The Role of Disaggregated Earnings for Stock Prices: Evidence from Listed Shipping Firms and Panel Tests

Nicholas Apergis(Corresponding author)

Department of Banking and Financial Management

University of Piraeus

80 Karaoli & Dimitriou, 18534 Piraeus, Greece

Tel: 0030-210-4142429   E-mail: napergis@unipi.gr

John Sorros

Department of Business Administration

University of Piraeus

80 Karaoli & Dimitriou, 18534 Piraeus, Greece

Tel: 0030-210-4142116   E-mail: sorros@unipi.gr

Abstract

This study examines the impact of disaggregated earnings on stock prices for listed shipping firms. It shows that operating and non-operating income from ships sales have higher power in explaining stock prices than operating earnings only as key information for future profitability, investment opportunities and firm valuation. The testing period is from 2000 to 2008. The methodologies are those of panel cointegration and panel causality tests. The empirical findings show that both types of earnings are positively related to stock prices, contrary to the belief that the non-operating component contains little information. The implications are very crucial, since managers may manipulate annual earnings by these non-operating activities.

JEL Classification: G10, M41, C33

Keywords: Operating income, Non-operating income, Shipping firms, Stock prices, Panel cointegration, Panel causality

1. Introduction

Operating income is considered one of the most important components of earnings (the other is operating cash flows), since it reflects the performance of a firm associated with its operational characteristics. At the same time, the strength of the operating income component is mainly attributed to the fact that this component incurs recurrent characteristics. Lipe (1986) argues that operating income has higher explanatory value vs the non-interest income for explaining stock prices. By contrast, non-operating income is treated as a transitory variable, which, according to Ronen and Sadan (1981), reveals low power information for potential investors.

Previous studies examined the association of total accounting earnings and financial aggregates, such as stock prices or returns (Ou and Penman, 1989; Lev and Thiagarajan, 1993; Fairfield *et al.*, 1996; Dechow *et al.*, 1999; Lee, 1999; Abarbanell and Bernard, 2000). These studies document that accounting earnings affect stock prices via the quality of information they reveal about the future of the firm. Abarbanell and Bushee (1997), however, investigate the link between disaggregated accounting earnings and stock prices. The identification as well as the explanation of this link is considered very valuable by investors and creditors as they pursue enough information that will help them to project future profitability, which, in turn, will provide additional profitability for investment opportunities as well as for the future value of firms. Fairfield *et al.* (1996) argue in favour of the forecast improvements arising from the earnings' disaggregation, while Fairfield *et al.* (2003) show explicitly that accruals as well as long-term net operating assets exert a negative impact on asset returns. Cooper *et al.* (2003) provide evidence that documents the fact that investment behaviour in banking stocks is fundamentally affected by information emerging from certain components of earnings, such as those in income from derivatives and non-interest income. Finally, Alam and Brown (2006) find that disaggregated earnings, also from the banking industry, have incremental predictive power vs aggregated earnings in terms of stock price predictability. They document, though, that operating components display the strongest impact on financial aggregates.

While the above studies exclusively focus on the U.S. financial markets, Chu (1997) shows that both operating and non-operating income affect stock returns for firms listed on the Taiwan capital market. He attributes this to the fact that many firms in this market manipulate their earnings in such a manner in order to dispose them in other markets, such as the real estate market.

All of the previous studies have examined how the above link is evaluated for all types of firms, i.e. mostly manufacturing and banking, except those in the shipping industry. These studies show that aggregated earnings are indeed considered a significant determinant of firm valuation (Ohlson, 1995) and debt financing (Watts and Zimmerman, 1986). The shipping industry literature has not addressed, however, whether certain earnings components convey information beyond aggregate earnings that is proven extremely valuable in assessing not only future earnings but also stock prices or returns. This paper comes forwards and in a novel way attempts to close the gap in this piece of literature by investigating whether both

types of income seem significant in determining financial aggregates, such as stock prices and returns.

Section 2 overviews the characteristics of the shipping industry and the goals of the study. Section 3 discusses the methodology employed, while Section 4 describes the data sample used. Section 5 presents the empirical results and, finally, concluding remarks are given in Section 6.

2. Shipping Industry

One of the most important industries in the service sector is the transportation industry, which has many particularities. In particular, one of these particularities has to do with the fact that income is directly affected by the cycle of the economy at the production and distribution level as well. The sensitivity of the sector becomes even higher by the fact that a typical service-sector firm must make significant investments in fixed assets. Moreover, firms that operate in this sector can divest part of their fixed assets through the sale of their transportation equipment. This is performed more easily compared to disinvestment in other fixed assets, such as buildings or machinery.

Non-operating income includes the non-operating and non-recurrent sources of gains or losses. Many international shipping listed firms have a considerable proportion of non-operating income against operating income. Shipping companies have a high level of fixed assets as compared to their total assets. It is common place for shipping companies to reduce the number of ships they handle through sales at prices higher than they were purchased, thus writing in their income statements non-operating profits. During the last six to seven years, shipping companies listed in the NYSE, disinvested a large part of their ships. Thus, seems very crucial to investigate whether both types of income can act as important information for potential investors or for regulators who investigate how managers are handling firms' profits with transparency.

3. Methodology

The focus of the empirical analysis is the impact of disaggregated earnings, i.e. operating vs non-operating earnings, on stock prices from the maritime industry. Our benchmark models planned to be tested in this study are expressed as:

Model 1: $SP_t = a_0 + a_1 OE_t + u_{1t}$

Model 2: $SP_t = b_0 + b_1 OE_t + b_2 SHIPSNOE_t + u_{2t}$

where SP is the variable of stock prices, OE is the variable of operating income, SHIPSNOE is the variable of non-operating income associated with ship sales, while u's are random variables, i.e. white noises with $N(0, \sigma^2)$. Parameters a_1, b_1, and b_2 are the response coefficients. According to theoretical expectations, parameters a_1 and b_1 are expected to be positive and statistically significant (Easton and Zmijewski, 1989), while for the coefficient b_2 the evidence is mixed. Finally, the empirical analysis will identify the relative strength between b_1 and b_2.

Given the presence of heterogeneity in both dynamics and error variances in the panel, the heterogeneous panel cointegration test advanced by Pedroni (1999, 2004), which allows for cross-section interdependence with different individual effects, is employed as follows:

Model 1: $SP_{it} = \alpha_{it} + \delta_{1i}t + a_1\,OE_{it} + u_{1it}$

Model 2: $SP_{it} = \beta_{it} + \delta_{2i}t + b_1\,OE_{it} + b_2\,SHIPNOE_{it} + u_{2it}$

where $i = 1,...,N$ for each firm in the panel and $t = 1,...,T$ refers to the time period. The parameters αi, βi and $\delta 1i$, $\delta 2i$ allow for the possibility of firm-specific fixed effects and deterministic trends, respectively. All variables are expressed in natural logarithms so the parameters of the model can be interpreted as elasticities. To test the null hypothesis of no cointegration, $\rho_i = 1$, the following unit root test is conducted on the residuals as follows:

$$u_{it} = \rho_i u_{it-1} + w_{it}$$

Pedroni (1999, 2004) proposes two sets of tests for cointegration. The panel tests are based on the within dimension approach (i.e. panel cointegration statistics) which includes four statistics: panel v-statistic, panel ρ-statistic, panel PP-statistic, and panel ADF-statistic. These statistics take into account common time factors and heterogeneity across countries. The group tests are based on the between dimension approach (i.e. group mean panel cointegration statistics) which includes three statistics: group ρ-statistic, group PP-statistic, and group ADF-statistic. These statistics are based on averages of the individual autoregressive coefficients associated with the unit root tests of the residuals for each country in the panel. All seven tests are distributed asymptotically as standard normal. Of the seven tests, the panel v-statistic is a one-sided test where large positive values reject the null hypothesis of no cointegration whereas large negative values for the remaining test statistics reject the null hypothesis of no cointegration.

4. Data

Our sample is derived from the Nasdaq. The sample includes 36 shipping companies and spans the period 2000-2008 on a quarterly basis due to the availability of data. Non-operating income includes only the income generated through ships sales and does not include any other component. Our sample was selected by certain criteria as: All firms are listed on, at least, one stock exchange market a quarter before the testing period and the availability of total, operating, non-operating and stock prices for the same quarter. The first criterion avoids the abnormal performance in the initially public offering period. Our data sample on operating and non-operating income could be tracked in the following sites:

www.nasdaq.com/asp/symbols.asp?exchange=SEC9&start=A&sort=0

www.nasdaq.com/asp/MasterDataEntry.asp?page=dynamic_charting

http://finance.yahoo.com/a/hp?s. Data on stock prices, the book value of common equity and the earnings before extraordinary items and discontinued operations was obtained from

Compustat. Throughout the paper, lower case letters denote variables expressed in natural logarithms, while the RATS (version 6.1) assisted the empirical analysis. Table 1 reports descriptive statistics for both operating and non-operating income. The results indicate that the two means are very close, strengthening the argument that a good number of shipping firms obtain a significant portion of their total revenues from non-operating activities. Moreover, Grubbs' test (1969) for outliers is also used to detect outliers in a data set. It is based on the assumption of normality. This test detects one outlier at a time, then, this outlier is expunged from the dataset and the test is iterated until no outliers are detected. The hypothesis of no outliers is rejected if the statistic is greater than the critical value of the t-distribution with N-2 degrees of freedom. The results, reported in Table 1, denote that no outliers are detected and, therefore, the empirical analysis followed will not be biased by the presence of such outliers.

Table 1. Descriptive Statistics (in logarithms)

	Mean	Variance	Kurtosis	Skewness
Operating Income	30.45	0.6775	-0.2014	0.1182
Non-Operating Income	29.48	4.5181	-0.3826	-0.6198
Grubbs' test = 6.75				

Critical value for t-test: 2.33

5. Empirical Analysis

5.1 Dynamic Heterogeneity

An issue that it is of major concern is the heterogeneity of the firms included in this data set. In particular, through time and across firms, the effects on stock prices-operating income–non-operating income relationship of the different structural operational framework established in each firm should be expected to be diverse.

In the statistical framework of this study we first test for heterogeneity and then by controlling for it through appropriate techniques (Holtz-Eakin, 1986; Holtz et al., 1985). The dynamic heterogeneity, i.e. variation of the intercept over firms and time, across a cross-section of the relevant variables can be investigated as follows. In the first step, an ADF(n) equation for each relationship in the panel is estimated; then, the hypothesis of whether regression parameters are equal across these equations is tested. Next, a similar test of parameter equality is performed by estimating a n-order autoregressive model for each of the relationships under investigation. Standard Chow-type F tests under the null of parameter equality across all relationships are also performed. Heterogeneity in cross-sectional parameters is indicated if the results reject the null hypothesis. Finally, homogeneity error variance across groups is also examined as another measure of dynamic heterogeneity. White's tests for group-wise heteroscedasticity are employed to serve this objective. The results of this procedure are reported in Table 2. The empirical findings indicate that the relationship under investigation in both types of models is characterized by heterogeneity of dynamics and error variance across groups, supporting the employment of panel analysis.

Table 2. Tests of Dynamic Heterogeneity across Groups

Specification	ADF(3)	AR(3)	White's Test
Model 1	19.81*	29.28*	58.77*
Model 2	18.77*	27.65*	52.39*
Model 3	26.78*	35.69*	67.82*

ADF(3) reports the parameter equality test (F-test) across all relationships in the panel. AR(3) displays the F-test of parameter equality conducted in a third-order autoregressive model of the relationships. White's test reports the White's test of equality of variances across the investigated relationships in the panel. Model contains only operating earnings, while Model 2 contains both types of earnings.

* denotes statistical significance at 1%.

5.2 Panel Unit Root Tests

There are a variety of panel unit root tests which include Maddala and Wu (1999), Breitung (2000), Hadri (2000), Choi (2001), Levin et al. (2002), Im et al. (2003), and Carrion-i-Silvestre et al. (2005), among others. Consider the following autoregressive specification:

$$y_{it} = \rho_i y_{it-1} + \delta_i X_{it} + \varepsilon_{it} \qquad (1)$$

where $i = 1,...,N$ for each firm in the panel; $t = 1,...,T$ refers to the time period; X_{it} represents the exogenous variables in the model including fixed effects or individual time trend; ρ_i are the autoregressive coefficients; and ε_{it} are the stationary error terms. If $\rho_i < 1$, y_{it} is considered weakly trend stationary whereas if $\rho_i = 1$, then y_{it} contains a unit root. The Breitung (2000), Hadri (2000), and Levin et al. (2002) panel unit root tests assume that the error terms, ε_{it}, are independently and normally distributed random variables for all $i's$ and $t's$ with mean zero and constant variance. This assumption implies that the coefficient of y_{it-1} is homogeneous across all cross-section units of the panel and individual processes are cross-sectionally independent. In the case of dynamic panel data models, the recognition of parameter heterogeneity is important in order to avoid potential biases which could emerge due to an improper specification.

In light of parameter heterogeneity, the Im et al. (IPS, 2003) panel unit root test is utilized which allows for heterogeneous autoregressive coefficients. Such heterogeneity could occur due to the different economic conditions and stages of economic development in each country. Im et al. (2003) suggest averaging the augmented Dickey-Fuller (ADF) unit root tests while allowing for different orders of serial correlation, $\varepsilon_{it} = \sum_{j=1}^{p_i} \varphi_{ij} \varepsilon_{it-j} + u_{it}$. Substitution of this expression into (1) yields

$$y_{it} = \rho_i y_{it-1} + \sum_{j=1}^{p_i} \varphi_{ij} \varepsilon_{it-j} + \delta_i X_{it} + \varepsilon_{it} \qquad (2)$$

where p_i represents the number of lags in the ADF regression. The null hypothesis is that each series in the panel contains a unit root ($H_0 : \rho_i = 1 \ \forall_i$). The alternative hypothesis is that at least one of the individual series in the panel is stationary ($H_0 : \rho_i < 1$). Im et al. (2003) specify a $t-bar$ statistic as the average of the individual ADF statistics as follows:

$$t - bar = \frac{1}{N} \sum_{i=1}^{N} t_{\rho_i} \qquad (3)$$

where t_{ρ_i} is the individual t-statistic for testing $H_0 : \rho_i = 1 \ \forall_i$ from (2). The $t-bar$ statistic is normally distributed under the null hypothesis with the critical values for given values of N and T provided by Im et al. (2003).

The LLC test allows heterogeneity of individual deterministic effects and heterogeneous serial correlation structure of the error terms assuming homogeneous first order autoregressive parameters. They also assume that both N and T tend to infinity but T increases at a faster rate, such that N/T→0. Thus, a procedure is developed that uses t-statistics of the estimator to evaluate the hypothesis that each individual time series contains a unit root against the alternative hypothesis that each time series is stationary. The test also makes use of the ADF regression methodology, with the ADF regression defined as:

            pi

Δyit = αi + γi yi,t-1 + Σbij∆yi,t-j + εit   (4)

            j=1

Moreover, the test implements a separate ADF regression for each country where the lag order is permitted to vary across individual countries. The appropriate lag order is chosen by allowing the maximum lag order and then uses the t-statistic for the coefficients of the lag terms to determine if a smaller lag order is preferred. Next, it runs two separate regressions, such as:

$$\text{pi}$$

$$\Delta y_{it} = a_i + \Sigma b_{ij} \Delta y_{i,t-j} + e_{it} \qquad\qquad (5)$$

$$j=1$$

and

$$\text{pi}$$

$$y_{i,t-1} = a_i + \Sigma b_{ij} \Delta y_{i,t-j} + v_{i,t-j} \qquad\qquad (6)$$

$$j=1$$

and we save the residuals (e_{it} and $v_{i,t-j}$). We divide the both of saved residuals by the regression standard error of the regression for normalization purposes and, next, we run the regression,

$$e_{it} = \rho\, v_{i,t-j} + \varepsilon_{it} \qquad\qquad (7)$$

with the null hypothesis being H0: $\rho_1 = ... = \rho_n = \rho = 0$ and the alternative hypothesis is H1: $\gamma_1 = .. = \gamma_n = \gamma < 0$ for all i. LLC show that the asymptotic properties of the regression estimators are a mixture of properties derived for stationary panel data and properties derived in unit roots testing. This test seems to have certain limitations, such as that it depends seriously upon the independence assumption across individual regressions and hence not applicable if cross sectional correlation is present. In addition, a limitation is associated with the fact that the autoregressive parameters are considered being identical across the panel regressions (see the above null hypothesis). However, this null hypothesis makes sense under some cases. As Maddala and Wu (1999) point out, the alternative hypothesis is too strong to be valid in any empirical case.

Maddala and Wu and Choi tests offer a strategy that seems to overcome the limitations of both LLC and Im et al. tests. They suggest a non-parametric Fisher-type test, which is based on a combination of the p-values of the t-statistics for a unit root in each cross-sectional unit (the ADF test). Pooling on the basis of the p-value is a common practice in meta-analysis. More specifically, the testing approach has the advantage of allowing for as much heterogeneity across units as possible. Under the hypothesis that the test statistics are continuous, the significance of p-values are independent in a uniform manner, e.g. they uniform (0,1) variables and -2logp has a chi-squared distribution with two degrees of freedom. Using the additive property of the chi-squared variables, the statistic

$$N$$

$\lambda = -2 \Sigma \log(p_i)$ is constructed, which has a chi-squared distribution with 2N degrees

$$i=1$$

of freedom. The advantage of this test is it does not require an infinite number of groups to be valid, we do not have to assume that all groups must have the same type of non-stochastic components, T is not necessarily assumed to be the same for all the cross-section units, its critical values are not sensitive to the choice of lag lengths in the ADF regressions, and finally, it does not have to assume that none of the groups have a unit root under the alternative hypothesis.

- The Carrion-i-Silvestre et al. test is associated with testing the residuals (ε) from the above individual OLS regression (4). Given those residuals, a Lagrange multiplier (LM) statistic is constructed defined as:

$$LM = 1/N \left(\sum_{i=1}^{N} \sum_{t=1}^{T} S_{it}^2 / T^2 / \bar{u}_0 \right) \qquad (8)$$

where Sit denotes the cumulative sum of the residuals, T is the number of time periods and ū0 is the average of individual estimators of the residual spectrum at frequency zero.

It is defined as, $\bar{u}_0 = \sum_{i=1}^{N} u_{i0}/N$. Under certain assumption the statistic: $\sqrt{N} (LM-\zeta)/\xi \rightarrow$

N(0,1), where $\zeta=1/15$ and $\xi=11/6300$. The panel unit root tests indicate all the variables are integrated of order one.

- The Hadri Lagrange multiplier (LM) test is closely related to that of the Carrion-i-Silvestre et al. test. It has the advantage of combining both stationary and non-stationary variables and permits a formulation for a residual-based LM test of stationarity. Hadri adopts the following representation,

$$yit = zit'\gamma + rit + \varepsilon it \qquad (9)$$

where zit is the deterministic component, rit is a random walk process defined as rit = ri,t-1 + uit, with uit→iid(0, σu2) and εit is a stationary process. The null hypothesis of trend stationarity corresponds to the hypothesis that the variance of the random walk is zero. The yit process from above can be written as yit = zit' γ + eit, where:

$$eit = \sum_{j=1}^{T} uij + \varepsilon it \qquad (10)$$

The residuals from the above regression (eit) are obtained. This time the statistic can be written as:

N   T

$$LM = 1/N \, \Sigma \, (\Sigma Sit2/T2/\sigma\varepsilon2) \qquad\qquad (11)$$

$$i=1 \quad t=1$$

where $\sigma\varepsilon2$ is the consistent Newey and West (1987) estimate of the long-run variance of residuals, while Sit and testing are defined as above. The LM statistic is consistent and has an asymptotic normal distribution as both T and N→∞. The main advantage of this test is that the moments of the asymptotic distribution are exactly derived, while the disturbance terms can be heteroskedastic across i. Finally, it is also possible to allow for serial dependence substituting the assumption that the errors εit are i.i.d. normally distributed over t with the assumption that they satisfy the strong mixing regularity conditions of Phillips and Perron (1988). In this case we replace $\sigma\varepsilon2$ by the long-run variance defined as:

$$N$$

$$\sigma2 = 1/N \, \Sigma \qquad limT-1 \, (SiT2) \qquad\qquad (12)$$

$$i=1 \qquad T\to\infty$$

A consistent estimator of the above variance is obtained using again the estimators provided by Newey and West (1994).

The results in Table 3 point out that the hypothesis that all the variables contain a unit root is accepted at the 1% significant level in all tests, suggesting that the log variables in our study are I(1).

Table 3. Panel Unit Root Tests

IPS Panel Unit Root Tests

Variables	Without Trend	With Trend
p	-1.23(3)	-1.44(3)
Δp	-4.47(2)*	-5.12(1)*
oe	-1.45(3)	-1.51(3)
Δoe	-4.36(2)*	-4.54(2)*
shipnoe	-1.31(3)	-1.82(3)
$\Delta shipnoe$	-4.84(2)*	-4.90(2)*
bv	-1.48(3)	-1.94(3)
Δbv	-5.67(1)*	-5.94(2)*
e	-1.18(4)	-1.57(4)
Δe	-5.07(2)*	-5.87(3)*

LLC Panel Unit Root Tests

Variables

p	-1.75
Δp	-9.23*
oe	-1.75
Δoe	-8.11*
shipnoe	-1.52
Δshipnoe	-10.71*
bv	-1.88
Δbv	-9.67*
e	-1.31
Δe	-9.29*

Handri (hom) Panel Unit Root Tests

Variables	
p	10.34*
Δp	1.45
oe	9.87*
Δoe	1.21
shipnoe	13.46*
Δshipnoe	1.79
bv	-16.53*
Δbv	-1.09
e	-18.11*
Δe	-1.52

Handri (het) Panel Unit Root Tests

Variables	
p	7.37*
Δp	0.71
oe	7.64*
Δoe	1.62
shipnoe	9.47*
Δshipnoe	1.06
bv	-14.88*

Δbv	-1.10
e	-15.78*
Δe	-1.24

Carrion-i-Silvestre *et al.* (no breaks, homogeneous)

Variables	
p	10.68*
Δp	2.36
oe	9.84*
Δoe	1.77
shipnoe	12.61*
Δshipnoe	3.11
bv	-13.42*
Δbv	-1.22
e	- 13.94*
Δe	-1.31

Carrion-i-Silvestre *et al.* (no breaks, heterogeneous)

Variables	
p	10.31*
Δp	2.25
oe	9.13*
Δoe	1.59
shipnoe	11.07*
Δshipnoe	2.84
bv	-11.39*
Δbv	-0.95
e	-12.53*
Δe	-1.11

Fisher-ADF

Variables	
p	18.95
Δp	97.74*

oe	21.36
Δoe	136.78*
shipnoe	16.77
Δshipnoe	91.28*

Fisher-PP

Variables

p	24.55
Δp	127.63*
oe	34.52
Δoe	178.94*
shipnoe	29.74
Δshipnoe	166.93*
bv	-29.45
Δbv	-169.48*
e	-35.64
Δe	-184.36*

p denote stock prices, oe denotes operating earnings, shipnoe denotes non-operating earnings, bv denotes book values and e denotes earnings before extraordinary items and discontinued operations. Numbers in parentheses are the augmented lags include the unit root test.

* denotes statistical significance at 1%

5.3 Panel Cointegration Tests

Given the presence of heterogeneity in both dynamics and error variances in the panel, the heterogeneous panel cointegration test advanced by Pedroni (1999, 2004), which allows for cross-section interdependence with different individual effects, is employed as follows. Based on the two model specifications presented above, Table 4 reports both the within and between dimension panel cointegration test statistics. All seven test statistics reject the null hypothesis of no cointegration at the 1 percent significance level.

Table 4. Panel Cointegration Tests

_____

Model 1

Panel v-stat	29.876013*
Panel rho-stat	-29.325922*
Panel pp-stat	-28.730904*
Panel adf-stat	-8.512296*
Group rho-stat	-29.254098*
Group pp-stat	-29.290071*
Group adf-stat	-8.805529*

Model 2

Panel v-stat	33.634472*
Panel rho-stat	-33.690884*
Panel pp-stat	-30.984412*
Panel adf-stat	-8.564588*
Group rho-stat	-32.477248*
Group pp-stat	-32.449065*
Group adf-stat	-7.983325*

Model 3

Panel v-stat	37.543812*
Panel rho-stat	-36.376690*
Panel pp-stat	-35.237611*
Panel adf-stat	-7.138944*
Group rho-stat	-35.094438*
Group pp-stat	-35.376698*
Group adf-stat	-7.156332*

Robustness tests: without years 2001 and 2007

Model 1

Panel v-stat	34.567726*
Panel rho-stat	-31.097831*
Panel pp-stat	-30.582219*
Panel adf-stat	-7.458322*
Group rho-stat	-31.908450*
Group pp-stat	-30.366712*
Group adf-stat	-7.368320*

Model 2

Panel v-stat	35.438851*
Panel rho-stat	-34.809437*
Panel pp-stat	-32.456711*
Panel adf-stat	-6.669232*
Group rho-stat	-33.547992*
Group pp-stat	-33.209085*
Group adf-stat	-6.562118*

Model 3

Panel v-stat	39.956328*
Panel rho-stat	-37.237709*
Panel pp-stat	-36.287442*
Panel adf-stat	-6.525782*
Group rho-stat	-38.343699*
Group pp-stat	-37.933622*
Group adf-stat	-8.990560*

* denotes statistical significance at 1%

Following Pedroni (2000), the fully modified OLS (FMOLS) methodological technique for heterogeneous cointegrated panels is followed. Table 5 displays the FMOLS results. In both models the coefficients are shown to be positive and statistically significant at the 1 percent significance level and given that the variables are expressed in natural logarithms, the coefficients are interpreted as elasticities. In particular, with respect to Model 1, the results display that 1 percent increase in operating income leads to 0.50 percent in stock prices, while with respect to Model 2, the results indicate that a 1 percent increase in operating income increases stock prices by 0.45 percent. A 1 percent increase in non-operating income increases stock prices by 0.37 percent. Finally, the adjusted R-square values show a significant incremental improvement, from 0.46 to 0.59, when both types of income are allowed to affect stock prices.

Table 5.   FMOLS Estimates

Model 1

p = 1.452 + 0.498 oe
   (4.47)*   (5.28)* $\bar{R}^2 = 0.46$

Model 2

p = 1.623 + 0.451 oe + 0.368 shipnoe
  (9.65)*   (5.07)*   (5.21)* $\bar{R}^2 = 0.59$

Model 3

p = 1.338 + 0.517 oe + 0.383 shipnoe + 0.186 bv + 0.279 e
  (7.45)*   (6.14)*   (5.14)*   (5.41)*   (4.94)* $\bar{R}^2 = 0.67$

Robustness tests: without years 2001 and 2007

Model 1

p = 2.571 + 0.466 oe
  (5.83)*   (6.37)* $\bar{R}^2 = 0.49$

Model 2

p = 1.882 + 0.414 oe + 0.395 shipnoe
   (8.73)*   (6.54)*   (6.88)* $\bar{R}^2 = 0.62$

Model 3

p = 1.258 + 0.471 oe + 0.363 shipnoe + 0.177 bv + 0.249 e
  (6.94)*   (5.66)*   (5.38)*   (5.08)*   (5.12)* $\bar{R}^2 = 0.64$

---

t-statistics are reported in parentheses. The definition of variables is similar to that in Table 2.
* denotes significance at 1%

5.4   Error Correction Estimates and Causality

Given that the variables under study are cointegrated, a panel vector error correction model is estimated to perform causality tests. Defining the lagged residuals from the cointegrating equations as the error correction terms, the following dynamic error correction models are estimated:

Model 1:

$$\Delta p_{it} = \alpha_{1j} + \sum_{k=1}^{q} \theta_{11ik} \Delta p_{it-k} + \sum_{k=1}^{q} \theta_{12ik} \Delta oe_{it-k} + \lambda_{1i}\varepsilon_{it-1} + u_{1it}$$

$$\Delta oe_{it} = \alpha_{2j} + \sum_{k=1}^{q} \theta_{21ik} \Delta p_{it-k} + \sum_{k=1}^{q} \theta_{22ik} \Delta oe_{it-k} + \lambda_{2i}\varepsilon_{it-1} + u_{2it}$$

Model 2:

$$\Delta p_{it} = \alpha_{1j} + \sum_{k=1}^{q} \theta_{11ik} \Delta p_{it-k} + \sum_{k=1}^{q} \theta_{12ik} \Delta oe_{it-k} + \sum_{k=1}^{q} \theta_{13ik} \Delta noe_{it-k} + \lambda_{3i} \varepsilon_{it-1} + u_{3it}$$

$$\Delta oe_{it} = \alpha_{2j} + \sum_{k=1}^{q} \theta_{21ik} \Delta p_{it-k} + \sum_{k=1}^{q} \theta_{22ik} \Delta oe_{it-k} + \sum_{k=1}^{q} \theta_{23ik} \Delta noe_{it-k} + \lambda_{4i} \varepsilon_{it-1} + u_{4it}$$

$$\Delta noe_{it} = \alpha_{3j} + \sum_{k=1}^{q} \theta_{31ik} \Delta p_{it-k} + \sum_{k=1}^{q} \theta_{32ik} \Delta oe_{it-k} + \sum_{k=1}^{q} \theta_{33ik} \Delta noe_{it-k} + \lambda_{5i} \varepsilon_{it-1} + u_{5it}$$

where Δ is the first-difference operator; k is the lag length set at three and five for Model 1 and Model 2, respectively; u s are serially uncorrelated error terms. In the stock returns equation, (Δp), short-run causality from operating income and non-operating income to stock returns is tested, respectively, based on: for Model 1, H0: $\theta 12ik=0 \forall ik$ and for Model 2, $H_0 : \theta_{12ik} = 0 \forall_{ik}$ and $\theta_{13ik} = 0 \forall_{ik}$ and H0: $\theta_{12ik} = \theta_{13ik} = 0 \forall_{ik}$, while the null hypothesis of no long-run causality is tested by examining the significance of the t-statistic for the coefficient of the error correction terms.

Table 6 reports the results of these causality tests. They indicate that both operating and non-operating income can determine stock returns in the short-run. In particular, Wald testing, with an F-value of 10.49 for operating income in Model 1 and F-values of 9.74 and 8.76 for operating and non-operating income, respectively, in Model 2 with statistically significant p-values at 1 percent in both cases. Moreover, Wald testing, with a significant at 1% F-value test of 16.39, indicates that both types of income jointly determine stock returns. In both models the sum of the lagged coefficients for short-run changes turns out to be positive (0.457 for operating income in Model 1, 0.416 for operating income in Model 2, 0.284 for non-operating income in Model 2 and, finally, 0.519 for both types of income in Model2). In the long-run, both error correction terms are negative and statistically significant. However, in Model 2 its size is larger, indicating stronger information content when both types of income are included.

Table 6. Panel Causality Test Results

Dependent Variable	Sources of Causation (Independent Variables)		
	Short-run		Long-Run
Model 1			
	Δp	Δoe	λ_1
Δp	----	10.49 (0.457)	-0.0349
		[0.03] [0.00]	[0.00]
$\bar{R}^2$	0.52		
LM	3.35[0.14]		
RESET	1.49[0.23]		

ARCH(1)	0.57[0.28]				
ARCH(4)	0.49[0.33]				
ARCH(8)	0.36[0.36]				
ARCH(12)	0.30[0.42]				

Model 2

	Δp	Δoe	$\Delta shipnoe$	$\Delta oe + \Delta shipnoe$	λ_3
Δp	----	9.74 (0.416)	8.76 (0.284)	16.39 (0.519)	-0.0586
		[0.02] [0.00]	[0.01] [0.00]	[0.00] [0.00]	[0.00]
$\bar{R}^2$	0.63				
LM	3.09[0.17]				
RESET	1.26[0.26]				
ARCH(1)	0.51[0.24]				
ARCH(4)	0.43[0.29]				
ARCH(8)	0.34[0.35]				
ARCH(12)	0.25[0.39]				

	t-value: Model 1 vs Model 2	t-value: Model 2 vs Model 1
J-test	-10.63*	-1.16
JA-test	-8.43*	-1.25

Robustness tests: without years 2001 and 2007

Model 1

	Δp	Δoe	λ_1
Δp	----	12.73 (0.428)	-0.0368
		[0.01] [0.00]	[0.00]
$\bar{R}^2$	0.55		
LM	3.09[0.19]		
RESET	1.24[0.28]		
ARCH(1)	0.51[0.34]		
ARCH(4)	0.43[0.39]		
ARCH(8)	0.37[0.35]		
ARCH(12)	0.33[0.39]		

Model 2

	Δp	Δoe	$\Delta shipnoe$	$\Delta oe + \Delta shipnoe$	λ_3
Δp	----	8.36 (0.386)	8.93 (0.307)	18.46 (0.559)	-0.0541
		[0.02] [0.00]	[0.00] [0.00]	[0.00] [0.00]	[0.00]
$\bar{R}^2$	0.68				
LM	2.78[0.12]				
RESET	1.16[0.22]				
ARCH(1)	0.43[0.28]				
ARCH(4)	0.41[0.32]				
ARCH(8)	0.30[0.39]				
ARCH(12)	0.24[0.40]				

	t-value: Model 1 vs Model 2	t-value: Model 2 vs Model 1
J-test	-14.02*	-1.05
JA-test	-8.81*	-1.13

Wald F-tests reported with respect to short-run changes in the independent variables. The sum of the lagged coefficients for the respective short-run changes is denoted in parentheses. λs represent the coefficient of the error correction term, while values in brackets denote p-values. LM is a serial correlation test, RESET is a functional misspecification test and ARCH is an ARCH test at 1, 4, 8, and 12 lags. Finally, J and JA tests are encompassing tests.
* denotes statistical significance at 1%

Looking at the models' overall performance as reported by a battery of diagnostic tests, the estimated equations satisfy certain econometric criteria, namely absence of serial correlation (LM test) and absence of functional misspecification (RESET test). ARCH tests were also applied to test the residual structure in the mean equations. The results accept the hypothesis that significant ARCH effects do not appear to exist.

Finally, to more formally test the adequacy of Model 2 vis-à-vis Model 1 and, in particular, to see if Model 2 can encompass Model 1 or *vice versa*, we make use both of the J and the JA tests, which are designed to settle encompassing model issues. The J-test, developed by Davidson and MacKinnon (1982), as well as the JA-test (Doran, 1993; with improvements, Fan and Li, 1995) are considered to be exact and robust tests when the two sets of regressors are near orthogonal. They produce tests of correct size even in small samples. Nevertheless, the J-test has been criticized that it is not exact and suffers from size distortions. It tends to over-reject a true hypothesis in small samples, which is not the case in our sample. It is highly recommended, though, that both tests must be applied simultaneously. The results, also reported in Table 6, show that both tests favour Model 2 against Model 1 at the 1% significance level. In other words, the explanation provided by Model 1 can be improved by incorporating information from Model 2.

Overall, the empirical findings demonstrate that the market, to assess financial aggregates, such as stock prices and/or returns, uses both types of prior period earning components, i.e. operating and non-operating income. The causality results of non-operating income tell that this component of earnings provides substantial information contents to investors. Therefore, non-operating income is viewed as a recurrent and an important piece of information for investors as well as for regulators in shipping stocks.

5.5 Robustness Tests

To test the validity of our results we employ the accounting-based valuation model to extend the analysis by examining the role of accounting information, such as equity book value and earnings, in firm's valuation (Hellstron, 2005; Callao *et al.*, 2006). In particular, the paper follows a model from the valuation perspective, i.e. evaluating the relation between firm value/return and accounting information (Ohlson, 1995). In this manner, we can determine whether more accounting variables reflect information that is used by investors to value

firm's equity. To this end, we augment the above model with these two additional accounting variables. The two additional accounting measures are reflected in the following model:

Model 3: $SP_{it} = \beta_{it} + \delta_{2i}t + b_1\ OE_{it} + b_2\ SHIPNOE_{it} + b_3\ BV_{it} + b_4\ E_{it}$

where BV is the book value of common equity and E is the earnings before extraordinary items and discontinued operations.

Once again heterogeneity tests (Table 2) support the employment of a panel analysis. Moreover, panel unit root tests support that both the BV and the E are characterized as I(1) variables (Table 3), while panel cointegration tests support the employment of an EC approach (Table 4). Table 5 reports the new FMOLS results. In this setting all coefficients are shown to be positive and statistically significant at the 1 percent significance level. Moreover, the results display that 1 percent increase in operating income leads to 0.52 percent in stock prices, a 1 percent increase in non-operating income increases stock prices by 0.38 percent, a 1 percent increase in the book value of the shipping firms leads to a 0.19 percent increase in stock prices, while a 1 percent increase in firms' earnings leads to a 0.28 percent increase in stock prices. The adjusted R-square values show an improvement, from 0.59 to 0.67, when the above variables as well as the book value and the earnings are allowed to affect stock prices. Finally, the following dynamic error correction model is estimated:

Model 3:

$$\Delta p_{it} = \alpha_{1j} + \sum_{k=1}^{q}\theta_{11ik}\Delta p_{it-k} + \sum_{k=1}^{q}\theta_{12ik}\Delta oe_{it-k} + \sum_{k=1}^{q}\theta_{13ik}\Delta noe_{it-k} \qquad +$$

$$\sum_{k=1}^{q}\theta_{14ik}\Delta bv_{it-k} + \sum_{k=1}^{q}\theta_{15ik}\Delta e_{it-k} + \lambda_{6i}\varepsilon_{it-1} + u_{6it}$$

$$\Delta oe_{it} = \alpha_{2j} + \sum_{k=1}^{q}\theta_{21ik}\Delta p_{it-k} + \sum_{k=1}^{q}\theta_{22ik}\Delta oe_{it-k} + \sum_{k=1}^{q}\theta_{23ik}\Delta noe_{it-k} \qquad +$$

$$\sum_{k=1}^{q}\theta_{24ik}\Delta bv_{it-k} + \sum_{k=1}^{q}\theta_{25ik}\Delta e_{it-k} + \lambda_{7i}\varepsilon_{it-1} + u_{7it}$$

$$\Delta noe_{it} = \alpha_{3j} + \sum_{k=1}^{q}\theta_{31ik}\Delta p_{it-k} + \sum_{k=1}^{q}\theta_{32ik}\Delta oe_{it-k} + \sum_{k=1}^{q}\theta_{33ik}\Delta noe_{it-k} \qquad +$$

$$\sum_{k=1}^{q}\theta_{34ik}\Delta bv_{it-k} + \sum_{k=1}^{q}\theta_{35ik}\Delta e_{it-k} + \lambda_{8i}\varepsilon_{it-1} + u_{8it}$$

Table 7 reports the results of these causality tests. They indicate again that both operating and non-operating income determine stock returns in the short-run. In particular, Wald testing, with an F-value of 10.49 for operating income in Model 1 and F-values of 9.74 and 8.76 for operating and non-operating income, respectively, in Model 2 with statistically significant p-values at 1 percent in both cases. Moreover, Wald testing, with a significant at 1% F-value test of 16.39, indicates that both types of income jointly affect stock returns. In both models

the sum of the lagged coefficients for short-run changes turns out to be positive (0.457 for operating income in Model 1, 0.416 for operating income in Model 2, 0.284 for non-operating income in Model 2 and, finally, 0.519 for both types of income in Model2). In the long-run, both error correction terms are negative and statistically significant. However, in Model 2 its size is larger, indicating stronger information content when both types of income are included.

Table 7. Robustness Causality Test Results

Dependent	Sources of Causation (Independent Variables)				
Variable	**Short-run**				**Long-Run**
Model 3					
	Δp	Δoe	$\Delta shipnoe$	$\Delta oe + \Delta shipnoe$	λ_3
Δp	----	11.28 (0.471) [0.00] [0.00]	9.19 (0.326) [0.00] [0.00]	20.68 (0.573) [0.00] [0.00]	-0.0663 [0.00]
$\bar{R}^2$	0.68				
LM	2.68[0.15]				
RESET	1.48[0.29]				
ARCH(1)	0.47[0.21]				
ARCH(4)	0.41[0.24]				
ARCH(8)	0.30[0.32]				
ARCH(12)	0.22[0.37]				

Notes: Similar to Table 6.

Looking at the model's overall performance as reported by a battery of diagnostic tests, the estimated equations satisfy again certain econometric criteria, namely absence of serial correlation (LM test) and absence of functional misspecification (RESET test). ARCH tests were also applied to test the residual structure in the mean equations. The results accept the hypothesis that significant ARCH effects do not appear to exist.

Moreover, the time period under study experienced two financial crises in the U.S., in 2001 and in 2007. To eliminate any possible impact of these two periods on our results, we redo the above analysis without including these two periods. The results appear in each table (Table 4 through Table 6). Our results validate the outlier test performed above, indicating, once again, that non-operating income retains its significant information content for the course of stock prices in the case of shipping firms. Overall, the robustness findings demonstrate that non-operating income tells that this component of earnings provides substantial information contents to investors. Therefore, non-operating income is viewed as a recurrent and an important piece of information for investors as well as for regulators in shipping stocks.

6. Concluding Remarks and Implications

This paper offers a panel methodological proposal to investigate the role of disaggregated income in determining the course of financial variables, such as stock prices, for the case of international listed shipping firms. To this end, a panel sample of 36 shipping firms is used over the time span 2000-2008. Through the methodology of panel cointegration and causality testing, the empirical findings show that disaggregating earnings, i.e. operating income and non-operating income, provide incremental information for stock prices. Our results support the professions' contention that non-operating items may often be equally useful to those who make investment, credit, or other decisions. These results suggest that the usefulness of accounting information is affected by the role of the information in the market. The implications of the results could be very crucial to regulators, since managers may manipulate firms' earnings through these non-operating activities.

References

Abarbanell, J. & Bernard, V. (2000). Is the US Stock Market Myopic?. *Journal of Accounting Research*, 38, pp. 221-243.

Abarbanell, J. & Bushee, B. (1997). Fundamental Analysis, Future Earnings, and Stock Prices. *Journal of Accounting Research*, 35, pp. 1-24.

Alam, P. & Brown, C. A. (2006). Disaggregated Earnings and the Prediction of ROE and Stock Prices: A Case of the Banking Industry. *Review of Accounting and Finance*, 5, pp. 443-463.

Breitung, J. (2000). The Local Power of Some Unit Root Tests for Panel Data. *Advances in Econometrics,* 15, pp. 161-177.

Callao, S., Cuellar, B. & Jarne, J. I. (2006). International Differences in Value Relevance of Accounting Data and Explaining Country Factors. *International Journal of Accounting, Auditing and Performance Evaluation*, 3, pp. 387-408.

Carrion-i-Silvestre, J. L., del Barrio-Castro, T. & Lopez-Bazo, E. (2005). Breaking the Panels: An Application to GDP per Capita. *Econometrics Journal*, 8, pp. 159-175.

Choi, I. (2001). Unit Root Tests for Panel Data. *Journal of International Money and Finance,* 20, pp. 249-272.

Chu, E. L. (1997). Impact of Earnings, Dividends and Cash Flows on Stock Returns: Case of Taiwan's Stock Market. *Review of Quantitative Finance and Accounting*, 9, pp. 181-202.

Cooper, M., Jackson III, W. & Patterson, G. (2003). Evidence of Predictability in the Cross-Section of Bank Stock Returns. *Journal of Banking and Finance*, 27, pp. 817-850.

Davidson, R. & MacKinnon, J. (1982). Some Non-Nested Hypothesis Tests and the Relations among Them. *Journal of Economic Studies*, 49, pp. 551-565.

Dechow, P., Hutton, A. & Sloan, R. (1999). An Empirical Assessment of the Residual Income Valuation Model. *Journal of Accounting and Economics*, 26, pp. 1-34.

Doran, H. (1993). Testing Non-Nested Models. *American Journal of Agricultural Economics*, 75, pp. 95-103.

Easton, P. & Zmijewski, M. E. (1989). Cross-Sectional Variation in Stock Market Response in Earnings Announcement. *Journal of Accounting and Economics*, pp. 117-141.

Fairfield, P., Whiscenant, J.. & Yohn, T. (2003). Accrued Earnings and Growth: Implications for Future Profitability and Market Mispricing. *The Accounting Review*, 78, pp. 353-371.

Fairfield, P., Sweeney, R. & Yohn, T. (1996). Accounting Classification and the Predictive Content of Earnings. *The Accounting Review*, 71, pp. 337-355.

Fan, Y. & Li, Q. (1995). Bootstrapping J-type Tests for Non-Nested Models. *Economics Letters*, 48, pp. 107-112.

Grubbs, F. (1969). Procedures for Detecting Outlying Observations in Samples. *Technometrics*, 11, pp. 1-21.

Hadri, K. (2000). Testing for Stationarity in Heterogeneous Panel Data. *Econometric Journal*, 3, pp. 148-161.

Hellstron, K. (2005). The Value relevance of Financial Accounting in a Transitional Economy: The Case of Czech Republic. Working Paper Series in Business Administration, University of Prague.

Holtz-Eakin, D. (1986). Testing for Individual Effects in Dynamic Models Using Panel Data. NBER Technical Paper Series", No. 57.

Holtz-Eakin, D., W. Newey, & H. Rosen (1985). Implementing Causality Tests with Panel Data with an Example from Local Public Finance. *NBER Technical Working Paper*, No. 48.

Im, K.S., M.H. Pesaran, & Y. Shin (2003) 'Testing for Unit Roots in Heterogeneous Panels. *Journal of Econometrics,* 115, pp. 53-74.

Lee, C. (1999). Accounting-Based Valuation: Impact on Business Practices and Research. *Accounting Horizons*, 13, pp. 413-425.

Lev, B. & Thiagarajan, S. (1993). Fundamental Information Analysis. *Journal of Accounting Research*, 31, pp. 190-215.

Levin, A., Lin, C. F. & Chu, C. (2002). Unit Root Tests in Panel Data: Asymptotic and Finite-Sample Properties. *Journal of Econometrics*, 108, pp. 1-24.

Lipe, R. C. (1986). The Information Contained in the Components of Earnings. *Journal of Accounting Research (Supplement)*, pp. 37-64.

Maddala, G. S. & Wu, S. (1999). A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test. *Oxford Bulleting of Economics and Statistics,* 61, pp. 631-652.

Newey, W. & West, K. (1994). Automatic Lag Selection in Covariance Matrix Estimation. *Review of Economic Studies*, 61, pp. 631-653.

Newey, W. & West, K. (1987). A Simple Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica*, 69, pp. 1519-1554.

Ohlson, J. (1995). Earnings, Book Values, and Dividends in Security Valuation. *Contemporary Accounting Research*, 11, pp. 661-687.

Ou, J. & Penman, S. (1989). Accounting Measurement, Price-Earnings Ratio, and the Information Content of Security Prices. *Journal of Accounting Research,* 27, pp. 111-144.

Pedroni, P. (2004). Panel Cointegration: Asymptotic and Finite Sample Properties of Pooled Time Series Tests with an Application to the PPP Hypothesis: New Results. *Econometric Theory*, 20, pp. 597-627.

Pedroni, P. (2000). Fully Modified OLS for Heterogeneous Cointegrated Panels. *Advanced in Econometrics*, 15, pp. 93-130.

Pedroni, P. (1999). Critical Values for Cointegration Tests in Heterogeneous Panels with Multiple Regressors. *Oxford Bulletin of Economics and Statistics*, 61, pp. 653-670.

Phillips, P. C. B. & Perron, P. (1988). Testing for a Unit Root in Time Series Regression. *Biometrika*, 75, pp. 335-346.

Ronen, J. & Sadan, S. (1981). *Smooth Income Number.* Addison Wesley Publishing Company.

Watts, R. & Zimmerman, J. (1986). *Positive Accounting Theory.* Prentice Hall, Englewood Cliffs, NJ.