form market earnings changes. The results in table 4 show that industrial prediction (IP) model is not superior to the simple regression model in predicting future aggregate earnings changes. Only 15 out of 40 quarters that IP model yields less prediction errors than simple regression model. Moreover, using T-test to compare squared prediction error from these models, I find that they are not statistically significantly different from each other (T-Statistic –0.89).

**Table 4** Summary of squared prediction errors from equal-weighted simple regression (EW) and industrial prediction (IP) model and the differences of squared prediction error from these two models. Sample used in industrial prediction (IP) model is aggregated based on Fama French industry classification scheme (10 industries).

| Prediction Quarter | Squared prediction error yield from simple regression | Squared prediction error yield from EW industrial prediction model | Difference in squared prediction error |
|---|---|---|---|
| 1995Q1 | 0.0000464 | 0.0000025 | 0.0000439 |
| 1995Q2 | 0.0000015 | 0.0000010 | 0.0000005 |
| 1995Q3 | 0.0000004 | 0.0000020 | (0.0000016) |
| 1995Q4 | 0.0000019 | 0.0000034 | (0.0000016) |
| 1996Q1 | 0.0000018 | 0.0000057 | (0.0000039) |
| 1996Q2 | 0.0000009 | 0.0000033 | (0.0000023) |
| 1996Q3 | 0.0000035 | 0.0000000 | 0.0000035 |
| 1996Q4 | 0.0000005 | 0.0000002 | 0.0000004 |
| 1997Q1 | 0.0000001 | 0.0000005 | (0.0000004) |
| 1997Q2 | 0.0000005 | 0.0000006 | (0.0000002) |
| 1997Q3 | 0.0000001 | 0.0000004 | (0.0000003) |
| 1997Q4 | 0.0000049 | 0.0000102 | (0.0000053) |
| 1998Q1 | 0.0000002 | 0.0000000 | 0.0000002 |
| 1998Q2 | 0.0000087 | 0.0000055 | 0.0000032 |
| 1998Q3 | 0.0000119 | 0.0000130 | (0.0000011) |
| 1998Q4 | 0.0000211 | 0.0000404 | (0.0000193) |
| 1999Q1 | 0.0000035 | 0.0000157 | (0.0000122) |
| 1999Q2 | 0.0000023 | 0.0000027 | (0.0000003) |
| 1999Q3 | 0.0000047 | 0.0000115 | (0.0000069) |
| 1999Q4 | 0.0000000 | 0.0000032 | (0.0000032) |
| 2000Q1 | 0.0000096 | 0.0000061 | 0.0000035 |
| 2000Q2 | 0.0000032 | 0.0000005 | 0.0000027 |
| 2000Q3 | 0.0000001 | 0.0000001 | 0.0000000 |
| 2000Q4 | 0.0000261 | 0.0000171 | 0.0000090 |
| 2001Q1 | 0.0000046 | 0.0000109 | (0.0000063) |
| 2001Q2 | 0.0000014 | 0.0000095 | (0.0000081) |
| 2001Q3 | 0.0000230 | 0.0000374 | (0.0000144) |
| 2001Q4 | 0.0000178 | 0.0000647 | (0.0000469) |
| 2002Q1 | 0.0000612 | 0.0000853 | (0.0000241) |
| 2002Q2 | 0.0000037 | 0.0000069 | (0.0000031) |

**Table 4** Summary of squared prediction errors from equal-weighted simple regression (Eq1) and industrial prediction (IP) model and the differences of squared prediction error from these two models. Sample used in industrial prediction (IP) model is aggregated based on Fama French industry classification scheme (10 industries). (Cont.)

| Prediction Quarter | Squared prediction error yield from simple regression | Squared prediction error yield from EW industrial prediction model | Difference in squared prediction error |
|---|---|---|---|
| 2002Q3 | 0.0000219 | 0.0000353 | (0.0000134) |
| 2002Q4 | 0.0000230 | 0.0000435 | (0.0000204) |
| 2003Q1 | 0.0000004 | 0.0000001 | 0.0000003 |
| 2003Q2 | 0.0000121 | 0.0000117 | 0.0000005 |
| 2003Q3 | 0.0000085 | 0.0000040 | 0.0000045 |
| 2003Q4 | 0.0000083 | 0.0000079 | 0.0000004 |
| 2004Q1 | 0.0000010 | 0.0000059 | (0.0000049) |
| 2004Q2 | 0.0000009 | 0.0000014 | (0.0000005) |
| 2004Q3 | 0.0000036 | 0.0000129 | (0.0000093) |
| 2004Q4 | 0.0000147 | 0.0000054 | 0.0000093 |

T-Stat: 0.89

### 5.2 Value-weighted aggregate earnings changes and 10 industries Fama French industry classification scheme

Similar to section 5.1, the industry earnings changes in this section are grouped based on Fama French 10 industries, however, unlike the previous section, I value weight them to form market earnings changes. The results in table 5 show that, 23 out of 40 quarters, industrial prediction (IP) model is more precise in predicting future market earnings changes. However, using T-test to compare squared prediction errors yield from these two models, I find that they are not statistically significantly different from each other (T-Statistic = 1.63).

**Table 5** Summary of squared prediction errors from value-weighted simple regression (Eq1) and industrial prediction (IP) model and the differences of squared prediction error from these two models. Sample used in industrial prediction (IP) model is aggregated based on Fama French industry classification scheme (10 industries).

| Prediction Quarter | Squared prediction error yield from simple regression | Squared prediction error yield from VW industrial prediction model | Difference in squared prediction error |
|---|---|---|---|
| 1995Q1 | 0.0000002 | 0.0000010 | (0.0000008) |
| 1995Q2 | 0.0000002 | 0.0000008 | (0.0000005) |
| 1995Q3 | 0.0000002 | 0.0000000 | 0.0000002 |
| 1995Q4 | 0.0000070 | 0.0000002 | 0.0000068 |
| 1996Q1 | 0.0000007 | 0.0000002 | 0.0000005 |
| 1996Q2 | 0.0000001 | 0.0000011 | (0.0000010) |
| 1996Q3 | 0.0000006 | 0.0000001 | 0.0000005 |
| 1996Q4 | 0.0000131 | 0.0000024 | 0.0000107 |
| 1997Q1 | 0.0000087 | 0.0000001 | 0.0000086 |
| 1997Q2 | 0.0000102 | 0.0000000 | 0.0000101 |
| 1997Q3 | 0.0000022 | 0.0000001 | 0.0000022 |
| 1997Q4 | 0.0000014 | 0.0000004 | 0.0000011 |
| 1998Q1 | 0.0000014 | 0.0000018 | (0.0000004) |
| 1998Q2 | 0.0000025 | 0.0000005 | 0.0000019 |
| 1998Q3 | 0.0000009 | 0.0000011 | (0.0000002) |
| 1998Q4 | 0.0000016 | 0.0000040 | (0.0000024) |
| 1999Q1 | 0.0000087 | 0.0000002 | 0.0000085 |
| 1999Q2 | 0.0000016 | 0.0000001 | 0.0000015 |
| 1999Q3 | 0.0000005 | 0.0000006 | (0.0000001) |
| 1999Q4 | 0.0000008 | 0.0000013 | (0.0000005) |
| 2000Q1 | 0.0000012 | 0.0000001 | 0.0000011 |
| 2000Q2 | 0.0000003 | 0.0000001 | 0.0000002 |
| 2000Q3 | 0.0000000 | 0.0000001 | (0.0000001) |
| 2000Q4 | 0.0000037 | 0.0000013 | 0.0000024 |
| 2001Q1 | 0.0000001 | 0.0000022 | (0.0000021) |
| 2001Q2 | 0.0000031 | 0.0000118 | (0.0000088) |
| 2001Q3 | 0.0000003 | 0.0000062 | (0.0000059) |
| 2001Q4 | 0.0000113 | 0.0000019 | 0.0000094 |
| 2002Q1 | 0.0000458 | 0.0000065 | 0.0000393 |
| 2002Q2 | 0.0000001 | 0.0000009 | (0.0000008) |

**Table 5** Summary of squared prediction errors from value-weighted simple regression (Eq1) and industrial prediction (IP) model and the differences of squared prediction error from these two models. Sample used in industrial prediction (IP) model is aggregated based on Fama French industry classification scheme (10 industries). (Cont.)

| Prediction Quarter | Squared prediction error yield from simple regression | Squared prediction error yield from VW industrial prediction model | Difference in squared prediction error |
|---|---|---|---|
| 2002Q3 | 0.0000107 | 0.0000058 | 0.0000049 |
| 2002Q4 | 0.0000072 | 0.0000003 | 0.0000069 |
| 2003Q1 | 0.0000090 | 0.0000011 | 0.0000079 |
| 2003Q2 | 0.0000008 | 0.0000039 | (0.0000031) |
| 2003Q3 | 0.0000000 | 0.0000089 | (0.0000089) |
| 2003Q4 | 0.0000344 | 0.0000375 | (0.0000031) |
| 2004Q1 | 0.0000508 | 0.0000020 | 0.0000488 |
| 2004Q2 | 0.0000022 | 0.0000148 | (0.0000126) |
| 2004Q3 | 0.0000041 | 0.0000002 | 0.0000040 |
| 2004Q4 | 0.0000152 | 0.0000039 | 0.0000113 |

T-Stat: 1.63

### 5.3 Equal-weighted aggregate earnings changes and 17 industries Fama French industry classification scheme

As another robustness check, sample in Table 6 is weighted equally based on 17 industries Fama French industry classification scheme. The results show that the industrial prediction (IP) model is not superior to the simple regression model in predicting future aggregate earnings changes. Only 13 out of 40 quarters that IP model yields lower squared prediction errors. Moreover, results from T-Test (T-Statistic = 1.31) suggests that squared prediction errors calculated from both models are not statistic significantly different from each other.

**Table 6** Summary of squared prediction errors from equal-weighted simple regression and industrial prediction (IP) model and the differences of squared prediction error from these two models. Sample used in industrial prediction (IP) model is aggregated based on Fama French industry classification scheme (17 industries).

| Prediction Quarter | Squared prediction error yield from simple regression | Squared prediction error yield from EW industrial prediction model | Difference in squared prediction error |
|---|---|---|---|
| 1995Q1 | 0.0000464 | 0.0000128 | 0.0000336 |
| 1995Q2 | 0.0000015 | 0.0000059 | (0.0000044) |
| 1995Q3 | 0.0000004 | 0.0000053 | (0.0000050) |
| 1995Q4 | 0.0000019 | 0.0000180 | (0.0000162) |
| 1996Q1 | 0.0000018 | 0.0000031 | (0.0000013) |
| 1996Q2 | 0.0000009 | 0.0000032 | (0.0000023) |
| 1996Q3 | 0.0000035 | 0.0000050 | (0.0000015) |
| 1996Q4 | 0.0000005 | 0.0000004 | 0.0000001 |
| 1997Q1 | 0.0000001 | 0.0000001 | (0.0000001) |
| 1997Q2 | 0.0000005 | 0.0000005 | (0.0000000) |
| 1997Q3 | 0.0000001 | 0.0000005 | (0.0000004) |
| 1997Q4 | 0.0000049 | 0.0000038 | 0.0000011 |
| 1998Q1 | 0.0000002 | 0.0000016 | (0.0000015) |
| 1998Q2 | 0.0000087 | 0.0000056 | 0.0000031 |
| 1998Q3 | 0.0000119 | 0.0000084 | 0.0000035 |
| 1998Q4 | 0.0000211 | 0.0000273 | (0.0000062) |
| 1999Q1 | 0.0000035 | 0.0000014 | 0.0000021 |
| 1999Q2 | 0.0000023 | 0.0000001 | 0.0000023 |
| 1999Q3 | 0.0000047 | 0.0000014 | 0.0000033 |
| 1999Q4 | 0.0000000 | 0.0000006 | (0.0000006) |
| 2000Q1 | 0.0000096 | 0.0000037 | 0.0000059 |
| 2000Q2 | 0.0000032 | 0.0000001 | 0.0000031 |
| 2000Q3 | 0.0000001 | 0.0000015 | (0.0000014) |
| 2000Q4 | 0.0000261 | 0.0000243 | 0.0000018 |
| 2001Q1 | 0.0000046 | 0.0000409 | (0.0000362) |
| 2001Q2 | 0.0000014 | 0.0000230 | (0.0000216) |
| 2001Q3 | 0.0000230 | 0.0000274 | (0.0000044) |
| 2001Q4 | 0.0000178 | 0.0000499 | (0.0000321) |
| 2002Q1 | 0.0000612 | 0.0000974 | (0.0000362) |
| 2002Q2 | 0.0000037 | 0.0000262 | (0.0000225) |

**Table 6** Summary of squared prediction errors from equal-weighted simple regression and industrial prediction (IP) model and the differences of squared prediction error from these two models. Sample used in industrial prediction (IP) model is aggregated based on Fama French industry classification scheme (17 industries). (Cont.)

| Prediction Quarter | Squared prediction error yield from simple regression | Squared prediction error yield from EW industrial prediction model | Difference in squared prediction error |
|---|---|---|---|
| 2002Q3 | 0.0000219 | 0.0000256 | (0.0000036) |
| 2002Q4 | 0.0000230 | 0.0000294 | (0.0000064) |
| 2003Q1 | 0.0000004 | 0.0000018 | (0.0000014) |
| 2003Q2 | 0.0000121 | 0.0000001 | 0.0000120 |
| 2003Q3 | 0.0000085 | 0.0000037 | 0.0000048 |
| 2003Q4 | 0.0000083 | 0.0000120 | (0.0000037) |
| 2004Q1 | 0.0000010 | 0.0000118 | (0.0000108) |
| 2004Q2 | 0.0000009 | 0.0000095 | (0.0000086) |
| 2004Q3 | 0.0000036 | 0.0000242 | (0.0000206) |
| 2004Q4 | 0.0000147 | 0.0000313 | (0.0000166) |

T-Stat: 1.31

### 5.4 Value-weighted aggregate earnings changes and 17 industries Fama French industry classification scheme

Even though Table 7 shows that 24 out of 40 quarters the squared prediction error calculated from value-weighted industrial prediction (IP) model are lower than the squared prediction error calculated from simple regression model, they are not statistically significantly different from each other (T-Statistic = 1.59). The results suggest that by aggregating industry earnings changes based on 17 industries Fama French algorithm value-weighted IP model is not superior to the value-weighted simple regression model in predicting future aggregate earnings changes.

**Table 7** Summary of squared prediction errors from value-weighted simple regression and industrial prediction (IP) model and the differences of squared prediction error from these two models. Sample used in industrial prediction (IP) model is aggregated based on Fama French industry classification scheme (17 industries).

| Prediction Quarter | Squared prediction error yield from simple regression | Squared prediction error yield from VW industrial prediction model | Difference in squared prediction error |
|---|---|---|---|
| 1995Q1 | 0.0000002 | 0.0000007 | (0.0000005) |
| 1995Q2 | 0.0000002 | 0.0000016 | (0.0000013) |
| 1995Q3 | 0.0000002 | 0.0000002 | 0.0000001 |
| 1995Q4 | 0.0000070 | 0.0000000 | 0.0000070 |
| 1996Q1 | 0.0000007 | 0.0000004 | 0.0000003 |
| 1996Q2 | 0.0000001 | 0.0000018 | (0.0000017) |
| 1996Q3 | 0.0000006 | 0.0000000 | 0.0000006 |
| 1996Q4 | 0.0000131 | 0.0000013 | 0.0000118 |
| 1997Q1 | 0.0000087 | 0.0000000 | 0.0000087 |
| 1997Q2 | 0.0000102 | 0.0000001 | 0.0000101 |
| 1997Q3 | 0.0000022 | 0.0000002 | 0.0000021 |
| 1997Q4 | 0.0000014 | 0.0000006 | 0.0000009 |
| 1998Q1 | 0.0000014 | 0.0000030 | (0.0000016) |
| 1998Q2 | 0.0000025 | 0.0000002 | 0.0000022 |
| 1998Q3 | 0.0000009 | 0.0000009 | 0.0000000 |
| 1998Q4 | 0.0000016 | 0.0000041 | (0.0000025) |
| 1999Q1 | 0.0000087 | 0.0000004 | 0.0000083 |
| 1999Q2 | 0.0000016 | 0.0000006 | 0.0000010 |
| 1999Q3 | 0.0000005 | 0.0000011 | (0.0000005) |
| 1999Q4 | 0.0000008 | 0.0000016 | (0.0000008) |
| 2000Q1 | 0.0000012 | 0.0000001 | 0.0000010 |
| 2000Q2 | 0.0000003 | 0.0000001 | 0.0000001 |
| 2000Q3 | 0.0000000 | 0.0000000 | 0.0000000 |
| 2000Q4 | 0.0000037 | 0.0000013 | 0.0000024 |
| 2001Q1 | 0.0000001 | 0.0000033 | (0.0000032) |
| 2001Q2 | 0.0000031 | 0.0000114 | (0.0000083) |
| 2001Q3 | 0.0000003 | 0.0000016 | (0.0000013) |
| 2001Q4 | 0.0000113 | 0.0000017 | 0.0000095 |
| 2002Q1 | 0.0000458 | 0.0000083 | 0.0000375 |
| 2002Q2 | 0.0000001 | 0.0000007 | (0.0000006) |

**Table 7** Summary of squared prediction errors from value-weighted simple regression and industrial prediction (IP) model and the differences of squared prediction error from these two models. Sample used in industrial prediction (IP) model is aggregated based on Fama French industry classification scheme (17 industries). (Cont.)

| Prediction Quarter | Squared prediction error yield from simple regression | Squared prediction error yield from VW industrial prediction model | Difference in squared prediction error |
|---|---|---|---|
| 2002Q3 | 0.0000107 | 0.0000100 | 0.0000007 |
| 2002Q4 | 0.0000072 | 0.0000001 | 0.0000071 |
| 2003Q1 | 0.0000090 | 0.0000036 | 0.0000054 |
| 2003Q2 | 0.0000008 | 0.0000024 | (0.0000016) |
| 2003Q3 | 0.0000000 | 0.0000006 | (0.0000006) |
| 2003Q4 | 0.0000344 | 0.0000346 | (0.0000002) |
| 2004Q1 | 0.0000508 | 0.0000110 | 0.0000398 |
| 2004Q2 | 0.0000022 | 0.0000157 | (0.0000134) |
| 2004Q3 | 0.0000041 | 0.0000007 | 0.0000034 |
| 2004Q4 | 0.0000152 | 0.0000037 | 0.0000115 |

T-Stat: 1.59

### 5.5 Equal-weighted aggregate earnings changes, 10 industries Fama French industry classification scheme and multiple regression model

Another possible regression model that can be used is multiple regression model which regress future market earnings changes on each industry current period earnings changes. In this case, each industry earnings changes is treated as an independent variable. The equation can be written as below;

$$\begin{aligned}\Delta\hat{X}_{mkt,k} = {} & \hat{a}_{0,t} + \hat{a}_{1,t}(\Delta X_{ind\ 1,\ k-1}) + \hat{a}_{2,t}(\Delta X_{ind\ 2,\ k-1}) \\ & + \hat{a}_{3,t}(\Delta X_{ind\ 3,\ k-1}) + \hat{a}_{4,t}(\Delta X_{ind\ 4,\ k-1}) \\ & + \hat{a}_{5,t}(\Delta X_{ind\ 5,\ k-1}) + \hat{a}_{6,t}(\Delta X_{ind\ 6,\ k-1}) \\ & + \hat{a}_{7,t}(\Delta X_{ind\ 7,\ k-1}) + \hat{a}_{8,t}(\Delta X_{ind\ 8,\ k-1}) \\ & + \hat{a}_{9,t}(\Delta X_{ind\ 9,\ k-1}) + \hat{a}_{10,t}(\Delta X_{ind\ 10,\ k-1}) \\ & + \varepsilon_t \qquad \ldots(5)\end{aligned}$$

In this section, 10 industries Fama French industry classification scheme is used to equal weight earnings changes among all firms in the same industry. In table 8 results show that 39 out of 40 quarters, the squared prediction error from simple regression is less than squared prediction error yield from multiple regression model. T-Statistic equals –3.93 also suggests that simple regression model is superior to multiple regression model.

**Table 8** Summary of squared prediction errors from equal-weighted simple regression and multiple regression model and the differences of squared prediction error from these two models. Sample used in multiple regression model is aggregated based on Fama French industry classification scheme (10 industries).

| Prediction Quarter | Squared prediction error yield from simple regression | Squared prediction error yield from EW multiple regression model | Difference in squared prediction error |
|---|---|---|---|
| 1995Q1 | 0.0000464 | 0.0007257 | (0.0006793) |
| 1995Q2 | 0.0000015 | 0.0003190 | (0.0003175) |
| 1995Q3 | 0.0000004 | 0.0004518 | (0.0004514) |
| 1995Q4 | 0.0000019 | 0.0003432 | (0.0003414) |
| 1996Q1 | 0.0000018 | 0.0006162 | (0.0006144) |
| 1996Q2 | 0.0000009 | 0.0012994 | (0.0012984) |
| 1996Q3 | 0.0000035 | 0.0000004 | 0.0000031 |
| 1996Q4 | 0.0000005 | 0.0001840 | (0.0001834) |
| 1997Q1 | 0.0000001 | 0.0000633 | (0.0000633) |
| 1997Q2 | 0.0000005 | 0.0001806 | (0.0001802) |
| 1997Q3 | 0.0000001 | 0.0000300 | (0.0000300) |
| 1997Q4 | 0.0000049 | 0.0015067 | (0.0015018) |
| 1998Q1 | 0.0000002 | 0.0010715 | (0.0010713) |
| 1998Q2 | 0.0000087 | 0.0015677 | (0.0015590) |
| 1998Q3 | 0.0000119 | 0.0037652 | (0.0037533) |
| 1998Q4 | 0.0000211 | 0.0054407 | (0.0054196) |
| 1999Q1 | 0.0000035 | 0.0030853 | (0.0030819) |
| 1999Q2 | 0.0000023 | 0.0001231 | (0.0001208) |
| 1999Q3 | 0.0000047 | 0.0005668 | (0.0005621) |
| 1999Q4 | 0.0000000 | 0.0003968 | (0.0003968) |
| 2000Q1 | 0.0000096 | 0.0008445 | (0.0008349) |
| 2000Q2 | 0.0000032 | 0.0000177 | (0.0000145) |
| 2000Q3 | 0.0000001 | 0.0001331 | (0.0001330) |
| 2000Q4 | 0.0000261 | 0.0004686 | (0.0004424) |
| 2001Q1 | 0.0000046 | 0.0000241 | (0.0000195) |
| 2001Q2 | 0.0000014 | 0.0000046 | (0.0000032) |
| 2001Q3 | 0.0000230 | 0.0005813 | (0.0005583) |
| 2001Q4 | 0.0000178 | 0.0027884 | (0.0027705) |
| 2002Q1 | 0.0000612 | 0.0111724 | (0.0111112) |
| 2002Q2 | 0.0000037 | 0.0001917 | (0.0001879) |

**Table 8** Summary of squared prediction errors from equal-weighted simple regression and multiple regression model and the differences of squared prediction error from these two models. Sample used in multiple regression model is aggregated based on Fama French industry classification scheme (10 industries). (Cont.)

| Prediction Quarter | Squared prediction error yield from simple regression | Squared prediction error yield from EW multiple regression model | Difference in squared prediction error |
|---|---|---|---|
| 2002Q3 | 0.0000219 | 0.0000553 | (0.0000334) |
| 2002Q4 | 0.0000230 | 0.0002831 | (0.0002600) |
| 2003Q1 | 0.0000004 | 0.0008403 | (0.0008399) |
| 2003Q2 | 0.0000121 | 0.0034572 | (0.0034451) |
| 2003Q3 | 0.0000085 | 0.0007474 | (0.0007390) |
| 2003Q4 | 0.0000083 | 0.0001607 | (0.0001524) |
| 2004Q1 | 0.0000010 | 0.0009458 | (0.0009448) |
| 2004Q2 | 0.0000009 | 0.0006380 | (0.0006371) |
| 2004Q3 | 0.0000036 | 0.0029887 | (0.0029851) |
| 2004Q4 | 0.0000147 | 0.0053700 | (0.0053554) |

T-Stat: –3.93

**5.6 Value-weighted aggregate earnings changes, 10 industries Fama French industry classification scheme and multiple regression model**

Similar to section 5.5, multiple regression and Fama French industry classification scheme are used. However, the sample in this section will be value weighted. The results in table 9 show that 39 out of 40 quarters, the squared prediction error from simple regression model is less than squared prediction error yield from multiple regression model and the differences are statistically significantly different from each other with the t statistic equal to –4.06.

**Table 9** Summary of squared prediction errors from value-weighted simple regression and multiple regression model and the differences of squared prediction error from these two models. Sample used in multiple regression model is aggregated based on Fama French industry classification scheme (10 industries).

| Prediction Quarter | Squared prediction error yield from simple regression | Squared prediction error yield from VW multiple regression model | Difference in squared prediction error |
|---|---|---|---|
| 1995Q1 | 0.0000002 | 0.0003893 | (0.0003891) |
| 1995Q2 | 0.0000002 | 0.0001067 | (0.0001064) |
| 1995Q3 | 0.0000002 | 0.0002333 | (0.0002331) |
| 1995Q4 | 0.0000070 | 0.0006762 | (0.0006692) |
| 1996Q1 | 0.0000007 | 0.0001251 | (0.0001244) |
| 1996Q2 | 0.0000001 | 0.0007967 | (0.0007965) |
| 1996Q3 | 0.0000006 | 0.0000372 | (0.0000366) |
| 1996Q4 | 0.0000131 | 0.0013247 | (0.0013117) |
| 1997Q1 | 0.0000087 | 0.0000064 | 0.0000023 |
| 1997Q2 | 0.0000102 | 0.0001798 | (0.0001697) |
| 1997Q3 | 0.0000022 | 0.0000076 | (0.0000054) |
| 1997Q4 | 0.0000014 | 0.0000018 | (0.0000004) |
| 1998Q1 | 0.0000014 | 0.0000030 | (0.0000015) |
| 1998Q2 | 0.0000025 | 0.0020258 | (0.0020234) |
| 1998Q3 | 0.0000009 | 0.0000357 | (0.0000349) |
| 1998Q4 | 0.0000016 | 0.0006645 | (0.0006629) |
| 1999Q1 | 0.0000087 | 0.0005550 | (0.0005463) |
| 1999Q2 | 0.0000016 | 0.0000748 | (0.0000731) |
| 1999Q3 | 0.0000005 | 0.0007572 | (0.0007567) |
| 1999Q4 | 0.0000008 | 0.0000016 | (0.0000007) |
| 2000Q1 | 0.0000012 | 0.0004941 | (0.0004930) |
| 2000Q2 | 0.0000003 | 0.0000151 | (0.0000149) |
| 2000Q3 | 0.0000000 | 0.0000075 | (0.0000075) |
| 2000Q4 | 0.0000037 | 0.0000872 | (0.0000835) |
| 2001Q1 | 0.0000001 | 0.0010545 | (0.0010544) |
| 2001Q2 | 0.0000031 | 0.0000847 | (0.0000816) |
| 2001Q3 | 0.0000003 | 0.0000312 | (0.0000309) |
| 2001Q4 | 0.0000113 | 0.0000222 | (0.0000109) |
| 2002Q1 | 0.0000458 | 0.0044534 | (0.0044076) |
| 2002Q2 | 0.0000001 | 0.0001673 | (0.0001672) |

**Table 9** Summary of squared prediction errors from value-weighted simple regression and multiple regression model and the differences of squared prediction error from these two models. Sample used in multiple regression model is aggregated based on Fama French industry classification scheme (10 industries). (Cont.)

| Prediction Quarter | Squared prediction error yield from simple regression | Squared prediction error yield from VW multiple regression model | Difference in squared prediction error |
|---|---|---|---|
| 2002Q3 | 0.0000107 | 0.0005873 | (0.0005766) |
| 2002Q4 | 0.0000072 | 0.0018751 | (0.0018679) |
| 2003Q1 | 0.0000090 | 0.0019335 | (0.0019245) |
| 2003Q2 | 0.0000008 | 0.0001029 | (0.0001021) |
| 2003Q3 | 0.0000000 | 0.0000855 | (0.0000855) |
| 2003Q4 | 0.0000344 | 0.0022046 | (0.0021702) |
| 2004Q1 | 0.0000508 | 0.0045893 | (0.0045385) |
| 2004Q2 | 0.0000022 | 0.0031248 | (0.0031226) |
| 2004Q3 | 0.0000041 | 0.0000230 | (0.0000188) |
| 2004Q4 | 0.0000152 | 0.0008696 | (0.0008544) |

T-Stat: –4.06

## 6. Conclusion

Brown and Ball (1967) document a relationship among earnings of an individual firm, earnings of the other firms in its industry and the earnings of all firms in the market. They find a high correlation between earnings from various industries and market earnings. They also find that earnings from different industries associate with market earnings in the different levels. Based on this finding, I hypothesize that the existing market earnings prediction model is not the most efficient model. In this analysis I aim to improve the random walk market earnings prediction model by using industry earnings changes to predict future market earnings changes. I find that a predictive ability of market earnings changes to predict future earnings changes can be improved by using industry earnings changes. However, this improvement is sensitive to the industry classification and aggregation schemes. To be specific, by using Fama French (either 10 or 17 industries list), simple regression model and IP model yield a similar results. In another word, the squared prediction errors from these two models are not statistically significantly different from each other during the analysis period. However, if SIC one-digit and value weighted schemes are used to aggregate earnings changes, industrial prediction model is more precise than simple regression model in predicting future market earnings. Using equal-weighted aggregation scheme on the other hand shows no improvement.

## References


### English

Bhojraj, S.; Lee,C. M. C., and Oler, D. (2003). What's My Line? A Comparison of Industry Classification Schemes for Capital Market Research. *Journal of Accounting Research, 41*, 745–74.

Brown, P., and Ball, R. (1967). Some Preliminary findings on the association between the earnings of a firm, its industry and the economy. Empirical research in accounting: Selected studies. *Supplement to Journal of Accounting research, 5*, 1–14.

Clarke, R. N. (1989). SICs as Delineators of Economic Markets. *Journal of Business, 62*, 17–31.

Cready, W M., and Gurun U. (2009). *Market Return and Aggregate Earnings News*. Working paper.

Dichev, I.D., and Tang, V. W. (2009). Earning Volatility and Earnings Predictability. *Journal of Accounting and Economics, 47*, 160-181

Fama, E., and French and K. (1997). Industry Costs of Equity. *Journal of Financial Economics, 43*, 153–193.

Finger, C. (1994). The ability of earnings to predict future earnings and cash flow. *Journal of Accounting Research, 32*, 210–223.

Foster, G. (1977). Quarterly Accounting Data: time-series Properties and Predictive-ability Results. *The Accounting Review*, 1–21.

King, Benjamin F. (1966). Market and Industry Factors in Stock Price Behavior. *Journal of Business, 39*, No. 1, 139–190.

Kothari, S. P., Lewellen, J. W., and Warner, J. B. (2005). Stock returns, aggregate earnings surprises, and behavioral finance. *Journal of Financial Economics, 79*, 537–568.