Aalto University
School of Economics

Cross-Predictability of Stock Returns: A Study of The Limited-Information Model

Finance

Master's thesis

Oliver Nilsson

2012

Department of Finance
Aalto University
School of Economics

Aalto University
School of Economics

Aalto University
School of Economics
Master's Thesis
Oliver Nilsson

Abstract
May 23, 2012

CROSS-PREDICTABILITY OF STOCK RETURNS: A STUDY OF THE LIMITED-INFORMATION MODEL

PURPOSE OF THE STUDY

The purpose of this study is to examine limited-information models which posit that returns on economically linked assets cross-predict each other, and to determine whether they provide a compelling explanation for stock return predictability in a time series. Furthermore, this paper investigates the model predictions concerning the effect of informed investors and investor geographic specialization on return cross-predictability. This paper also investigates self-financing trading strategies that capitalize on return cross-predictability effects.

DATA

This paper analyzes two samples which include all publicly listed companies traded in Eurozone and EU27 countries over the time period ranging from January 2000 to December 2009. The accounting and stock market data used in this paper are from Worldscope and Thomson One Banker databases, respectively. In addition, consolidated Eurostat input-output tables for years 2000 and 2005 are used to identify customer and supplier industries for each sample company.

RESULTS

The empirical evidence in this paper shows that previous-month supplier industry returns cross-predict stock- and industry-level returns. On the other hand, previous-month returns in customer industries exhibit only weak cross-predictability effects, particularly in the Eurozone. In addition, the results show that the magnitude of return cross-predictability is negatively related to the number of informed investors. Furthermore, this paper is able to provide new empirical evidence indicating that the magnitude of return cross-predictability is positively related to the geographic dispersion of informative signals diffusing from related industries. Finally, the results show that cross-predictability effects can be economically significant: self-financing trading strategies based on return cross-predictability are able to generate mean annual abnormal returns of up to 9.7%.

KEYWORDS

Limited-information model, return cross-predictability, gradual information diffusion, investor specialization, market segmentation, informed investor, input-output table

Aalto-yliopiston kauppakorkeakoulu
Pro Gradu -tutkielma
Oliver Nilsson

Tiivistelmä
23. toukokuuta 2012

CROSS-PREDICTABILITY OF STOCK RETURNS: A STUDY OF THE LIMITED-INFORMATION MODEL

TUTKIELMAN TAVOITTEET
Tutkimuksen tavoitteena on tarkastella rajallisen informaation -mallien kykyä selittää osaketuottojen ennustettavuutta. Työssä tutkitaan, onko osaketuottoja mahdollista ennustaa taloudellisesti linkittyneiden osakkeiden historiallisten tuottojen perusteella. Lisäksi tutkimuksessa selvitetään informoituneiden sijoittajien sekä sijoittajien maantieteellisen erikoistumisen vaikutusta tuottojen ennustettavuuteen. Tutkimuksessa tarkastellaan myös osaketuottojen ennustettavuuteen perustuvien sijoitusstrategioiden kannattavuutta.

LÄHDEAINEISTO
Lähdeaineisto koostuu euro- ja EU27-alueella vuoden 2000 tammikuusta vuoden 2009 joulukuulle asti julkisen kaupankäynnin kohteena olleista osakkeista. Tilinpäätöstiedot on kerätty Worldscope-tietokannasta ja osakemarkkinatiedot Thomson One Banker -tietokannasta. Lisäksi tutkimuksessa käytetään Eurostatin konsolidoituja panos-tuotostaulukoita vuosilta 2000 ja 2005 asiakas- ja tavarantoimittajatoimialojen määrittämiseksi näytteessä oleville yrityksille.

TULOKSET
Tutkimuksen tulokset osoittavat, että tavarantoimittajatoimialojen edeltävän kuukauden tuotot ennustavat osake- ja toimialatason tuottoja. Sen sijaan asiakastoimialojen edeltävän kuukauden tuottoihin perustuvasta osaketuottojen ennustettavuudesta löytyy ainoastaan heikkoa näyttöä, erityisesti euroalueella. Lisäksi tutkimustulokset viittaavat negatiiviseen yhteyteen osaketuottojen ennustettavuuden ja markkinoilla olevan informaation määrän välillä. Tuottojen ennustettavuus vaikuttaa myös olevan positiivisesti korreloitunut asiakas- ja tavarantoimittajatoimialoihin liittyvän informaation maantieteellisen hajautuneisuuden kanssa. Tutkimuksen tulokset osoittavat, että havaittu osaketuottojen ennustettavuus on taloudellisesti merkittävä ilmiö: ennustettavuuteen perustuvat omarahoitteiset sijoitusstrategiat tuottavat jopa 9,7% ylisuuria tuottoja.

AVAINSANAT
Rajallisen informaation -malli, tuottojen ennustettavuus, informaation diffuusio, sijoittajien erikoistuminen, markkinoiden segmentoituminen, informoitunut sijoittaja, panos-tuotostaulukko

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 INTRODUCTION

1.1 BACKGROUND

The basic paradigm of asset pricing which posits that asset prices are determined by risk is in vigorous flux (Hirshleifer, 2001). According to Hirshleifer, the assumption of fully rational investors in asset pricing theory is being subsumed by a broader approach that builds upon the psychology of investors. In this approach, expected returns on securities are driven by both risk and investor misvaluation.

In their article, Huberman and Regev (2001) provide one of the cleanest examples of investor misvaluation and its impact on asset prices. They examine the stock-market behavior of EntreMed, a biotechnology firm that featured in a New York Times front-page story announcing a potential breakthrough in the firm's cancer-curing drug development. The salient piece of news caught the attention of investors and caused EntreMed's stock price to soar from 12 dollars to 85 dollars over night. Remarkable about Huberman and Regev's finding is that the New York Times story contained essentially no real news. The substance of the story had been published five months earlier, in a scientific journal Nature as well as in various popular newspapers, including the New York Times itself. The authors interpret EntreMed's stock-price behavior as suggestive of the existence of two types of investors in EntreMed across whom information gradually diffuses and causes the stock price to respond to news with a delay: a small group of experts who read publications like Nature and a larger group of generalists who gather information from salient sources such as the front page of the New York Times.

The circumstantial evidence by Huberman and Regev, although admittedly interesting, comprises only a fraction of the growing empirical evidence that has emerged to challenge the traditional asset pricing theories. Early works by DeBondt and Thaler (1985, 1987) and Chopra, Lakonishok and Ritter (1992) document that stock returns experience long-term reversals in periods of three to five years. Subsequent studies by Jegadeesh and Titman (1993, 2001) and Rouwenhorst (1998, 1999) extend the evidence on return predictability by reporting that stock returns experience short- and medium-term continuation in the run of three to twelve months. Furthermore, another strand of literature documents that returns on certain stocks lead those of other stocks (see e.g., Lo and MacKinlay, 1990a; Chordia and Swaminathan, 2000). Importantly, many of these studies also suggest that trading strategies capitalizing on the observed return predictability effects are able to generate abnormal returns that cannot be explained by classical asset-pricing theories.

Motivated by the empirical findings on return predictability, new behavioral asset pricing models and theories are being developed to supplement traditional models, such as the Capital Asset Pricing Model (Sharpe, 1964; Lintner, 1965). A common thread between these new models is that they strive to explain asset price behavior by relaxing some of the stringent assumptions of the efficient market hypothesis (EMH). A prominent behavioral model by Hong and Stein (1999) builds upon the gradual information diffusion hypothesis[1] and emphasizes the interactions between two groups of investors which cause asset returns to exhibit predictability. Limited-information models by Hong, Torous, and Valkanov (2007) and Menzly and Ozbas (2010) further extend the gradual information diffusion hypothesis of Hong and Stein by incorporating investor specialization into their models. Both authors provide evidence that investor specialization, and the resulting informational segmentation of markets, are important determinants of information diffusion in the markets and consequently, asset return predictability. Essentially, the informational segmentation of markets between groups of investors, sketched out by Huberman and Regev (2001) in the case of EntreMed, is also at the core of the limited-information models.[2] While the limited-information models build on several ingredients, such as dispersed information about fundamentals (Hayek, 1945) and limits to arbitrage (e.g., Shleifer and Summers 1990), the analogy to the case of EntreMed is clear: investors specialize in their information gathering activities, which renders the markets informationally segmented, causing value-relevant information to diffuse slowly across the markets and returns on economically related assets to cross-predict each other.

Behavioral models, such as the limited-information models, which attribute the observed asset mispricings to underlying processes of possibly less than fully rational investors and interactions between these investors, have sparked a considerable amount of interest and debate among the academics. Even though academics have become more receptive to psychological explanations, proponents of the classical asset-pricing theory argue that the underlying evidence is not yet compelling enough to support a wider application of behavioral asset pricing models. Particularly, empirical research on return predictability is often overshadowed by accusations of data mining and spurious results (Hirshleifer, 2001). Moreover, according to Hong and Stein (2007) the enduring appeal of classical asset-pricing theory is based on a consensus around a foundational framework that lends itself to effective theorizing and modeling. Conversely, they note that while a lot has

[1] Gradual information diffusion refers to an important feature of the assets markets in which certain pieces of value-relevant information arrive in the hands of some investors before others as a result of either the technology of information distribution or investor specialization (Hong and Stein, 2007).

[2] Huberman and Regev's (2001) example of EntreMed describes how gradual diffusion of information caused by investor specialization can result in asset prices reacting with a lag to publicly available information. It is useful to note that while Huberman and Regev's example addresses investor specialization and return predictability within a single company, the limited-information models extend this impact of investor specialization to the context of multiple economically related units such as industries.

already been achieved, a similar broad consensus on key concepts is still missing in the field of behavioral asset pricing which aspires to understand the origins of market mispricings.

Authors have taken very different approaches in attempting to explain asset prices based on investor psychology and, as a result, there exists a large collection of empirical facts and competing one-off models (Hong and Stein, 2007). The previously mentioned limited-information models represent the new frontline of behavioral models which build on the "disagreement" between investors.[3] According to Hong and Stein (2007) this class of heterogeneous-agent models collectively holds the promise of being able to deliver comprehensive insight on theoretical work in asset pricing. Thus, the purpose of this study is to extend the empirical evidence on the limited-information model by Menzly and Ozbas (2010) and simultaneously provide further evidence also on other behavioral models that strive to explain return predictability based on the gradual information diffusion hypothesis of Hong and Stein (1999). The creation of a fully satisfying behavioral asset-pricing model requires vigorous research and providing more empirical evidence on the most prominent behavioral models and theories contributes to this end. As Barberis and Thaler (2002) point out, the scientific way to evaluate competing theories is by conducting empirical tests on the novel predictions that the theories make.

1.2 RESEARCH PROBLEM AND OBJECTIVES

The research problem that this paper addresses is whether limited-information models provide a compelling explanation for stock return predictability in a time-series. To answer this, the main objective of this paper is to study the validity of the limited-information model by Menzly and Ozbas (MO, 2010). MO derive a limited-information model that shows how the specialization of investors in their information gathering activities, and the resulting informationally segmented markets, have significant effects on the formation of prices. More specifically, they argue that gradual diffusion of information from economically related industries is a pervasive feature of markets and leads to cross-predictability in asset returns.[4]

In order to examine the validity of the MO model, this paper studies the testable predictions that the model makes using a cross-country sample. More specifically, this paper provides empirical evidence on return cross-predictability between economically linked assets and the effect of informed investors and investor specialization on return cross-predictability. It is important to note

[3] Disagreement models are a class of heterogeneous-agent models which build on mechanisms that create disagreement between investors concerning asset values. These disagreement mechanisms are (i) gradual diffusion of information, (ii) limited attention and (iii) heterogeneous priors. Heterogeneous-agent model is a broad term that refers to models in which agents are assumed to be heterogeneous in terms of behavior. This can be contrasted with the traditional assumption of representative agents which assumes that all agents behave identically. (Hong and Stein, 2007.)

[4] See subsection 2.2.2.2 for a detailed description of the limited-information model of Menzly and Ozbas (2010).

that the Menzly and Ozbas (2010) model builds on the gradual information diffusion model of Hong and Stein (1999) and the limited-information model by Hong, Torous, and Valkanov (2007). Therefore, a test of the MO model can also be seen as an indirect test of these two models. The main objective of this study can be divided into three sub-objectives as explained below.

First, this paper tests whether return cross-predictability exists between economically linked units as suggested by the Menzly and Ozbas (2010) limited-information model. In order to test for return cross-predictability, this paper first defines economic linkages between industries based on Eurostat input-output tables (see subsection 4.3) and provides preliminary evidence on the correlation of firm (industry) fundamentals along the supply chain. Thereafter, this paper provides empirical evidence on return cross-predictability by examining the predictive power of previous-month supplier and customer industry returns over stock (industry) returns using two Europe-wide samples comprising of Eurozone and EU27 countries. The rationale for studying return cross-predictability is to understand whether the MO model provides a compelling explanation for asset price formation also in a non-US cross-country sample.

Second, this paper studies the effect of informed investors and investor specialization on the magnitude of return cross-predictability. The limited-information model of Menzly and Ozbas (2010) predicts that the magnitude of cross-predictability should be lower (higher) when the number of informed investors following a stock is higher (lower). In addition, the MO model builds on the presumption that investor specialization leads to informationally segmented markets and consequently, cross-predictability in asset returns. In order to provide more evidence on the fundamental drivers of investor specialization and investor specialization being the source of return cross-predictability, this paper studies the relation between investor specialization along geographic boundaries and the magnitude of the cross-predictability effect. This aspect of investor specialization in the context of limited-information models has not been studied in earlier research.

Finally, this paper studies the economic significance of return cross-predictability by constructing self-financing trading strategies that capitalize on the observed cross-predictability effects. This paper also examines the robustness of the self-financing trading strategy returns on well-known return factors, namely the four-factor model risk factors — market, SMB, HML, and MOM by Fama and French (1993) and Carhart (1997), in order to determine whether the reported returns are abnormal.

1.3 MOTIVATION AND CONTRIBUTION OF THE STUDY

The overall aspiration of this paper is to provide further insight on the role of "disagreement" models in theoretical asset pricing. More specifically, the contribution of this study is to extend the empirical evidence on behavioral models that strive to explain return predictability based on the information diffusion hypothesis by studying the limited-information model of Menzly and Ozbas (2010) and its predictions.[5] In addition, this paper provides new insight on the role of investor geographic specialization, and the resulting market segmentation, which remains untested in earlier academic literature concerning return predictability based on gradual diffusion of information.

The methodology used in this paper is mainly consistent with Menzly and Ozbas (2010). However, from a sample point of view there are notable differences as this study uses cross-country European data over a time period that is different from the original study. Given that there is only a limited number of previous studies on limited-information models with many of them focusing on US data[6], this paper provides a geographically and temporally out-of-sample testing of the model by Menzly and Ozbas (2010). For example, Chui, Titman and Wei (2010) point out that, cross-country tests on international markets offer a robustness check on results acquired from excessively mined US data. Moreover, a parsimonious theoretical model should be capable of explaining patterns in different contexts (Daniel, Hirshleifer and Subrahmanyam, 1998) and be invariant to changes in the economic environment (Zin, 2002).

A central contribution of this study is to provide empirical evidence on a novel aspect of investor specialization in the context of limited information models, namely investor specialization along geographic boundaries. More specifically, using a self-constructed measure that captures the geographic dispersion of information diffusing from an industry's customer and supplier industries, this paper provides new evidence on geographic boundaries being a fundamental driver of investor specialization and thus, an important determinant of the magnitude of return cross-predictability. In a similar vein, this paper is able to provide further evidence on investor specialization being the primary source of information diffusion and return cross-predictability. The cross-country sample used in this paper allows examining this novel aspect of investor specialization and its impact on the magnitude of return cross-predictability.

[5] Other prominent models which build upon gradual information diffusion hypothesis include Hong and Stein (1999) and Hong, Torous, and Valkanov (2007).

[6] For limited-information model studies on US market see, Hong, Torous and Valkanov (2007), Cohen and Frazzini (2008) and Menzly and Ozbas (2010). Hong et al. (2007) also report results for eight largest non-US stock markets. An exception to US based studies is a working paper by Shahrur, Becker and Rosenfeld (2009) who use purely international data in their study. See subsection 2.1.4 for more information on all of these studies.

The practical motivation of this study is to provide evidence on the profitability trading strategies that capitalize on return cross-predictability in Europe. For example, Menzly and Ozbas (2010) study US stocks over a period from 1973 to 2005 and report mean annual excess returns of 8.7% from cross-predictability based trading strategies. In addition, Shahrur, Becker and Rosenfeld (2009) use international data and document annual excess returns of 15% on a similar strategy that buys and sells stocks based on previous-month customer industry returns. Furthermore, despite the controversy and disagreement on the sources profits of return predictability based strategies, Menzly and Ozbas show that cross-predictability strategies are also used in practice by providing evidence of long/short equity hedge funds engaging in trading strategies that exploit cross-predictability effects.

The contribution of this paper is thus threefold: (i) provide an out-of-sample test of the model by Menzly and Ozbas (2010) and the underlying gradual information diffusion hypothesis (Hong and Stein, 1999), (ii) provide new insight into investor specialization in gradual information diffusion literature by studying the impact of investor geographic specialization on return cross-predictability and (iii) provide evidence on the profitability of cross-predictability based trading strategies in Europe.

1.4 MAIN FINDINGS AND LIMITATIONS OF THE STUDY

The empirical evidence in this paper provides partial support to the limited-information models and the underlying gradual information diffusion hypothesis. Previous-month returns in supplier industries cross-predict stock returns in both Eurozone and EU27 samples after controlling for factors known to predict stock returns.[7] The supplier industry cross-predictability effects are also robust to adjusting stock returns for country-level differences in overall stock market development and the exclusion of outliers. These findings are consistent with the limited-information models and earlier empirical evidence by Menzly and Ozbas (2010).

On the other hand, previous-month returns in customer industries exhibit only weak cross-predictability effects in both samples. More specifically, customer industry cross-predictability effects are statistically significant only when outliers are eliminated. Furthermore, the results suggest that customer industry cross-predictability effects are considerably weaker in the Eurozone sample than in the EU27 sample. This novel empirical evidence, which is among the most interesting findings in this paper, poses a challenge to the limited-information models which predict that returns on economically related assets cross-predict each other. Moreover, they are not in line

[7] The control variables used are short-term reversals (see e.g., Jegadeesh; 1990; Lehmann, 1990), momentum (see e.g., Jegadeesh and Titman, 1993), industry momentum (Moskowitz and Grinblatt, 1999) and excess return on the country-specific market portfolio.

with earlier empirical evidence which documents statistically significant customer industry cross-predictability effects in the US (Menzly and Ozbas; 2010) and in international markets (Shahrur, Becker and Rosenfeld, 2009).[8]

In addition, using firm size and analyst coverage as proxies for the number of informed investors, this paper shows that the magnitude of return cross-predictability is negatively related to the level of information in the market. More specifically, stocks with the highest number of informed investors exhibit no cross-predictability effects at conventional levels while stocks with the lowest number of informed investors tend to exhibit two to three times stronger cross-predictability effects than the average stock. These findings are consistent with the limited-information model predictions and previous empirical evidence by Menzly and Ozbas (2010) and Sharur et al. (2009).

A central contribution of this paper is to provide empirical evidence on a novel aspect of investor specialization in the context of limited information models, namely the influence of investor specialization along geographic boundaries. More specifically, using a self-constructed measure to proxy for geographic dispersion of related industry information, this paper shows that the magnitude of return cross-predictability is positively related to the geographic dispersion of informative signals from related industries. The findings in this paper suggest that geographic boundaries may be a fundamental driver of investor specialization and hence, an important determinant of the magnitude of cross-predictability effects. The role of investor geographic specialization studied in this paper remains untested in previous literature concerning limited-information models and information diffusion. The findings are in line with the limited-information model hypothesis of investor specialization, and the resulting market segmentation, being the source of return predictability.

Finally, this paper documents that self-financing trading strategies that capitalize on the cross-predictability effects are able to generate abnormal returns also in a European cross-country sample. The trading strategies that buy (sell) industries based on previous-month related industry returns are able to generate up to 9.7% mean annual excess returns during the sample period. Furthermore, these returns exhibit low exposure to the four-factor model risk factors — market, SMB, HML, and MOM (Fama and French, 1993; Carhart, 1997). The findings in this paper may have practical implications for European portfolio managers regarding construction of investment strategies. The

[8] It is worth noting that also Menzly and Ozbas (2010) find that customer industry cross-predictability effects are weaker than supplier industry cross-predictability effects, which is consistent with the findings in this study. Moreover, it should be noted that Sharur, Becker and Rosenfeld (2009) omit supplier industry returns in their regression model which is not consistent with the limited-information models which posit that returns in both supplier and customer industries cross-predict stock returns. Thus, their model possibly suffers from omitted variable bias as customer industry returns take up some of the explanatory power of the omitted supplier industry returns and hence, appear to be statistically significant.

findings are also in line with Menzly and Ozbas (2010) who find that similar trading strategies yield up to 8.7% mean annual excess returns in the US over a period from 1973 to 2005. In addition, Shahrur, Becker and Rosenfeld (2009) document mean annual excess returns of 15% on an international trading strategy that exploits customer industry cross-predictability effects.

It is also important to note the limitations of this study when interpreting the results. First of all, the Eurostat input-output framework, which used in this study, classifies industries based on the General Industrial Classification of Economic Activities within the European Communities revision 1.1 (NACE rev 1.1). In the analysis, each sample firm is assigned to a NACE industry based on primary SIC codes that are retrieved from Thomson One Banker database.[9] Firms that operate in multiple different industries pose a conceptual problem as the primary SIC codes assign each firm to only one industry. Therefore, an important limitation of this study is the ability of primary SIC codes to assign each firm into an industry that correctly represents its main business activities and economic exposures. However, empirical evidence by Menzly and Ozbas (2010) alleviates this concern as their results show that cross-predictability results are similar across single- and multi-segment firms.

Another limitation of this study is the limited sample time period that covers ten years ranging from January 2000 to December 2009. A longer time period would, for example, allow examining whether the magnitude of cross-predictability effect persists through time and different economic cycles. However, the choice of sample time period is due to data quality considerations concerning the consolidated Eurostat input-output tables (see subsection 4.3). In addition, data availability on European listed companies is better in more recent years, which also supports the choice of a more recent sample period.

Furthermore, this paper uses the four-factor model (Fama and French, 1993; Carhart, 1997) to examine the risk factor exposure of returns from cross-predictability based trading strategy that are implemented across countries. It is important to note that it is outside the scope of this paper to conduct an analysis on the applicability of the four-factor model in the cross-country sample used in this study. For a discussion on global and local aspects concerning the four-factor model see e.g., Fama and French (2011).

Finally, as this study uses cross-country data to examine the limited-information models, one potential concern is that unobserved country-specific differences are affecting the test results.

[9] Primary SIC code represents a company's business activity with the largest percentage of sales revenue. The primary SIC codes are converted into NACE rev 1.1 using a concordance table retrieved from Eurostat. See subsection 4.2 for a description of the conversion process.

However, the use of European data should alleviate this potential drawback as Europe comprises a relatively developed market area with a high degree of interdependence between different countries. In addition, the use of cross-country data ensures a sufficient sample size, provides out-of-sample evidence on the results obtained using US data and allows investigating the influence of investor specialization along geographic boundaries.

1.5 STRUCTURE OF THE STUDY

This study is organized as follows. Section 2 covers the related academic literature. Section 3 develops the hypotheses studied in this paper based on the limited-information models and earlier empirical evidence. Section 4 describes the data and sample used in this study. Section 5 presents the methodology. Section 6 presents and discusses the results of this study. Section 7 concludes and presents suggestions for future research.

2 LITERATURE REVIEW

This section aims to provide a brief review of the existing academic literature and empirical evidence concerning return predictability. First, key empirical findings on stock return autocorrelation, momentum and stock-return cross-autocorrelation are reviewed. Thereafter, the most prominent traditional and behavioral explanations for return cross-autocorrelation are presented along with supporting and contradicting empirical evidence.

An important caveat for the reader to acknowledge is that this section does not aim to provide an exhaustive review on return predictability and its suggested explanations. Rather, this section focuses on the most prominent research and empirical evidence put forth in existing academic literature that coincide with the gradual information diffusion model (Hong and Stein, 1999) and the limited-information models (Hong, Torous and Valkanov, 2007; Menzly and Ozbas, 2010).[10]

2.1 RETURN PREDICTABILITY

The reason why return predictability has sparked such considerable interest among the academics lies in its contradiction with the efficient market hypothesis (EMH). EMH posits that security prices fully reflect all relevant information available to the market and that no investment strategy should earn average returns that are greater than warranted for the risk taken. On the other hand, empirical evidence suggests that returns exhibit predictable patterns which implies that it may be possible to make abnormal returns in the market only by conditioning on past price information, thereby appearing to defy even the weak-form of efficient market hypothesis.

[10] For an extensive review of studies on the predictability of common stock returns, see Hawawini and Keim (1995).

This section reviews the empirical evidence on return predictability. The section is organized as follows. Subsection 2.1.1 reviews the empirical evidence on individual stock autocorrelation, more specifically short-term and long-term return reversals. Subsection 2.1.2 presents literature on individual stock momentum. Subsection 2.1.3 covers literature on lead-lag effects, that is, cross-autocorrelation in stock returns. Subsection 2.1.4 introduces the literature on return cross-autocorrelation based on economic linkages.

2.1.1 Individual Stock Return Autocorrelation

Purpose of this subsection is to provide an overview of main academic research and findings related to individual stock return autocorrelation. The two specific types of return autocorrelation covered here are short-term and long-term return reversals.

Individual stock return autocorrelation is a time-series phenomenon which refers to correlation between a stock's own past and future returns (Lewellen, 2002). Return autocorrelation is an interesting concept as it defies the theory of random walks, a building block of the efficient market hypothesis, which posits that successive stock price changes are independent and identically distributed random variables (Fama, 1965). Interestingly, there is evidence from almost a half-a-century ago suggesting that individual stock returns exhibit negative serial correlation over short time horizons. Fama (1965) is one of the first to document that individual stock returns tend to exhibit negative autocorrelation over the short-term. However, he does not consider the magnitude of the observed serial correlation to be economically significant and concludes that his results are supportive of the random walk hypothesis.

In addition to the seminal work by Fama (1965), more recent studies also find evidence on negative return autocorrelation in the short-term. For example, French and Roll (1986) report significant negative serial correlation in daily stock returns. More specifically, they study daily returns on NYSE and AMEX listed stocks over the period of 1963 to 1982 and find that individual stock prices exhibit negative serial correlations between lags of two to 12 days. They interpret their results as supportive of the noise trade hypothesis in which the process of trading introduces noise into stock returns and causes possible investor overreaction. However, French and Roll point out that it is difficult to determine economic significance of the phenomena since the estimates are small in magnitude. Furthermore, Lo and MacKinlay (1988) note that while short-term autocorrelation is in contradiction with the random walk hypothesis, it does not necessarily imply that the stock markets are inefficient.

Inspired by the findings on short-term reversals both Lehmann (1990) and Jegadeesh (1990) investigate the economic significance of the negative short-term autocorrelations. More specifically, both authors examine the profitability of contrarian trading strategies[11] and find evidence of significant negative serial correlation in stock returns (i.e., short-term reversals). Jegadeesh (1990) documents profits of about 2% per month from a contrarian strategy that buys and sells stocks based on their previous-month returns and holds them for one month. In a similar vein, Lehmann (1990) constructs a strategy with weekly portfolio formation and holding periods which is also able to generate significant abnormal returns. Both authors interpret their results as suggestive of an inefficient market where stock prices overreact to information. However, in subsequent article, Jegadeesh and Titman (1991) point out that contrarian strategies are transaction intensive and based on short-term price movements. Based on this, they conclude that the apparent profits reported in earlier studies may be compensation for carrying inventory risk and may reflect the presence of short-term price pressure or illiquidity in the market rather than being suggestive of market inefficiencies.

Perhaps a more prominent strand of return reversal literature has focused on long run return reversals. More specifically, there is empirical evidence suggesting that stocks that perform the worst (best) over a three to five year period tend to perform well (poorly) over the subsequent three to five year period. One of the most influential papers on long-term reversals is by DeBondt and Thaler (1985) who study market efficiency by examining whether people's tendency to "overreact" to unexpected news affects stock prices. They study contrarian trading strategies that buy past losers (defined as companies with the lowest stock returns over the past three to five years) and sell past winners (defined as companies with the highest stock returns over the past three to five years). Using a sample of NYSE traded stocks from 1926 to 1982, DeBondt and Thaler find that 36 months after portfolio formation the prior losers outperform prior winners by about 25%. They interpret the systematic price reversals as supportive of the overreaction hypothesis and as evidence of substantial market inefficiencies.

Subsequent papers argue that De Bondt and Thaler's results are subject to various methodological problems (see e.g., Chan, 1988; Ball and Kothari, 1989). In order to address some of the doubts that their initial results on long run return reversals sparked, DeBondt and Thaler (1987) re-evaluate their overreaction hypothesis finding that it is robust to two alternative hypothesis concerning firm size and differences in risk. In addition, Chopra, Lakonishok and Ritter (1992) provide further

[11] Contrarian strategies are strategies that capitalize on the presence of negative autocorrelation by buying losers (securities that have performed poorly in the past) and selling winners (securities that have performed well in the past) (Lo and MacKinlay, 1990a).

support to DeBondt and Thaler's (1985, 1987) findings by studying the overreaction hypothesis using data on NYSE listed companies over the period of 1926 to 1986. Constructing portfolios based on prior five-year returns, they find evidence of an economically important overreaction effect. More specifically, they find that extreme prior losers tend to outperform extreme prior winners by 5-10% percent per year during the subsequent five years following portfolio formation. Chopra et al. argue that their results are robust to tax-loss selling effects and mismeasurement of risk. Furthermore, they find that the overreaction effect is pronounced for smaller firms which they interpret as evidence of overreaction being more prevalent among individuals than institutional investors.

Complementing the findings from US based studies, Richards (1997) examines long-term return reversals in international markets. He studies winner-loser reversals in national stock market indices using data on 16 countries[12] over the period of 1970 to 1995. Richards finds evidence of prior losers outperforming prior winners at longer horizons and argues that the reversals are not driven by only small markets and cannot be explained by differences in risk. In a more recent study, Bhojraj and Swaminathan (2006) examine 38 national stock indices and find that past 6-month losers (defined as national stock indices with the lowest returns) outperform past 6-month winners (defined as national stock indices with the highest returns) two to three years after portfolio formation. They also rule out differences in risk as the source of the return reversals.

Finally, it is worth noting that further evidence on long-term return reversals is also provided in studies that examine stock return momentum (see e.g., Jegadeesh and Titman, 1993; Chan, Jegadeesh and Lakonishok, 1996; Rouwenhorst, 1998; Chui, Titman and Wei, 2000). These and several other studies on momentum are covered in the next subsection.

2.1.2 Individual Stock Momentum

This subsection reviews literature and evidence on individual stock return momentum. While stock return autocorrelation that was reviewed in the previous subsection is a time-series phenomenon, stock return momentum is a cross-sectional phenomenon (Lewellen, 2002). More specifically, individual stock momentum, or stock return continuation, refers to the tendency of past winners to continue outperform past losers in medium-term horizon in the cross-section of stock returns.

Since its discovery, individual stock momentum has been a source of ongoing debate among academics. Several researchers have shown that stocks that perform the best (worst) over a three to

[12] The sample countries included in the study by Richards (1977) are Australia, Austria, Canada, Denmark, France, Germany, Hong Kong, Italy, Japan, the Netherlands, Norway, Spain, Sweden, Switzerland, the United Kingdom, and the United States. The national stock market indices used in the study are the Morgan Stanley Capital International (MSCI) total return indices.

12 month period tend to continue perform well (poorly) over the subsequent three to 12 months. Moreover, there is historical evidence from various markets showing that momentum strategies that capitalize on this phenomenon have been profitable. In a seminal article Jegadeesh and Titman (1993) examine market efficiency using data on NYSE and AMEX listed stocks over a period ranging from 1965 to 1989. They analyze relative strength strategies that select stocks based on their return over the past one to four quarters and then hold the stocks over an equivalent time period. Jegadeesh and Titman find that strategies that buy stocks with high returns (winners) over the previous three to 12 months and sell stocks with low returns (losers) over the same time period earn positive returns of about 1% per month over the subsequent three- to 12-month holding periods. They argue that the profitability of the strategies is not attributable to systematic risk or lead-lag effects[13] that result from delayed price reactions to common factors. Rather, they claim that the observed excess returns are due to delayed price reaction to firm-specific information.

As mentioned earlier, empirical research on return predictability is often overshadowed by accusations of data mining and the findings by Jegadeesh and Titman (1993) provide no exception. It is noteworthy that the data mining accusations by proponents of the efficient market hypothesis may be well-founded. After all, as noted by Barberis and Thaler (2002), if stocks are sorted and ranked in enough different ways, it is more than likely that some striking but completely spurious cross-sectional differences in average returns are found. However, an effective way to reduce concerns regarding spurious results caused by data mining is to perform out-of-sample tests. Therefore, in order to address the data mining issue and provide further support to their previous findings, Jegadeesh and Titman (2001) reexamine their original trading strategy using data from NYSE, AMEX and NASDAQ stocks reaching from 1965 to 1998. Their results show that the profitability of momentum strategies has persisted in the US in the 1990s, and that past winners continue to outperform past losers by about the same magnitude as in the earlier period.

In addition to research conducted on the US equity markets, momentum strategies have been found to generate statistically significant returns also in international stock markets. Rouwenhorst (1998) examines momentum in 12 European countries[14] in the period of 1978 to 1995 and documents that an internationally diversified relative strength portfolio of past winners outperforms a portfolio of losers by more than 1% per month. In addition, he finds that momentum is present in all sample countries for both small and large firms, although it is stronger for smaller firms. Moreover, momentum lasts on average for approximately one year after which there is evidence of the return

[13] Lead-lag effects are covered in subsection 2.1.3.

[14] The sample countries included in the study by Rouwenhorst (1998) are Austria, Belgium, Denmark, France, Germany, Italy, The Netherlands, Norway, Spain, Sweden, Switzerland and the United Kingdom.

continuation effect at least partially reversing in the second year after portfolio formation. As Rouwenhorst points out, his findings are remarkably similar to those by Jegadeesh and Titman (1993) in the US market. In a subsequent study, Rouwenhorst (1999a) examines momentum strategies in 20 emerging markets[15]. His findings indicate that also emerging markets exhibit momentum and that factors driving cross-sectional differences in expected stock returns in emerging stock markets are qualitatively similar to those in developed stock markets.

Even further evidence on the existence of momentum in international markets is provided by Chui, Titman and Wei (2000) who study momentum strategies in eight Asian markets over a time period ranging from 1972 to 2000.[16] They find that the momentum strategies are highly profitable when implemented across the sample stock markets excluding Japan. Interestingly, they also find that the distinction between civil and common law countries provides a perfect indicator of whether a market exhibits momentum prior to the financial crisis.[17] Chui et al. opine that the absence of momentum in civil law countries might be due to weaker enforcement of security laws in these countries which in turn enables more frequent market manipulation that potentially offsets the momentum effect.

In their related study Chui, Titman and Wei (2010) further expand the literature on momentum in international markets by examining cross-country differences in momentum profits. They find that momentum strategies with global data provide even higher Sharpe ratios than reported in previous studies using US data. In addition, they study the impact of country-specific differences on the profitability of momentum strategies. More specifically, they study the impact of culture, financial market development and institutional quality, and information uncertainty on momentum strategy returns. Chui et al. measure cultural differences using Hofstede's (2001) individualism index, which they argue is related to oveconfidence and self-attribution biases, and find that individualism is positively associated with magnitude of momentum profits. Regarding the two other groups of country-specific factors, they find that some of these variables have significant explanatory power on momentum profits, but their effect is largely subsumed by the individualism factor.

---

[15] The sample countries included in the study by Rouwenhorst (1999a) are Argentina, Brazil, Chile, Colombia, Greece, Indonesia, India, Jordan, Korea, Malaysia, Mexico, Nigeria, Pakistan, Philippines, Portugal, Taiwan, Thailand, Turkey, Venezuela and Zimbabwe.

[16] The sample countries included in the study by Chui, Titman and Wei (2000) are Hong Kong, Indonesia, Japan, Malaysia, Singapore, Taiwan and Thailand. It is noteworthy that there is some variation in the sample periods between different countries due to data availability issues.

[17] In the study, Hong Kong, Malaysia, Singapore and Thailand are defined as common law countries and Indonesia, Japan, Korea and Taiwan as civil law countries. The distinction between different legal origins is based on La Porta, Lopez-de-Silanes, Shleifer and Vishny (1998, 2000)

To complete the discussion on momentum, it is useful to mention a closely related pattern of return continuation around earnings announcements, that is, earnings momentum or post-earnings announcement drift. Earnings momentum refers to the finding that firms reporting unexpectedly high earnings tend to outperform firms reporting unexpectedly low earnings (Chordia and Shivakumar, 2006). Ball and Brown (1968) were among the first to document that, following earnings announcements, cumulative abnormal returns continue to drift up for firms with good earnings news and down for firms with bad earnings news. In subsequent articles, Bernard and Thomas (1989, 1990) confirm the robustness of Ball and Brown's findings and provide more evidence on the sluggish market response around earnings announcement. More specifically, they find evidence of stock prices failing to reflect fully the implications of current earnings for future earnings by studying past earnings momentum and future returns of NYSE and AMEX listed stocks over a period from 1974 to 1986.

Chan, Jegadeesh and Lakonishok (1996) relate the evidence on price momentum to earnings momentum by investigating whether the profitability of momentum strategies is entirely due to market's underreaction to earnings-related information on medium-horizons. Analyzing their sample of NYSE, AMEX and NASDAQ listed stocks over the period of 1977 to 1993 they find that both past return and past earnings surprise predict large drifts in future returns after controlling for the other. Furthermore, price momentum effect appears to be stronger and longer-lived than the earnings momentum effect. Chan et al. also show that momentum strategies are profitable even among larger stocks and their profitability cannot be explained by the Fama-French three-factor model. When reflecting on the underlying causes of momentum, they conclude that the phenomenon may be driven by the market's gradual response to new information which is evidenced by sluggish revisions of analyst forecasts.

Chordia and Shivakumar (2006) also study the relation between price and earnings momentum. Following standard practice in earnings momentum literature, they construct a zero-investment portfolio that buys the highest earnings surprise portfolio and shorts the lowest earnings portfolio. The authors find that their zero-investment portfolio captures the price momentum phenomenon in both time-series and cross-sectional asset pricing tests. They conclude that momentum and post-event drift appear to be manifestations of the same underlying mechanism, in other words price momentum anomaly is a manifestation of the earnings momentum anomaly.

2.1.3 Stock Return Cross-Autocorrelation

This subsection reviews literature and empirical evidence on stock return cross-autocorrelations also known as lead-lag effects. The major difference to the phenomena covered in the previous sections, namely individual stock autocorrelation and individual stock momentum, is that in cross-autocorrelation the future returns of a stock are correlated with past returns of other stocks and not the stock's own past returns. Thus, lead-lag effect in the equity markets refers to the tendency of some firms' stock prices to exhibit a delayed reaction to price innovations of other firms (Hou, 2007). It is worth noting also that lead-lag effects are asymmetric e.g., returns on large firms lead return on small firms, but not vice versa (Lo and MacKinlay, 1990a).

The first and most influential article on stock return cross-autocorrelations is by Lo and MacKinlay (1990a) who study the profitability of contrarian trading strategies and the overrreaction hypothesis by De Bondt and Thaler (1985). Analyzing a sample of NYSE and AMEX listed stocks Lo and MacKinlay find that positive autocorrelations in portfolio returns are due to positive cross-autocorrelations between individual stock returns. More specifically, they find a pronounced lead-lag structure in which the returns of large-capitalization stocks almost always lead those of smaller stocks but not vice versa. This lead-lag effect appears to account for majority of the observed profitability of contrarian trading strategies which Lo and MacKinlay interpret as evidence against stock market overreaction being the only explanation for the profitability of contrarian strategies. The authors note that their empirical results on the lead-lag effect may be partially a symptom of nonsynchronous trading or thin trading, but point out that these are unlikely to fully explain the observed cross-autocorrelations.

The findings by Lo and MacKinlay (1990a) are striking as they undermine the overreaction hypothesis used to explain observed profitability of contrarian strategies. Prior to their article, the standard explanation for abnormal returns from short-horizon contrarian strategies was negative serial correlation in individual stock returns, which was interpreted as evidence of stock market overreaction to information (see e.g., De Bondt and Thaler, 1985, 1987; Jegadeesh, 1990). In an attempt to provide a better understanding of the various sources of profits of contrarian strategies, Jegadeesh and Titman (1995) examine the contribution of stock price overreaction and delayed reaction to the profitability of contrarian strategies. Their results indicate that stock prices on average react with a delay to common factors, but overreact to firm-specific information. The authors conclude that the delayed reaction to common information gives rise to the size-related lead-lag relation in stock returns found by Lo and MacKinlay (1990a), but contributes only little to contrarian profits which are mainly driven by the reversal of the firm-specific component of returns.

An interesting aspect of the cross-autocorrelation finding by Lo and MacKinlay (1990a) is that it implies a complex process of information transmission between firms. In their article, Lo and MacKinlay suggest that the lead-lag effects are due to the tendency of smaller stocks to adjust more slowly to common information but offer no explanation for why size may be an important determinant of the speed of adjustment. Brennan, Jegadeesh and Swaminathan (1993) provide more evidence on the economic basis for the cross-autocorrelation in equity returns by examining the effect of informed investors on the speed with which a firm's stock price adjusts to new common information. Controlling for firm size and using analyst coverage as a proxy for informed investors, Brennan et al. find that stocks with higher analyst following tend to lead stocks with lower analyst following. They interpret their findings as suggestive of many-analyst firms adjusting faster to common information than few-analyst firms.

In a similar vein, Badrinath, Kale and Noe (1995) study transmission of information between securities in an attempt to explain the economic basis for the cross-autocorrelation in equity returns. The authors argue that firm size per se may have little economic significance for information transmission across firms. Instead, they suggest that the lagged information transmission underlying cross-autocorrelation of stock returns is mainly due to markets being informationally segmented. Using data on institutional ownership, they show that returns on high-institutional ownership stocks lead returns on low-institutional ownership stocks by up to two months and that this result holds even after controlling for firm size. Badrinath et al. attribute their findings to information set-up costs and legal restrictions surrounding the investment activity of institutional portfolio managers which cause the returns on low-institutional ownership stocks to slowly adjust to price innovations in high-institutional ownership stocks. Based on their findings, the authors suggest that different levels of institutional interest in stocks may be the primary path for information transmission in the markets and thus, induce cross-autocorrelation in equity returns.

Chordia and Swaminathan (2000) also study the speed of adjustment hypothesis as an explanation for the lead-lag effects. Using data on NYSE and AMEX traded stocks over the period of 1963 to 1996 they find that trading volume is a significant determinant of the cross-autocorrelation observed in stock returns. More specifically, they find that the daily and weekly returns on high volume portfolios predict low volume portfolio returns even after controlling for firm size. Furthermore, they show that the observed lead-lag effect is robust to stock own autocorrelation and market microstructure biases such as nonsynchronous trading and thin trading. Chordia et al. conclude that their findings are consistent with the speed of adjustment hypothesis which posits that lead-lag effects are due to some stocks adjusting more slowly (quickly) to common information.

Building on the previous empirical research, Hou (2007) argues that if the lead-lag effect is driven by slow diffusion of common information, it should be largely an intra-industry phenomenon as common information is often industry-specific. His sample consists of NYSE, AMEX and NASDAQ listed stocks over a time period from 1963 to 2001. In order to test his hypothesis, Hou examines industry portfolios constructed by assigning stocks into 12 industries based on four-digit SIC codes and finds evidence of lead-lag effects being predominantly an intra-industry phenomenon. He also documents that larger firms and firms with higher analyst coverage, institutional ownership and trading volume lead other firms within an industry, thus providing further support for the findings by Lo and MacKinlay (1990a), Brennan et al. (1993), Badrinath et al. (1995) and Chordia and Swaminathan (2000). Furthermore, Hou finds that high market share firms lead firms with low market shares and firms with lower analyst dispersion lead firms with higher analyst dispersion. He concludes that the results are mainly consistent with the hypothesis that information gradually diffuses across firms.

2.1.4 Stock Return Cross-Autocorrelation and Economic Links

The limited-information models, studied in this paper, are designed to explain return cross-predictability between economically linked assets. In order to provide the reader with an understanding of the findings underlying the limited-information models, this subsection briefly reviews literature and empirical evidence on return predictability between economically related assets. The main difference to the lead-lag literature reviewed in the previous section is that instead of focusing on stocks with different characteristics leading or lagging one another, this strand of literature focuses on return predictability that is caused by closely specified economic links between assets. In addition to economic links, another important ingredient in this literature is the gradual diffusion of information across markets caused by investors' limited information-processing capabilities and investor specialization.

In their article, Cohen and Frazzini (2008) provide a case example of return predictability based on a customer–supplier link that clearly illustrates the rationale underlying the limited-information models. The authors examine the stock market behavior of two companies, namely Coastcast Corporation, a leading manufacturer of golf club heads, and its major customer Callaway Golf, a retail company that specializes in golf equipment. In 2001, Callaway Golf had been a major customer of Coastcast for almost ten years and accounted for 50% of Coastcast's total sales. When Callaway announced in June 2001 that its second quarter earnings projections would be $50 million less than previously anticipated, the firm's mean EPS forecast was revised down to 35 cents (previous estimate was 70 cents per share) and its share price fell 30%. Surprisingly, the negative news about Callaway's future earnings had no immediate impact on Coastcast's share price or EPS forecasts. Cohen and Frazzini note that even though investors eventually woke up to the news and both Coastcast's share price and EPS forecasts fell dramatically, an investor who would have reacted to Callaway's announcement of slowing demand by shorting Coastcast would have earned returns of 20% over the subsequent two months. The pattern of return predictability between economically linked assets illustrated by Cohen and Frazzini also underlies the limited-information models. More specifically, some investors, due to limited information-processing capabilities and specialization on a subset of assets, fail to process informative signals from related firms and adjust their demand accordingly, which leads to cross-predictability of returns.

Motivated by the example of Callaway and Coastcast, Cohen and Frazzini (2008) investigate whether cross-asset return predictability between economically related stocks is a pervasive phenomenon in the stock markets. More specifically, they focus on customer–supplier links between US firms using COMPUSTAT's customer information database to identify principal customers for their sample firms over a period from 1980 to 2004. In order to test for return predictability, the authors construct a monthly long-short trading strategy that buys (sells) firms whose customers had the most positive (negative) returns in the previous month. Their trading strategy yields abnormal returns of 1.55% per month. Cohen and Frazzini refer to this return predictability as "customer momentum" and show that their results are robust to own autocorrelation, industry momentum, lead-lag effects and cross-industry momentum. They also investigate the impact of mutual fund joint holdings of customer and supplier firms on customer momentum and find that return predictability is significantly greater (weaker) when inattention is likely to be greater (smaller). They take their findings as evidence that investor limited attention can lead to return predictability across assets.

Motivated by models on limited investor information-processing capacity and market segmentation, Hong, Torous, and Valkanov (2007) construct a limited-information model (see subsection 2.2.2.3) that explains cross-asset return predictability. The authors study the speed of adjustment hypothesis laid out in previous studies on lead-lag effects by examining whether gradual diffusion of information across asset markets leads to return predictability between economically linked assets. In order to provide evidence on their model, they analyze returns on industry portfolios in US stock market and eight largest stock markets outside the US.[18] Hong et al. find that several important

[18] The eight non-US stock markets in the study by Hong, Torous and Valkanov (2007) are Australia, Canada, France, Germany, Japan, Netherlands, Switzerland and the United Kingdom. It is important to note that there are differences in sample periods for these countries. In addition, the authors were unable to obtain the same set of control variables for these markets that they used for the US market. Also Eleswarapu and Tiwari (1996) investigate the ability of industry returns to predict the future market returns. Their

industries such as commercial real estate, petroleum, metal, retail, financial, and services predict the future returns on the overall stock market. Moreover, they also note that an industry's predictive power is strongly correlated with its ability to forecast economic indicators. They interpret these results as supportive of their model and that gradual diffusion of information across the market is the source of cross-asset return predictability. Although Hong et al. focus on industry portfolios and the broad market index, they also suggest that cross-asset return predictability may exist in various contexts between economically linked assets.

In an article closely related to Hong et al. (2007), Menzly and Ozbas (2010) also study return cross-predictability focusing on economic boundaries defined by industries.[19] Menzly and Ozbas use Benchmark Input-Output Surveys of the Bureau of Economic Analysis (BEA) to define supplier and customer industries for their sample of NYSE, AMEX and NASDAQ listed stocks over the period from 1963 to 2005. Testing for return predictability, they find that lagged returns on supplier and customer industries cross-predict returns at both industry- and stock-level for up to a year. In addition, they construct a self-financing trading strategy that buys (sells) industries with the highest (lowest) previous-month related industry returns. Importantly, the trading strategy is able to generate annual premiums as high as 8.7% which cannot be explained by the Fama-French-Carhart four-factor model. Menzly and Ozbas conclude that their findings are consistent with the gradual information diffusion hypothesis by Hong and Stein (1999) and argue that informational segmentation of the markets due to investor and/or information producer specialization is the underlying reason for the slow diffusion of information across markets.

In addition to Menzly and Ozbas (2010), also Shahrur, Becker and Rosenfeld (2009) use BEA input-output surveys to determine industry relatedness. Instead of focusing on the whole supply chain, Shahrur et al. identify customer relations between industries using an international sample containing 22 developed markets[20] over the period of 1995 to 2007. In the spirit of Cohen and Frazzini (2008), they construct a trading strategy that consists of buying (selling) supplier firms with the highest (lowest) customer industry returns in the previous month. Their trading strategy is able to generate up to 15% annual abnormal returns. Furthermore, the predictability effect appears to be more pronounced for smaller stocks, supplier industries with dispersed sales and higher

---

results indicate that the stock returns of basic industry, construction, consumer durables, food and tobacco, and textiles and trade industry portfolios are significant predictors of the return on the equally weighted NYSE index

[19] Menzly and Ozbas (2010) note that cross-predictability effects are different from lead-lag effects. In the latter case, there exists one market that either always leads or always lags another market whereas with cross-predictability, a market can sometimes lead and sometimes lag another related market depending on where the information originates.

[20] The sample countries included in the study by Shahrur, Becker and Rosenfeld (2009) are Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Hong Kong, Ireland, Italy, Japan, Netherlands, New Zealand, Norway, Portugal, Singapore, Spain, Sweden, Switzerland, and the United Kingdom.

relationship-specific investments with their customers, and in markets that are less financially integrated with the world. In addition, Shahrur et al. test for cross-sectional variation and find that the customer-supplier return predictability is not driven by small and illiquid stocks. The authors conclude that their results are consistent with gradual diffusion of information being the source of return predictability.

2.2 EXPLANATIONS FOR RETURN CROSS-AUTOCORRELATION

The objective of this subsection is to briefly present the most prominent explanations and models suggested to explain return cross-autocorrelation. It is worth noting that some of the explanations reviewed also apply to other forms of return predictability. Subsection 2.2.1 presents the traditional explanations that attempt to explain return cross-autocorrelation based on exposure to economically meaningful risk factors. Subsection 2.2.2 reviews literature and evidence on behavioral models that strive to explain return predictability based on investor psychology and limits to arbitrage.

2.2.1 Traditional Explanations

This subsection briefly reviews the traditional explanations for the stock return cross-autocorrelation patterns. According to Chordia and Swaminathan (2000), these explanations can be divided into two groups: (i) the first group of explanations argues that cross-autocorrelations are due to time-varying expected returns while (ii) the other group argues that autocorrelations and cross-autocorrelations result from various market microstructure biases. First, the arguments of group (i) are covered along with empirical evidence. Then the arguments of group (ii) are reviewed.

As mentioned earlier, some authors interpret the evidence on asymmetric cross-autocorrelations, first documented by Lo and MacKinlay (1990a), to be consistent with the gradual information diffusion hypothesis. However, there is another group of academics who claim that these cross-autocorrelation patterns can also be consistent with an alternative hypothesis based on time-varying expected returns. More specifically, proponents of the latter explanation argue that variation in risk factors, such as past market returns, past size returns, or interest rate spreads can induce variation in short-horizon risk premiums (Boudoukh, Richardson and Whitelaw, 1994). This explanation is appealing because it posits that returns from investment strategies that exploit times-series patterns in returns are not abnormal and thus, do not violate the efficient market hypothesis.

For example, Conrad and Kaul (1988) argue that predictable components of returns could reflect changes in expected returns, that is, portfolio autocorrelation is caused by changing expected returns instead of market inefficiencies. The authors study ten size-based portfolios over the 1962 to 1985 period and provide empirical evidence on their hypothesis by showing that expected returns explain up to 26% of the variance in realized returns. Moreover, they document that the magnitude of this proportion is inversely related to firm size. In a more recent article, Mech (1993) tests the time-varying expected returns hypothesis by Conrad and Kaul. Contradictory to the earlier results, Mech finds evidence of portfolio autocorrelation being a result of mispricing, rather than autocorrelation of expected returns. Based on his findings, he concludes that portfolio return autocorrelation is inconsistent with a strict interpretation of the efficient market hypothesis.

Another explanation that is related to the time-varying expected returns hypothesis posits that cross-autocorrelations are simply a restatement of own autocorrelations and contemporaneous correlations between assets. For example, Boudoukh, Richardson and Whitelaw (1994) argue that the cross-autocorrelations between big firms and small firms are the result of the high own autocorrelations of small firms combined with the high contemporaneous correlation between big and small firms. Under this explanation, the lagged returns of large firms serve as noisy proxies for the lagged returns of small firms and the lead-lag effect will disappear once controlled for lagged small-firm returns. However, Chordia and Swaminathan (2000) reject the explanation of Boudoukh et al. (1994) by showing that returns on high volume portfolios lead returns on low trading volume portfolios even after controlling for own autocorrelations. In addition, Boudoukh et al. are unable to provide compelling empirical evidence on the sources of own autocorrelation that they document. However, they argue that market microstructure biases, particularly nonsynchronous trading, are important determinants of the magnitude of autocorrelations in portfolio returns.

As mentioned earlier, the second group of explanations argues that return autocorrelations arise from market microstructure biases. Particularly, nonsynchronously sampled data is an often cited source of return cross-autocorrelation under this group of explanations. Nonsynchronous sampling refers to measurement errors in data arising from nonsynchronous trading and thin trading problems which may cause spurious autocorrelation and cross-autocorrelation. More specifically, nonsynchronous trading refers to the fact that listed securities are traded at different intervals which causes their prices to be reported at distinct random intervals. In a time-series analysis, the nonsynchronous trading problem arises when prices of distinct securities are mistakenly assumed to be sampled simultaneously, that is, when nonsynchronous prices are treated as if they were observed at the same time (Lo and MacKinlay, 1990a). Fisher (1966) and Scholes and Williams (1977) were among the first to point out that nonsynchronous trading can induce positive autocorrelation in stock returns. In later studies, Boudoukh et al. (1994) confirm these findings by showing that nonsynchronous trading may induce return autocorrelation when actual returns are not autocorrelated. Thin trading or nontrading is closely related to nonsynchronous trading and refers to the fact that some stocks are more heavily traded than others which causes the prices of thinly traded stocks to react to news with a lag. For example, Lo and MacKinlay (1990b) show that lead-lag effects are a natural consequence of thin trading. Thin trading problem has essentially same implications as the nonsynchronous trading problem and it arises when the prices of thinly traded securities are mistakenly considered to be sampled simultaneously.

Empirical research on cross-autocorrelations has addressed the nonsynchronous and thin trading problems. For example, Lo and MacKinley (1990a) show that although their lead-lag finding could be the result of infrequent trading in small stocks, this interpretation would require unrealistically high nontrading frequencies. In addition, Chordia and Swaminathan (2000) and Hou (2007) both address the nonsynchronity issue in their articles. Both papers use weekly returns to avoid the confounding microstructure influences and show that nonsynchronous trading cannot be the only source of the cross-autocorrelation findings. Chordia and Swaminathan (2000) also impose a trading requirement on stocks to be included in their sample for further robustness. Finally, also Menzly and Ozbas (2010) provide empirical evidence that return cross-predictability is not driven by nonsynchronous or thin trading problems. They address the problem of nonsynchronous sampling by using monthly return data and further excluding stocks without a closing price in the previous month. They also exclude stocks with market capitalizations below the 20th NYSE percentile to address the possibility that thin markets might be driving their results.

Another market microstructure related explanation argues that transaction-costs may result in slow price adjustment of stocks (Mech, 1993). This explanation argues that even though return predictability violates market efficiency, it does not necessarily imply that investors are less than rational if transaction costs cause them to postpone their investment decisions. In his article, Mech (1993) develops a transaction-cost model which predicts that stock prices adjust faster when changes in valuation are large in relation to the bid-ask spread. He also provides empirical evidence that supports his model's predictions in cross-sectional tests but not in time-series tests.

2.2.2 Behavioral Explanations

As mentioned earlier, the purpose of this literature review is to set the limited-information models, into a proper context. With this objective in mind, subsection 2.2.2.1 briefly presents evidence on investor specialization and market segmentation, which is an important building block of limited-information models.[21] Subsection 2.2.2.2 covers the gradual information diffusion model and the

[21] Investor psychology, particularly limited information-processing capacity of investors, is another important ingredient in the limited-information models. However, covering this literature is outside the scope of this study. For an extensive review on investor psychology and asset pricing, see Hirshleifer (2001).

limited-information models. Finally, subsection 2.2.2.3 presents the previous empirical evidence on these models.

Behavioral finance is a rather recent approach to financial markets that has emerged, at least to some extent, in response to the difficulties faced by the traditional approach. Prior to looking at the models and empirical evidence, it is useful to understand the main concepts behind behavioral finance. Behavioral finance argues that some financial phenomena can be better understood by relaxing the assumption of fully rational economic agents[22]. The field has two distinct fronts: (i) limits to arbitrage, which argues that it can be difficult for rational traders to arbitrage away the dislocations caused by less rational traders; and (ii) psychology, which sets out the various deviations from full rationality in order to understand the origins of market mispricings. (Barberis and Thaler, 2002.)

As mentioned earlier, there exists a large variety of behavioral models that entertain different approaches to explaining the specific nature of the patterns of predictability. The gradual information diffusion model and the limited information models presented in this subsection comprise only a fraction of all these models. Furthermore, it is important to note that the patterns of stock return predictability discussed in this literature review constitute only one part of behavioral finance. For a more complete overview of the field of behavioral finance, see Barberis and Thaler (2002).

2.2.2.1 Investor Specialization and Market Segmentation

This subsection briefly presents evidence on investor specialization and market segmentation which are important building blocks of limited-information models.

One of the first theoretical approaches to segmented markets is the limited participation model of Merton (1987). In his model, investors face fixed costs to acquiring information concerning asset returns which causes them to trade only a limited number of stocks that they have information on. As a result, stocks that are less recognized by investors have a smaller investor base (neglected stocks) and trade at a discount stemming from limited risk sharing. Among other things Merton also recognizes that information can diffuse at different speeds among investors and that information diffusion may be incomplete which can lead to an empirically significant effect on asset returns. More specifically, the stock's idiosyncratic risk and investor base emerge as additional factors in explaining returns. Since Merton (1987) there has been subsequent literature on segmented markets

[22] By rationality, two things are considered. Firstly, when agents receive new information, they update their beliefs correctly as described by Bayes' law. Secondly, given their beliefs, agents make normatively acceptable choices, in the sense that they are in line with Savage's notion of Subjective Expected Utility. (Barberis and Thaler, 2002.)

and limited market participation providing more evidence on investor specialization (see e.g., Allen and Gale, 1994).

In addition to papers examining investor specialization across industries and stocks, there is also a growing literature on the importance of geography in investor specialization. Many researchers have provided evidence on the phenomenon known as home bias, in which investors shun foreign stocks in their portfolios. Perhaps the most prominent study which documents the home country bias is by French and Poterba (1991). They estimate that US, Japan, and UK investors hold 93%, 98%, and 82% of their equity investments, respectively, in their domestic markets, and argue that these numbers are inconsistent with standard models of asset allocation.

Initial explanations for the home country bias have focused on barriers to international investment and the effect of national borders in discouraging investment abroad. However, it is useful to note that home bias is not only restricted to international capital markets, but there exists home bias even within national borders. For example, Coval and Moskowitz (1999) report that the typical equity portfolio of a US investment manager consists of stocks of firms that are located 100 miles closer to the manager's office than the average US firm. They interpret this as evidence that geographic proximity, in addition to national borders, is an important driver of home country bias. In their subsequent study, Coval and Moskowitz (2001) attempt to shed more light on the underlying determinants of home bias. Examining the geographic link between mutual fund investments and performance, they find evidence of fund managers earning substantial abnormal returns in geographically proximate investments. Based on these findings, they argue that geographic proximity is inversely related to the cost of information acquisition and suggest that investors have an informational advantage to trading local securities.

In a related study Massa and Simonov (2006) examine individual portfolio holdings in Sweden and also find evidence supporting informational advantage of local investors over nonlocal investors. They suggest that investors deliberately tilt their portfolio towards stocks that are most closely related to them due to local information that they have on these stocks. However, contradictory to the previous findings of Coval and Moskowitz (2001) and Massa and Simonov (2006), Seasholes and Zhu (2010) find that portfolios of local holdings do not generate abnormal performance and conclude that on average individuals do not have value-relevant information about local stocks. Based on their findings, the authors argue that informational advantage concerning local stocks does not seem to be key driver of individual investors making local investment choices. Even though the root causes of the home bias effect remain unknown, some authors suggest it may be part of a larger phenomenon in which investors exhibit preference for familiar companies (see e.g., Huberman, 2001; Grinblatt and Keloharju (2001).

Another strand of literature that is closely related to investor specialization examines the specialization of equity analysts. For example, Kini et al. (2003) study analyst specialization across the world and find that analysts specialize among both countries and industries. More specifically, using a sample of 7,106 analysts across 45 countries and 11 sectors over the period 1995-1996, they find that 86% of analysts in their sample are single-country analysts whereas 39% focus on a single sector. Furthermore, they find that this specialization can be explained by economies of scale in information acquisition and production. Importantly, the results by Kini et al. indicate that specialization along geographic boundaries is more pervasive than specialization along economic boundaries (e.g., industries) among equity analysts.

In a related study, Bolliger (2004) examines analyst specialization in 14 European countries[23] over the period of 1988–1999 and finds that European financial analysts cover an average of 1.5 different countries and 2.4 industries. This finding confirms the results by Kini et al. by showing that geographic specialization of analysts is pervasive also in Europe. In addition, Bolliger also observes a negative link between analyst forecast accuracy and the degree of international diversification of analyst portfolios. Based on this result, he hypothesizes that having to interpret signals from multiple countries is a difficult task for analysts. Similar to Bolliger also Malloy (2005) finds that geographically proximate analysts possess an information advantage over other analysts, and that this advantage translates into better performance. Importantly, the above finding suggests that it is more difficult for informed investors to process all relevant information when it needs to be gathered from a geographically dispersed area.

2.2.2.2 Gradual Information Diffusion and Limited-Information Models

This subsection presents the limited-information models of Hong, Torous and Valkanov (HTV, 2007) and Menzly and Ozbas (MO, 2010) that strive to explain return cross-autocorrelation. Also the two behavioral models underlying these limited-information models namely the positive feedback trading model of De Long, Shleifer, Summers and Waldmann (DSSW, 1990) and the gradual information diffusion model of Hong and Stein (HS, 1999) are briefly presented. First, this subsection provides an overview of the models and the linkages between the different models. Then each model is explained in more detail.

---

[23] The sample countries in the study by Bolliger (2004) include Austria, Belgium, Denmark, Finland, France, Ireland, Italy, Germany, Netherlands, Norway, Spain, Sweden, Switzerland and the United Kingdom.

The gradual information diffusion model of HS (1999) is the most important behavioral model underlying the limited-information models of HTV (2007) and MO (2010). Common feature of all three models is that they strive to explain time-series return predictability based on gradual diffusion of information. Furthermore, all three models attribute the diffusion of information on some type of market segmentation which leads to disagreement between groups of investors concerning asset values. While the HS model does not specify the source of information diffusion, the limited-information models attribute the gradual diffusion of information to investor specialization, and resulting segmentation of markets. Therefore, empirical evidence on the validity of limited-information models is also evidence on the information diffusion hypothesis of HS. The DSSW (1990) model preceded the HS (1999) model and both models have similarity in their concepts. Thus also the DSSW model is briefly covered in this subsection.

De Long, Shleifer, Summers and Waldmann (1990) were among first to formally model how the actions of irrational investors can affect asset prices. Their model challenges the general view that rational investors are able to dampen the fluctuations caused by so-called noise traders. In their article, De Long et al. show how rational speculation can actually be de-stabilizing in the presence of noise traders with positive feedback trading strategies (i.e., buying when prices increase and selling when prices fall) if actions by rational speculators trigger positive feedback trading. DSSW conclude that their positive feedback trading model is consistent with empirical findings such as stock price overreaction to news, price bubbles, positive correlation of asset returns in the short-term and long run reversals. For example, their model posits that positive feedback traders respond to past price increases by entering the market and bidding up prices which causes positive short-term return correlation. The subsequent reversal in the long run occurs when prices eventually return to fundamentals.

Positive feedback trading type of trend-chasing behavior also plays a central role in the model of Hong and Stein (1999) which attributes stock return predictability to gradual diffusion of information. More specifically, HS develop a dynamic model of a single asset in which information gradually diffuses across investors, some of whom are unable to extract information from prices. As stated earlier, the HS model builds upon the empirical findings on momentum and long-term reversals and aspires to form a unified behavioral model that captures both phenomena.[24] Main difference to DSSW (1990) is that while the positive-feedback traders in the DSSW model are

[24] Another behavioral model by Daniel, Hirshleifer and Subrahmanyam (DHS, 1998) also strives to provide an integrated theory that explains several anomalies, including momentum and reversals, based on positive feedback trading. Their model attributes the observed return predictability patterns to investor overconfidence and self-attribution biases.

extremely irrational, the trend-chasing momentum traders in HS model are nearly rational and are able to take advantage of the other group of traders.

The gradual information diffusion model by Hong and Stein (1999) builds upon empirical findings on momentum and reversals. The model emphasizes interaction between two boundedly rational groups of investors: newswatchers and momentum traders. Bounded rationality here implies that each group of investors can process only a subset of the publicly available information. Newswatchers make forecasts based on private information that they observe about future fundamentals and cannot condition on current or past prices. On the other hand, momentum traders can only make simple predictions based on the past price changes. Moreover, HS model assumes that private information diffuses across the newswatcher population.

In their article, HS (1999) show that when only newswatchers are active in the market there is underreaction to news and stock prices exhibit momentum which is caused by the slow diffusion of information across the market. Importantly, in this setting, there is only underreaction but never overreaction. However, when momentum traders are added to model, they strive to take advantage of the underreaction with a simple arbitrage strategy. As a result, the initial underreaction to news becomes weaker but stock prices begin to exhibit overreaction in the longer run. This overreaction follows from momentum traders' inability to directly condition on news about fundamentals in order to determine whether or not they are buying at a price above or below the long-run equilibrium. It is important to note, that the gradual diffusion of information about fundamentals is the main driver that explains both underreaction and overreaction in the model.

The limited-information processing capacity of investors from the HS model also underlies the limited-information model by Hong, Torous and Valkanov (2007). While the HS model strives to explain momentum and reversals, the HTV model is designed to generate return cross-autocorrelation between economically linked assets. Put differently, the HTV model aims to explain why information diffuses across markets and causes cross-serial correlation of returns whereas the HS model focuses on information diffusion across investors within the same market and the resulting own autocorrelation of returns. In order to meet its objective, the HTV model introduces investor specialization and market segmentation that drive information diffusion and cause returns in markets with correlated fundamentals to exhibit cross-predictability. Importantly, HTV argue that investor limited participation is a pervasive feature of financial markets and, in addition to limited information-processing capacities of investors, a central reason for why information diffuses slowly across markets.

The basic setup of the HTV model is that there are two markets X and Y which have specialized investors. The model makes two central assumptions regarding investor behavior: (i) the limited market participation assumption which posits that investors participate in either market X or market Y and (ii) the limited information-processing capacity assumption which says that investors in market X cannot process information pertaining to market Y, and vice versa. HTV argue that the assumption (i) can be motivated by exogenous reasons such as taxes or regulations or alternatively, by fixed cost of participation in each market. However, HTV point out that this assumption is not crucial to their model as long as there are limits to arbitrage, which ensure that some cross-predictability will remain. Assumption (ii) implies that investors have limited cognitive capabilities and therefore, have hard time processing information from asset markets in which they do not participate. One motivation for this assumption is that information from other markets is less salient. Alternatively, investors already have their hands full trying to figure out the market that they are in and thus, do not process information from other markets in a timely fashion. In a sense, assumption (ii) in HTV model is similar to the bounded rationality assumption by HS (1999). To synthesize, the HTV model posits that investors specialize in markets about which they receive informative signals but due to limited-information processing capacities are unable to efficiently capture corresponding signals from other markets that they do not specialize in and, as a result, returns on fundamentally correlated markets exhibit serial cross-correlation. HTV note that their model can easily be augmented to simultaneously generate own autocorrelations and cross-serial correlations if the HS (1999) assumption of gradual information diffusion within a market is included in the model.

Inspired by HS (1999) and HTV (2007) models, Menzly and Ozbas (2010) construct their own limited-information model which is the one studied in this paper. Similar to HTV, their model strives to explain how the specialization of investors in their information gathering efforts leads to informationally segmented markets and subsequently, cross-predictability in asset returns. Also in the MO model, dispersed information diffuses slowly across markets with correlated fundamentals and causes returns on economically linked assets to cross-predict each other. The MO model formally extends the HTV model into two directions (i) by introducing uninformed investors and (ii) by relaxing the assumption that informed investors invest only in the market about which they acquire informative signals (assumption (i) in HTV model). This setup allows Menzly and Ozbas to study the joint behavior of stock returns and informed trade across related markets.

First, MO present how returns exhibit autocorrelation in their model by considering a single asset market. They introduce two investor groups, uninformed investors and informed investors, whose actions cause returns to exhibit predictability based on publicly available information.[25] In their model, informed investors receive informative signals about the eventual cash flows on an asset and adjust their demand for the asset correspondingly. On the other hand, uninformed investors do not receive the informative signal and at least some of them are also incapable to infer the signal from publicly available information. This failure of some uninformed investors to process publicly available information is due to either their limited information-processing capabilities or the costs of processing information that they face. In this setting, MO show that returns exhibit predictability based on publicly available information. When an informative signal arrives, informed investors adjust their demand but the information is not completely incorporated in the prices. This partial adjustment of prices is due to limited risk-bearing capacity of informed investors and a failure by uninformed investors to recover the informative signal from the observed price in order to adjust their demand. As a result of this failure by some uninformed investors, returns on assets exhibit positive autocorrelation.

After establishing the basis for autocorrelation of returns in their model, MO turn to return cross-predictability. They introduce two markets in which both informed and uninformed investors are able to invest. MO further note that in order to have return cross-predictability with two asset markets, two additional assumptions are needed: i) the two markets need to have correlated fundamentals, which causes an informative signal in one market to have value-relevant information about the eventual payoff in the other market, and (ii) the two markets need to be informationally segmented as informed investors specialize along market boundaries in their information gathering activities. Informed investors specialize in one of the markets for which, at some point in time, they receive informative signals about the eventual cash flows. The specialization of informed investors causes an informative signal originating in one of the markets to be received only by those investors who specialize in that market, whereas those investors who specialize in the other market are not able to recover this signal. This results in return cross-predictability as informative signals with cross-market content are incorporated into prices only partially. In other words, prices exhibit cross-predictability for the same reasons that they exhibit autocorrelation.

2.2.2.3 Empirical Evidence on Gradual Information Diffusion and Limited-Information Models

This subsection presents results from empirical tests of the gradual information diffusion and limited-information models in order to provide insight on the validity of these models.

[25] This distinction is very much in the spirit of the HS (1999) model which has two groups of boundedly rational investors whose interactions cause return predictability. MO (2010) suggest that the distinction between uninformed and informed investors is caused by skill-based differences in information acquisition and processing costs. These differences result in heterogeneous beliefs between investors and lead to a situation in which some investors with information acquisition and processing costs below a certain threshold choose to become informed while others choose to remain uninformed.

As mentioned earlier, Huberman and Regev (2001) provide one of the cleanest examples of gradual information diffusion and investor specialization in the stock market. In their article, the authors examine the stock-market behavior of a single biotechnology firm, EntreMed, around the announcement of value-relevant information concerning the company's future prospects. More specifically, they find that a front-page story in the New York Times that announced a potential breakthrough in the development of EntreMed's new cancer-curing drugs caused the firm's stock price to soar from 12 dollars to 85 dollars over night. Remarkable about their finding is that the New York Times story contained essentially no real news as the substance of the story had been published five months earlier in the scientific journal Nature, as well as in various other media including the Times itself.

Furthermore, Huberman and Regev show that while the earlier news stories were also associated with jumps in both EntreMed's stock price and trading volume, these increases weren't nearly as dramatic as the ones induced by the front-page Times story more than five months later. In addition, their results suggest that although there was evidence of short-run overreaction, the impact of the front-page Times story on EntreMed's stock price was to a large extent permanent. The authors interpret the evidence as suggestive that there are two groups of investors in EntreMed: a small group of experts who read publications like Nature and a larger group of generalists who receive their information from sources such as the front page of the New York Times. In this setting, information gradually diffuses as a result informational segmentation of markets caused by investor specialization and the stock price responds to news with a delay. The evidence by Huberman and Regev, although circumstantial, is consistent with both the gradual information diffusion hypothesis of Hong and Stein (1999) and the limited-information models.

In another article, Jegadeesh and Titman (2001) test the empirical predictions of the Hong and Stein (1999) gradual information diffusion model which posits that momentum profits are due to delayed overreactions that are eventually reversed.[26] In order to study the HS model, the authors construct momentum portfolios based on stocks' past 6 month performance and analyze the return patterns over a post-holding period of 60 months using US data over 1965 to 1998. Jegadeesh and Titman document that returns on momentum portfolios are negative in the 13 to 60 months following portfolio formation which is consistent with the gradual information diffusion model. They also contrast their finding of negative post-holding period returns with the presumption of Conrad and Kaul (1998), who claim that momentum profits are due to cross-sectional variation in expected

[26] Jegadeesh and Titman also study the predictions of two other behavioral models, namely a model by Barberis, Shleifer and Vishny (1998) and a model by Daniel, Hirshleifer and Subrahmanyam (1998). Both models have similar predictions to the Hong and Stein (1999) model concerning momentum and reversals.

returns which causes returns on momentum portfolios to be positive on average in any post-ranking period. In addition, Jegadeesh and Titman point out that the negative post-holding period returns depend on the composition of the sample e.g., sample period and whether small stocks are included or not. Based on their results, they conclude that the HS model at least partially explains the momentum effect.

Also Hong, Lim and Stein (2000) study the gradual information diffusion model of HS (1999). They divide stocks based on firm size and analyst coverage into different groups for which the speed of information diffusion should vary and test the HS model prediction that stocks with slower information diffusion should exhibit stronger momentum. Using a sample of NYSE, AMEX and NASDAQ listed firms over 1976 to 1996, Hong et al. find that the profitability of momentum strategies declines sharply with firm size once one moves past the very smallest stocks. Moreover, they find that controlling for size, momentum strategies are more profitable among stocks with low analyst coverage. Finally, they also document that low-coverage stocks react more sluggishly to bad news, suggesting that particularly negative information diffuses gradually across the investing public.

Hong, Lim and Stein (2000) interpret their results as supportive of the HS (1999) model in which information diffuses gradually across the market. They also suggest that the slow diffusion of negative information is due to managers' lower information disclosure incentives when bad news, as opposed to good news, needs to be published to the market. The lower disclosure incentives of managers cause the role of outside analysts to be pronounced when the news is bad. Lesmond, Schill and Zhou (2004) question the interpretation of Hong et al. by arguing that the variables used to proxy for the information diffusion speed are highly correlated with trading costs. They claim that the observed size effect by Hong et al. is actually a manifestation of a price friction effect rather than evidence supportive of the gradual information diffusion hypothesis. However, Lesmond et al. are unable to provide any conclusive evidence of analyst coverage being related to transaction costs.

Also Yalcin (2008) tests the gradual information diffusion model of HS (1999) by studying the model prediction which posits that reversal strategy returns should be more pronounced among stocks with slower information diffusion. Yalcin's sample comprises of NYSE, AMEX and NASDAQ listed stocks over 1980 to 2004. Using size-adjusted analyst coverage as a proxy for the speed of information diffusion, he shows that the profitability of contrarian strategies declines monotonically with increasing rates of information diffusion. Furthermore, he shows that the results are robust to choice of time period and sample composition and continue to hold even after controlling for the four-factor model risk factors. Yalcin concludes that his findings are consistent with the HS (1999) model.

While there is a vast amount of empirical research on the HS (1999) model, the evidence on the more recent limited-information models still remains rather scarce. In their article, Hong, Torous and Valkanov (2007) also provide empirical evidence on their limited-information model by showing that positive cross-autocorrelation exists between economically linked units.[27] Analyzing a sample of 34 value-weighted industry portfolios for the years 1946 to 2002 in the US stock market, HTV show that returns on several important industries predict the stock market. More specifically, they document that 14 out of 34 industries including commercial real estate, petroleum, financial and services are able to predict the stock market movements by one month. Furthermore, they show that these results are robust to the inclusion of well-known market predictors such as excess market returns, inflation and the market dividend yield. HTV also report that an industry's predictive power is strongly correlated with its ability to forecast indicators of economic activity such as industrial production growth. They interpret these results as evidence that markets incorporate information contained in industry returns about their fundamentals with a lag and that information diffuses gradually across asset markets. The authors also extend their analysis to eight largest non-US stock markets and find remarkably similar patterns. HTV conclude that their findings are consistent with the gradual information diffusion hypothesis and their limited-information model predictions. They also suggest that cross-asset return predictability should exist in many contexts beyond industry portfolios and the broad market index between sets of assets that have correlated payoffs.

Similar to Hong, Torous and Valkanov (2007), also Menzly and Ozbas (2010) provide empirical evidence on their limited-information model predictions. In their article, MO analysze a sample consisting of NYSE, AMEX and NASDAQ listed stocks over a sample period from 1973 to 2005. Using a series of Benchmark Input-Output Surveys of the Bureau of Economic Analysis (BEA), they identify supplier and customer industries for each stock in their sample and test for cross-predictability effects between these economically linked industries. More specifically, they show that lagged returns on supplier and customer industries predict stock (industry) returns from one month up to 12 months. MO provide further evidence on their limited-information model predictions by showing that the magnitude of cross-predictability effects declines with the number of informed investors as proxied by analyst coverage and institutional ownership. Moreover, they show that changes in institutional ownership at the stock level are positively related to changes in

[27] Hong, Torous and Valkanov (2007) point out that this cross-autocorrelation is different from the previously studied lead-lag effects (see e.g., Lo and MacKinlay, 1990a; Chordia and Swaminathan, 2000) in which stocks of different characteristics lead or lag one another.

institutional ownership in related industries which implies that informed investors (proxied by institutional owners) exploit the cross-market content of their signals. Finally, they show that trading strategies that exploit the cross-predictability effects are able to generate annual premiums as high as 8.7% which cannot be explained by traditional risk factors. Menzly and Ozbas conclude that their findings are consistent with their limited-information model and the gradual information diffusion hypothesis by Hong and Stein (1999). Furthermore, they argue that informational segmentation of the markets due to investor and/or information producer specialization is the underlying reason for the slow diffusion of information across markets.

As mentioned earlier in subsection 2.1.4, also Cohen and Frazzini (2008) and Shahrur, Becker and Rosenfeld (2009) provide empirical evidence on return predictability between economically linked assets. Although not direct tests of the limited-information models, also these results are supportive of gradual information diffusion and investor specialization being the sources of cross-autocorrelation in returns.

3 HYPOTHESIS BUILDING

This section builds on existing literature and presents the hypotheses for this study. The hypotheses are aimed to test the study objectives described in subsection 1.1.

3.1 HYPOTHESIS 1: INDUSTRY RELATEDNESS

The first hypothesis concerns the asset relatedness assumption of the limited-information models. In their articles, Hong, Torous and Valknanov (2007) and Menzly and Ozbas (MO, 2010) both derive their own limited-information models. Basically, both limited-information models predict that asset returns exhibit cross-predictability when two central assumptions are fulfilled: (i) the sets of assets such as stocks in different industries or market segments have correlated fundamentals and (ii) markets are informationally segmented as informed investors, to some degree, specialize along these boundaries in their information-gathering activities (MO, 2010). Under these assumptions value-relevant information diffuses slowly across the markets and causes returns on economically linked units to cross-predict each other.

Empirical evidence on assumption (ii) is provided in the literature review in subsection 2.2.2.1 and providing further proof of this assumption is out of the scope of this study. However, assumption (i) requires a closer examination in order to verify the validity of the empirical design for testing cross-predictability in this paper. As mentioned, this paper uses Eurostat consolidated input-output tables to determine industry relatedness. Since there is no earlier evidence on using Eurostat IO-tables in cross-predictability analysis, it is important to confirm that they provide a meaningful description of industry relatedness for both Eurozone and EU27 samples. Furthermore, there are differences between the Benchmark Input-Output Surveys of the Bureau of Economic Analysis (BEA) that are used in previous studies on return cross-predictability (Shahrur, Becker and Rosenfeld, 2009; Menzly and Ozbas, 2010) and the Eurostat IO-tables. For example, the BEA IO-tables use different industry classifications compared to the Eurostat input-output tables.[28] In addition, there are differences in the granularity of the tables: the BEA IO-tables contain 85 different industry accounts whereas the Eurostat tables contain 59 different industry accounts.

Another reason for studying industry relatedness in this paper is that, unlike previous research on the limited-information models, this paper examines return cross-predictability in a cross-country sample with data on the Eurozone and EU27 countries. Therefore, it needs to be verified that factors, such as national borders, do not blur the economic links between industries. For example, national borders might affect the way in which economic shocks are transmitted across industries in different countries. Importantly, earlier academic literature that compares within-country correlations of business cycles to cross-country correlations identifies a relation between national borders and business cycles, known as the border effect (see e.g., Wynne and Koo, 2000; Clark and van Wincoop, 2001). More specifically, the border effect refers to substantially lower business cycles correlations across countries than within countries. For example, Clark and van Wincoop (2001) document that business cycles in the US Census regions are substantially more synchronized than those of European countries. They also find that the lower level of trade between European countries, in comparison to US regions, accounts for majority of the observed border effect.

Another strand of literature decomposes the sources of within-country and cross-country fluctuations of business cycles into common, country-specific, region-specific, and industry-specific components. Clark and Shin (1998) review this literature and note that, in general, the evidence indicates that European business cycles are less synchronized than those of US regions. They also report that country-specific and idiosyncratic components are responsible for more than two-thirds of total variation in business cycles across countries and show that across US regions the common component is much larger than in Europe. They attribute these findings to the border effect and argue that economic borders that divide nations are greater than economic borders that separate

[28] The Eurostat input-output tables are based on the General Industrial Classification of Economic Activities within the European Communities revision 1.1 (NACE rev 1.1) whereas BEA input-output tables are based on the Standard Industrial Classification (SIC).

regions within nations.[29] In other words, lower economic borders are associated with a lower importance of country-specific disturbances and greater importance of common and industry-specific factors as drivers of business cycles.

There remains a great amount of uncertainty as to whether business cycles within Europe are converging or diverging. On one hand, integration in Europe is increasing as fiscal policies are becoming more coordinated and barriers to cross-border flows of goods, capital and labor are being removed (Clark and van Wincoop, 2001). Particularly, the formation of the Economic and Monetary Union (EMU) and the adoption of the single currency may have lowered economic borders between the euro countries and thus, increased business cycle correlations. On the other hand, for example, Clark and Shin (1998) question the divergence of business cycles within the euro area and argue that despite the creation of EMU, it is likely that the economic borders between EMU-member nations still remain higher than those in the US. De Haan, Inklaar and Jong-A-Pin (2008) provide an extensive review of the literature on business cycle synchronization in the Economic and Monetary Union. They note that business cycles in the euro area have experienced periods of both convergence and divergence and there is no monotone movement towards the emergence of a European business cycle.

Despite the mixed empirical evidence, the recent European sovereign debt crisis has clearly shown that European countries are in very different economic positions. While the Eurozone's southern periphery countries have struggled with out-of-control government debts and high unemployment rates, other euro countries, most importantly Germany, have weathered the crisis much better. Practical observations and empirical evidence on the border effect in Europe highlight the need to verify the limited-information model assumption of correlated fundamentals in the European cross-country sample. The main concern here is that if national borders significantly decrease the correlations between industries across European countries, it may be that return cross-predictability does not exist because the units of analysis are not economically linked. To address this concern, an industry relatedness analysis is performed similar to Menzly and Ozbas (2010). Given the findings by Menzly and Ozbas, it is expected that industry fundamentals in Europe are positively correlated, as defined by the consolidated Eurostat input-output tables.

[29] Factors such as independent monetary and fiscal policies of different nations and restrictions on labor, trade and capital flows create economic borders between nations. Also cultural differences and language strengthen economic borders between nations. This suggests that the US states, which are regions within a nation, are likely to have lower economic borders than the European nations studied in this paper. Even though each US state determines their local fiscal policies they are also affected by national monetary and fiscal policies. Moreover, there are practically no restrictions on trade and capital flows within the US and also cultural differences are small. (Clark and Shin, 1998.)

Building upon the previous literature and the limited information models, the first hypothesis of this study is:

H1: Industry fundamentals are positively correlated over and above the market as measured by firm-, industry- and market-level measures of profitability.

3.2 HYPOTHESIS 2: CROSS-PREDICTABILITY OF STOCK RETURNS

The second hypothesis concerns cross-predictability of stock returns in a cross-country sample.

Based on the objectives stated in subsection 1.1 this paper aims to test whether the limited-information models appear valid in a cross-country sample. In order to reach this objective, this paper examines cross-predictability effects using data on the Eurozone and EU27 countries. Even though the focus is on the limited-information model by Menzly and Ozbas (2010), it should be noted that performing an out-of-sample test of the cross-predictability effect provides insight on the validity of the limited-information models and information diffusion hypothesis in general.

As mentioned earlier, the limited-information model by MO (2010) posits that there are two types of investors, informed and uninformed. Informed investors specialize in one market for which they receive informative signals about the eventual cash flows on assets whereas uninformed investors do not receive any informative signals. In addition, at least some uninformed investors fail to process information due to either limited information-processing capabilities or costs to processing information. The interaction of these two investors generates return cross-predictability between economically linked assets based on publicly available information. In their article, MO provide empirical evidence on their model's predictions by showing that lagged returns in supplier and customer industries cross-predict stock (industry) returns. Furthermore, they show that this cross-predictability effect is economically significant and lasts from one month up to 12 months. More empirical evidence on return cross-predictability is provided by Hong, Torous and Valkanov (2007) who show that several important industries predict the movements of stock markets in multiple countries. Consistent with their model, HTV also document that an industry's predictive ability is strongly correlated with its propensity to forecast indicators of economic activity.[30]

The above findings suggest that cross-predictability effects should exist, as predicted by the limited-information models, also in non-US samples. However, the use of European cross-country data in this paper may induce differences compared to the previous studies. As noted earlier, national borders are potentially a significant factor that may impact the correlation of industry fundamentals

[30] For a detailed description of the limited-information models by Hong, Torous and Valkanov (2007) and Menzly and Ozbas (2010), see subsection 2.2.2.2.

across countries. In addition, to the possible effect of national borders on industry correlations, borders may also have an impact on the way how price innovations are transmitted across industries in different countries. In other words, the cross-predictability effect in this study may differ from the one reported in earlier studies due to country-specific factors. Relevant country-specific factors that may affect stock returns are, for example, differences in transaction costs across stock exchanges and differences in information costs.

Academic literature on stock market comovements has documented that equity markets of many developed and economically integrated economies move to a larger extent independently of each other in terms of returns and volatility.[31] According to Rouwenhorst (1999b), the explanations for the low correlations between country index returns can be divided into three groups: (i) home bias, (ii) country effects and (iii) industry effects. The first explanation attributes the low correlations to investors' tendency to hold disproportionately large amounts of domestic shares in their portfolios which may cause country portfolios to reflect, at least partly, the different sentiments of domestic investors. The second explanation states that country-specific factors such as regional economic shocks, local monetary and fiscal policies and differences in institutional structures cause global economic shocks to have different effects on companies in different countries. In other words, there are country-specific factors that drive equity returns in different countries. The third explanation argues that the low country correlations are driven by differences in industrial structures between countries rather than country-specific differences. For example, Switzerland has a large banking sector and thus, the Swiss country index is imperfectly correlated with the Swedish stock market index which is heavy on the basic industries. (Heston and Rouwenhorst, 1994.)

In a sense, the stock market comovement literature is closely related to the business cycle literature discussed in the previous section. However, as the analysis in the stock market comovement literature is performed on the level of stock returns, reviewing it is important since there are several forces, other than the business cycle, that may affect stock returns across countries. Thus, while providing evidence on the correlation of industry fundamentals may satisfy the limited-information model assumption, the observed cross-predictability effect may still be different in a cross-country sample as compared to a single country sample. For example, if country effects dominate industry effects or if differences in investor sentiment drive a wedge between the returns of companies that are in the same industry but located in different countries, the cross-predictability effect may be weaker in a cross-country sample. This point is highlighted by Heston and Rouwenhorst (1994)

[31] Grubel (1968), Levy and Sarnat (1970) and Solnik (1974) were among the first to document the low correlations between stock market returns in different countries.

who note that country-specific components of return variation may still dominate any industry effects even though the actual correlation between industries might be high.

A vast amount of empirical evidence from stock market comovement literature suggests that industry effects play a little role in explaining the low country correlations of stock returns (see e.g., Heston and Rouwenhorst, 1994; Beckers, et al., 1992). For example, Heston and Rouwenhorst (1994) study the European stock markets and find that country effects dominate industry effects even in European countries that are economically strongly integrated. Even though the industry portfolios in their sample are strongly positively correlated, the industrial structure explains only a small portion of the correlation of country index returns. Griffin and Karolyi (1998) later confirm the findings by Heston and Rouwenhorst by showing that less than 4% of the variation in country index returns can be attributed to their industrial composition. Furthermore, Griffin and Karolyi also document cross-sectional differences in the variances of industry effects for industry indexes. More specifically, they find that traded goods-industries tend to have higher industry effects than nontraded-goods industries.[32] In addition, Bekaert, Hodrick and Zhang (2009) study international asset return comovements by using a linear factor model which they argue is a better model than the widely-applied Heston–Rouwenhorst (1994) dummy variable model. Also their results suggest that country factors dominate industry factors in Europe.

It is worth noting that the empirical evidence on country effects' dominance on international stock return comovements is not unanimous. Roll (1992) was the first one to provide evidence that a significant part of the international stock market comovement can be explained by the industry compositions of the national stock market indices. Although his results have been questioned in later studies, there exists a number of more recent articles which claim that industry factors have become more dominant (see e.g,, Cavaglia, Brightman and Aked, 2000; and Baca, Garbe, and Weiss, 2000). For example, Ferreira and Ferreira (2006) study the Economic and Monetary Union countries over the period 1975 to 2001 and find evidence that industrial effects have become similar in magnitude to country effects in the post-euro period. They also document similar results for non-EMU countries. Also Brooks and Del Negro (2004) provide evidence that industry effects have become increasingly important in Europe. To summarize, based on the earlier research, it remains unclear whether national borders have an impact on the cross-predictability of stock returns.

[32] Traded-goods industries are industries that produce goods which are traded internationally such as automobiles, computers, office equipment, pharmaceuticals and semi-conductors. Nontraded-goods industries are industries that produce goods which are not traded internationally such as media, heavy construction, plantations, conglomerates and real estate. (Griffin and Karolyi, 1998.)

In addition to country and industry effects, researchers have also studied the exposure of the European equity markets to the US equity market. For example, Baele (2005) investigates the magnitude and the time-varying nature of volatility spillovers from aggregate European and US equity market indices to 13 local European equity markets. He documents that, while the relative importance of the regional European market has increased, the US equity market continues to be the dominating influence in European equity markets. He also finds some evidence of contagion effects from the US market to several local European markets in times of high equity market volatility. In another study, Fratzcher (2002) provides evidence that while the USA is the dominant market outside the Eurozone, it is no longer the only dominant market within the Eurozone. His results indicate that the euro area market has become the dominant market for individual Euro area countries since the mid-1990s. The importance of the US equity market to the European equity market may influence the return cross-predictability results obtained in this paper, particularly if the US market has a varying impact on equity markets in different European countries. However, it is extremely difficult to predict the possible influence of the US equity market on the cross-predictability effect.

Based on the above literature and empirical evidence on the cross-predictability effect, the second hypothesis of this study is divided into two parts:

H2.1: Lagged supplier industry returns cross-predict stock (industry) returns in a cross-country European sample.

H2.2: Lagged customer industry returns cross-predict stock (industry) returns in a cross-country European sample.

3.3 HYPOTHESIS 3: EFFECT OF INFORMED INVESTORS AND GEOGRAPHIC SPECIALIZATION

The third hypothesis concerns the magnitude of cross-predictability effects and how it varies with the number of informed investors and geographic dispersion of information.

The limited-information model by Menzly and Ozbas (2010) posits that asset returns exhibit cross-predictability because information diffuses slowly across informationally segmented markets due to the specialization of informed investors. Importantly, the information-impounding demand by informed investors is an important factor that affects the magnitude of return cross-predictability. More specifically, the information-impounding demand increases with the number of informed investors which leaves little left to predict as informative signals are being more or less fully incorporated into prices. Thus, an important prediction of the MO model is that the magnitude of the cross-predictability effect is negatively related to the number of informed investors and the level of information in the market. (MO, 2010.)

Earlier research provides empirical evidence that when the amount of informed investors increases (decreases), the magnitude of return predictability decreases (increases). For example, Hong, Lim, and Stein (2000) study the profitability of momentum strategies using firm size as a proxy for the amount of informed investors and find that the profitability of momentum strategies declines sharply as firm size increases. They interpret this as evidence supporting the gradual information diffusion hypothesis by Hong and Stein (1999) which also underlies the limited-information model of Menzly and Ozbas (2010). In addition, several researchers have used analyst coverage as a proxy for the amount of informed investors. For example, Brennan, Jegadeesh and Swaminathan (1993), Yalcin (2008) and Menzly and Ozbas (2010) all use analyst coverage as a proxy for the speed of information flow and find evidence that supports the gradual information diffusion hypothesis.

As mentioned earlier, limited-information models posit that investors specialize in their information gathering activities only on a subset of assets for which they receive informative signals about eventual cash flows. This specialization of investors in turn leads to informationally segmented markets and consequently, cross-predictability in asset returns (Menzly and Ozbas, 2010).[33] In their article MO study the effect of investor specialization along economic boundaries defined by industries and the resulting cross-predictability effects. However, earlier academic literature shows that, in addition to specializing along economic boundaries, investors also specialize along geographic boundaries in their information gathering activities. For example, Kini, et al. (2003) use international data on analyst following and document that geographic specialization among analysts is even more pervasive than specialization along industries. Similarly, Bolliger (2004) studies analyst specialization in 14 European countries over the period of 1988–1999 and finds that European financial analysts cover on average 1.5 different countries and 2.4 industries. Furthermore, both Bolliger (2004) and Malloy (2005) find evidence that analysts have a harder time processing all relevant information when it needs to be gathered from a geographically dispersed area. Finally, there are several studies that document investors tendency to shun foreign stocks also known as home bias (see e.g., French and Poterba, 1991; Coval and Moskowitz, 1999; Grinblatt

[33] Equilibrium prices exhibit cross-predictability as informed investors who receive informative signals and adjust their demand for the risky asset accordingly do not completely make up for the lack of adjustment in uninformed demand due to limited risk-bearing capacity.

and Keloharju, 2001). The home bias effectively leads to geographic specialization in investors' information gathering activities as domestic stocks are preferred.[34]

Based on prior academic literature, it is evident that investor specialization along geographic boundaries is a prevailing feature of the stocks markets. Therefore, it is possible that geographic boundaries are a fundamental driver of investor specialization and thus, an important determinant of the magnitude of return cross-predictability. This is an indirect implication of the limited-information models which posit that investor specialization causes information to diffuse slowly across the markets and consequently, asset returns to exhibit cross-predictability. More specifically, increased geographic dispersion of a firm's supplier and customer industries causes the informative signals from these related industries to diffuse from a geographically more dispersed area. Since investors are geographically specialized in their information gathering activities, they have a harder time capturing and processing these geographically scattered signals. In other words, investors are more likely to either omit, process slowly and/or misinterpret[35] related industry signals when they diffuse from a geographically more dispersed area. To summarize, investor specialization along geographic boundaries may decrease (increase) the information-impounding demand of informed investors, and increase (decrease) the magnitude of the cross-predictability effect, when information diffuses from more (less) geographically dispersed related industries.

Based on the previous literature and empirical evidence, the third hypothesis of this study can be divided into three parts:

H3.1: Magnitude of the cross-predictability effect decreases (increases) as firm size increases (decreases).

H3.2: Magnitude of the cross-predictability effect decreases (increases) as analyst coverage increases (decreases).

H3.3: Magnitude of the cross-predictability effect increases (decreases) as the geographic dispersion of supplier and customer industries increases (decreases).

[34] See subsection 2.2.2.1 for more on market segmentation and investor and analyst specialization.
[35] The higher likelihood of erratic interpretations may follow, for example, from the difficulty to incorporate country-specific factors into the security analysis (Bolliger, 2004) and/or differences in accounting practices between countries (Capstaff, Paudyal and Rees, 2001).

3.4 HYPOTHESIS 4: ECONOMIC SIGNIFICANCE AND EXPOSURE TO RISK FACTORS

The fourth hypothesis concerns the economic significance of return cross-predictability and its robustness to well-known risk factors.

As mentioned earlier, Menzly and Ozbas (2010) show that cross-predictability based trading strategies are able to generate mean annual excess returns of 8.7% in the US. Furthermore, Shahrur, Becker and Rosenfeld (2009) document mean annual excess returns of 15% on a similar trading strategy that buys (sells) stocks based on previous-month customer industry returns. Finally, Cohen and Frazzini (2008) report annual returns of 18.6% on a strategy that exploits US firm-level customer-supplier data to define strong economic links between firms. In order to test the economic significance of the cross-predictability effect in European markets, this paper constructs self-financing trading strategies that buy (sell) industries based on previous-month related industry returns.

Previous empirical evidence also shows that returns from cross-predictability based trading strategies are fairly orthogonal to well-known risk factors (Shahrur et al., 2009; Menzly and Ozbas, 2010). Therefore, the monthly returns generated by the trading strategies in this paper are regressed on the four-factor model risk factors – market, SMB, HML, and MOM (Fama and French, 1993; Carhart, 1997), in order to determine whether cross-predictability based returns are abnormal or not. The motivation for testing the exposure of trading strategy returns to common risk factors is to address a potential concern that is often related to findings on time-series return predictability, namely that the returns can be explained by risk.

Thus, the fourth hypothesis of this study can be divided into two parts.[36]

H4.1: The cross-predictability effect is economically significant and self-financing trading strategies that capitalize on cross-predictability effects are able to generate abnormal returns.

H4.2: The cross-predictability based trading strategy returns are explained by the four-factor model.

4 DATA

This section provides an overview of the data used in the study. The section is organized as follows. Subsection 4.1 reviews the data collection and screening process. Subsection 4.2 introduces the industry classifications and their matching procedure. Subsection 4.3 presents the Eurostat input-

[36] Hypothesis 4.2 is a counter hypothesis for Hypothesis 4.1.

output framework and the transformation of consolidated industry-by-industry tables. Subsection 4.4 reports the sample characteristics.

4.1 DATA COLLECTION AND SCREENING

The sample used in this paper consists of publicly listed stocks from EU27 countries over the period ranging from January 2000 to December 2009.[37] First, a complete list of publicly traded companies for each country is retrieved from the Thomson One Banker database using the Screening and Analysis – Company Screener function.[38] Then the obtained list is screened for duplicates based on Thomson entity keys. Finally, required data items are retrieved for all companies in the screened list. In the analyses, the EU27 sample is further divided into Eurozone sample which is examined separately.[39] Moreover, the analysis performed in this paper requires two separate data sets to be gathered: (1) accounting data for the industry relatedness analysis and (2) stock market data for the cross-predictability analysis. Due to differences in data availability, the sample composition varies slightly across the two data sets. The required data items for both analyses are explained below in more detail.

For the industry relatedness analysis, annual accounting data items are retrieved for the sample period in order to calculate return on assets (ROA) figures. The firm-level accounting data obtained include operating income, depreciation and amortization and total assets. All accounting data is retrieved from Worldscope database to ensure the consistency of the data across the sample (Worldscope data items *OperatingIncome, DepreciationDepletionandAmortization* and *TotalAssets*, respectively). Several accounting figures from each three category were randomly chosen and crosschecked from company financial statements to verify the validity of the data. In addition, returns on assets over (less than) 100% are excluded from the sample.

For the cross-predictability analysis, monthly stock market data is retrieved for the sample period. Monthly end-of-month return data is obtained from Thomson One Banker database as return indices (Thomson One Banker data item *TotalReturnRaw*) which are used to calculate lognormal monthly returns. Monthly stock returns of over (less than) 100% are excluded from the sample. Monthly

[37] See subsection 4.4 for a detailed description of sample characteristics.

[38] In creating the list of companies, only two search criteria are used in the Thomson Company Screener: isNa(tf.PrivateIndicator) and IsInList(tf.CountryCode, ""). This search yields a complete list of all publicly listed companies. In addition, country codes, Thomson entity keys, SIC codes and primary SIC codes are retrieved for all companies within the same search.

[39] EU 27 countries include Austria, Belgium, Bulgaria, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden and United Kingdom. However, Romania is not included in the final sample due to data availability issues. Eurozone countries include Austria, Belgium, Cyprus, Estonia, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Malta, Netherlands, Portugal, Slovakia, Slovenia and Spain. Motivation for studying both the Eurozone and EU27 is two-fold. First of all, conducting the analyses on two partially different samples allows estimating the robustness of the results. Secondly, there are differences between the samples which may affect the results, for example Eurozone sample contains fewer countries with a higher degree of integration (e.g., common currency and monetary policy) than EU27.

market capitalization data used to calculate value-weighted monthly returns on market and industry portfolios is also retrieved from Thomson One Banker database (Thomson One Banker data item *MarketCap*). Market capitalization is calculated as the number of shares outstanding times share price. Random lookups were performed on the market capitalization figures to ensure their validity.

In addition, as a part of the cross-predictability analysis, data on analyst following is retrieved for the sample companies. The analyst following data comprises of monthly numerical count of analyst EPS estimates included in the mean estimate for a company during a given month. The data on analyst EPS estimates is retrieved from the Institutional Brokers' Estimate System (I/B/E/S) detail database (I/B/E/S data item *EPS#ofEstimates*). The monthly numbers of analyst EPS estimates are used to construct stock-level analyst coverage measure that proxies for the amount of informed investors.

Finally, the one-month Euro Interbank Offered Rate (Euribor) is retrieved for the sample period and used as the risk-free rate in the study.[40] The risk-free rate used in related studies conducted in the US is often the one-month Treasury bill rate (see e.g., Hong, Torous and Valkanov, 2007) and the objective is to use a comparable rate.

4.2 INDUSTRY CLASSIFICATIONS

In order to study the limited-information models, individual firms need to be allocated into industries. The industry classification system used in this paper is the General Industrial Classification of Economic Activities within the European Communities revision 1.1 (NACE rev 1.1) classification system which comprises of 59 separate industry accounts. The choice of industry classifications is dictated by the Eurostat input-output framework which is also based on NACE rev 1.1 industry classifications. The input-output tables are used to determine industry relatedness in this study (see subsection 4.3 for more information on Eurostat input-output framework).

The Thomson One Banker database that is used to collect the sample in this study does not provide NACE codes for firms. To circumvent this problem, primary Standard Industrial Classification (primary SIC) codes[41] are retrieved from the database for each sample firm and then converted into NACE rev 1.1 using a concordance table retrieved from Eurostat.[42] The procedure used to assign NACE codes to sample firms is as follows. First, primary SIC codes are retrieved from Thomson One Banker database for the list of publicly traded companies that was formed using the Company Screener function. Then, each primary SIC code is matched with a NACE rev 1 code using a

---

[40] Euribor rates: http://www.euribor-rates.eu/

[41] Primary SIC code represents a company's business activity with the largest percentage of sales revenue.

[42] Eurostat RAMON: http://ec.europa.eu/eurostat/ramon/index.cfm?TargetUrl=DSP_PUB_WELC

Eurostat concordance table. Finally, the NACE rev 1 codes are matched with NACE rev 1.1 codes using another concordance table to take into account possible changes in the updated industry definitions. This procedure is followed out of necessity as there are no concordance tables that directly convert SIC codes into NACE rev 1.1 codes. Appendix 1 presents the correspondences between NACE rev 1.1 and SIC codes that are used in this study.

Due to the above described matching procedure, some individual firms are excluded from the final sample because either they do not have a primary SIC code or their primary SIC code corresponds to multiple NACE rev 1.1 industry accounts. In addition, consistent with Menzly and Ozbas (2010), closed-end funds and real estate investment trusts are excluded from the sample. The corresponding SIC codes for these industries are 6798 and 6726, respectively. Finally, some industry accounts are excluded from the sample due to insufficient number of firms with available data within the industry.[43] More detailed description of individual industry account sample characteristics is provided in subsection 4.4.

4.3 EUROSTAT INPUT-OUTPUT FRAMEWORK

This subsection provides a brief overview of the Eurostat input-output framework which is used in this study to define industry relatedness. For a detailed description of how the input-output tables are used in the formation of customer and supplier portfolio, see subsection 5.1.

As mentioned earlier, this study uses Eurostat symmetric input-output tables (IO-tables) to identify customer and supplier industries for a given stock and industry.[44] More specifically, this paper uses consolidated Eurostat IO-tables for the aggregate Eurozone and EU27 which provide a detailed picture of the interdependent structure of the European economy by reporting the amount of inter-industry flows of goods and services.[45] The Eurostat input-output-tables are part the European System of Accounts (ESA95) framework which consists of three types of tables: supply tables, use tables and symmetric input-output tables. The supply and use tables (SUT) constitute the core of the ESA95 framework. SUTs are matrices that show the production processes and transactions for particular products and industries by product and industry: the supply table reports where goods and services are produced whereas the use table shows where they are used in intermediate

---

[43] Industry accounts with less than five firms for each month (year) during the sample period are excluded from the sample in the return cross-predictability analysis (industry fundamentals analysis). Minimum industry size of five firms per industry is used also in prior studies (see, e.g., Fan and Lang, 2000; Hong, Torous, and Valkanov, 2007; Ling, Chan, Dasgupta and Gao, 2011).

[44] For other studies that use input-output tables to define industry relatedness, see e.g., Fan and Lang (2000) and Kale and Shahrur (2007).

[45] Eurostat launched a set of projects aimed to combine individual supply and use tables of EU27 countries into aggregated EU-level tables. The outcome of the project was consolidated supply and use tables for EU27and the Eurozone which are transformed in this study to symmetric input-output tables. The use of consolidated EU-level input-output tables provides an accurate description of industry relations within EU. For more information on consolidated EU-level tables, see EUROSTAT (2011).

consumption, final consumption, gross capital formation and export. The SUTs as such serve mainly for statistical purposes, but when transformed into symmetric input-output tables, they provide valuable insight on, for example, the interdependencies between industries. (EUROSTAT, 2008.)

The symmetric Eurostat input-output tables used in this study are industry-by-industry matrices which show the flow of goods and services between industries. Basically, the IO-tables combine the supply and use tables into a single table with identical classification of industries applied to both rows and columns. Since consolidated EU-level industry-by-industry IO-tables are not readily available in the Eurostat database, this paper transforms consolidated EU-level SUTs into EU-level industry-by-industry IO-tables. It is important to note that compiling input-output tables on the basis of supply and use tables is an analytical step, which requires various assumptions. This paper follows EUROSTAT (2008) and uses a transformation method known as fixed product sales structure transformation model, (also referred to as industry-by-industry variant of the industry technology assumption in literature) to transform the SUTs into industry-by-industry IO-tables.[46] This transformation procedure is based on the assumption that each product has its own specific sales structure, irrespective of the industry where it is produced. The transformed input-output table for EU27 with few selected industries is presented in Appendix 2 in order to provide the reader with a better understanding of the tables.[47]

The ESA95 framework requires that each EU-member country transmits their individual country-level supply and use tables annually and their input-output tables every five years. For the purposes of this study, industry-by-industry input-output tables are constructed for both EU27 and the Eurozone for years 2000 and 2005.[48] The IO-tables are used on a rolling basis to measure supplier and customer relations between industries. This approach is adopted primarily to improve measurement accuracy because each table provides a historical snapshot and thus, may be inadequate for describing the structure of the European economy for the entire sample period. Hence, in the analysis industry relations data from the 2000 Eurostat IO-table is used between 2000 and 2004 and data from 2005 Eurostat IO-table is used between 2005 and 2009.

---

[46] An alternative transformation method for constructing industry-by-industry input-output tables would require the assumption of fixed industry sales structure. However this model is rejected due to its unrealistic character of the alternative assumption of fixed industry sales structure. In addition, the fixed product sales structure model that is used in this paper is also widely applied by countries such as Denmark, Hungary, the Netherlands, and Finland. (EUROSTAT 2008.)

[47] For more information on the ESA95 framework and a detailed description of the transformation process, see EUROSTAT (2008).

[48] Years 2000 and 2005 are chosen based on data quality considerations. EUROSTAT (2011) notes that the data situation for European countries is best for years 2000 and 2005 due to the five-yearly data transmission of symmetric input-output tables required by ESA95 framework.

4.4 SAMPLE CHARACTERISTICS

This subsection presents the general characteristics of the Eurozone and EU27 samples examined in this paper. Both samples consist of publicly listed companies traded over January 2000 to December 2009. It is important to note that as the companies are not required to exist over the entire sample period, the realized sample size varies at any point in time. Furthermore, the analysis in this paper can be divided into two sections which use different samples: industry relatedness analysis and cross-predictability analysis. This subsection first discusses the characteristics of the samples used in the industry relatedness analysis. Then the sample characteristics of the cross-predictability samples are reviewed.

Table 1 shows the average country-level return on assets (ROA) statistics for all sample countries along with supplementary figures. As can be seen, average firm profitability in the Eurozone and EU27 samples is approximately the same during the sample period. However, there is notable variation in average ROAs between individual countries as the average firm profitability ranges from 1.57% (Cyprus) to 11.20% (Hungary). In addition, there appears to be some variation in average firm size across the sample countries. Particularly, Latvia seems to have a low median firm size (€8.67 million) which may partially be explained by the limited number of Latvian firms with available accounting data in the Worldscope database.

Table 1 also includes the total number of different firms and industries included in the sample by country. The Eurozone sample consists of 2,946 listed firms from 47 different industries and the EU27 sample contains 5,636 listed firms from 51 different industries. It is useful to note that large countries dominate the sample. More specifically, France and Germany represent approximately 23% and 24% of the Eurozone sample, respectively while the United Kingdom is the single largest country representing 28% of the EU27 sample. The country-specific differences between the sample sizes are due to industries with less than five firms with available data for the whole sample period being excluded from the samples. For more detailed information, Appendix 3 presents the annual country-level ROAs and sample size statistics for each sample country by year. The rather large variation in country-level ROAs, apparent in the appendix, suggests that country-specific differences such as nonsynchronized business cycles may have an impact on individual firm profitability and thus, may need to be taken into account in the industry relatedness analysis.[49]

[49] See subsection 5.2 for a description of country-specific adjustments made to individual firms' return on assets figures.

Table 1: Descriptive Statistics for Industry Relatedness Sample

This table presents country-level return on assets (ROA) statistics for all sample countries. The figures are reported for both Eurozone and EU27 samples separately. All data items are retrieved from Worldscope database and the industries are based on NACE rev 1.1 industry classifications. Mean ROA is the simple average of annual country-level ROAs calculated over the the whole sample period ranging from year 2000 to 2009. Annual country-level ROA is calculated by weighting all individual firm ROAs with firm assets for a given country and year. Median total assets is the median value of total assets calculated over all sample companies within a country and expressed in millions of euros. Number of firms is the total number of different firms and number of industries is the total number of different industries included in the sample for a given country. These figures are slightly different for the Eurozone and EU27 samples due to industries with less than five firms being exlcuded from the samples.

				Eurozone		EU27	
Country	Eurozone	Mean ROA	Median Total Assets (€M)	No. of Firms	No. of Industries	No. of Firms	No. of Industries
Austria	Yes	5.25 %	211.69	96	31	99	33
Belgium	Yes	6.46 %	198.11	131	34	136	35
Cyprus	Yes	1.57 %	46.72	81	22	85	23
Estonia	Yes	11.10 %	112.14	14	9	14	9
Finland	Yes	11.12 %	132.52	105	33	108	35
France	Yes	4.88 %	88.35	688	47	703	50
Germany	Yes	3.31 %	81.74	780	45	797	48
Greece	Yes	4.48 %	102.83	272	38	278	40
Ireland	Yes	5.66 %	178.64	57	23	61	26
Italy	Yes	6.48 %	365.08	256	41	268	44
Luxembourg	Yes	6.31 %	550.14	33	17	35	18
Malta	Yes	2.26 %	154.07	13	7	13	7
Netherlands	Yes	7.95 %	363.58	167	31	174	33
Portugal	Yes	4.92 %	348.99	61	22	62	23
Slovakia	Yes	6.32 %	64.72	14	9	14	9
Slovenia	Yes	7.37 %	294.72	30	18	32	19
Spain	Yes	5.71 %	561.53	148	33	150	34
Bulgaria	No	-	-	-	-	-	-
Czech Republic	No	10.65 %	246.12	-	-	24	12
Denmark	No	2.21 %	138.21	-	-	167	33
Hungary	No	11.20 %	56.85	-	-	37	19
Latvia	No	6.47 %	8.67	-	-	25	14
Lithuania	No	8.64 %	62.08	-	-	35	17
Poland	No	7.15 %	36.22	-	-	346	43
Romania	No	-	-	-	-	-	-
Sweden	No	5.42 %	43.05	-	-	389	44
United Kingdom	No	5.78 %	65.95	-	-	1 584	50
Eurozone	-	4.82 %	137.11	2 946	47	-	-
EU27	-	4.83 %	94.35	-	-	5 636	51
Min	-	1.57 %	8.67	13	7	13	7
Max	-	11.20 %	561.53	780	47	1 584	50

In addition to reviewing the sample characteristics by country, it is also useful to present the industry-level statistics as these are important variables in the analysis. Appendix 4 shows the average ROAs by industry and related industries along with industry size information for the industry relatedness sample. It can be seen that, industry profitability does not vary greatly between the Eurozone and EU27 samples but there is significant variation between industries. The most profitable industries in the samples are 23 (Manufacture of coke, refined petroleum products and nuclear fuels) and 10 (Mining of coal and lignite; extraction of peat) with average return on assets of 21.7% and 19.27%, respectively. On the other hand the least profitable industry (73 Research and development) has an average ROA of -10.6% for the sample period. It is also worth noticing that the sample size varies greatly across industries with the largest industry comprising of 348 (611) firms and the smallest comprising of 11 (5) companies in the Eurozone (EU27) sample.

Table 2 provides general sample statistics for the cross-predictability sample by country. A closer examination of the figures shows that average monthly stock market returns and firm size (measured by median market capitalization) appear to vary considerably across the sample countries. More specifically, Bulgaria has the highest average monthly stock market returns (21.7% p.a.) while Cyprus has the lowest average stock market returns (-20.91% p.a.) during the sample period. Interestingly, Spain appears to have the largest companies in both the industry relatedness and cross-predictability samples. The average monthly stock market returns are negative for both Eurozone and EU27 which reflects the challenging economic situation in Europe during the sample period. According to the Business Cycle Dating Committee of the Centre for Economic Policy Research, which monitors the euro area business cycle, the Eurozone experienced a prolonged pause in the growth of economic activity in the early 2000s followed by a recession which began in the first quarter of 2008.[50]

Table 2 also shows that the cross-predictability sample is of similar size as the industry relatedness sample. More specifically, the Eurozone sample consists of 2,925 listed firms from 45 different industries and the EU27 sample contains 5,464 listed firms from 49 different industries. The larger countries also dominate the cross-predictability sample: firms listed in France and Germany account for over 50% of the Eurozone sample while the United Kingdom is the single largest country representing almost 27% of the EU27 sample. It is important to remember that these figures are only approximations as the realized sample size varies between months ranging from 1,364 to 2,382 (2,290 to 5,464) firms in the Eurozone (EU27) sample. For more detailed information on the annual country-specific stock market returns and sample size statistics, see Appendix 5.

Similar to the industry relatedness sample, industry level statistics are presented in Appendix 6 for the cross-predictability sample. An examination of the average monthly stock returns for industries and their customer and supplier industries reveals that there is significant variation in average monthly returns across different industries in the sample period. Interestingly, industry 16 (Manufacture of tobacco products) has the highest average monthly stock returns (19.36% p.a.) during the sample period. A comparison with the ROA figures also reveals that the industry has been among the most profitable industries in the EU27 sample. Industry 30 (Manufacture of office

[50] CEPR Euro Area Business Cycle Dating Committee: http://www.cepr.org/data/dating/

machinery and computers) appears to have been the worst investment in the sample period with an average monthly stock return of -27% in both Eurozone and EU27 samples. Similar to the industry relatedness analysis, the sample size varies across industries also in the cross-predictability sample: the largest industry consists 346 (607) firms whereas the smallest industry consists 9 (9) companies in the Eurozone (EU27) sample.

Table 2: Descriptive Statistics for Cross-Predictability Sample

This table presents country-level stock returns for all sample countries. The figures are reported for both Eurozone and EU27 samples separately. All data items are retrieved from Thomson One Banker database and the industries are based on NACE rev 1.1 industry classifications. Mean monthly market return is the simple average of the monthly value-weighted stock market returns for a given country calculated over the whole sample period ranging from year 2000 to 2009. Market return figures are annualized. Median market capitalization is the median firm market value calculated over all sample companies within a country. Market values are expressed in millions of euros. Number of firms is the total number of different firms and number of industries is the total number of different industries included in the sample for a given country. These figures are different for the Eurozone and EU27 samples due to industries with less than five firms being exlcuded from the samples.

				Eurozone		EU27	
Country	Eurozone	Mean Monthly Market Return	Median Market Capitalization (€M)	No. of Firms	No. of Industries	No. of Firms	No. of Industries
Austria	Yes	3.70 %	125.61	101	30	101	30
Belgium	Yes	-2.81 %	149.01	132	33	138	35
Cyprus	Yes	-20.91 %	12.67	85	22	86	23
Estonia	Yes	13.46 %	36.45	13	8	13	8
Finland	Yes	4.45 %	121.70	69	29	69	29
France	Yes	-1.53 %	59.58	648	44	656	47
Germany	Yes	-7.15 %	46.63	843	44	851	46
Greece	Yes	-13.80 %	53.26	284	38	287	40
Ireland	Yes	-5.17 %	240.85	55	24	59	26
Italy	Yes	-4.31 %	231.18	267	39	273	42
Luxembourg	Yes	-13.25 %	315.50	32	14	33	15
Malta	Yes	1.55 %	67.50	12	8	12	8
Netherlands	Yes	-4.21 %	240.88	163	31	169	33
Portugal	Yes	-1.71 %	150.48	65	23	65	23
Slovakia	Yes	16.71 %	17.74	15	10	15	10
Slovenia	Yes	9.50 %	172.76	31	18	31	18
Spain	Yes	2.41 %	546.94	110	30	111	31
Bulgaria	No	21.70 %	6.65	-	-	159	34
Czech Republic	No	13.30 %	159.79	-	-	24	11
Denmark	No	1.66 %	71.28	-	-	174	33
Hungary	No	4.86 %	40.62	-	-	36	20
Latvia	No	8.38 %	26.82	-	-	9	7
Lithuania	No	-1.73 %	27.06	-	-	29	17
Poland	No	0.50 %	29.71	-	-	242	36
Romania	No	-	-	-	-	-	-
Sweden	No	-2.84 %	44.81	-	-	368	42
United Kingdom	No	-1.61 %	67.30	-	-	1454	48
Eurozone	-	-4.12 %	80.85	2925	45	-	-
EU27	-	-2.67 %	67.99	-	-	5464	49
Min	-	-20.91 %	6.65	12	8	9	7
Max	-	21.70 %	546.94	843	44	1454	48

5 METHODOLOGY

This section presents the methodology used to test the hypotheses of this study. Since one objective of this paper is to provide an out-of-sample test of the limited-information model by Menzly and Ozbas (2010), the methodology used is mainly consistent with their article. This section is organized as follows. Subsection 5.1 shows the formation of industry, customer and supplier portfolios. Subsection 5.2 presents the ROA regressions which are used to examine correlation of fundamentals across the supply chain. Subsection 5.3 introduces the methodology used to analyze stock-and industry level cross-predictability effects. Subsection 5.4 presents the methodology used to evaluate the economic significance of cross-predictability effects and exposure to risk factors.

5.1 FORMATION OF INDUSTRY, CUSTOMER AND SUPPLIER PORTFOLIOS

First this subsection explains the formation of industry portfolios and calculation of industry portfolio returns. Then the formation of customer and supplier portfolio based on the Eurostat input-output tables and calculation of customer and supplier industry portfolio returns is presented.

As explained earlier in subsection 4.2, each sample firm is assigned to an industry based on NACE rev 1.1 classification system. Industry portfolios are formed based on these industry assignments, and the value-weighted monthly industry returns are calculated for each industry portfolio. Following Menzly and Ozbas (2006), the value-weighted industry portfolio returns are calculated as in Equation (1) below.

$$r_{i,t} = \sum_{j_1 \in i}^{n_{i,t-1}} \frac{M_{j_1,t-1}}{\sum_{j_1 \in i}^{n_{i,t-1}} M_{j_1,t-1}} r_{j_1,t} \tag{1}$$

where, $r_{i,t}$ is the value-weighted portfolio return for industry i in month t, $n_{i,t-1}$ is the number of firms in industry i in month $t-1$, $M_{j_1,t-1}$ is the market capitalization of firm j belonging to industry i at the end of month $t-1$ and $r_{j_1,t}$ is the stock return of firm j in month t. All returns are in excess of risk-free rate observed at the beginning of month t.[51]

After calculating industry portfolio returns, two separate portfolios are formed for each industry. More specifically, one portfolio composed of supplier industries and another composed of customer industries are constructed for each industry using the flow of goods and services data from the Eurostat input-output-tables as portfolio weights. Following the methodology by Menzly and Ozbas (2010), the share of an industry's total purchases from other industries is used to calculate supplier industry returns and the share of an industry's total sales to other industries is used to calculate

[51] The risk-free rate used in this study is the one-month Euribor (see subsection 4.1).

customer industry returns. According to Menzly and Ozbas, this portfolio weighting scheme has the desired properties for the purpose of testing limited-information models as it (i) identifies the set of economically related supplier and customer industries for a given industry, and (ii) reflects the relative economic importance of related industries as proxied by the amount of inter-industry trade. The formal calculation for supplier industry portfolio returns is described in Equation (2) and the formal calculation for customer industry portfolio returns is described in Equation (3) below.

$$r_{i,t}^{supplier} = \sum_{k_1 \neq i}^{n_k} \frac{C_{k_1,i}}{\sum_{k_1 \neq i}^{n_k} C_{k_1,i}} r_{k_1,t} \tag{2}$$

where $r_{i,t}^{supplier}$ is the return on industry *i*'s supplier industries in month *t* weighted by the flow of goods and services into industry i, n_k is the number of industry *i's* supplier industries excluding industry *i,* $C_{k_1,i}$ is the amount of industry *i*'s purchases from supplier industry *k* and $r_{k_1,t}$ is the value-weighted portfolio return on supplier industry *k* in month *t* as calculated in Equation (1). All returns are in excess of risk-free rate observed at the beginning of month *t*.

$$r_{i,t}^{customer} = \sum_{k_1 \neq i}^{n_k} \frac{C_{k_1,i}}{\sum_{k_1 \neq i}^{n_k} C_{k_1,i}} r_{k_1,t} \tag{3}$$

where $r_{i,t}^{customer}$ is the return on industry *i*'s customer industries in month *t* weighted by the flow of goods and services out of industry i, n_k is the number of industry *i's* customer industries excluding industry *i,* $C_{k_1,i}$ is the amount of industry *i*'s sales to customer industry *k* and $r_{k_1,t}$ is the value-weighted portfolio return on customer industry *k* in month *t* as calculated in Equation (1). All returns are in excess of risk-free rate observed at the beginning of month *t*.

5.2 INDUSTRY RELATEDNESS ALONG THE SUPPLY CHAIN

This subsection presents the methodology used in the industry relatedness analysis. First, the calculation of firm-, industry- and market-level return on assets is presented. Then the fixed effects panel regression used to study correlation of fundamentals along the supply chain is presented. The methodology used is consistent with Menzly and Ozbas (2010).

As mentioned earlier, two important assumptions are required to obtain return cross-predictability in a limited-information model: (i) firms in different industries or market segments have correlated fundamentals and (ii) markets are informationally segmented as informed investors, to some degree, specialize along these boundaries in their information-gathering activities (Menzly and Ozbas, 2010). Evidence on the latter assumption is provided in the literature review in subsection 2.2.2.1 and providing further proof of this assumption is outside the scope of this study. However, empirical evidence on assumption (i) is provided in order to verify the empirical design of this paper. More specifically, this is confirmation is performed to ensure the validity of the consolidated Eurostat input-output tables in describing industry relatedness and to address the possibility that border effects may cause industry fundamentals to be insufficiently correlated (see subsection 3.1).

In the particular empirical setting used in this study, the correlated fundamentals assumption means that firms in a given industry need to have correlated fundamentals with firms in their supplier and customer industries. To test whether firms along the supply chain have correlated fundamentals, firm-, industry- and market-level measures of profitability are constructed and used in a fixed-effects panel regression. The calculation of firm-level return on assets is shown in Equation (4) below.

$$ROA_{j,t} = \frac{OI_{j,t} + DA_{j,t}}{TA_{j,t}} \tag{4}$$

where, $ROA_{j,t}$ is the return on assets of firm j in year t, $OI_{j,t}$ is the operating income of firm j in year t, $DA_{j,t}$ is the depreciation, depletion and amortization of firm j in year t and $TA_{j,t}$ is the total assets of firm j in year t. Industry- and market-level ROA are calculated by aggregating the above firm-level ROAs with a portfolio approach using firm assets as portfolio weights. The industry- and market-level profitability calculation are shown below in Equations (5) and (6) for both market level and industry level ROAs, respectively.

$$ROA_t^{market} = \sum_{i=1}^{n_t} \frac{TA_{j,t}}{\sum_{i=1}^{n} TA_{j,t}} ROA_{j,t} \tag{5}$$

where, ROA_t^{market} is the aggregated market-level ROA for year t, n_t is the total number of firms in year t, $TA_{j,t}$ is the total assets of firm j in year t and $ROA_{j,t}$ is the firm-level return on assets for firm j from equation (4).

$$ROA_{i,t} = \sum_{j_1 \epsilon\, i}^{n_{i,t}} \frac{TA_{j_1,t}}{\sum_{j_1 \epsilon\, i}^{n_{i,t}} TA_{j_1,t}} ROA_{j_1,t} \tag{6}$$

where, $ROA_{i,t}$ is the aggregated industry-level ROA in year t for industry i, $n_{i,t}$ is the number of firms in industry i in year t, $TA_{j_1,t}$ is the total assets for firm j belonging to industry i in year t and $ROA_{j_1,t}$ is the firm-level return on assets for firm j from Equation (4). Industry-level ROAs are used to further calculate supplier and customer industry ROAs for each industry i. This calculation is done by weighting the industry-level ROAs of supplier and customer industries with the flow of goods and services to and from the industries in question. This approach is similar to the calculation of supplier and customer industries returns in Equations (2) and (3). The ROA calculation for supplier and customer industries is shown in Equation (7) below.

$$ROA_{i,t}^{supplier\ (customer)} = \sum_{k_1 \neq i}^{n_k} \frac{c_{k_1,i}}{\sum_{k_1 \neq i}^{n_k} c_{k_1,i}} ROA_{k_1,t} \quad (7)$$

where, $ROA_{i,t}^{supplier\ (customer)}$ is the aggregated return to assets in year *t* for industry *i*'s supplier (customer) industries weighted by the flow of goods and services into (out of) industry i, n_k is the number of industry *i's* supplier (customer) industries excluding industry i, $C_{k_1,i}$ is the amount of industry *i*'s purchases from (sales to) industry k and $ROA_{k_1,t}$ is the return on assets of industry k in year *t* from Equation (6). The ROA figures received from Equations (4) to (7) are used in a fixed effects panel regression that is presented in Equation (8) below. Two specifications of Equation (8) are performed, one with firm-level ROA as dependent variable and another with industry-level ROA as dependent variable.

$$ROA_{j(i),t} = \propto_i + \beta^{market} ROA_t^{market} + \beta^{supplier} ROA_{j(i),t}^{supplier} + \beta^{customer} ROA_{j(i),t}^{customer} \quad (8)$$

$$+ \varepsilon_{j(i),t}$$

where, the dependent variable $ROA_{j(i),t}$ is the return on assets of firm *j* (industry *i*) in year *t* depending on the specification, β^{market} is the coefficient on market-level ROA, ROA_t^{market} is the contemporaneous market-level ROA from equation (5), $\beta^{supplier}$ is the coefficient on supplier-industry ROA for firm *j* (industry *i*), $ROA_{j(i),t}^{supplier}$ is the supplier-industry ROA for firm *j* (industry *i*) in year *t* from equation (7), $\beta^{customer}$ is the coefficient on customer-industry ROA for firm *j* (industry *i*), $ROA_{j(i),t}^{customer}$ is the customer-industry ROA for firm *j* (industry *i*) in year *t* from Equation (7) and $\varepsilon_{j(i),t}$ is the error term from the regression in year *t*.

As mentioned in section 4.4, there is considerable variation in the annual ROA figures across sample countries. Therefore, a possible concern is that country-specific differences in average profitability are driving the results, particularly in the firm-level ROA panel regressions where the independent variables are country-specific. In order to address the possibility that country-specific factors such as varying economic conditions are confounding the regression results, the firm-level ROAs are adjusted for country-specific differences in overall profitability in the firm-level specification of Equation (8). The indexation formula used in this study is shown in Equation (9).

$$ROA_{j,t}^{indexed} = \frac{1+ ROA_{j,t}}{1+ROA_{c,t}^{market}} - 1 \quad (9)$$

where, $ROA_{j,t}^{indexed}$ is the indexed return on assets of firm j in year t, $ROA_{j,t}$ is the return on assets of firm j in year t and $ROA_{c,t}^{market}$ is the aggregated market-level ROA in year t for country c in which firm j is listed. Appendix 3 contains the annual country-level return on assets used in Equation (9).

In panel data sets, such as the one studied here, the residuals may be correlated across observations which can cause OLS standard errors to be biased (Petersen, 2009). More specifically, the firm effect and/or time effect may introduce bias in the standard errors which, in turn, may lead to incorrect t-statistics and unjustified rejection (acceptance) of the null hypothesis.[52] Following Petersen (2009), the regression standard errors are adjusted for both firm and time effects, to address the possibility that correlated residuals are biasing the standard errors and thus, the t-statistics. Firm effect is taken into account parametrically by using fixed effects regression with fixed firm effects. The time effect is addressed by adjusting the standard errors for clustering by year.[53]

5.3 CROSS-PREDICTABILITY OF STOCK RETURNS

This section presents the methodology used in this paper to test for cross-predictability of stock returns. The section is organized as follows. Subsection 5.3.1 presents the regressions used to examine stock- and industry-level cross-predictability effects. Subsection 5.3.2 shows the approach used to investigate the effect of informed investors on return cross-predictability. Subsection 5.3.3 illustrates the methodology used to study the impact of geographic specialization on cross-predictability.

5.3.1 Stock- and Industry-Level Cross-Predictability Effects

This subsection, first presents the Fama-MacBeth (1973) regression used to test for both stock- and industry-level return cross-predictability. Then the calculation of the Fama-MacBeth regression coefficients, standard errors and t-statistics is shown.

In order to test for stock- and industry-level return cross-predictability, Fama-MacBeth (FM, 1973) regressions of stock (industry) returns on lagged related industry returns and control variables are

[52] Firm effect and time effect are two general forms of residual dependence in finance applications. The firm effect refers to the possibility that the residuals of a given firm are correlated across years for a given firm (time-series dependence). The time effect refers to the possibility that the residuals of a given year are correlated across different firms (cross-sectional dependence). (Petersen, 2009.)

[53] Menzly and Ozbas (2010) use double-clustering by year and stock (industry) in their regression. However, according to Thompson (2011), double-clustering works well only in samples with more than 25 observations on both firms and time periods. Since the sample used in this study contains only 10 years, double-clustering is not appropriate. However, addressing two sources of correlation is possible by parametrically estimating one of the dimensions and clustering by the other, as done in this paper. Clustering by time is chosen because it is more important to cluster along the dimension with fewer observations as it should eliminate most of the bias (Thompson, 2011).

performed. According to Petersen (2009) FM regression is an alternative way to estimate the regression coefficients and standard errors when the residuals cannot be assumed to be independent. Basically, the FM approach requires running *T* monthly cross-sectional regressions for the entire sample period and then taking an average of the monthly coefficients in order to obtain an estimate of the FM regression coefficients. Standard deviations of the monthly cross-sectional regression coefficients are used to calculate the standard errors for the FM regression coefficients. (Cochrane, 2000.)

Following Menzly and Ozbas (2010), Equations (10), (11) and (12) show the general form regression used to test for stock- and industry-level return cross-predictability. Based on Fama and MacBeth (1973), Equations (13), (14) and (15) present the calculation of the FM regression coefficients, standard errors and *t*-statistics, respectively.

$$r_{j(i),t} = \propto_t + \lambda_t^{supplier} r_{j(i),t-1}^{supplier} + \lambda_t^{customer} r_{j(i),t-1}^{customer} + \Lambda_t Z_{j(i),t-1} + \varepsilon_{j(i),t} \tag{10}$$

where, $r_{j(i),t}$ is the return on stock *j* (industry *i*)[54] in month *t*, $\lambda_t^{supplier}$ is the coefficient on the lagged return on stock *j's* (industry *i*'s) portfolio of supplier industries, $r_{j(i),t-1}^{supplier}$ is the lagged return on stock *j*'s (industry *i*'s) portfolio of supplier industries[55] in month *t-1*, $\lambda_t^{customer}$ is the coefficient on the lagged return on stock *j's* (industry *i*'s) portfolio of customer industries, $r_{j(i),t-1}^{customer}$ is the lagged return on stock *j*'s (industry *i*'s) portfolio of customer industries[56] in month *t-1*, $Z_{j(i),t-1}$ is a vector of lagged control variables known to predict $r_{j(i),t}$ and it is explained below in Equations (11) and (12), and $\varepsilon_{j(i),t}$ is the error term from the regression in month *t*. For the stock-level specification the vector of lagged control variables $Z_{j,t-1}$ is shown in Equation (11) and for the industry-level specification $Z_{i,t-1}$ is presented in Equation (12) below.

$$Z_{j,t-1} = \Lambda_t^{streversal} r_{j,t-1}^{streversal} + \Lambda_t^{mtcontinuation} r_{j,t-2:t-12}^{mtcontinuation} + \Lambda_t^{industrymom} r_{j,t-1}^{industrymom} \tag{11}$$

$$+\lambda_t^{market} r_t^{market}$$

where, $\Lambda_t^{\text{streversal}}$ is the coefficient on the short-term reversal variable and $r_{j,t-1}^{streversal}$ is the return on stock *j* in the previous month *t – 1* which accounts for the effect of short-term reversal documented by, for example, Jegadeesh (1990) and Lehmann (1990), $\Lambda_t^{mtcontinuation}$ is the coefficient on the medium-term continuation variable and $r_{j,t-2:t-12}^{mtcontinuation}$ is the return on stock *j*

[54] Industry-level return used here is the value-weighted industry portfolio return calculated in equation (1).
[55] Supplier industry portfolio return used here is calculated in equation (2).
[56] Customer industry portfolio return used here is calculated in equation (3).

over the 11 months covering months $t-12$ through $t-2$ to control for medium-term continuation (i.e., momentum) at the stock level[57] documented by, for example, Jegadeesh and Titman (1993), $\Lambda_t^{\text{industrymom}}$ is the coefficient on the industry level momentum variable and $r_{j,t-1}^{industrymom}$ is the return on the industry in which stock j is a member in month *t-1* which accounts for industry momentum documented by Moskowitz and Grinblatt (1999), λ_t^{market} is the coefficient on the market return and r_t^{market} is the country specific market return in month *t*.[58]

$$Z_{i,t-1} = r_{j,t-1}^{industrymom} \tag{12}$$

where, $r_{j,t-1}^{industrymom}$ is the return on industry *i* in month *t-1*. This is the only lagged control variable in the industry-level specification. It is important to note that the explanatory variables of interest in Equation (10) are $r_{j(i),t-1}^{supplier}$ and $r_{j(i),t-1}^{customer}$ which are the lagged supplier and customer industry returns expected to have predictive power over individual stock and industry returns. Equation (13) shows the calculation of the Fama-MacBeth regression coefficients from the monthly cross-sectional regression coefficients introduced in Equation (10).

$$\hat{\lambda}_{FM} = \sum_{t=1}^{T} \frac{\hat{\lambda}_t}{T} \tag{13}$$

where $\hat{\lambda}_{FM}$ is the Fama-MacBeth regression coefficient estimate, *T* is the number of months in the sample and $\hat{\lambda}_t$ is the monthly cross-sectional regression coefficients from Equation (10). The calculation of Fama-MacBeth standard errors is shown in Equation (14).

$$s.e._{\hat{\lambda}_{FM}} = \sqrt{\frac{1}{T} * \sum_{t=1}^{T} \frac{(\hat{\lambda}_t - \hat{\lambda}_{FM})^2}{T-1}} \tag{14}$$

where, $s.e._{\hat{\lambda}_{FM}}$ is the standard error of the Fama-MacBeth regression coefficient, *T* is the number of months in the sample, $\hat{\lambda}_t$ is the monthly cross-sectional regression coefficient from Equation (10) and $\hat{\lambda}_{FM}$ is the average of the monthly regression coefficients from Equation (13). The standard errors from Equation (14) are used to calculate the *t*-statistics for the Fama-MacBeth regression coefficients as shown in Equation (15).

[57] Following Menzly and Ozbas (2010), this variable is divided by 11 in the regression to maintain comparability with other variables.

[58] More specifically, r_t^{market} is the excess return on the country-specific broad market portfolio $(R_m - R_f)$ and it is calculated as the value-weighted return on all the stocks included in a country's sample less the risk-free rate. Market return r_t^{market} is included only in the stock level specification of Equation (10). This follows from the Fama-MacBeth method used which involves *T* monthly cross-sectional regressions to be performed. Given the monthly regressions, any variable which does not vary across firms within a month cannot be estimated (Petersen, 2009). This is not a problem with the stock-level specification since country specific market returns can be used, but the industry-level specification does not allow the use of Eurozone (EU27) market returns and thus, they are omitted from the industry-level regressions.

$$t = \frac{\hat{\lambda}_{FM}}{s.e._{\hat{\lambda}_{FM}}} \tag{15}$$

where t is the t-statistic of the Fama-MacBeth regression coefficient, $\hat{\lambda}_{FM}$ is the the average of the monthly regression coefficient from Equation (13) and $s.e._{\hat{\lambda}_{FM}}$ is the standard error of the Fama-MacBeth coefficient from Equation (14). The t-statistics obtained for each coefficient are interpreted in the usual manner from the t-distribution table (Fama and MacBeth, 1973). The degrees of freedom are T-1, where T is the number of months in the sample, in other words, the number of regressions performed which in this study equals 120 (10 years equals 120 months).

5.3.2 The Effect of Informed Investors

An important prediction of limited-information models is that the magnitude of cross-predictability effect should be negatively related to the amount of informed investors (Menzly and Ozbas, 2010). In order to test this prediction, this paper employs two different stock-level proxies for the amount of informed investors, namely firm size and analyst coverage.

Firm size as a proxy for the supply of information to the market has been used in earlier studies, for example, by Hong, Lim, and Stein (2000) and Zhang (2006). In their paper, Hong et al. they argue that information about large firms gets out faster because there are e.g., fixed costs to information gathering and large companies release more information. In addition, Chopra, Lakonishok and Ritter (1992) and Yalcin (2008) point out that larger firms are predominantly owned by institutional investors and thereby, it is likely that firm size is correlated with the amount of informed investors. Also the lead-lag literature has studied the impact of firm size. For example, Lo and MacKinlay (1990a) and Hou (2007) find that large firms lead smaller firms. Hou attributes this effect to the slower diffusion of information related to small firms. The firm size measure used in this paper for a given stock in month t is the stock's end-of-month market capitalization in month t-1. It is important to note that firm size can also capture a variety of other factors such as cross-stock differences in arbitrage costs and thus, may not constitute a clean test of the effect of informed investors (Hong, Lim, and Stein, 2000). For example, Lesmond, Schill and Zhou (2004) argue that firm size is highly correlated with transaction costs and hence, the results might be confounded by the price friction effect.

For robustness reasons, and to address the potential concerns associated with firm size, another proxy for the amount of informed investors is used, namely the analyst coverage. Analyst coverage has been used as a proxy for the role of informed investors in prior research, for example, by Brennan, Jegadeesh and Swaminathan (1993), Yalcin (2008) and Menzly and Ozbas (2010) who all find evidence supportive of the gradual information diffusion hypothesis. The stock-level analyst coverage measure used in this study is based on analyst EPS forecast data received from I/B/E/S detail database. More specifically, analyst coverage measure for a given stock in month t is the numerical count of EPS estimates included in the mean estimate for the stock in month *t*-1. The reason for choosing EPS forecasts is that they are the most commonly available type of analyst forecast and thus, restrict the size of the sample as little as possible (Menzly and Ozbas, 2010). It is worth noticing that Lesmond, Schill and Zhou (2004) argue that analyst coverage might be related to transaction costs which may have an influence on the results. However, they are unable to produce any conclusive evidence.

The stock-level Fama-Macbeth regression shown in Equation (10) is augmented in order to test whether the magnitude of cross-predictability effects based on lagged returns in supplier and customer industries declines (increases) as the number of informed investors increases (decreases). The augmented stock-level regression is shown in equation (16). The reason why stock-level returns are chosen over industry-level returns to examine the effect of informed investors on cross-predictability is that they allow preserving the degree of cross-sectional heterogeneity of information conditions and investor types at the stock level (Menzly and Ozbas, 2010).

$$r_{j,t} = \propto_t + \sum_{h=1}^{5} \lambda_t^h A_{t-1}^h \, r_{j,t-1}^{composite} + \varepsilon_{j,t} \tag{16}$$

where, $r_{j,t}$ is the return on stock j in month t, λ_t^h is the coefficient on lagged related industry returns, A_{t-1}^h is an indicator variable equal to one if the firm size (level of analyst coverage) for stock j in month $t-1$ places the stock in the h^{th} quintile and zero otherwise, $r_{j,t-1}^{composite}$ is the return on stock j's portfolio of supplier and customer industries in month $t-1$ calculated as the simple average of $r_{j,t-1}^{supplier}$ and $r_{j,t-1}^{customer}$, and $\varepsilon_{j,t}$ is the error term from the regression in month t. The composite variable $r_{j,t-1}^{composite}$ is used here as it reduces the number of parameters to be estimated while being model-justified (Menzly and Ozbas, 2010).

Basically, Equation (16) ranks stocks each month t into five quintiles based on their market capitalization (level of analyst coverage) measured in month $t-1$. Smallest (lowest analyst coverage) stocks are allocated into quintile 1 while largest (highest analyst coverage) stocks are allocated into quintile 5. The underlying rationale is to see whether there are differences in the predictive power of lagged returns from supplier and customer industries between different quintiles that are supposed to have different amounts of informed investors.

5.3.3 The Effect of Investor Geographic Specialization

The limited-information models posit that investor specialization causes information to diffuse gradually across markets and consequently, asset prices to exhibit cross-predictability. In order to further study the determinants of investor specialization and the impact of investor specialization, this paper examines the effect of geographic boundaries on the magnitude of return cross-predictability. More specifically, this paper tests whether investor specialization along geographic boundaries has an impact on the speed of information diffusion and thus, the magnitude of the cross-predictability effect. In order to investigate this untested aspect of investor specialization, an approach similar to the one used to study the effect of informed investors is adopted. In other words, sample stocks are allocated into different quartiles[59] based on an industry-level measure that captures the degree of geographic dispersion of an industry's supplier and customer industries.

Before introducing the methodology in more detail, it is useful to briefly go through the rationale underlying the tests of investor geographic specialization. In order to better understand the possible impact of related industry geographic dispersion on the magnitude of the cross-predictability effect, it is useful to think of the related industries as the channel of information flow through which shocks are transmitted and reflected as (delayed) price responses in economically related stocks. The more (less) geographically dispersed this channel is, the slower (faster) is the diffusion of relevant information because the informative signals from related industries are processed by fewer (more) investors who are specialized along geographic boundaries. This slower (faster) diffusion of information in turn is reflected in less (more) information impounding demand by informed investors and consequently, stronger (weaker) cross-predictability effects. In other words, because investors specialize along geographic boundaries in their investment activities as discussed in subsection 2.2.2.1, they are likely to process informative signals (in a timely manner) only on a geographically limited subset of assets. This implies that information diffusion from supplier and customer industries is slower when these industries become more geographically dispersed. Consequently, this slower rate of information diffusion causes industries with more (less) geographically dispersed related industries to exhibit stronger (weaker) cross-predictability effects.

To further clarify the tests presented in this subsection, an analogy can be made between the investor specialization test in this paper and Shahrur, Becker and Rosenfeld (2009) who investigate the impact of sales dispersion on the magnitude of customer industry cross-predictability effects.

---

[59] Quartiles are used instead of quintiles because the $Geodisp_j$ allocation is performed on the industry level. Using an industry-level measure provides a rather rough measure as opposed to firm-level measures used to analyze the impact of informed investors and therefore, the use of quintiles does not provide more information. In addition, using quartiles allows the number of different industries represented within a quartile to remain sufficiently high.

Similar to this paper, Shahrur et al. also use an industry-level measure to define industry characteristics that they hypothesize are related to slower rate of information diffusion. In their article, they find that suppliers with more dispersed sales experience stronger return cross-predictability effects from customer industries. They interpret this finding as evidence of slower information diffusion caused by relevant information diffusing from multiple sources as sales dispersion increases. In other words, when a supplier deals with a larger number of customers, investors need to assess the effect of news from multiple customers which is more challenging and results in stronger return cross-predictability. Although the methodology used in this paper is different from Shahrur et al., the simple analogy of multiple source shocks being more difficult for investors to process and incorporate in their investment decisions still applies. However, instead of studying supplier sales dispersion like Shahrur et al., this paper focuses on sources of shocks that are geographically dispersed. In order to do this, a self-constructed industry-level measure that proxies for the geographic dispersion of informative signals from both supplier and customer industries is developed. The calculation of the measure is shown in Equation (17) below.

$$Geodisp_{i,t} = \frac{D_{i,t-1}^{countries}}{D_{i,t-1}^{large}} \tag{17}$$

where, $Geodisp_{i,t}$ is the measure of geographic dispersion of industry *i's* supplier (customer) industries in month *t*, $D_{i,t-1}^{countries}$ is a measure of overall geographic dispersion of stock j's supplier (customer) industries and $D_{i,t-1}^{large}$ is a measure of geographic concentration of industry *i's* largest supplier (customer) industry. More detailed description of $D_{i,t-1}^{countries}$ and $D_{i,t-1}^{large}$ is provided below in Equations (18) and (19), respectively.

$$D_{i,t-1}^{countries} = \sum_{k_1 \neq i}^{n_i} \frac{C_{k_1,i}}{\sum_{k_1 \neq i}^{n_k} C_{k_1,i}} N_{k_1,t-1} \tag{18}$$

where, $D_{i,t-1}^{countries}$ is the number of different countries represented in industry *i*'s supplier (customer) industry portfolio in month $t-1$ weighted by the flow of goods and services specified in the Eurostat input-output tables, n_i is the number of supplier (customer) industries for industry *i* excluding industry *i*, $C_{k_1,i}$ is the amount of industry *i*'s purchases from (sales to) supplier (customer) industry k and $N_{k_1,t-1}$ is the number of different countries represented in supplier (customer) industry *k* in month $t-1$.[60]

[60] In other words, the weighting scheme here is similar to the one used to calculate related industry returns in Equations (2) and (3). The number of different countries represented by each supplier (customer) industry is weighted by the inter-industry flow of goods

As mentioned in subsection 2.2.2.1, analysts and investors focus largely on their domestic markets and/or into a few countries. The purpose of $D_{i,t-1}^{countries}$ is to take into account the effect of national borders in investment and information gathering activities. Basically, the higher the number of different countries represented in a given stock's related industry portfolio, the more geographically dispersed the stock's related industries can be considered. However, this measure alone is not sufficient to capture the geographic dispersion of a given stock's related industries. As an illustration, think of industry i's supplier industry portfolio that contains 100 stocks from 20 European countries. Based on the $N_{i,t-1}^{countries}$ variable the supplier industry portfolio might seem to be geographically highly dispersed. However, a closer examination might reveal that 80 of the 100 related industry stocks are actually listed in just one country. In this case the $N_{i,t-1}^{countries}$ variable alone would clearly provide an overestimate of the geographic dispersion of industry i's supplier industries. Therefore, to address this concern, $Geodisp_{i,t}$ also needs to take into account the possibility that the related industry stocks are concentrated into only a few countries. This is captured by $D_{i,t-1}^{large}$ as explained below in Equation (19).

$$D_{i,t-1}^{large} = \frac{N_{i,t-1}^{high}}{N_{i,t-1}^{large}} \tag{19}$$

where, $D_{i,t-1}^{large}$ is the measure of geographic dispersion in industry i's largest supplier (customer) industry, $N_{i\,t-1}^{high}$ is the number of sample stocks from industry i's largest supplier (customer) industry that are listed in the country with the highest number of industry i's largest supplier (customer) industry stocks in month t – 1 and $N_{i,t-1}^{large}$ is the total number of sample stocks included in industry i's largest supplier (customer) industry portfolio in month t – 1.[61] In other words, $D_{i,t-1}^{large}$ takes into consideration the possibility that the related industry stocks might be concentrated into only a few countries. The interpretation of this measure is that the smaller the relative number of supplier (customer) industry stocks in the country with the highest number of these stocks, the less

---

and services reported in the Eurostat Input-Output tables in order to take into account the relative importance of different industries for industry i. Similar to the calculation of related industry returns in subsection 5.1, the Eurostat input-output tables are used on a rolling basis also in this calculation. Thus, industry relatedness data from the 2000 Eurostat IO-table is used between years 2000 and 2004 and data from 2005 Eurostat IO-table is used between years 2005 and 2009 to calculate the $Geodisp_{j,t}$ measure. The country of a given stock is based on Thomson One Banker country indicator data item. In order to ensure the representativeness of the $Geodisp_{j,t}$ measure, all stocks listed in EU27 that have been assigned to a NACE industry and have available return data for a given month are used to calculate the measure.

[61] $D_{i,t-1}^{large}$ is calculated only for industry i's largest supplier (customer) industry defined as the industry with which industry i has the most trade. The rationale is that the largest supplier (customer) industry is likely to have the greatest influence on industry i's performance and consequently, is likely to be the most closely followed industry by informed investors.

geographically concentrated the supplier (customer) industry is. Thus, $D_{i,t-1}^{large}$ has a lower (higher) value when the geographic dispersion in the largest supplier (customer) industry is higher (lower).

Based on the above description, the interpretation of $Geodisp_{i,t}$ measure is that the more (less) geographically dispersed industry i's related industries are, the higher (lower) is the value of $Geodisp_{i,t}$. After the monthly $Geodisp_{i,t}$ measures have been calculated for each stock, an augmented stock-level regression similar to equation (16) is performed to test whether geographic dispersion of supplier and customer industries has an impact on the predictive power of lagged returns from these industries.[62] The augmented regression used is shown below in Equation (20).

$$r_{j,t \in i} = \propto_t + \sum_{h=1}^{4} \lambda_t^h A_{t-1}^h \, r_{j,t-1}^{composite} + \varepsilon_{j,t} \qquad (20)$$

where, $r_{j,t}$ is the return on stock j belonging to industry i in month t, A_{t-1}^h is an indicator variable equal to one if the $Geodisp_{i,t}$ measure for industry i in month t places stock j in the h^{th} quartile[63] and zero otherwise, $r_{j,t-1}^{composite}$ is the return on stock j's portfolio of supplier and customer industries in month $t-1$, calculated as the simple average of $r_{j,t-1}^{supplier}$ and $r_{j,t-1}^{customer}$, and $\varepsilon_{j,t}$ is the error term from the regression in month t.

5.4 ECONOMIC SIGNIFICANCE OF RETURN CROSS-PREDICTABILITY

Since one objective of this paper is to determine the economic significance of return cross-predictability, this section presents a self-financing trading strategy that capitalizes on the cross-predictability effects. This section is organized as follows. Subsection 5.4.1 presents the self-financing trading strategy. Subsection 5.4.2 presents the four-factor model by Fama and French (1993) and Carhart (1997). Subsection 5.4.3 shows the methodology used to estimate the exposure of the monthly self-financing strategy returns on the four-factor model risk factors.

5.4.1 Self-Financing Trading Strategy

This subsection presents the self-financing trading strategies used to estimate the economic significance of return cross-predictability. The trading strategy constructed in this paper is consistent with Menzly and Ozbas (2010).

---

[62] Even though the $Geodisp_{i,t}$ measure is calculated at the industry-level, the augmented regression is performed using stock returns as the dependent variables in order to have a sufficient number of observations per quartile.

[63] While $Geodisp_{i,t}$ measure is calculated separately for each industry's supplier and customer industries, the measure that is used to allocate industries into quartiles is the simple average of these two figures. This follows from the use of composite returns in the augmented regression.

The self-financing trading strategies employed in this study involve buying and selling industries. The rationale for focusing on industry portfolios instead individual stocks is that the industry portfolios are value-weighted portfolios of stocks which should address potential concerns that thin markets could be driving the results or that transactions costs could render the trading strategies obsolete (Menzly and Ozbas, 2010). The trading strategies used are based on sorting industries at the beginning of each month into five bins according to previous-month returns in supplier and customer industries. Industries with previous-month related industry returns in the bottom quintile are allocated to the first bin, industries with previous-month related industry returns in the second quintile to the second bin, and so forth. After sorting industries in this manner, value-weighted portfolios are constructed for each of the five bins and value-weighted returns on these portfolios for the ensuing 1-month period are calculated. The calculation of the value-weighted returns on the sorted portfolios is shown in Equation (21).

$$r_{q,t} = \sum_{i_1 \in q}^{n_{q,t-1}} \frac{M_{i_1,t-1}}{\sum_{i_1 \in q}^{n_{q,t-1}} M_{i_1,t-1}} r_{i_1,t} \tag{21}$$

where, $r_{q,t}$ is the value-weighted return on quintile q portfolio in month t, $n_{q,t-1}$ is the number of industries in quintile q in month $t-1$, $M_{i_1,t-1}$ is the market capitalization of industry i at the end of month $t-1$ and $r_{i_1,t}$ is the value-weighted industry portfolio return on industry i in month t from Equation (1). All returns are in excess of risk-free rate observed at the beginning of month t. The actual self-financing trading strategy involves buying the high portfolio (industries with previous-month related industry returns in the top quintile) and selling the low portfolio (industries with previous-month related industry returns in the bottom quintile). Three specifications of the trading strategy are tested:

i. a strategy that sorts industries based on previous-month returns in supplier industry portfolio;
ii. a strategy that sorts industries based on previous-month returns in customer industry portfolio;
iii. a strategy that sorts industries based on previous-month returns in composite industry portfolio.

The related industry portfolio returns used for sorting industries in strategies (i), (ii) and (iii) are $r_{i,t-1}^{supplier}$, $r_{i,t-1}^{customer}$ and $r_{i,t-1}^{composite}$, respectively. The calculation of the monthly self-financing strategy returns is shown in Equation (22) below.

$$r_t^{hi-lo} = r_t^{hi} - r_t^{lo} \tag{22}$$

where, r_t^{hi-lo} is the return on the self-financing trading strategy portfolio in month t, r_t^{hi} is the return on the high portfolio (highest previous-month related industry returns) in month t and r_t^{lo} is the return on the low portfolio (lowest previous-month related industry returns) in month t.

5.4.2 Four-Factor Model Estimation

This subsection first presents the four-factor model which is an extension of the Fama-French (FF, 1993) three-factor model by Carhart (1997). Then the calculation of each four-factor model risk factor is explained.

The underlying idea in the Fama-French (1993) three factor model is that covariance with the market returns is an insufficient measure for risk. More specifically, it is argued that while market risk factor can explain the average excess returns[64] on stocks, it is not able to explain the differences in average excess returns in a cross-section of stocks. In order to explain, the differences in average returns across stocks, FF argue that two additional factors are needed, that is, size and book-to-market equity (B/M). Thus, their suggestion is that all three factors – market, size and B/M- are required when evaluating a portfolio's performance.

Fama and French (1996) note that their three-factor model is not able to explain cross-sectional variation in momentum-sorted portfolio returns. Motivated by this finding, Carhart (1997) investigates the persistence of mutual fund performance and finds that Jegadeesh and Titman's (1993) stock return momentum explains performance persistence among mutual managers. Based on his results, Carhart constructs a four-factor model by extending the FF three-factor model with an additional factor that captures the Jegadeesh and Titman's one-year momentum effect.

The four-factor model argues that the expected return on a portfolio in excess of the risk-free rate can be explained by the sensitivity of its returns to four-factors: the excess return on a market portfolio, the return on SMB (small minus big) portfolio, the return on a HML (high minus low) portfolio and the return on a MOM (past winners minus past losers) portfolio. Equation (23) presents the four-factor model.

$$r_{i,t} - R_{f,t} = a_i + b_i(R_{m,t} - R_{f,t}) + s_i SMB_t + h_i HML_t + m_i MOM_t + \varepsilon_{i,t} \tag{23}$$

where, $r_{i,t}$ is the return on portfolio i in time t, $R_{f,t}$ the risk-free rate in time t, $(R_{m,t} - R_{f,t})$ is the excess return on the market portfolio in time t, SMB_t is the difference between the return on a

[64] Excess return here means the difference between a stock's return and the risk-free rate.

portfolio of small stocks and the return on a portfolio of large stocks in time t, HML_t is the difference between the return on a portfolio of high book-to-market stocks and the return on a portfolio of low book-to-market stocks in time t, MOM_t is the difference between the return on a portfolio of past winner stocks and the return on a portfolio of past loser stocks in time t and $\varepsilon_{i,t}$ is the error term from the regression in time t.

After establishing a general understanding of the four-factor model, it is time to explain how the four factors – market, SMB, HML and MOM factors – are estimated in practice. Proceeding from left to right in Equation (23) and following Fama and French (1993), the market factor is calculated as the difference between the monthly return on the broad market portfolio and the monthly risk-free rate.[65] The market return is calculated as the value-weighted return of all the stocks included in sample.

Similarly, following FF (1993), the SMB portfolios are constructed by sorting all sample stocks at the end of June each year *t* based on size, which is measured by a company's market capitalization at the end of June of the same year. The median market value for the sample is used to divide stocks into two groups, small (S) and big (B). In order to form the HML portfolios, another independent sort is performed at the end of June each year *t* based on the book-to-market equity ratio of each company. The B/M-ratio used in the sort is calculated at the end of the previous December in year *t* – 1 and stocks with negative book-to-market equity are excluded from the sample. The B/M sorted stocks are split into three groups: bottom 30 percent of companies compiles the low B/M (L) group, the middle 40 percent the medium B/M (M) group and the top 30 percent the high B/M (H) group. (Fama and French, 1993.)

After the sort, six portfolios (S/L, S/M, S/H, B/L, B/M and B/H) are constructed from the intersections of the two size and three B/M groups. For example, group B/L includes stocks that were ranked in the higher size category (B) and in the low (L) B/M category. Monthly value-weighted returns are computed for the six portfolios from July of year *t* to June of year *t+1*, after which the portfolios are reformed in June *t+1*. (FF, 1993.) Following this procedure, the portfolio returns are calculated for the entire sample period from 2000 to 2009. It is important to note that for a stock to be included in one of the six portfolios, it is required to have data available for both size and B/M ratio.

[65] As mentioned earlier, the risk-free rate used in this study is the one-month Euribor.

Returns on the previously described six portfolios form the basis for calculating the SMB and HML factors. SMB factor is the monthly difference between the simple average of returns on the three small stock portfolios (S/L, S/M, S/H) and the simple average of returns on the three large stock portfolios (B/L, B/M, B/H). Thus, SMB can be interpreted as mimicking the size related risk factor in returns. Six size-B/M -portfolios are used as opposed to using only two size portfolios (small and big) because this procedure yields a return pattern that is largely free from the influence of B/M. In other words, using difference in returns between portfolios of small and large stocks with approximately the same weighted-average book-to-market equity allows focusing on the different returns on small and large stocks. (FF, 1993.)

HML factor is the monthly difference between the simple average of returns on the two high B/M portfolios (S/H, B/H) and the simple average of returns on the two low B/M portfolios (S/L, B/L). HML mimics the risk factor in returns related to book-to-market equity, in other words, risk factors between value stocks and growth stocks. Since the HML is calculated as the difference between returns on high and low B/M portfolios with approximately the same weighted average size, the obtained return pattern should be largely free of the size factor in returns. (FF, 1993.)

Finally, following Carhart (1997), the MOM portfolios are constructed each month by sorting all stocks in the sample based on their past returns. More specifically, past eleven month returns from month $t - 2$ to $t - 12$ are used to split stocks into two groups: top 30 percent of companies with the highest past returns form the winner (W) group and the bottom 30 percent of companies with the lowest past returns form the loser (L) group. MOM factor is calculated, each month, as the equal-weighted average of returns on the winner portfolio minus equal-weighted average of returns on the loser portfolio.

5.4.3 Return Factor Exposure

Given an understanding of the four-factor model and estimation of its risk factors, this subsection presents the methodology used to estimate the exposure of monthly self-financing trading strategy returns on the four-factor model risk factors. The methodology is based on Menzly and Ozbas (2010).

In order to investigate whether the self-financing trading strategies are able to generate abnormal returns, the monthly returns generated by the strategies are regressed on commonly known risk factors. More specifically, the monthly returns are regressed on the four-factor model risk factors – market, SMB, HML, and MOM. This addresses the potential concern that the returns are not abnormal because they contain a significant amount of systematic risk and thus, are highly exposed to already well-known return factors. The regression is shown in Equation (24) below.

$$r_t^{hi-lo} = \alpha_t + \lambda_t^{market} r_t^{market} + \lambda_t^{HML} HML_t + \lambda_t^{SMB} SMB_t + \lambda_t^{MOM} MOM_t + \varepsilon_t \quad (24)$$

where, r_t^{hi-lo} is the return on the self-financing trading strategy portfolio in month *t* from Equation (22), r_t^{market}, HML_t, SMB_t and MOM_t are the Fama-French-Carhart (1993, 1997) four-factor model risk factors[66] and λ_t^{market}, λ_t^{HML}, λ_t^{SMB} and λ_t^{MOM} are their coefficients, respectively, and ε_t is the error term from the regression in month *t*.

6 EMPIRICAL RESULTS

This section presents the empirical findings of the paper. The section proceeds in the same order as the methodology section and is organized as follows. First, subsection 6.1 presents the empirical evidence from the industry relatedness analysis. Then subsection 6.2 discusses the results from the stock- and industry-level return cross-predictability analysis. Thereafter, subsection 6.3 introduces the results from tests concerning the effect of informed investors. Then subsection 6.4 presents the empirical evidence on the effect of investor specialization along geographic boundaries. Finally, subsection 6.5 reviews the self-financing trading strategy results.

6.1 EMPIRICAL EVIDENCE ON INDUSTRY RELATEDNESS

This subsection presents the empirical evidence on industry relatedness along the supply chain. The results reported in this subsection are from the firm- and industry-level panel regressions presented in subsection 5.2.

As mentioned earlier, the underlying assumption in the limited-information models is that firms in a given industry have correlated fundamentals with firms in their supplier and customer industries (Menzly and Ozbas, 2010). It is important to note that without correlated fundamentals return cross-predictability should not exist. Table 3 reports the results from panel regressions of firm (industry) returns on assets (ROA) on contemporaneous market-wide ROA and related industry ROAs. Panel A reports the results for the Eurozone sample and panel B contains the results for the EU27 sample. Results for the stock-level specification are presented in columns 1 to 4 whereas results for the industry-level specification are shown in column 5.

Firm-level returns on assets appear to be positively correlated with contemporaneous ROA in supplier and customer industries over and above the market-wide ROA as evidenced by the

[66] For a detailed description of four-factor model risk factors, see subsection 5.4.2.

statistically significant coefficients on $ROA^{supplier}$ and $ROA^{customer}$ in column 1 for both Eurozone and EU27 samples. More specifically, the coefficient on $ROA^{supplier}$ is 0.797 (1.349) and the coefficient on $ROA^{customer}$ is 1.123 (0.956) in the Eurozone (EU27) sample. Importantly, both explanatory variables are statistically significant at the 1% level in both samples. This result is in line with Menzly and Ozbas (2010) who use a similar empirical design on US data and find that firm-level ROAs are positively correlated along the supply chain.

A somewhat surprising result from column 1 is that the coefficient on ROA^{market} is not statistically significant in the EU27 sample. One possible explanation for the weaker explanatory power of aggregate market profitability in the EU27 sample is related to the calculation of the ROA^{market} variable and differences in sample composition. More specifically, compared to the Eurozone sample, the EU27 sample contains more countries that on average have a lower degree of economic, monetary, and financial integration than the Eurozone countries. The higher number of countries and weaker economic linkages may induce stronger border effects which in turn is likely to reduce the explanatory power of aggregate market profitability (see subsection 3.1). In other words, firm-level ROAs in the EU27 sample may be more driven by country-specific factors than ROA^{market} which is calculated over all sample countries by weighting individual firm ROAs with firm assets. It is important to note that while country-specific factors may cause ROA^{market} to be statistically insignificant in the EU27 sample, the coefficients on $ROA^{supplier}$ and $ROA^{customer}$ are still statistically significant. This finding further confirms that the Eurostat input-output tables provide a meaningful description of industry relatedness in both cross-country samples.

Column 2 presents the results for the firm-level regression with indexed firm ROAs as the dependent variable.[67] When individual firm ROAs are adjusted for country-specific differences in average profitability, the magnitude of ROA^{market} increases considerably and the variable becomes statistically significant at the 1% level in both samples. This finding supports the interpretation that country-specific factors cause ROA^{market} to have a weaker explanatory power in column 1. However, the key takeaway from column 2 is that the magnitude of $ROA^{customer}$ declines and the magnitude of $ROA^{supplier}$ increases in both samples when indexed ROAs are used. Based on these results, it appears that firm fundamentals exhibit a stronger positive correlation with supplier industries than customer industries. Nevertheless, $ROA^{customer}$ remains positive and significant at the 1% and 10% levels for the Eurozone and EU27 samples, respectively.

[67] As mentioned earlier in subsection 5.2, annual firm-level ROAs are adjusted for country-specific differences in average profitability in order to address the concern that country-specific factors such as varying economic conditions are confounding the regression results.

Table 3: Industry Relatedness along the Supply Chain

This table presents regressions of annual firm (industry) return on assets (ROA) on contemporaneous market-wide ROA and related customer and supplier industry ROAs. Panel A contains the results for the Eurozone sample and panel B contains the results for the EU27 sample. Columns 1 to 4 present results for the stock-level specification and column 5 reports results for the industry-level specification. The industries are based on NACE rev 1.1 industry classifications. All data items for the regression are retrieved from Worldscope database. Individual firm ROA is calculated as operating income less depreciation and amortization divided by total assets. Industry ROAs are calculated by weighting individual firm ROAs with firm assets for all companies within an industry. ROA^{market} is calculated by weighting individual firm ROAs with firm assets for all sample companies. $ROA^{supplier\ (customer)}$ is calculated by weighting the industry ROAs of supplier (customer) industries by the inter-industry flow of goods and services reported in the Eurostat Input-Output tables. Stock- and industry-level specifications include stock and industry fixed effects, respectively. t-statistics are reported in parentheses. The standard errors used to calculate the t-statistic are adjusted for firm (industry) effects and clustering by year. ***, **, or * indicates that the coefficient estimate is different from zero at the 1%, 5%, or 10% level, respectively.

Panel A: Eurozone	(1)	(2)	(3)	(4)	(5)
Constant	-0.104***	-0.092***	-0.091***	-0.077***	0.004
	(-4.63)	(-4.14)	(-4.25)	(-3.81)	(0.32)
ROA^{market}	0.351**	1.282***	0.212	1.119***	0.116
	(2.56)	(3.63)	(1.63)	(3.46)	(1.44)
$ROA^{supplier}$	0.797***	1.091***	1.005***	1.275***	0.449**
	(4.44)	(6.01)	(5.49)	(7.56)	(2.86)
$ROA^{customer}$	1.123***	0.813***	0.804***	0.462**	0.531***
	(5.22)	(3.89)	(3.68)	(2.61)	(3.36)
R^2	0.594	0.595	0.623	0.630	0.840
No. of observations	19 779	19 779	19 779	19 779	470
Winsorized (5^{th} and 95^{th} percentiles)	No	No	Yes	Yes	No
Indexed firm level ROA	No	Yes	No	Yes	No
Panel B: EU27	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**
Constant	-0.142***	-0.131***	-0.137***	-0.122***	-0.014
	(-7.38)	(-8.63)	(-6.83)	(-7.72)	(-1.04)
ROA^{market}	0.049	1.433***	-0.096	1.343***	0.058
	(0.37)	(5.96)	(-0.76)	(5.49)	(1.43)
$ROA^{supplier}$	1.349***	1.643***	1.495***	1.759***	0.458***
	(5.7)	(6.75)	(6.31)	(7.63)	(6.61)
$ROA^{customer}$	0.956***	0.504*	0.790***	0.289	0.734***
	(5.57)	(2.01)	(4.47)	(1.27)	(6.39)
R^2	0.573	0.637	0.670	0.672	0.901
No. of observations	38 764	38 764	38 764	38 764	510
Winsorized (5^{th} and 95^{th} percentiles)	No	No	Yes	Yes	No
Indexed firm level ROA	No	Yes	No	Yes	No

Columns 3 and 4 present the results from the firm-level regressions with winsorized firm ROAs as dependent variables. The winsorization is performed at the 5^{th} and 95^{th} percentiles for each sample year separately to ensure that outliers are not driving the results. Similar to column 2, regressions with winsorized values reveal that firm fundamentals appear to be more strongly correlated with supplier industries than customer industries. This is evidenced by the smaller coefficient on $ROA^{customer}$ compared to $ROA^{supplier}$ in columns 3 and 4 for both samples. Furthermore, $ROA^{customer}$ remains no longer statistically significant at conventional levels in the EU27 sample when firm ROAs are indexed and winsorized in column 4. On the other hand, $ROA^{supplier}$ remains positive and statistically significant at the 1% level across all specifications. Overall, these findings indicate that firm fundamentals may be more strongly correlated with supplier industries than customer industries. This result is in line with the findings by Menzly and Ozbas (2010), who document that industry-level ROAs are less correlated with customer industries than with supplier industries.

Column 5 reports the results from the industry-level specification with industry ROAs as the dependent variable. Overall, the industry-level results are in line with the firm-level results and provide further evidence that industry fundamentals are positively correlated along the supply chain. Furthermore, the industry-level results should be more robust to the possible influence of country-specific factors on firm profitability because the dependent variables are calculated across all sample countries by weighting individual firm ROAs with firm assets. In the Eurozone (EU27) sample the coefficient on $ROA^{supplier}$ is 0.449 (0.458) which is statistically significant at the 5% (1%) level. Similarly, the coefficient on $ROA^{customer}$ is 0.531 (0.734) and statistically significant at the 1% (1%) level in the Eurozone (EU27) sample. Coefficient on $ROA^{supplier}$ has the same magnitude in both samples, whereas industry ROAs exhibit a slightly stronger positive correlation with customer industries in the EU27 sample.

As a summary, the industry relatedness analysis results confirm that firms along the supply chain have positively correlated fundamentals as evidenced by the positive and statistically significant coefficients on both $ROA^{supplier}$ and $ROA^{customer}$ across different specifications and samples. The magnitude of $ROA^{supplier}$ and $ROA^{customer}$ varies slightly across samples and specifications. However, the results suggest that firm fundamentals are more strongly correlated with supplier industries than customer industries in both samples. Overall, the results confirm the validity of the empirical design for testing return cross-predictability in both Eurozone and EU27 samples. Importantly, the Eurostat input-output tables seem to provide a meaningful description of industry relatedness in both cross-country samples. Based on the results reported in this subsection, hypothesis 1 is accepted.

6.2 EMPIRICAL EVIDENCE ON RETURN CROSS-PREDICTABILITY

This subsection presents the empirical evidence from return cross-predictability analysis. The results reported in this subsection are from the stock- and industry-level regressions presented in subsection 5.3.1.

Table 4 reports the time-series averages of each regression coefficient obtained from monthly cross-sectional regressions of stock (industry) returns on lagged related industry returns and control variables. Panel A reports the results for the Eurozone sample and Panel B contains the results for the EU27 sample. Results for the stock-level specification are presented in columns 1 to 4 along with corresponding t-statistics. Similarly, the results for the industry-level specification are shown in columns 5 and 6.

The results in column 1 confirm that previous-month returns in supplier industries cross-predict stock-level returns as evidenced by the statistically significant coefficient on $r^{supplier}$ in both samples. More specifically, the estimated Fama-MacBeth coefficient on $r^{supplier}$ is 0.141 (0.144) and statistically significant at the 5% (5%) level in the Eurozone (EU27) sample. Importantly, the sign on $r^{supplier}$ is positive which is consistent with firm and industry fundamentals being positively correlated with supplier industries as shown in subsection 6.1. Furthermore, the supplier industry cross-predictability effect appears to be robust to sample composition as evidenced by the same magnitude of $r^{supplier}$ in both Eurozone and EU27 samples. In addition, the supplier industry cross-predictability effect is comparable to known medium-term continuation effects as confirmed by the coefficients on $r^{mtcontinuation}$ (0.133) and $r^{industrymom}$(0.094) in the Eurozone sample. The corresponding coefficients for the EU27 sample are 0.131 and 0.99, respectively. Overall, these results are in line with the limited-information models and empirical evidence by Menzly and Ozbas (2010) who report a statistically significant supplier industry cross-predictability effect using US data. The coefficient on $r^{supplier}$ in their article is 0.114, which is comparable to the cross-predictability effect documented in this paper.

Contrary to the limited-information model prediction, lagged returns in customer industries do not appear to significantly cross-predict stock returns as evidenced by the statistically insignificant coefficients on $r^{customer}$ in column 1. Moreover, this finding is not consistent with Menzly and Ozbas (2010) and Shahrur, Becker and Rosenfeld (2009) who find that previous-month customer returns significantly cross-predict stock-level returns. Regardless of the statistical significance, the customer industry cross-predictability effect appears to be weaker than the supplier industry cross-predictability effect as evidenced by the smaller coefficient on $r^{customer}$ compared to $r^{supplier}$ in both samples. Importantly, this finding is consistent with Menzly and Ozbas (2010) who analyze US data and also report a weaker cross-predictability effect from customer industries in their article. In fact, the coefficient on $r^{customer}$ (0.074) in the EU27 sample has the same magnitude as $r^{customer}$ (0.071) documented by Menzly and Ozbas, albeit the latter one is statistically significant. Furthermore, $r^{customer}$ has a positive sign in column 1 for both samples, which is consistent with the positively correlated customer industry fundamentals reported in the previous subsection. An interesting finding in column 1 is that while the supplier industry cross-predictability effect has the same magnitude in both samples, the customer industry cross-predictability effect appears to be considerably weaker in the Eurozone sample than in the EU27 sample.

Despite the similarities to previous findings, the statistically insignificant coefficient on $r^{customer}$ reported in this paper poses a challenge to the limited-information models which predict that asset returns exhibit cross-predictability as a result of two conditions: (i) the assets have correlated fundamentals and (ii) the markets are informationally segmented. The industry relatedness analysis in subsection 6.1 confirms that assumption (i) holds as firms (industries) appear to have correlated fundamentals with their supplier and customer industries over and above the market. However, it should be noted that firm fundamentals appear to be more strongly correlated with supplier industries than customer industries which may contribute to the weaker customer cross-predictability effects observed in Table 4. In addition, the statistically significant coefficient on $r^{supplier}$ in Table 4 and the previous literature on analyst and investor specialization (see subsection 2.2.2.1) suggest that also assumption (ii) should hold at least to some degree also in the European sample used in this paper. Therefore, the statistically insignificant coefficients on $r^{customer}$ reported in column 1 are likely to be explained by other factors which are discussed below.

One possible explanation for the statistically insignificant customer industry cross-predictability effects in column 1 as opposed to the findings by Menzly and Ozbas (2010) is the potentially larger importance of domestic customers to US companies than to European companies. A comparison of exports-to-Gross Domestic Product (GDP) –ratios for US and Europe reveals that European industries export a larger fraction of their output outside of their domestic (i.e., European) market than US firms export outside of their domestic (i.e., US) market. More specifically, Eurostat statistics show that extra-EU export of goods and services comprised approximately 15.5% of the 2010 GDP in the European Union,[68] while the corresponding exports-to-GDP –ratio for the United

[68] Eurostat: http://epp.eurostat.ec.europa.eu/portal/page/portal/globalisation/indicators

States was only 8.6%.[69] Therefore, the relatively larger exposure of European firms to non-domestic customers may decrease the explanatory power of $r^{customer}$ in the sample because relevant price signals from out-of-sample (non-European) customers are omitted from the customer industry portfolios that are used to predict stock returns in this paper.[70] However, as the signal omission explanation could also apply to supplier industries, it needs to be addressed why $r^{supplier}$ receives a statistically significant coefficient in both samples. One possible explanation is that non-domestic suppliers are relatively less important than non-domestic customers to European firms and thus, signals from supplier industries are sufficiently captured in both samples. Some support for this interpretation is provided by the results in subsection 6.1 which indicate that industry fundamentals in both samples are less correlated with customer industries than supplier industries.

In order to provide more insight on the signal omission explanation, an augmented regression similar to Equation (20) is performed. In the regression industries are allocated into quartiles based on each industry's exports-to-total output –ratios obtained from the Eurostat input-output tables.[71] More specifically, industries with the lowest (highest) level of out-of-sample exports are allocated into the 1st (4th) quartile. Similar to the geographic specialization analysis, the export regression is also performed using stock-level returns as dependent variables to ensure a sufficient amount of observations per quartile. In order to distinguish the possible impact of smaller exposure to out-of-sample customers on the customer cross-predictability effect, the explanatory variable used in the augmented regressions is $r^{customer}$. In other words, the purpose of the augmented regression is to see whether industries that are more (less) exposed to out-of-sample customers (as measured by their exports-to-total output –ratio) exhibit weaker (stronger) customer cross-predictability effects.

The results from the augmented monthly cross-sectional regressions of stock returns on lagged related industry returns interacted with exports-to total output –ratio are provided in Appendix 7. Overall, the results in columns 1 to 4 do not suggest that cross-predictability effects from customer industries would be weaker (stronger) for industries with more (less) exposure to out-of-sample customers. This is evidenced by the lack of a clear pattern in the magnitude of $r^{customer}$ across the quartiles. To the extent that the exports-to-total output –ratio is a good proxy for the magnitude of omitted signals from customer industries, the results in Appendix 7 do not support the signal

---

[69] European Commission: http://trade.ec.europa.eu/doclib/docs/2006/september/tradoc_113465.pdf

[70] This signal omission follows from the empirical design of used in this paper. Since the related industry returns are calculated on a sample of European-only firms, this effectively omits all related industry price signals from outside of Europe and thus, their possible predictive power on stock returns.

[71] Exports are defined as extra-Eurozone exports and extra-EU27 exports for the Eurozone and EU27 samples, respectively. The export figures are obtained from the consolidated Eurostat input-output tables used in this study to determine industry relatedness. Exports data from year 2000 input-output tables are used for years 2000 to 2004 after which year 2005 IO-tables are used for years 2005 to 2009.

omission explanation. However, a likely explanation is that the exports-to-total output –ratios are a very rough proxy for the magnitude of omitted informative customer industry signals and thus, provide an unclean test of the hypothesized relation. Hence, it is difficult to make definitive conclusions from the observed results and more granular firm-level information would be required to obtain a better insight. However, this task is left for future research as it is outside the scope of this study.

It is important to note that the signal omission explanation applies only to differences between results obtained from US data and the results in this paper. On the other hand, Sharur, Becker and Rosenfeld (2009) find a statistically significant customer industry cross-predictability effect using international (non-US) data over a sample period that covers years 1995 to 2007 and thus, the signal omission explanation is not likely to apply to their findings. However, Sharur et al. incorporate only customer industry returns in their regression model which they use to study the cross-predictability effects. The omission of supplier industries from the regression model is not consistent with the limited-information models which posit that returns in both supplier and customer industries cross-predict stock returns. Furthermore, given that the industry relatedness analysis in this paper and the article by Menzly and Ozbas (2010) confirm that industries have correlated fundamentals with both their customer and supplier industries, it is likely that the model of Shahrur et al. suffers from omitted variable bias. In other words, $r^{customer}$ takes up some of the explanatory power of the omitted $r^{supplier}$ variable and thus, appears to be statistically significant in their regression. This concern is further highlighted by the strong correlation between $r^{customer}$ and $r^{supplier}$ found in this paper (correlation coefficient of approximately 0.96 in both samples).

Another explanation for the observed insignificant customer industry cross-predictability observed in column 1 is related to the sample composition in this paper. As mentioned in subsection 3.2, since this study uses cross-country data, it is possible that country effects and/or differences in investor sentiment dominate stock return correlations between industries in different countries and thus, have an impact on the cross-predictability effect. Country effects and differences in investor sentiment may particularly contribute to the observed insignificant coefficient on $r^{customer}$ as customer industries are known to be associated with weaker cross-predictability effects than supplier industries (Menzly and Ozbas, 2010). However, providing evidence on the influence of these factors is outside the scope of this study.

Table 4: Stock- and Industry-Level Cross-Predictability Effects

This table presents Fama-MacBeth coefficient estimates calculated as time-series averages from monthly cross-sectional regressions of stock (industry) returns on lagged related industry returns and control variables. Panel A contains the results for the Eurozone sample and panel B contains the results for the EU27 sample. Columns 1 to 4 present results for the stock-level specification and columns 5 and 6 report results for the industry-level specification. The industries are based on NACE rev 1.1 industry classifications. All data items for the regression are retrieved from Thomson One Banker database. Industry returns are calculated by weighting individual firm returns with firm market values. $r^{supplier\ (customer)}$ is the lagged return on stock j 's (industry i 's) portfolio of supplier (customer) industries in month $t-1$ and is calculated by weighting the industry returns of supplier (customer) industries by the inter-industry flow of goods and services reported in the Eurostat Input-Output tables. $r^{streversal}$ is the previous-month return on stock j , $r^{mtcontinuation}$ is the return on stock j over the 11 months covering months $t-12$ through $t-2$, $r^{indsutrymom}$ is the previous-month return on stock j 's industry and r^{market} is the excess return on the country-specific broad market portfolio in month t . $r^{composite}$ is the related industry return in month $t-1$, calculated as the simple average of $r^{supplier}$ and $r^{customer}$. All return variables are in excess of the risk-free rate defined as the 1-month Euribor. R^2 is calculated as the average of the R^2 collected from the cross-sectional regressions in the first step of the Fama-MacBeth procedure. t-statistics are reported in parentheses. ∗∗∗, ∗∗, or ∗ indicates that the coefficient estimate is different from zero at the 1%, 5%, or 10% level, respectively.

Panel A: Eurozone	(1)	(2)	(3)	(4)	(5)	(6)
Constant	0.000	0.005	-0.002	-0.002	-0.003	-0.002
	(-0.11)	(1.15)	(-0.52)	(-0.87)	(-0.58)	(-0.47)
$r^{supplier}$	0.141**	0.139**	0.097**	0.149***	0.084	-
	(2.40)	(2.35)	(2.04)	(2.79)	(1.20)	-
$r^{customer}$	0.023	0.025	0.047	0.080*	0.134**	-
	(0.40)	(0.43)	(1.02)	(1.70)	(2.13)	-
$r^{streversal}$	-0.058***	-0.059***	-0.035***	-0.029***	-	-
	(-6.71)	(-6.89)	(-5.35)	(-4.17)	-	-
$r^{mtcontinuation}$	0.133***	0.148***	0.111***	0.187**	-	-
	(3.96)	(4.46)	(4.06)	(2.40)	-	-
$r^{industrymom}$	0.094***	0.097***	0.078***	0.086***	0.059**	0.058**
	(5.76)	(5.79)	(5.80)	(5.13)	(2.12)	(2.11)
r^{market}	0.626***	-	0.545***	-	-	-
	(18.96)	-	(20.14)	-	-	-
$r^{composite}$	-	-	-	-	-	0.231***
	-	-	-	-	-	(2.62)
R^2	0.074	0.038	0.086	0.032	0.129	0.112
T	120	120	120	120	120	120
Winsorized (5th and 95th percentiles)	No	No	Yes	Yes	No	No
Indexed stock-level return	No	Yes	No	Yes	No	No
Panel B: EU27	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Constant	-0.003	0.002	-0.004	-0.003	0.000	0.000
	(-0.68)	(0.47)	(-1.07)	(-1.05)	(-0.09)	(-0.01)
$r^{supplier}$	0.144**	0.143**	0.108**	0.155***	0.189**	-
	(2.52)	(2.44)	(2.38)	(3.09)	(2.50)	-
$r^{customer}$	0.074	0.082	0.089**	0.135***	0.085	-
	(1.40)	(1.49)	(2.12)	(2.87)	(1.12)	-
$r^{streversal}$	-0.037***	-0.037***	-0.024***	-0.019***	-	-
	(-4.87)	(-5.16)	(-4.18)	(-3.14)	-	-
$r^{mtcontinuation}$	0.131***	0.143***	0.110***	0.132**	-	-
	(4.46)	(4.96)	(4.51)	(1.98)	-	-
$r^{industrymom}$	0.099***	0.103***	0.080***	0.0905***	0.067**	0.067**
	(5.88)	(5.97)	(5.84)	(5.44)	(2.44)	(2.43)
r^{market}	0.606***	-	0.516***	-	-	-
	(23.39)	-	(26.03)	-	-	-
$r^{composite}$	-	-	-	-	-	0.281***
	-	-	-	-	-	(2.96)
R^2	0.056	0.030	0.068	0.026	0.146	0.120
T	120	120	120	120	120	120
Winsorized (5th and 95th percentiles)	No	No	Yes	Yes	No	No
Indexed stock-level return	No	Yes	No	Yes	No	No

The deliberation on country-specific differences and industry correlations leads to another conclusion, namely that new information from related industries may be at least partially country-specific. For example, Shahrur, Becker and Rosenfeld (2009) who find a statistically significant customer industry cross-predictability effect suggest that new information about customers may be country-specific.[72] Since the related industry returns in this paper are calculated across all sample countries, they do not correctly capture the country-specific component of related industry signals. Therefore, if information from customer industries contains stronger country-specific components than information diffusing from supplier industries, the weaker cross-predictability effects from customer industries documented in this paper may partially be due to the inability of customer industry returns to sufficiently capture these country-specific components. This explanation could be tested by constructing related industry portfolios at the country-level for the whole sample. Unfortunately, the country-specific approach would drastically reduce the number of countries and industries included in this study because several countries would not have enough listed firms to form meaningful country-specific industry portfolios. Also the composition of the related industry portfolios would vary greatly by country, which might distort the comparability of the results across countries. Given the sample limitations, providing empirical evidence on the country-specificity of related industry information is left to future research.

Another possible explanation for the statistically insignificant customer industry cross-predictability effect observed in this paper compared to Menzly and Ozbas (MO, 2010) is related to differences in the sample periods. In their article, MO use a long sample period ranging from 1973 to 2005 whereas this paper uses relatively recent data from a sample period reaching from 2000 to 2009. Since the amount of international trade has increased significantly from 1973 to the present, also the economic linkages across industries in different countries are likely to have become stronger. For example, Kaltenhaeuser (2003) provides evidence that industry returns in both Europe and US have become more exposed to international sector-specific shocks in the late 90's. Based on this, it is possible that firms and industries in MO's sample have significantly less exposure to international markets during the early part of their sample period as compared to firms in this study. This leads to a relatively stronger domestic component in related industry information which may increase the importance of information from domestic customer industries and thus, result in greater customer industry cross-predictability effects. Thus, the sample period selection may at least partially contribute to the significant coefficient on $r^{customer}$ reported by Menzly and Ozbas in their article.

[72] In a sense this explanation is a counter hypothesis to the signal omission explanation which posits that industries are global and unless return information from a sufficiently large geographic area is used, the cross-predictability effect is likely to be weak in the tests.

Unfortunately, the authors do not divide their sample into sub-periods to see whether the customer industry cross-predictability persists throughout their entire sample period.

The overall weaker explanatory power of $r^{customer}$ as opposed to $r^{supplier}$ observed in this paper and by Menzly and Ozbas (2010) could be related to investor information gathering and processing activities. The weaker customer industry cross-predictability effects suggest that informed investors impound customer industry signals more efficiently into stock prices than signals from supplier industries. This could be due to, for example, investors finding it more intuitive to pay attention to information that diffuses from customer industries as opposed to supplier industries. Or alternatively, major customer relationships may be more salient sources of information to investors than important supplier relationships and thus, news from customer industries are more closely monitored. In addition, information from supplier industries may be associated with higher information processing costs than information from customer industries. For example, investors may find it easier to evaluate the news content of customer industry signals due the more salient economic linkage or more straightforward cause-and-effect relation. The above reasons could cause supplier related information, as opposed to customer related news, to be processed by a fewer investors which in turn would result in slower diffusion of information from supplier industries and consequently, stronger cross-predictability effects.

The previous explanations have attempted to explain the differences observed in this paper compared to earlier research and also addressed the overall weaker explanatory power of lagged customer industry returns as opposed to lagged supplier industry returns. However, another interesting finding in column 1, namely the observation that customer industry cross-predictability effects are considerably weaker in the Eurozone sample than in the EU27 sample, has not yet been addressed. One possible explanation for the smaller coefficient on $r^{customer}$ in the Eurozone sample is related to the level of financial integration in the sample countries. Sharur, Becker and Rosenfeld (2009) who study customer cross-predictability effects using international data suggest that the degree to which a particular market is financially integrated with the world also affects the magnitude of the cross-predictability effect. More specifically, they find evidence that the ability of customer industry returns to predict stock returns is weaker in countries whose financial markets are more integrated with the world. The authors argue that higher financial integration with other markets may result in more investors exploiting the supplier-customer lead-lag effect, which in turn causes the effect to be weaker. For example, Baele (2005) notes that the Eurozone countries have undergone a period of significant economic, monetary, and financial integration which suggests an increase in the degree of market integration. Therefore, it is possible that the higher level of market integration among the Eurozone countries, compared to EU27 countries, causes the customer industry cross-predictability effect to be weaker as more investors exploit the phenomenon.

Another explanation that is unrelated to the amount of informed investors exploiting the customer industry cross-predictability effect concerns the ease with which investors' are able to gather and process related industry information. For example, as mentioned earlier, analysts have a hard time incorporating country-specific factors into their security analysis (Bolliger, 2004) which is further complicated by differences in accounting practices between countries (Capstaff, Paudyal and Rees, 2001). Hence, given the highly integrated markets in the Eurozone, it could be that investors in the Eurozone, compared to EU27 countries, suffer from these types of investment analysis problems to a lesser degree. Thus, investors may find it easier to gather and process cross-country related industry information in Eurozone which in turn results in stronger information-impounding demand and informative signals being incorporated faster into prices. This may particularly have an impact on the customer industry cross-predictability effect as it is known to be weaker than the supplier industry cross-predictability effect (Menzly and Ozbas, 2010)

Finally, another explanation attributes the weaker customer industry cross-predictability effects observed in the Eurozone sample, compared to the EU27 sample, to differences in sample composition. More specifically, some individual countries may be contributing to the observed differences in customer industry cross-predictability between the samples. For example, the United Kingdom represents a significant fraction of the EU27 sample. If customer industry cross-predictability effects in the United Kingdom are particularly strong for some reason, they are likely to contribute to the magnitude of the overall customer industry cross-predictability observed in the EU27 sample. However, performing country-by-country analysis of return cross-predictability is outside the scope of this study and hence, it is left for future research to determine the country-specific differences in cross-predictability effects.

Similar to the industry relatedness analysis in the previous subsection, column 2 in Table 4 presents the results for the stock-level regression with indexed stock returns as the dependent variable.[73] Adjustment of stock returns to country-specific differences in stock market development does not yield important new findings. More specifically, $r^{supplier}$ continues to have a positive sign and remains statistically significant in both samples. Similarly, $r^{customer}$ continues to have a positive

[73] The return indexation is performed as in Equation (9) using individual stock returns and country-specific market returns instead of returns on assets. Also note that r^{market} is left out of the specification with indexed stock returns. This is due to country-specific market returns being used to adjust individual stock returns and thus, their effect is already taken into account in the dependent variable.

sign but remains statistically insignificant. The negligible differences between the two specifications support the view that country-specific factors are not strongly influencing the results. Also similar to the previous subsection, columns 3 and 4 contain the results from stock-level regressions with winsorized stock returns as dependent variables. The winsorization process is performed at the 5th and 95th percentiles for each sample month separately in order to ensure that outliers are not driving the results. The most important finding from columns 3 and 4 is that $r^{customer}$ becomes statistically significant in the EU27 sample when stock returns are winsorized. Furthermore, $r^{customer}$ in the Eurozone sample becomes statistically significant at 10% level in column 4. Interestingly, the results in columns 3 and 4 suggest that outliers are influencing the regression results and that cross-predictability effects may exist also from customer industries. The magnitude of the coefficient on $r^{customer}$ increases in both specifications and samples but remains smaller than $r^{supplier}$ which is consistent with Menzly and Ozbas (2010).

Columns 5 and 6 present the results from the industry-level regressions in which industry returns are regressed on lagged related industry returns. Menzly and Ozbas (2010) argue that industry-level returns can be considered economically more important than stock-level returns because they are value-weighted. Moreover, they conjecture that if cross-predictability effects also survive the industry aggregation, then the associated premiums may represent compensation for taking on undiversifiable risk, in which case they could be a permanent feature of stock returns. In addition, Menzly and Ozbas point out that, according to limited-information models, economically related assets should exhibit cross-predictability regardless of the unit of analysis. Hence, the cross-predictability effect should also exist at the industry-level as long as the aggregation process does not result in the complete elimination of market segmentation. Based on these arguments, the industry-level regression can be seen as more than a robustness test.

In the light of the stock-level results, the findings in column 5 appear surprising as the coefficient on $r^{customer}$ (0.134) becomes significant at the 5% level whereas the coefficient on $r^{supplier}$ (0.084) becomes statistically insignificant in the Eurozone sample. This result is puzzling as one would expect $r^{supplier}$ to receive a statistically significant coefficient and conversely, $r^{customer}$ to have a considerably weaker explanatory power than $r^{supplier}$. However, the results from the EU27 sample are more in line with the stock-level results as $r^{supplier}$ (0.189) is significant at the 5% level and $r^{customer}$ (0.067) is statistically insignificant. A likely explanation for the results in column 5 is that multicollinearity causes the regression coefficients to be erratic.[74] The likelihood that multicollinearity is affecting the results is evidenced by two observations: (i) the high correlation between $r^{supplier}$ and $r^{customer}$ which rises up to 0.59 in both samples and (ii) the small sample size which is only 45 (49) observations in the Eurozone (EU27) sample for each monthly regression.[75] Therefore, in order to alleviate multicollinearity, the regression specification in column 6 combines $r^{supplier}$ and $r^{customer}$ into $r^{composite}$ which is the simple average of the two explanatory variables.[76] Results in column 6 show that the coefficient on $r^{composite}$ is 0.231 and 0.281 in the Eurozone and EU27 samples, respectively. Both coefficients are statistically significant at the 1% and indicate that return cross-predictability exists also at the industry-level which is consistent with Menzly and Ozbas (2010). Importantly, the magnitude of $r^{composite}$ reported in Table 4 is comparable to the combined industry-level coefficients on $r^{supplier}$ (0.113) and $r^{customer}$ (0.075) reported by Menzly and Ozbas..

In addition to lagged related industry returns, the regressions in Table 4 contain several control variables that are known to predict stock returns. Overall, all the control variables have expected signs and are statistically significant either at the 1% or 5% level across different specifications and samples. More specifically, $r^{streversal}$ which controls for short-term reversals (see subsection 2.1.1) at the stock-level has a negative and significant coefficient in each regression. This implies that stock returns tend to exhibit negative autocorrelation in the short-term which is in line with previous empirical evidence on short-term reversals (see e.g., Lehmann, 1990; Jegadeesh, 1990). $r^{mtcontinuation}$ which controls for the momentum effect in stock returns (see subsection 2.1.2) has a positive and significant coefficient across different specifications and samples. This result is also consistent with earlier research that documents the tendency of past winners to continue outperform past losers in medium-term horizon (see e.g., Jegadeesh and Titman, 1993). Similarly, $r^{industrymom}$ which controls for industry-level momentum is positive and statistically significant in all regressions and samples which is in line with previous empirical evidence (see e.g., Moskowitz and

[74] According to Dougherty (1992, pp. 133) multicollinearity refers to high correlation between independent variables that causes an econometric model to become unsatisfactory. He points out that multicollinearity is present in almost every regression model but it becomes a problem when combined with one or more factors that also determine the variances of the regression coefficients. Thus, when combined with, for example, a small number of observations, multicollinearity can produce erratic regression results. Multicollinearity may, for example, inflate standard errors and cause *t*-statistics to be too low as well as cause unexpected changes in the coefficient magnitudes (Dougherty, 1992, pp. 133).

[75] It is important to keep in mind that the Fama-MacBeth method used to calculate the regression coefficients and *t*-statistics for the return cross-predictability regressions requires monthly regressions to be conducted on the sample variables. Thus, this procedure reduces the sample size down to the number of different industries in the industry-level regression. More specifically, the Eurozone sample contains only 45 observations (industries) and the EU27 sample 49 observations (industries) per monthly regression. The small sample size is not a problem with the firm-level specification as the number of firms is considerably higher than the number of industries.

[76] According to Dougherty (1992, pp. 133) combining the correlated explanatory variables should mitigate the effects of multicollinearity.

Grinblatt, 1999). Finally, r^{market} has a positive and significant coefficient in each regression which is consistent with the Capital Asset Pricing Model prediction (Sharpe, 1964; Lintner, 1965).

As a summary, the stock-level results in Table 4 provide partial support to the limited-information models. More specifically, both $r^{supplier}$ and $r^{customer}$ remain consistently positive across different regression specifications and samples which is in line with the limited-information models which posit that economically linked assets exhibit cross-predictability effects that are of the same sign as the correlation between the assets' fundamentals. Importantly, the coefficient on $r^{supplier}$ is statistically significant and has the same magnitude across different specifications and samples confirming that previous-month supplier industry returns cross-predict stock returns. This finding is consistent with the limited-information models and earlier empirical evidence. On the other hand, the empirical evidence on customer industry cross-predictability effects in this paper is more mixed and difficult to interpret. Similar to Menzly and Ozbas (2010), this paper finds that customer industry cross-predictability effects are weaker than supplier industry cross-predictability effects. However, the coefficient on $r^{customer}$ is statistically significant only after elimination of outliers which is only partially in line with earlier research that documents a statistically significant cross-predictability effect also for customer industries. Possible explanations for the statistically insignificant customer industry cross-predictability are related to omission of informative customer industry signals, country effects and/or differences in investor sentiment, country-specific components of related industry information and differences in sample period. An interesting finding is also that the customer industry cross-predictability effect is considerably weaker in the Eurozone sample than in the EU27 sample. Possible explanations for this are related to the higher level of integration among Eurozone countries which may cause more investors to exploit the customer industry cross-predictability effect and/or investors to process cross-country related industry information more efficiently. Alternatively, individual countries may be contributing to the observed differences.

Results from the industry-level regression in column 5 are likely to suffer from multicollinearity and need to be interpreted carefully. However, combining the independent variables into $r^{composite}$ in column 6 yields a positive and significant cross-predictability effect that should be free of multicollinearity. This result confirms that cross-predictability effects also survive the industry aggregation. Overall, the findings in this subsection partially confirm the main prediction of the limited-information models which posits that lagged returns in related industries cross-predict firm- and industry-level returns: the results in Table 4 show that previous-month returns in supplier industries cross-predict stock returns and are also indicative of a weaker cross-predictability effect associated with customer industries. In a similar vein, the results also support the gradual information diffusion hypothesis of Hong and Stein (1999). Based on the findings in this subsection, hypothesis 2.1 is accepted and hypothesis 2.2 is partially accepted with the caveat that further evidence is needed to evaluate the magnitude of the customer industry cross-predictability effect. All in all, the limited-information model by Menzly and Ozbas (2010) provides a partially compelling explanation for stock return predictability in a cross-country sample.

6.3 EMPIRICAL EVIDENCE ON THE EFFECT OF INFORMED INVESTORS

This subsection presents empirical evidence on the limited-information models prediction which posits that the magnitude of return cross-predictability is negatively related to the level of information in the market, that is, the amount of informed investors. The results reported in this subsection are from the augmented regressions presented in subsection 5.3.2. First, this subsection reviews the results from augmented regressions with market capitalization as a proxy for the amount of informed investors. Then the results from a similar regression using analyst coverage as a proxy for the amount of informed investors are presented.

Table 5 reports the time-series averages from augmented monthly cross-sectional regressions of stock returns on lagged related industry returns interacted with firm size. Each month t stocks are ranked into five quintiles based on firm size which is measured as the firm market capitalization in month $t-1$. Smallest firms are allocated into quintile 1 while largest firms are allocated into quintile 5. The explanatory variable $r^{composite}$ is the simple average of $r^{supplier}$ and $r^{customer}$. Panel A reports the results for the Eurozone sample and Panel B contains the results for the EU27 sample. Furthermore, Appendix 8 presents the descriptive statistics for each firm size quintile. The main observation from these statistics is that different countries and industries appear to be well represented in each quintile. This suggests that unobserved country- or industry-specific differences should not be driving the differences between the quintiles.

The results in column 1 confirm that small firms exhibit stronger cross-predictability effects than large firms. This is evidenced by the magnitude of $r^{composite}$ which declines sharply in both samples when moving down from the 1st quintile (smallest market capitalization firms) to the 5th quintile (largest market capitalization firms). Furthermore, while the cross-predictability effect in the 1st quintile is significant at the 5% level, the largest firm size quintiles exhibit no cross-predictability effects at conventional levels. Importantly, the magnitude of the cross-predictability effect declines monotonically across the size quintiles in the Eurozone sample. Also in the EU27 sample the cross-predictability effect declines monotonically once past the quintile with the smallest firms. Overall, the findings in column 1 are consistent with Shahrur, Becker and Rosenfeld (2009) who classify stocks into four groups based on market value and find customer industry cross-predictability effects in all size portfolios except the large-cap portfolio. To the extent that firm size provides a good proxy for the amount of informed investors, the findings in column 1 are consistent with the limited-information model prediction that the magnitude of the cross-predictability effect decreases with the number of informed investors.

As mentioned, the cross-predictability effect declines monotonically across the quintiles only in the Eurozone sample in column 1. In several other specifications, including the EU27 sample, the cross-predictability effect starts to decline monotonically once one moves past the 1st quintile (smallest firms). This observation could be due to stocks in the smallest size quintile reacting with a longer than 1-month lag to price innovations from related industries, for example, due to lower trading frequencies. The result is also in line with Hong, Lim and Stein (2000) who study the profitability of momentum trading strategies. Similar to this paper, they find that the profitability of momentum strategies declines sharply with firm size only once one moves past the very smallest stocks. Hong et al. attribute their finding to thin markets, that is, more limited investor participation in the smallest capitalization stocks.

Similar to the previous sections, columns 2 to 4 present the regression results using indexed and/or winsorized stock returns as dependent variables. In general, adjustment of stock returns to country-specific differences in stock market development and elimination of outliers do not significantly affect the main findings from column 1. In other words, smaller firm quintiles continue to exhibit strong and statistically significant cross-predictability effects whereas the largest firm quintile exhibits no cross-predictability effects. However, an interesting observation from columns 2 to 4 is that the differences in the magnitude of the cross-predictability effect across the middle quintiles (2nd, 3rd and 4th quintiles) become less pronounced, particularly in the EU27 sample. This is evidenced by the negligible differences in the magnitude of the cross-predictability effect between the 2nd and 3rd quintiles in Panel B columns 2 to 4. A likely explanation for the observed small differences between the middle quintiles is that market capitalization is a noisy proxy for the amount of informed investors. As mentioned in subsection 5.3.2, firm size may capture multiple firm-specific features that possibly confound the results between different size quintiles and thus, it may not constitute a clean test of the effect of informed investors on the cross-predictability effect. In order to address this concern, another proxy for the amount of informed investors, namely analyst coverage, is constructed.

Table 5: Firm Size and Cross-Predictability Effects

This table presents Fama-MacBeth coefficient estimates calculated as time-series averages from augmented monthly cross-sectional regressions of stock returns on lagged related industry returns interacted with firm size. Panel A contains the results for the Eurozone sample and panel B contains the results for the EU27 sample. The industries are based on NACE rev 1.1 industry classifications. All data items are retrieved from Thomson One Banker database. Each month t stocks are ranked into five quintiles based on firm size which is measured as the firm market capitalization in month $t-1$. Stocks with the lowest market capitalizations are allocated into quintile 1 and stocks with the highest market capitalizations are allocated into quintile 5. $r^{composite}$ is the related industry return in month $t-1$ and is calculated as the average of $r^{supplier}$ and $r^{customer}$. R^2 is calculated as the average value of the R^2s collected from the cross-sectional regressions in the first step of the Fama-MacBeth procedure. t-statistics are reported in parentheses. ∗∗∗, ∗∗, or ∗ indicates that the coefficient estimate is different from zero at the 1%, 5%, or 10% level, respectively.

Panel A: Eurozone	(1)	(2)	(3)	(4)
Constant	-0.005	0.002	-0.007	0.001
	(-1.21)	(0.41)	(-1.57)	(0.24)
$r^{composite} \times Rank_{t-1}$ (1st quintile - low)	0.267**	0.251**	0.212**	0.192**
	(2.18)	(2.09)	(2.21)	(2.11)
$r^{composite} \times Rank_{t-1}$ (2nd quintile)	0.246**	0.258**	0.196**	0.225***
	(2.29)	(2.49)	(2.16)	(2.63)
$r^{composite} \times Rank_{t-1}$ (3rd quintile)	0.196**	0.226**	0.168**	0.203**
	(2.01)	(2.30)	(2.07)	(2.49)
$r^{composite} \times Rank_{t-1}$ (4th quintile)	0.183**	0.213**	0.158**	0.196***
	(2.10)	(2.41)	(2.13)	(2.65)
$r^{composite} \times Rank_{t-1}$ (5th quintile - high)	0.048	0.096	0.038	0.090
	(0.48)	(0.99)	(0.42)	(1.06)
R^2	0.012	0.012	0.016	0.015
T	120	120	120	120
Winsorized (5th and 95th percentiles)	No	No	Yes	Yes
Indexed stock return	No	Yes	No	Yes

Panel B: EU27	(1)	(2)	(3)	(4)
Constant	-0.005	0.000	-0.006	-0.001
	(-0.97)	(-0.04)	(-1.28)	(-0.26)
$r^{composite} \times Rank_{t-1}$ (1st quintile - low)	0.229**	0.242**	0.204**	0.218**
	(2.19)	(2.30)	(2.51)	(2.71)
$r^{composite} \times Rank_{t-1}$ (2nd quintile)	0.272***	0.317***	0.228***	0.256***
	(3.00)	(3.69)	(3.05)	(3.66)
$r^{composite} \times Rank_{t-1}$ (3rd quintile)	0.262***	0.322***	0.235***	0.276***
	(2.63)	(3.38)	(2.86)	(3.62)
$r^{composite} \times Rank_{t-1}$ (4th quintile)	0.250***	0.283***	0.239***	0.263***
	(2.80)	(3.25)	(3.23)	(3.63)
$r^{composite} \times Rank_{t-1}$ (5th quintile - high)	0.085	0.136	0.083	0.125
	(0.88)	(1.46)	(0.98)	(1.54)
R^2	0.010	0.010	0.013	0.012
T	120	120	120	120
Winsorized (5th and 95th percentiles)	No	No	Yes	Yes
Indexed stock return	No	Yes	No	Yes

Table 6 reports the time-series averages from augmented monthly cross-sectional regressions of stock returns on lagged related industry returns interacted with analyst coverage. More specifically, each month t stocks are ranked into five quintiles based on analyst coverage which is measured as the numerical count of EPS estimates included in the mean EPS estimate for the stock in month $t-1$. Firms with the lowest (highest) level of analyst coverage are allocated into quintile 1 (5). Panel A presents the results for the Eurozone sample and Panel B contains the results for the EU27 sample. The interpretation of the analyst coverage quintiles is the same as with firm size quintiles: the 1st (5th) quintile contains the lowest (highest) number of informed investors and thus, should exhibit the strongest (weakest) cross-predictability effects according to the limited-information model prediction.

The results in column 1 for the EU27 sample show that the cross-predictability effect decreases monotonically across the quintiles, once one moves past the 1st quintile. This result is consistent with the findings by Menzly and Ozbas (2010) who also document that the magnitude of the cross-predictability effect declines monotonically once past the stocks with the lowest level of analyst following. Importantly, the sharp drop in the magnitude of $r^{composite}$ from the 1st quintile (lowest analyst coverage) to the 5th quintile (highest analyst coverage) in both EU27 and Eurozone samples is supportive of the limited-information models which posit that cross-predictability decreases with the amount of informed investors. More specifically, stocks in the lowest analyst coverage quintiles tend to exhibit two to three times stronger cross-predictability effects than the average stock, which also is in line with Menzly and Ozbas. Overall, these results in column 1 are supportive of the limited-information models with the caveat that the analyst coverage measure provides a good proxy for the amount of informed investors.

A surprising observation from Table 6 is that $r^{composite}$ mainly remains statistically insignificant in the Eurozone sample across specifications. This finding is likely to be explained by the small number of observations per quintile in the Eurozone sample as compared to the EU27 sample. In order to verify this, Appendix 9 presents descriptive statistics for each analyst coverage quintile. It can be seen that apart from the 1st quintile, the average number of observations per quintile is particularly low in the Eurozone sample. The small sample sizes are due to the limited availability of data from I/B/E/S detail database. On the other hand, the relatively large number of observations in the 1st quintile in both samples is caused by a disproportionately high number of sample firms with only one analyst following them. Appendix 9 also shows that different countries and industries appear to be well represented across the analyst coverage quintiles. This decreases the likelihood that the observed differences between quintiles could be attributed to unobserved country- or industry-specific factors. Another important observation from the descriptive statistics is the positive relation between analyst coverage and market capitalization. It seems that larger firms have more analysts following them, which is in line with Bhushan (1989) who finds that firm size is strongly correlated with analyst following. This suggests that the results reported in Table 6 are not completely free of unobserved firm size effects. However, the correlation between the analyst coverage measure and firm market capitalization is low (approximately 0.15 in both samples). Therefore, unobserved firm size effects are likely to have only minor influence on the differences between the analyst coverage quintiles.

Overall, the Eurozone sample in Panel A provides a less clear pattern across different analyst coverage quintiles than the EU27 sample. Since the coefficient on $r^{composite}$ remains statistically insignificant in each quintile, it is difficult to draw definitive conclusions from the observed pattern. However, one possible explanation for the less clear pattern is that analyst coverage provides a better proxy for the amount of informed investors in the EU27 sample than in the Eurozone sample. For example, Clement, Rees and Swanson (2000) find that analysts issue forecasts more frequently in UK than in Germany. In addition, Capstaff, Paudyal and Rees (2001) document that analyst EPS forecasts are most accurate in UK and the Netherlands in a sample of eight European countries. Findings in both articles are suggestive of a more pronounced role of analysts in UK. Therefore, it could be that analyst coverage provides a better proxy for the amount of informed investors in the EU27 sample than in the Eurozone sample due to the large number of UK firms in the prior sample.

Similar to the previous sections, columns 2 to 4 provide the results from regressions with indexed and/or winsorized stock returns as dependent variables. Overall, the results from these regressions are consistent with those in column 1: the magnitude of $r^{composite}$ declines sharply from the 1st quintile to the 5th quintile across different specifications in both samples. This suggests that the cross-predictability effect is negatively correlated with the level of analyst coverage which is in line with the limited-information model prediction. Importantly, the pattern of monotonic decline once past the 1st quintile in the EU27 sample continues to hold across all specifications. Likewise, the more obscured pattern for the Eurozone sample persists and the coefficients remain mainly statistically insignificant.

Table 6: Analyst Coverage and Cross-Predictability Effects

This table presents Fama-MacBeth coefficient estimates calculated as time-series averages from augmented monthly cross-sectional regressions of stock returns on lagged related industry returns interacted with analyst coverage. Panel A contains the results for the Eurozone sample and panel B contains the results for the EU27 sample. The industries are based on NACE rev 1.1 industry classifications. Return data is retrieved from Thomson One Banker database. Each month t stocks are ranked into five quintiles based on the level of analyst coverage in month $t-1$. Stocks with the lowest level of analyst coverage are allocated into quintile 1 and stocks with the highest level of analyst coverage are allocated into quintile 5. Analyst coverage measure is retrieved from I/B/E/S detail database and is defined as the numerical count of EPS estimates included in the mean EPS estimate for the stock in month $t-1$. $r^{composite}$ is the related industry return in month $t-1$ and is calculated as the average of $r^{supplier}$ and $r^{customer}$. R^2 is calculated as the average value of the R^2s collected from the cross-sectional regressions in the first step of the Fama-MacBeth procedure. t-statistics are reported in parentheses. ∗∗∗, ∗∗, or ∗ indicates that the coefficient estimate is different from zero at the 1%, 5%, or 10% level, respectively.

Panel A: Eurozone	(1)	(2)	(3)	(4)
Constant	-0.007	0.000	-0.008*	-0.002
	(-1.31)	(0.04)	(-1.75)	(-0.45)
$r^{composite} \times Rank_{t-1}$ (1^{st} quintile - low)	0.178	0.167	0.176*	0.170*
	(1.56)	(1.54)	(1.89)	(1.96)
$r^{composite} \times Rank_{t-1}$ (2^{nd} quintile)	0.163	0.178	0.162	0.165
	(1.13)	(1.32)	(1.44)	(1.54)
$r^{composite} \times Rank_{t-1}$ (3^{rd} quintile)	0.215	0.200	0.187*	0.190*
	(1.50)	(1.43)	(1.68)	(1.75)
$r^{composite} \times Rank_{t-1}$ (4^{th} quintile)	0.139	0.171	0.127	0.159*
	(1.18)	(1.53)	(1.36)	(1.79)
$r^{composite} \times Rank_{t-1}$ (5^{th} quintile - high)	0.079	0.061	0.094	0.089
	(0.74)	(0.62)	(1.02)	(1.07)
R^2	0.013	0.013	0.015	0.014
T	120	120	120	120
Winsorized (5^{th} and 95^{th} percentiles)	No	No	Yes	Yes
Indexed stock return	No	Yes	No	Yes

Panel B: EU27	(1)	(2)	(3)	(4)
Constant	-0.008	-0.002	-0.008	-0.003
	(-1.28)	(-0.33)	(-1.55)	(-0.60)
$r^{composite} \times Rank_{t-1}$ (1^{st} quintile - low)	0.287**	0.284***	0.233***	0.233***
	(2.54)	(2.64)	(2.61)	(2.74)
$r^{composite} \times Rank_{t-1}$ (2^{nd} quintile)	0.340**	0.371***	0.268**	0.291***
	(2.50)	(2.80)	(2.57)	(2.84)
$r^{composite} \times Rank_{t-1}$ (3^{rd} quintile)	0.317**	0.293**	0.260***	0.252***
	(2.53)	(2.42)	(2.65)	(2.71)
$r^{composite} \times Rank_{t-1}$ (4^{th} quintile)	0.213*	0.231**	0.184**	0.199**
	(1.88)	(2.11)	(2.05)	(2.32)
$r^{composite} \times Rank_{t-1}$ (5^{th} quintile - high)	0.080	0.065	0.096	0.087
	(0.67)	(0.58)	(0.99)	(0.97)
R^2	0.010	0.010	0.011	0.010
T	120	120	120	120
Winsorized (5^{th} and 95^{th} percentiles)	No	No	Yes	Yes
Indexed stock return	No	Yes	No	Yes

As a summary, the results in this subsection show that the magnitude of the cross-predictability effect declines with the amount of informed investors. This finding is supportive of the limited-information models which predict that cross-predictability effects decrease with the number of informed investors whose information-impounding demand suppresses the predictability. More specifically, using both firm size and analyst coverage as proxies for the number of informed investors, this paper shows that the cross-predictability effects disappear in the quintile with the highest number of informed investors. This is evidenced by the large spreads between the 1st and 5th quintiles and the statistically insignificant coefficient on $r^{composite}$ in the 5th quintile across all specifications and both samples. Importantly, the cross-predictability effect either declines monotonically across quintiles or once one moves past the 1st quintile in several specifications. Overall, the results reported in this subsection are in line with the findings by Menzly and Ozbas (2010) who use analyst coverage and institutional ownership as proxies for the amount of informed investors and Shahrur, Becker and Rosenfeld (2009) who study the impact of size on the customer industry cross-predictability effect. Moreover, the results are supportive of limited-information models to the extent that firm size and the number of analysts can be considered good proxies for the number of informed investors. Based on the results and analysis in this section, Hypotheses 3.1 and 3.2 are accepted.

6.4 EMPIRICAL EVIDENCE ON THE EFFECT OF INVESTOR GEOGRAPHIC SPECIALIZATION

This subsection presents empirical evidence on the relation between investor specialization along geographic boundaries and the magnitude of the cross-predictability effect. The results reported in this subsection are from the augmented regression presented in subsection 5.3.3.

Table 7 presents the time-series averages from augmented monthly cross-sectional regressions of stock returns on lagged related industry returns interacted with $Geodisp_j$ measure. Each month t stocks are ranked into quartiles based on the geographic dispersion of their related industries which is measured by the self-constructed $Geodisp_j$ variable in month t – 1. Firms with the least (most) geographically dispersed related industries in the sample are allocated into quartile 1 (4). Panel A reports the results for the Eurozone sample and panel B contains the results for the EU27 sample. It is important to note that the interpretation of the $Geodisp_j$ quartiles is opposite to the firm size and analyst coverage quintiles: the 1st (4th) quartile contains the firms with the least (most) geographically dispersed related industries and thus, should exhibit the weakest (strongest) cross-predictability effects according to Hypothesis 3.3.

The results in column 1 indicate that the cross-predictability effect increases with the geographic dispersion of related industries as evidenced by the sharp increase in the magnitude of $r^{composite}$ from the 1st quartile (lowest related industry geographic dispersion) to the 4th quartile (highest related industry geographic dispersion) in both samples. More specifically, the cross-predictability effect increases monotonically across the quartiles until it peaks in the 3rd quartile. Despite the small drop, the magnitude of the cross-predictability effect in the 4th quartile remains higher than in the 1st quartile in both samples. Importantly, the finding that $r^{composite}$ increases across quartiles is consistent with Hypothesis 3.3 which posits that the magnitude of the cross-predictability effect increases with the geographic dispersion in related industries. Similar to previous sections, columns 2 to 4 present the regression results with indexed and/or winsorized dependent variables. The indexation of stock returns and elimination of outliers do not seem to affect the main findings. Overall, the findings in Table 1 suggest that the magnitude of return cross-predictability is positively related to the geographic dispersion of informative signals from related industries.

One possible explanation for the observed decrease in the cross-predictability effect between the 3rd and 4th quartiles is related to the $Geodisp_j$ measure that is used to allocate firms into different quartiles. First of all, $Geodisp_j$ is calculated at the industry level and thus, does not accurately capture the true geographic dispersion of an individual firm's related industries. Secondly, since $Geodisp_j$ measures the level of geographic dispersion of related industries based on the sample composition, it is dependent on the representativeness of the sample. In other words, if the geographic dispersion of industries in the sample is not representative of the underlying population, then neither is the $Geodisp_j$ measure.[77] Based on these limitations, it is possible that all individual firms are not allocated correctly into quartiles. For example, there may be firms in the 4th quintile which actually have less geographically dispersed related industries than some of the firms in the 3rd quintile and vice versa. Hence, the limitations of the $Geodisp_j$ measure may at least partially contribute to the smaller coefficient on $r^{composite}$ in the 4th quartile.

Furthermore, the geographic dispersion of supplier and customer industries, as measured by $Geodisp_j$, may not accurately proxy for geographic dispersion of informative signals from these industries. In other words, all geographic dispersion of related industries may not be informationally relevant from an investor's stand point. For example, a firm may have exposures to large geographic markets through its business operations, but still be most reliant on only a fraction of its overall market. Thus, from an investors' point of view it might be sufficient to focus only on the

[77] This problem is alleviated by calculating $Geodisp_j$ from an unscreened sample that covers the whole EU27 area.

firm's key market areas in information gathering activities, which effectively reduces the geographic dispersion of informative signals. Put otherwise, increases in related industry geographic dispersion may not always result in increases in the geographic dispersion of informative signals. This is possibly a concern since $Geodisp_j$ measures the geographic dispersion of related industries which is thought to proxy for the geographic dispersion of the informative signals from these industries. Thus, it could be that the difference in geographic dispersion of informative signals between the 3rd and 4th quartiles is actually negligible. This deliberation clearly highlights the difficulty of constructing a measure that accurately captures the true geographic dispersion of informative signals from related industries. Regardless of the possible limitations of the $Geodisp_j$ measure, it should still provide a sufficient estimate of the average geographic dispersion of informative signals from related industries.

The observed pattern between the 3rd and 4th quartiles may also be attributable to the cost function associated with processing geographically dispersed information.[78] It is possible that the costs of information processing faced by investors do not increase continuously with the geographic dispersion of information as assumed in Table 7. Instead, the relation between geographic dispersion of industries and costs to information processing may rather resemble a step function that does not grow monotonically. The more indirect relation could be due to, for example, synergies associated with gathering and processing information from a geographically dispersed but otherwise integrated market area such as the Nordic countries. In this case, identifying the actual cost function would be required to determine the true costs an investor faces when gathering information from a geographically dispersed area. Thus, it is possible that even though firms in the 4th quartile have more geographically dispersed related industries, they may also contain more information processing synergies than firms in the 3rd quartile which results in the observed drop in the magnitude of $r^{composite}$. Another possible explanation is that the marginal effect of geographic dispersion of informative signals on the magnitude of the cross-predictability effect becomes zero after a certain level of geographic dispersion is reached. This could be due to, for example, limits to investor geographic specialization.

Furthermore, it is possible that unobserved country effects may be contributing to the decline in the cross-predictability effect in the 4th quartile. As mentioned in subsection 3.2, earlier academic research has documented that country-specific factors are important in explaining low correlations of stock market returns across countries. Since related industry signals, by definition, diffuse from a

---

[78] Here, costs to information processing refer to cognitive limitations as well as actual costs that an investor faces when gathering and processing information from related industries.

wider variety of different countries when geographic dispersion in related industries increases, it is possible that unobserved country effects have a greater influence on the results in the high geographic dispersion quartiles. Thus, more pronounced country-specific factors in the 4th quartile may cause the coefficient on $r^{composite}$ to have a smaller magnitude. Another related concern is that since the allocation into $Geodisp_j$ quartiles is performed on industry-level, the importance of country effects may vary between industries in different quartiles. For example, Griffin and Karolyi (1998) note that traded-goods industries tend to have higher industry effects than nontraded-goods industries.[79] The concern here is that if non-traded goods (traded-goods) industries are concentrated into certain quartiles, then country effects (industry effects) may have a disproportionately large influence in these quartiles. This in turn may cause cross-predictability effects to be weaker (stronger) in quintiles with relatively more non-traded goods (traded goods) industries as country-effects (industry effects) are likely to dominate industry return correlations.

In order to assess this possibility, Appendix 10 shows the frequencies with which a given industry enters a $Geodisp_j$ quartile during the sample period. As expected certain industries are concentrated into certain quartiles throughout the sample period. This follows from the fact that even though the allocation into quartiles is performed each month, the geographic dispersion of related industries changes slowly. Appendix 10 also reports whether an industry is considered a traded-goods industry or a nontraded-goods industry.[80] A closer examination reveals that non-traded goods industries are equally present in all quartiles with the exception that the 2nd quartile contains a larger fraction of nontraded-goods industries. These findings suggest that country effects have an equal chance of dominating industry return correlations within all quartiles excluding the 2nd quartile which may exhibit stronger country effects. However, since related industry signals arrive from a geographically more concentrated area in the 2nd quartile, also the influence of country effects is likely to be less severe than suggested by the industry composition. Overall, it seems unlikely that country effects would significantly bias the results between different quartiles in Table 7. This is also supported by results from regressions with country-adjusted returns as dependent variables (column 2) which are similar to those in column 1.

---

[79] Traded-goods industries are defined as industries which produce goods that are traded internationally such as automobiles, computers, pharmaceuticals and resources (e.g., oil and coal) whereas nontraded-goods industries are industries that produce goods which are not traded internationally such as media, heavy construction, plantations, conglomerates and real estate (Griffin and Karolyi, 1998).

[80] The division into nontraded- and traded-goods industries is based on Griffin and Karolyi (1998). It is important to note that the division shown in this paper is only indicative as NACE industries do not entirely match the descriptions by Griffin and Karolyi.

Table 7: Investor Geographic Specialization and Cross-Predictability Effects

This table presents Fama-MacBeth coefficient estimates calculated as time-series averages from augmented monthly cross-sectional regressions of stock returns on lagged related industry returns interacted with $Geodisp_{j,t}$ variable. Panel A contains the results for the Eurozone sample and panel B contains the results for the EU27 sample. The industries are based on NACE rev 1.1 industry classifications. All data items are retrieved from Thomson One Banker database. Each month t stocks are ranked into quartiles based their industry's $Geodisp_{j,t}$ variable which measures the geographic dispersion of related industries in month $t-1$. Industries with the lowest level of geographic dispersion are allocated into quartile 1 and industries with the highest level of geographic dispersion are allocated into quartile 4. $r^{composite}$ is the related industry return in month $t-1$ and is calculated as the average of $r^{supplier}$ and $r^{customer}$. R^2 is calculated as the average value of the R^2 collected from the cross-sectional regressions in the first step of the Fama-MacBeth procedure. t-statistics are reported in parentheses. ***, **, or * indicates that the coefficient estimate is different from zero at the 1%, 5%, or 10% level, respectively.

Panel A: Eurozone	(1)	(2)	(3)	(4)
Constant	-0.005	0.002	-0.007	0.001
	(-1.16)	(0.38)	(-1.53)	(0.24)
$r^{composite} \times Rank_{t-1}$ (1st quartile - low)	0.161*	0.185**	0.139*	0.171**
	(1.86)	(2.19)	(1.91)	(2.47)
$r^{composite} \times Rank_{t-1}$ (2nd quartile)	0.286**	0.322***	0.195**	0.236**
	(2.49)	(2.89)	(1.99)	(2.59)
$r^{composite} \times Rank_{t-1}$ (3rd quartile)	0.309***	0.352***	0.269***	0.296***
	(3.03)	(3.36)	(3.08)	(3.44)
$r^{composite} \times Rank_{t-1}$ (4th quartile -high)	0.223*	0.258**	0.187*	0.220**
	(1.67)	(2.04)	(1.76)	(2.20)
R^2	0.009	0.011	0.009	0.010
T	120	120	120	120
Winsorized (5th and 95th percentiles)	No	No	Yes	Yes
Indexed stock return	No	Yes	No	Yes

Panel B: EU27	(1)	(2)	(3)	(4)
Constant	-0.005	0.000	-0.006	-0.001
	(-0.86)	(0.04)	(-1.20)	(-0.16)
$r^{composite} \times Rank_{t-1}$ (1st quartile - low)	0.152*	0.183**	0.152**	0.175***
	(1.85)	(2.37)	(2.34)	(2.86)
$r^{composite} \times Rank_{t-1}$ (2nd quartile)	0.294**	0.328**	0.272**	0.291***
	(2.18)	(2.51)	(2.44)	(2.73)
$r^{composite} \times Rank_{t-1}$ (3rd quartile)	0.431***	0.481***	0.379***	0.412***
	(3.67)	(4.06)	(3.93)	(4.25)
$r^{composite} \times Rank_{t-1}$ (4th quartile -high)	0.327***	0.377***	0.300***	0.328***
	(2.77)	(3.28)	(3.21)	(3.68)
R^2	0.008	0.008	0.009	0.010
T	120	120	120	120
Winsorized (5th and 95th percentiles)	No	No	Yes	Yes
Indexed stock return	No	Yes	No	Yes

As reported in subsection 6.4, firm size has an impact on the magnitude of the cross-predictability effect. Hence, an important question is whether the results in Table 7 are driven by firm size related effects rather than geographic specialization of investors. To investigate the possibility that firm size might be driving the results, Appendix 11 provides descriptive statistics of the $Geodisp_j$ quartiles. It can be seen that apart from the 1st quartile which has the highest median market capitalization in both samples, the differences in firm size between different quartiles are small. Importantly, the cross-predictability effect in the 1st $Geodisp_j$ quartile is considerably weaker than in the corresponding size quintiles in Table 5 (the 3rd and 4th size quintiles). Therefore, larger firm size alone is unlikely to explain the smaller coefficient on $r^{composite}$ observed in the 1st quartile for both samples. Furthermore, the correlation between average industry market capitalization and the $Geodisp_j$ measure is low (-0.14 and -0.24 in the Eurozone and EU27 samples, respectively) which suggest that firm size is not driving the results in Table 7. This is not surprising since $Geodisp_j$ allocation is performed on the industry level and thus, should be related only coincidentally to firm size.

To summarize, it appears that investor specialization along geographic boundaries and the resulting informational segmentation of markets have an impact on the formation of prices. Results in Table 7 show that the magnitude of the cross-predictability effect is positively related to the geographic dispersion of related industries as evidenced by the coefficient on $r^{composite}$. To the extent that $Geodisp_j$ provides is a good proxy for the geographic dispersion of related industry information, the results suggests that information from geographically scattered sources diffuses more slowly across the markets due to investors being geographically specialized in their information gathering activities and thus, being able to process information only on a geographically limited subset of assets. This finding is consistent with the limited-information models which posit that specialization of investors in their information gathering activities leads to informationally segmented markets and consequently, cross-predictability in asset returns (Menzly and Ozbas, 2010). Overall, the results suggest that geographic boundaries may be a fundamental driver of investor specialization.

It is important to note that the results are not unequivocal as the magnitude of the cross-predictability effect declines in the 4th quartile (highest level of related industry geographic dispersion). Possible explanations for this are related to limitations of the $Geodisp_j$ measure, relation of information processing costs and geographic dispersion of informative signals, marginal effect of geographic dispersion on the magnitude of the cross-predictability effect and country effects. Overall, the results in Table 7 need to be interpreted with caution due to the limitations of the $Geodisp_j$ measure. Based on the results reported in this subsection, Hypothesis 3.3 is partially accepted with the caveat that further evidence is still required to determine the exact impact of investor geographic specialization on the cross-predictability effect.

6.5 EMPIRICAL EVIDENCE ON ECONOMIC SIGNIFICANCE OF RETURN CROSS-PREDICTABILITY

This subsection analyzes the economic significance of cross-predictability effects. First, the profitability of self-financing trading strategies that exploit return cross-predictability documented in subsection 6.2 is analyzed. Then the exposure of the trading strategy returns to well-known risk factors is examined.

Table 8 presents the mean and standard deviation of monthly excess returns on value-weighted industry portfolios formed on the basis of previous-month related industry returns. At the beginning of each month t industries are sorted into five bins based on returns in supplier and customer industries in month t – 1. Industries with previous-month related industry returns in the lowest (highest) quintile are allocated to the 1st (5th) bin. The mean annualized returns are reported for each bin separately in Table 8 along with standard deviations and Sharpe ratios. The last column contains the returns on a self-financing trading strategy that buys the high portfolio (highest previous-month related industry returns) and sells the low portfolio (lowest previous-month related industry returns). The self-financing strategy is constructed over the sample period ranging from January 2000 to December 2009. Panel A reports the results for the Eurozone sample and panel B presents the results for the EU27 sample.

It can be seen from Table 8 that the mean excess returns across the five portfolios are negative in both samples which reflects the challenging economic situation in Europe during the sample period. Looking at the Eurozone sample in panel A, it is evident that the mean excess returns in the high portfolio are higher than in the low portfolio regardless of whether industries are sorted based on $r^{supplier}$, $r^{customer}$ or $r^{composite}$. Moreover, there is a visible positive trend in mean excess returns across the five portfolios when sorting on supplier industry and composite industry returns. Sorting on customer industry returns produces a less clear pattern. These observations are in line with the results in subsection 6.2 which indicate that the customer industry cross-predictability effect is weaker than the supplier industry cross-predictability effect in the Eurozone. Interestingly, panel A shows that sorting industries based on $r^{composite}$ produces the highest spread between the high and low portfolios: the mean excess return over the next month is –2.5% for the high portfolio and –11.2% for the low portfolio. Consequently, a self-financing strategy that capitalizes on this return difference between the high and low portfolios yields the best trading strategy in panel A with annual mean excess returns of 9.7% and a Sharpe ratio of 0.713. Similar self-financing trading strategies for industries sorted on $r^{supplier}$ and $r^{customer}$ yield annual excess returns of 4.7% and 3.2% during the sample period, respectively.

Table 8: Self-Financing Trading Strategy Returns

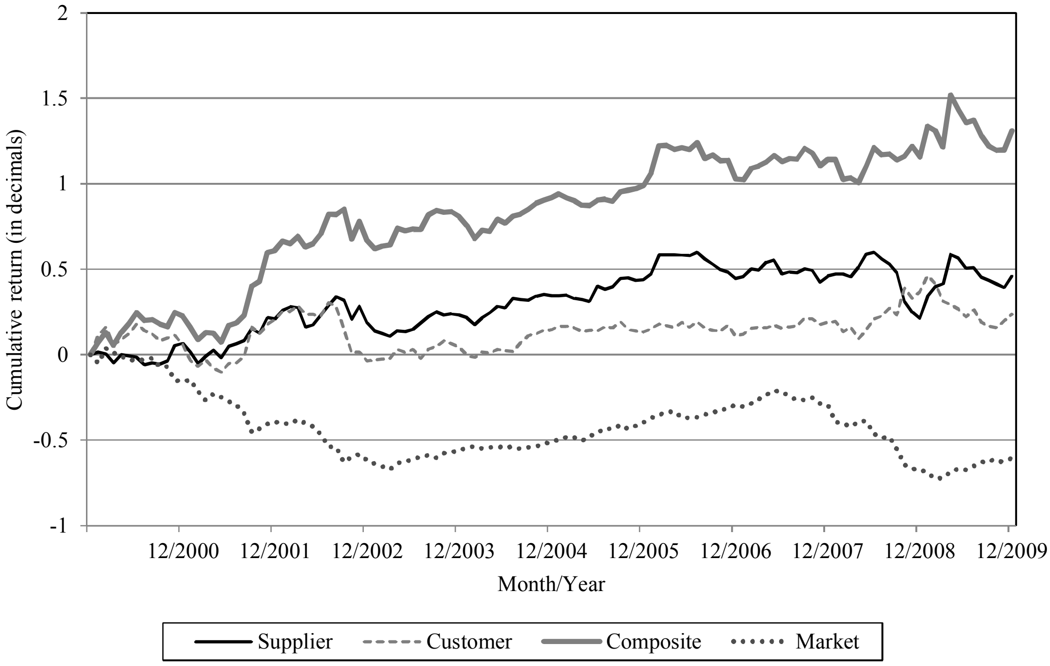
This table presents the mean and standard deviation of monthly excess returns on value-weighted industry portfolios formed on the basis of previous-month related industry returns. Each month t industries are sorted into five bins based on returns on their supplier ($r^{supplier}$), customer ($r^{customer}$) and composite ($r^{composite}$) industry portfolios in month $t-1$. The self-financing trading strategy in the last column consists of buying the high (5) portfolio (highest previous-month related industry returns) and selling the low (1) portfolio (lowest previous-month related industry returns). The self-financing trading strategy is constructed for the sample period ranging from January 2000 to December 2009. All return figures are annualized and in excess of the risk-free rate. Corresponding Sharpe ratios are reported for each trading strategy.

Panel A: Eurozone						
Industries sorted on $r^{supplier, t-1}$	Low (1)	(2)	(3)	(4)	High (5)	H – L
Mean Return	-0.080	-0.078	-0.032	-0.035	-0.036	0.047
Standard Deviation	0.200	0.188	0.188	0.18	0.216	0.132
Sharpe Ratio	-0.398	-0.417	-0.168	-0.192	-0.165	0.360
Industries sorted on $r^{customer, t-1}$	Low (1)	(2)	(3)	(4)	High (5)	H – L
Mean Return	-0.082	-0.092	-0.069	-0.034	-0.052	0.032
Standard Deviation	0.209	0.216	0.195	0.198	0.182	0.146
Sharpe Ratio	-0.394	-0.425	-0.356	-0.171	-0.288	0.220
Industries sorted on $r^{composite, t-1}$	Low (1)	(2)	(3)	(4)	High (5)	H – L
Mean Return	-0.112	-0.032	-0.065	-0.035	-0.025	0.097
Standard Deviation	0.212	0.190	0.192	0.188	0.202	0.136
Sharpe Ratio	-0.529	-0.166	-0.339	-0.184	-0.124	0.713
Panel B: EU27						
Industries sorted on $r^{supplier, t-1}$	Low (1)	(2)	(3)	(4)	High (5)	H – L
Mean Return	-0.072	-0.068	-0.048	-0.013	-0.015	0.062
Standard Deviation	0.184	0.202	0.175	0.185	0.181	0.134
Sharpe Ratio	-0.392	-0.339	-0.272	-0.073	-0.082	0.458
Industries sorted on $r^{customer, t-1}$	Low (1)	(2)	(3)	(4)	High (5)	H – L
Mean Return	-0.073	-0.080	-0.037	-0.044	-0.008	0.070
Standard Deviation	0.213	0.190	0.199	0.192	0.166	0.182
Sharpe Ratio	-0.343	-0.424	-0.186	-0.232	-0.049	0.384
Industries sorted on $r^{composite, t-1}$	Low (1)	(2)	(3)	(4)	High (5)	H – L
Mean Return	-0.038	-0.076	-0.062	-0.030	0.014	0.053
Standard Deviation	0.197	0.185	0.179	0.195	0.167	0.141
Sharpe Ratio	-0.191	-0.413	-0.344	-0.155	0.084	0.380

Panel B reports the same portfolio returns for the EU27 sample. It can be seen that the high portfolios have higher mean excess returns than the low portfolios also in the EU27 sample regardless of the sorting method. Moreover, all three sorting methods produce a clear positive trend in mean excess returns across the five portfolios. Interestingly, sorting on $r^{composite}$ in the EU27 sample yields the highest mean returns in both the high (1.4%) and low (-3.8%) portfolios and produces an inferior self-financing trading strategy when compared to sorting on $r^{supplier}$ or $r^{customer}$. Sorting on $r^{composite}$ seems to perform extremely well in the high portfolio: the $r^{composite}$ sorted high portfolio in panel B is the only portfolio that is able to produce positive mean excess returns during the sample period. However, given that sorting separately on $r^{supplier}$ and $r^{customer}$ in panel B produces low portfolio returns which are comparable to those in panel A, it is difficult to estimate what causes $r^{composite}$ to yield such high returns also in the low portfolio. Sorting industries based on supplier industry returns yields a superior self-financing trading strategy in panel B as evidenced by the highest Sharpe ratio (0.46). Furthermore, the self-financing strategy that sorts industries based on $r^{customer}$ is clearly more profitable in the EU27 than in the Eurozone sample as evidenced by the higher Sharpe ratio. This finding is consistent with results in subsection 6.2 which suggest that customer industry cross-predictability effects are stronger in the EU27 sample.

Figure 1 shows the cumulative excess returns from the three self-financing trading strategies plotted against the cumulative excess market returns. Panel A presents the cumulative returns for the Eurozone sample and panel B contains the cumulative returns for the EU27 sample. Overall, the cumulative excess return on the broad market portfolio is negative over the sample period in both samples. The bursting of the dot-com bubble explains the downturn observable in market returns at the beginning of the sample period. Thereafter, the stock market seems to have experienced a period of recovery which ended in 2007 at the wake of the global financial crisis. Despite the poor market returns, all cross-predictability based trading strategies are profitable for most of the sample period. Interestingly, the only trading strategy that appears to have been profitable during the entire sample period is the strategy that sorts industries based on $r^{composite}$ in the Eurozone sample. Panel A clearly illustrates the superiority of this trading strategy in the Eurozone sample compared to other strategies and the EU27 sample. The trading strategies that sort industries based on $r^{supplier}$ and $r^{customer}$ perform consistently in both samples.

Panel A: Eurozone



Panel B: EU27

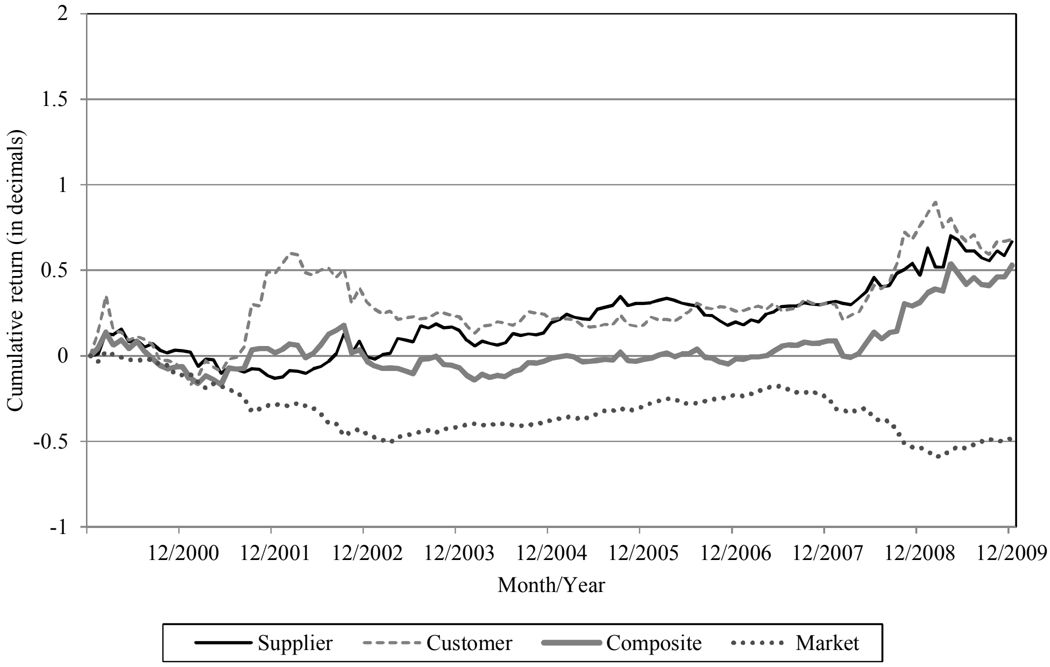


Figure 1: Cumulative Self-Financing Trading Strategy Returns

This figure presents the cumulative excess returns from the self-financing trading strategies plotted against the cumulative excess return on the broad market portfolio. The trading strategies buy (sell) industries with the highest (lowest) previous-month returns in related industries over the sample period from January 2000 to December 2009. Panel A contains the returns for the Eurozone sample and panel B contains the returns for the EU27 sample. All data items are retrieved from Thomson One Banker database. Cumulative returns are in decimals and in excess of the risk-free rate.

Overall, the results in Table 8 show that profits from self-financing trading strategies that capitalize on cross-predictability effects can be economically significant which is in line with earlier research. Menzly and Ozbas (2010) report mean annual excess returns on similar strategies that buy and sell industries based on previous-month related industry returns. More specifically, they analyze US data over a period ranging from 1973 to 2005 and find that sorting industries on $r^{composite}$ produces a superior trading strategy with a Sharpe ratio of 0.66 and mean annual excess return of 8.7%. Furthermore, they find that sorting on $r^{supplier}$ produces a better strategy than sorting on $r^{customer}$ which is consistent with the findings in this paper. In addition, Shahrur, Becker and Rosenfeld (2009) report annual returns of 15% on a strategy that buys industries based on previous-month customer industry returns. Compared to results in Table 8, their trading strategy returns appear surprisingly high. However, the high returns may be explained by different sample composition and a different sample period that ranges from 1995 to 2007 and thus, mostly avoids the global financial crisis. Moreover, Cohen and Frazzini (2008) report annual returns of up to 18.6% on a strategy that buys and sells stocks based on previous-month customer returns. They use more precise US firm-level customer-supplier data to determine firm relatedness. This allows them to define stronger economic links between sample firms which is likely to contribute to the higher returns on their strategy. Hence, their results are not directly comparable to those observed in this paper.

To address a concern that the trading strategy profitability is driven by a small number of industries entering the trading strategy in an excessive way, Appendix 12 presents industry inclusion probabilities for the high and low portfolios. There appears to be some heterogeneity in the frequencies with which different industries enter the trading strategy as evidenced by certain industries entering the high and low portfolios almost half the time. As a comparison, if the industries had identical inclusion probabilities, they would enter the trading strategy approximately 2% of the time (one divided by the number of industries in the sample). However, the observed heterogeneity is attributable to the small sample size (only 120 months) which makes it only likely that the inclusion probabilities deviate from their theoretical value. Furthermore, the inclusion probabilities appear to be correlated meaning that when an industry enters either trading strategy portfolio with a high probability, it also enters in the opposite portfolio with a higher probability. The correlation of inclusion probabilities varies between 0.5 and 0.8 depending on the sorting method. Given the limited sample size, one cannot draw definitive conclusions on whether the trading strategy profits are driven by only a few industries.

The remainder of this subsection examines whether the returns reported in Table 8 are exposed to well-known risk factors. In order to analyze this, Table 9 reports the return factor exposure of monthly returns from the above discussed self-financing trading strategies that buy (sell) industries with high (low) previous-month returns in related industries. More specifically, monthly returns from the self-financing trading strategies presented in Table 8 are regressed on the four-factor model risk factors – excess market return, SMB, HML and MOM factors (Fama and French, 1993; Carhart, 1997).[81] Panel A contains the results for the Eurozone sample and panel B presents the results for the EU27 sample.

It can be seen in panel A that the monthly trading strategy returns appear to have fairly low exposure to traditional risk factors in the Eurozone sample. Only HML factor in the $r^{supplier}$ based trading strategy has a coefficient that is statistically significant at the 10% level. This finding is somewhat inconsistent with Menzly and Ozbas (2010) who find that returns from cross-predictability trading strategies are orthogonal to the HML factor. Furthermore, the MOM factor remains statistically insignificant in the Eurozone sample which is not in line with Menzly and Ozbas who find that returns from trading strategies based on previous-month supplier and composite industry returns are positively and significantly exposed to the MOM factor. They attribute their finding to the MOM factor proxying for sustained information diffusion from related industries. Also Shahrur, Becker and Rosenfeld (2009) document a positive and significant coefficient on the MOM factor associated with trading strategy returns based on $r^{customer}$. The low exposure of the trading strategy returns to the market factor reported in Table 9 is in line with Menzly and Ozbas who note that this is due to the long-short nature of the trading strategies. Importantly, the supplier and composite industry based strategies in panel A are able to produce positive and statistically significant alphas. This is mainly consistent with Menzly and Ozbas who report statistically significant and positive alpha for all three self-financing trading strategies. Moreover, the alphas are consistent with the cross-predictability analysis results in subsection 6.2 which indicate only weak customer industry cross-predictability effects in the Eurozone.

[81] See subsection 5.4.2 for a detailed description of each factor.

Table 9: Return Factor Exposure

This table presents the return factor exposure of monthly excess returns from self-financing trading strategies that buy (sell) industries with highest (lowest) previous-month returns in related industries. Panel A contains the results for the Eurozone sample and panel B contains the results for the EU27 sample. All data items are retrieved from Thomson One Banker database. Monthly trading strategy returns are regressed on the monthly value-weighted excess market return and Fama-French-Carhart HML, SMB and MOM factors. *t*-statistics are reported in parentheses. ***, **, or * indicates that the coefficient estimate is different from zero at the 1%, 5%, or 10% level, respectively.

Panel A: Eurozone			
Trading Strategy Based on	$r^{supplier,t-1}$	$r^{customer,t-1}$	$r^{composite,t-1}$
Alpha	0.007*	0.000	0.009**
	(1.82)	(0.03)	(2.28)
r^{market}	0.033	-0.023	-0.051
	(0.40)	(-0.25)	(-0.58)
HML	0.203*	-0.076	0.067
	(1.88)	(-0.64)	(0.59)
SMB	0.114	0.116	0.122
	(0.88)	(0.81)	(0.89)
MOM	-0.076	0.154	-0.046
	(-0.66)	(1.20)	(-0.37)
R^2	0.064	0.058	0.015
No. of Observations	120	120	120

Panel B: EU27			
Trading Strategy Based on	$r^{supplier,t-1}$	$r^{customer,t-1}$	$r^{composite,t-1}$
Alpha	0.008**	0.008*	0.006
	(2.18)	(1.73)	(1.59)
r^{market}	-0.063	0.219**	0.037
	(-0.79)	(2.22)	(0.46)
HML	-0.064	-0.185**	-0.161**
	(-0.88)	(-2.06)	(-2.17)
SMB	0.009	-0.004	-0.068
	(0.06)	(-0.02)	(-0.44)
MOM	-0.282**	-0.348**	-0.168
	(-2.13)	(-2.12)	(-1.24)
R^2	0.043	0.187	0.079
No. of Observations	120	120	120

Panel B which contains the trading strategies implemented in the EU27 sample provides somewhat different insights than the Eurozone sample. More specifically, the HML factor is statistically significant but negative in columns 2 and 3, which is different from the Eurozone results. However, this result is consistent with Shahrur, Becker and Rosenfeld (2009) who also report a negative and statistically significant exposure of monthly returns from customer industry strategies to the HML factor. Furthermore, the statistically significant exposure of the trading strategy returns to the market factor in column 2 is surprising but does not affect the alpha which remains statistically significant at the 10% level. As mentioned earlier, Menzly and Ozbas (2010) and Shahrur et al. both report a positive and statistically significant coefficient on the MOM factor whereas, in panel B columns 1 and 2, the coefficient on MOM is negative and statistically significant. This difference may be due to return factors being calculated across sample countries in this paper. Overall, the self-financing trading strategy returns in the EU27 sample exhibit more exposure to traditional risk factors than the returns in the Eurozone sample. Regardless, the trading strategies based on previous-month supplier and customer industry returns produce positive and statistically significant alphas in the EU27 sample.

To summarize the findings in this subsection, it appears that self-financing trading strategies based on cross-predictability effects are economically significant and able to generate abnormal returns. This is evidenced by the low exposure of the monthly trading strategy returns to well-known risk factors reported in Table 9. Sorting industries based on $r^{composite}$ yields the highest returns (9.7%) and Sharpe ratio (0.71) in the Eurozone sample while sorting on $r^{supplier}$ produces the highest Sharpe ratio (0.46) and annual mean excess return of 6.2% in the EU27 sample. Results in Table 8 confirm that, in terms of trading strategy profitability, supplier industry cross-predictability effects are more important than customer industry cross-predictability effects. This finding is consistent with the results by Menzly and Ozbas (2010) and the findings in subsection 6.2. Based on the results in this subsection, Hypothesis 4.1 is accepted and consequently, Hypothesis 4.2 is rejected.

7 CONCLUSION

This section provides a conclusion of the research and results in this paper. First, subsection 7.1 summarizes the main findings of the study. Thereafter, subsection 7.2 presents suggestions for future research.

This paper aspired to provide further insight on the role of “disagreement” models in theoretical asset pricing. More specifically, the aim of this study was to extend the empirical evidence on the limited-information model of Menzly and Ozbas (2010) and simultaneously provide further evidence on the gradual information diffusion hypothesis by Hong and Stein (1999). Analyzing a European cross-country sample this paper provided geographically and temporally out-of-sample evidence on the limited-information model predictions concerning return cross-predictability between economically linked assets and the effect of informed investors. Furthermore, this paper provided new empirical evidence on geographic boundaries being a fundamental driver of investor specialization and investor specialization being the source of return cross-predictability. The role of investor geographic specialization studied in this paper remains untested in previous literature concerning limited-information models and information diffusion. Finally, this paper provided evidence on the profitability of return cross-predictability based trading strategies in Europe.

This paper analyzed a sample consisting of publicly listed stocks from Eurozone and EU27 countries over a period ranging from January 2000 to December 2009. First, preliminary evidence on the correlation of firm (industry) fundamentals along the supply chain was provided by studying the relation of firm (industry) returns on assets and related industries ROAs. Then empirical evidence on stock- and industry-level return cross-predictability was provided using monthly cross-sectional regressions of stock (industry) returns on previous-month related industry returns and control variables.[82] Thereafter, evidence on limited-information model prediction regarding the effect of informed investors was provided using augmented monthly cross-sectional regressions of stock returns on lagged related industry returns interacted with firm size and analyst coverage. Furthermore, new evidence on the relation between investor geographic specialization and return cross-predictability was provided using a self-constructed measure that proxies for geographic dispersion of related industry information. Finally, the economic significance of cross-predictability effects was estimated by constructing monthly self-financing trading strategies that buy (sell) industries based on previous-month related industry returns and testing their exposure to the four-factor model risk factors.

7.1 SUMMARY OF THE MAIN RESULTS

Table 10 provides a summary of the key findings in this paper. The table is organized by hypothesis and the results are reported separately for both samples. Overall, the findings in this paper provide partial support for the limited information models and the underlying gradual information diffusion hypothesis.

The preliminary analysis on industry relatedness confirms the validity of the empirical design used in this paper. More specifically, using a fixed effects panel regression, the findings indicate a positive and statistically significant relation between firm-level (industry-level) ROAs and contemporaneous ROAs in supplier and customer industries over and above the market in both samples. These findings verify that the limited-information model assumption of correlated fundamentals is fulfilled as Eurostat input-output tables provide a meaningful description of industry relatedness in both samples. In addition, the results are in line with Menzly and Ozbas (2010) who use a similar empirical design on US data and find that firms (industries) along the

[82] The control variables used were short-term reversals (see e.g., Jegadeesh; 1990; Lehmann, 1990), momentum (see e.g., Jegadeesh and Titman, 1993), industry momentum (Moskowitz and Grinblatt, 1999) and excess return on the country-specific market portfolio.

supply chain have positively correlated fundamentals. Interestingly, the stock-level results in this paper also suggest that individual firm profitability may be more strongly correlated with supplier industry profitability than customer industry profitability.

Analysis of return cross-predictability provides mixed support for the limited-information models which posit that returns on economically linked assets cross-predict each other. Monthly cross-sectional regressions of stock returns on lagged related industry returns and control variables show a consistently positive and statistically significant relation between previous-month supplier industry returns and stock returns. This supplier industry cross-predictability effect is robust to adjusting stock returns for country-level differences in stock market development and the exclusion of outliers. Moreover, the supplier industry cross-predictability effects have similar magnitude in both Eurozone and EU27 samples. These findings are consistent with the limited-information models and earlier empirical evidence by Menzly and Ozbas (2010).

Contrary to lagged supplier industry returns, previous-month returns on customer industries do not appear to significantly cross-predict stock returns until the elimination of outliers. Furthermore, the customer industry cross-predictability effect is considerably weaker in the Eurozone sample than in the EU27 sample. These findings are inconsistent with the limited-information models and earlier empirical evidence by Menzly and Ozbas (2010) and Shahrur, Becker and Rosenfeld (2009) who both document statistically significant customer industry cross-predictability effects. However, it is worth noting that also Menzly and Ozbas find that the cross-predictability effect from customer industries is weaker than the cross-predictability effect from supplier industries. Moreover, the customer industry cross-predictability effect (0.071) in their article and the one reported in this paper for the EU27 sample (0.074) has the same magnitude, albeit the latter is not statistically significant.

The finding of weaker customer industry cross-predictability effects in the Eurozone, compared to the EU27 countries, is among the most interesting findings in this study. One explanation could be that some individual non-Eurozone countries exhibit stronger customer industry cross-predictability effects and thus, contribute to the observed stronger customer industry cross-predictability in the EU27 sample. Another explanation attributes the observed weaker customer industry cross-predictability effect to the higher level of integration among the Eurozone countries compared to EU27 countries. More specifically, it is hypothesized that higher integration causes either (i) more informed investors to exploit cross-predictability effects in their investment activities or (ii) investors finding it easier to process customer industry information in Eurozone due to e.g., less country-specific differences having to be incorporated into analyses. Consistent with this explanation, Sharur, Becker and Rosenfeld (2009) find evidence that the ability of customer industry returns to predict stock returns is weaker in countries whose financial markets are more integrated with the world. They argue that higher financial integration with other markets may result in more investors exploiting the return predictability phenomenon, which in turn causes the cross-predictability effect to be weaker.

This paper also examines the limited-information model prediction concerning negative relation between the number of informed investors and the magnitude of the cross-predictability effect. More specifically, using both firm size and analyst coverage as proxies for the amount of informed investors, the findings in this paper confirm that the magnitude of cross-predictability declines (increases) as the number of informed investors in the market increases (decreases). Importantly, in several specifications the cross-predictability effect either declines monotonically across all quintiles or once one moves past the quintile with the lowest number of informed investors. Moreover, there are large spreads in the magnitude of return cross-predictability between stocks with the lowest and highest number of informed investors: stocks in the lowest firm size (analyst coverage) quintiles tend to exhibit two to three times stronger cross-predictability effects than the average stocks whereas stocks in the highest firm size (analyst coverage) quintiles exhibit no cross-predictability effects. Overall, these findings support the limited-information models and are in line with previous empirical evidence by Menzly and Ozbas (2010) and Sharur, Becker and Rosenfeld (2009).

A central contribution of this study is to provide empirical evidence on a novel aspect of investor specialization in the context of limited-information models, namely the influence of investor specialization along geographic boundaries. Using a self-constructed industry-level measure to proxy for geographic dispersion of related industry information, this paper shows that the magnitude of cross-predictability effects is positively related to the geographic dispersion of informative signals from related industries. The results imply that information from geographically scattered sources diffuses more slowly due to investors being geographically specialized in their information gathering activities and thus, being able to process information only on a geographically limited subset of assets. Overall, these findings suggest that geographic boundaries may be a fundamental driver of investor specialization and hence, an important determinant of the magnitude of return cross-predictability. The results are in line with the limited-information model hypothesis of investor specialization being the source of gradual diffusion of information and consequently, return predictability.

Table 10: Summary of Key Findings

Hypothesis	Description	Key findings: Eurozone sample	Key findings: EU27 sample
H1	Industry fundamentals are positively correlated over and above the market	**Strong support**. Positive and statistically significant association between firm-level (industry-level) ROAs and contemporaneous ROA in supplier and customer industries over and above the market.	**Strong support**. Positive and statistically significant association between firm-level (industry-level) ROAs and contemporaneous ROA in supplier and customer industries over and above the market.
H2.1	Lagged supplier industry returns cross-predict stock and industry returns.	**Strong support**. Consistently positive and statistically significant relation between previous-month supplier industry returns and stock returns across different specifications after controlling for previously documented return predictability effects. Positive and statistically significant relation between previous-month composite industry returns and industry-level returns.	**Strong support**. Consistently positive and statistically significant relation between previous-month supplier industry returns and stock returns across different specifications after controlling for previously documented return predictability effects. Positive and statistically significant relation between previous-month composite industry returns and industry-level returns.
H2.2	Lagged customer industry returns cross-predict stock and industry returns.	**Weak support**. Consistently positive but mainly statistically insignificant relation between previous-month customer industry returns and stock returns across different specifications. Significant at the 10% level after indexation of stock returns and elimination of outliers. Positive and statistically significant relation between previous-month composite industry returns and industry-level returns.	**Moderate support**. Consistently positive relation between previous-month customer industry returns and stock returns across different specifications. Statistically significant after elimination of outliers. Positive and statistically significant relation between previous-month composite industry returns and industry-level returns.
H3.1	Magnitude of the cross-predictability effect decreases as firm size increases.	**Strong support**. Magnitude of cross-predictability effect decreases monotonically as firm size increases. Highest market capitalization firms exhibit no cross-predictabilty at conventional levels across different specifications.	**Strong support**. Magnitude of cross-predictability effect decreases monotonically as firm size increases once past the smallest firms. Highest market capitalization firms exhibit no cross-predictabilty at conventional levels across different specifications.
H3.2	Magnitude of the cross-predictability effect decreases as analyst coverage increases.	**Mixed support**. Statistically insignificant cross-predictability effects across most quintiles and specifications. Highest analyst coverage stocks exhibit weakest cross-predictability but no clear pattern across quintiles. The low statistical significance is likely a result of small sample size.	**Strong support**. Magnitude of cross-predictability effect decreases monotonically as analyst coverage increases once past firms with the lowest level of analyst coverage. Highest analyst coverage firms exhibit no cross-predictabilty at conventional levels across different specifications.
H3.3	Magnitude of the cross-predictability effect increases as the geographic dispersion of related industries increases.	**Moderate support**. Magnitude of cross-predictability effect increases monotonically with geographic dispresion of related industries excluding quartile with the highest level of geographic dispersion. Large spreads between the low and high quartiles.	**Moderate support**. Magnitude of cross-predictability effect increases monotonically with geographic dispresion of related industries excluding quartile with the highest level of geographic dispersion. Large spreads between the low and high quartiles.
H4.1	Self-financing trading strategies that capitalize on the cross-predictability effect are able to generate abnormal returns.	**Strong support**. Returns from self-financing trading strategies that capitalize on cross-predictability effects exhibit low exposure to well-known risk factors. Positive alphas associated with strategies that sort industries based on previous-month supplier and composite industry returns.	**Strong support**. Returns from self-financing trading strategies that capitalize on cross-predictability effects exhibit moderate exposure to well-known risk factors. Positive alphas associated with strategies that sort industries based on previous-month supplier and customer industry returns.
H4.2	The cross-predictability based trading strategy returns are explained by the four-factor model (counter hypothesis for H4.1).	**No support.** Returns from cross-predictability based trading strategies exhibit low exposure to well-known risk factors. Only HML factor receives a statistically significant and positive coefficient in one strategy.	**No support.** Returns from cross-predictability based trading strategies exhibit some exposure to HML, MOM and market factors. However, alphas are positive in two of three trading strategies.

Finally, this paper analyzes the economic significance of return cross-predictability by constructing self-financing trading strategies that capitalize on cross-predictability effects. Overall, trading strategies based on the cross-predictability effects appear to be economically significant and able to generate abnormal returns. Sorting industries based on $r^{composite}$ yields the highest annual mean return (9.7%) and Sharpe ratio (0.71) in the Eurozone sample while sorting on $r^{supplier}$ produces the highest Sharpe ratio (0.46) in the EU27 sample. In comparison, Menzly and Ozbas (2010) analyze the US stock market and report mean annual excess returns of 8.7% on similar strategies that exploit cross-predictability effects. Furthermore, monthly returns from the trading strategies exhibit fairly low exposure to well-known risk factors, namely market, HML, SMB and MOM factors (Fama and French, 1993; Carhart, 1997) and the strategies are able to yield statistically significant and positive alphas. The differences in trading strategy profitability documented in this paper confirm that supplier industry cross-predictability effects are economically more important than customer industry cross-predictability effects. This is also in line with earlier empirical evidence by Menzly and Ozbas and the stronger supplier industry cross-predictability effect reported in this paper.

7.2 SUGGESTIONS FOR FUTURE RESEARCH

A central contribution of this paper is to provide new empirical evidence on the relation between investor geographic specialization and return cross-predictability. Previous academic literature and the novel findings in subsection 6.4 suggest that geographic boundaries may be a fundamental driver of investor specialization and thus, an important determinant of the magnitude of the cross-predictability effect. Therefore, investor specialization along geographic boundaries and return cross-predictability also offers an interesting direction for future research. The channel of information flow that was studied in this paper, namely delayed price responses to shocks that originate in related firms in supplier and customer industries, is likely to provide a suitable methodology for studying the effect of geographic specialization on return predictability also in future. Hence, the focus of future research should be on constructing firm-level measures that accurately capture the geographic dispersion of relevant information that diffuses from related industries. As mentioned earlier, the measure of geographic dispersion used in this paper has its limitations and thus, creating a more accurate measure of geographic dispersion should be the first priority. For example, firm-level data on the geographic dispersion of a firm’s sales might provide a sufficiently accurate description of a firm’s geographic exposure and hence, allow a more detailed estimation of the impact of geographic dispersion on the customer industry cross-predictability effect. An obvious drawback with this line of study is that accurate firm-level data is likely to be difficult and laborious to obtain.

This paper also provides puzzling empirical evidence showing that the magnitude of customer industry cross-predictability effect is significantly weaker in the Eurozone than in the EU27 countries. In order to provide more insight on the validity of limited-information models, future research should focus on determining what factors are driving differences in return cross-predictability between different countries. Basically, this would require modifying the empirical design used in this paper by shifting the cross-predictability analysis to country-level instead of using a cross-country sample. Particularly, the focus should be on country-specific factors that may drive differences between the cross-predictability effects across countries. Factors worth studying are provided, for example, by Chui, Titman and Wei (2010) who study the impact of individualism on the profitability of momentum strategies in different countries. They use several country-specific factors that are mainly related to the information environment of a country's financial market. These factors can be broadly divided into three groups: (i) average firm characteristics in a country (ii) country-specific factors related to information uncertainty of a financial market and (iii) variables related to the financial market development and integrity of countries' financial markets.

A good starting point for future research could be to conduct an analysis similar to Chui, Titman and Wei (2010), but with focus on return cross-predictability instead of momentum. Basically, this would require constructing self-financing trading strategies, similar to the ones in this paper, on country-level and then, regressing the monthly country-specific trading strategy returns on various country-specific variables that are suggested to explain return predictability. The findings by Chui et al. suggest that particularly individualism, common language, transaction costs, median firm size and stock market volatility might be relevant factors in explaining differences between countries. A potential problem associated with this type of study is the availability of data. In order to construct a country-specific trading strategy, each sample country is required to have a sufficiently high number of different firms and industries with available data. Without sufficient data it is not possible to form meaningful industry portfolios and have representative related industry portfolios for each country. Also obtaining reliable input-output data on multiple countries may be challenging.

REFERENCES

Allen, F., Gale, D., 1994. Limited market participation and volatility of asset prices. *American Economic Review* 84, 933–955.

Baca, S.P., Garbe, B.L., Weiss, R.A., 2000. The rise of sector effects in major equity markets. *Financial Analysts Journal*, 35–40.

Badrinath, S.G., Kale, J.R., Noe, T.H., 1995. Of shepherds, sheep, and the cross-autocorrelations in equity returns. *Review of Financial Studies* 8, 401–430.

Baele, L., 2005. Volatility spillover effects in European equity markets. *Journal of Financial and Quantitative Analysis* 40, 373–402.

Ball, R., Kothari, S.P., 1989. Nonstationary expected returns: Implications for tests of market efficiency and serial correlation in returns. *Journal of Financial Economics* 25, 2551–74.

Ball, R., Brown, P., 1968. An empirical evaluation of accounting income numbers. *Journal of Accounting Research* 6, 159–177.

Barberis, N., Thaler, R., 2002. A survey of behavioral finance. *Handbook of the Economics of Finance*, edited by George Constantinides, Milt Harris and Rene Stulz.

Barberis, N., Shleifer A., Vishny, R., 1998. A model of investor sentiment. *Journal of Financial Economics* 49, 307–343.

Beckers, S., Grinold, R., Rudd, A., Stefek. D., 1992. The relative importance of common factors across the European equity markets. *Journal of Banking and Finance* 16, 75–95.

Bekaert, G., Hodrick R.J., Zhang, X., 2009. International stock return comovements. *Journal of Finance* 64, 2591-2626.

Bernard, V.L., Thomas, J.K., 1989. Post-earnings announcement drift: delayed price response or risk premium. *Journal of Accounting Research* 27, 1–35.

Bernard, V.L., Thomas, J.K., 1990. Evidence that stock prices do not fully reflect the implications of current earnings for future earnings. *Journal of Accounting Economics* 13, 305–340.

Bhojraj, S., Swaminathan, B., 2006. Macromomentum: Returns Predictability in International Equity Indices. *Journal of Business* 79, 429–451.

Bhushan, R., 1989. Firm characteristics and analyst following. *Journal of Accounting and Economics* 11, 255–274.

Bolliger, G., 2004. The characteristics of individual analysts' forecasts in Europe. *Journal of Banking & Finance* 28, 2283–2309.

Boudoukh, J., Richardson M.P., Whitelaw, R.F., 1994. A tale of three schools: Insights on autocorrelations of short horizon stock returns. *Review of Financial Studies* 7, 539–573.

Brennan, M.J., Jegadeesh N., Swaminathan, B., 1993. Investment analysis and the adjustment of stock prices to common information. *Review of Financial Studies* 6, 799–824.

Brooks, R., Del Negro, M., 2004. The rise in comovement across national stock markets: market integration or IT bubble?. *Journal of Empirical Finance* 11, 659–680.

Capstaff, J., Paudyal, K., Rees, W., 2001. A Comparative analysis of earnings forecasts in Europe. *Journal of Business Finance and Accounting* 28, 531–562.

Carhart, M., 1997. On persistence in mutual fund performance. *Journal of Finance* 52, 57–82.

Cavaglia, S., Brightman, C., Aked, M., 2000. The increasing importance of industry factors. *Financial Analysts Journal* 41–54.

Chan, K.C., 1988. On the contrarian investment strategy. *Journal of Business* 61, 147–163.

Chan, K., Covrig, V., Ng, L., 2005. What determines the domestic bias and foreign bias? Evidence from mutual fund equity allocations worldwide. *Journal of Finance* 60, 1495–1534.

Chan L.K., Jegadeesh N., Lakonishok, J., 1996. Momentum strategies. *Journal of Finance* 51, 1681–1713.

Chopra, N., Lakonishok, J., Ritter, J.R., 1992. Measuring abnormal performance: Do stocks overreact?. *Journal of Financial Economics* 31, 235–268.

Chordia, T., Shivakumar, L., 2006. Earnings and price momentum. *Journal of Financial Economics* 80, 627–656.

Chordia, T., Swaminathan, B., 2000. Trading volume and cross-autocorrelations in stock returns. *Journal of Finance* 55, 913–935.

Clark, T.E., Shin, K., 1998. The sources of fluctuations within and across countries. SSRN working paper.

Clark, T.E., van Wincoop, E., 2001. Borders and Business Cycles. *Journal of International Economics* 55, 59–85.

Clement, M., Rees, L., Swanson, E.P., 2000. The influence of corporate governance, culture, and income predictability on the characteristics that distinguish superior analysts. Working Paper, The University of Texas at Austin and Texas A&M University.

Cochrane, J.H., 2001. Asset pricing. Princeton University Press, Princeton, NJ.

Cohen, L., Frazzini, A., 2008. Economic links and predictable returns. *Journal of Finance* 63, 1977–2011.

Conrad, J., Kaul, G., 1988. Time variation in expected returns. *Journal of Business* 61, 409–425.

Conrad, J., Kaul, G., 1998. An anatomy of trading strategies. *Review of Financial Studies* 11, 489–519.

Coval, J.D., Moskowitz, T.J., 1999. Home bias at home: Local equity preference in domestic portfolios. *Journal of Finance* 54, 2045–2073.

Coval, J.D., Moskowitz, T.J., 2001. The geography of investment: Informed trading and asset prices. *Journal of Political Economy* 4, 811–841.

Chui, A., Titman, S., Wei, J., 2000. Momentum, legal systems and ownership structure: An analysis of Asian stock markets. Working paper, University of Texas at Austin.

Chui, A., Titman, S., Wei, J., 2010. Individualism and Momentum around the World. *Journal of Finance* 65, 361-392.

Daniel, K., Hirshleifer D., Subrahmanyam, A., 1998. A theory of overconfidence, self-attribution and security market under- and overreactions. *Journal of Finance* 53, 1839–1886.

De Bondt, W., Thaler, R., 1985. Does the stock market overreact?. *Journal of Finance* 40, 793–805.

De Bondt, W., Thaler, R., 1987. Further evidence on investor overreaction and stock market seasonality. *Journal of Finance* 42, 557–581.

De Haan, J., Inklaar, R., Jong-A-Pin, R., 2008. Will business cycles in the Euro area converge? A critical survey of empirical research. *Journal of Economic Surveys*, 22, 234–273.

De Long, J., Shleifer, B.A., Summers L.H., Waldmann, R.J., 1990. Positive feedback investment strategies and destabilizing rational speculation. *Journal of Finance* 45, 379–395.

Dougherty, C., 1992. Introduction to Econometrics. Oxford University Press, New York.

Eleswarapu, V.R., Tiwari, A., 1996. Business cycles and stock market returns: Evidence using industry-based portfolios. *Journal of Financial Research* 19, 121–134

EUROSTAT, 2008. Eurostat Manual of Supply, Use and Input-Output Tables. Eurostat Methodologies and Working Papers. Luxembourg: Eurostat.

EUROSTAT, 2011. Creating consolidated and aggregated EU27 Supply, Use and Input-Output Tables, adding environmental extensions (air emissions), and conducting Leontief-type modelling to approximate carbon and other 'footprints' of EU27 consumption for 2000 to 2006. Luxembourg: Eurostat.

Fama, E.F., 1965. The behavior of stock market prices. *Journal of Business* 38, 34–105.

Fama, E.F., 1998. Market efficiency, long-term returns and behavioral finance. *Journal of Financial Economics* 49, 283–306.

Fama, E.F., French, K.R., 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33, 3–56.

Fama, E.F., French, K.R., 1996. Multifactor explanations of asset pricing anomalies. *Journal of Finance* 51, 55-84.

Fama, E.F., French, K.R., 2011. Size, value and momentum in international stock returns. SSRN working paper.

Fama, E.F., MacBeth J., 1973. Risk, return and equilibrium: Empirical tests. *Journal of Political Economy* 81, 607–636.

Fan, J., Lang, L., 2000. The measurement of relatedness: An application to corporate diversification. *Journal of Business* 73, 629–660.

Ferreira, M.A., Ferreira, M.A., 2006. The importance of industry and country effects in the EMU equity markets. *European Financial Management* 12, 341–373.

Fratzscher, M., 2002. Financial market integration in Europe – on the effects of EMU on stock markets. *International Journal of Finance and Economics* 7, 165–193.

French, K.R., Poterba, M.J., 1991. Investor Diversification and International Equity Markets. *American Economic Review*, 81, 222–226.

French, K.R., Roll, R., 1986. Stock return variances: The arrival of information and reaction of traders. *Journal of Financial Economics* 17, 5–26.

Griffin J.M., Karolyi, A.G., 1998. Another look at the role of industrial of markets for international diversification strategies. *Journal of Financial Economics* 50, 351–373.

Grinblatt, M., Keloharju, M., 2001. How distance, language, and culture influence stockholdings and trading. *Journal of Finance* 56, 1053–1073.

Grubel, H.G., 1968. Internationally diversified portfolios: welfare gains and capital flows. *American Economic Review* 58, 1299–1314.

Hayek, F., 1945. The use of knowledge in society. *American Economic Review* 35, 519– 530.

Hawawini, G., Keim, D.B., 1995. On the predictability of common stock returns: world-wide evidence, in R. Jarrow et al., (Ed.), *Finance, Handbooks in OR and MS*, Vol. 9, Amsterdam.

Heston, S.L., and Rouwenhorst, G.K., 1994. Does industrial structure explain the benefits of international diversification?. *Journal of Financial Economics* 36, 3–27.

Hirshleifer, D., 2001. Investor psychology and asset pricing. *Journal of Finance* 56, 1533–1579.

Hofstede, G., 2001. Culture's Consequences: Comparing Values, Behaviors, Institutions, and Organizations across Nations, Sage Publication, Beverly Hills, CA.

Hong, H., Lim, T., Stein, J., 2000. Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies. *Journal of Finance* 55, 265–295.

Hong, H., Stein, J., 1999. A unified theory of underreaction, momentum trading, and overreaction in asset markets. *Journal of Finance* 54, 2143–2184.

Hong, H., Stein, J., 2007. Disagreement and the stock market. *Journal of Economic Perspectives* 21, 109–128.

Hong, H., Stein, J., Yu, J., 2007. Simple forecasts and paradigm shifts. *Journal of Finance* 62, 1207–1242.

Hong, H., Torous, W., Valkanov, R., 2007. Do industries lead stock markets?. *Journal of Financial Economics* 83, 367–396.

Hou, K., 2007. Industry information diffusion and the lead-lag effect in stock returns. *Review of Financial Studies* 20, 1113–1138.

Huberman, G., 2001. Familiarity breeds investment. *Review of Financial Studies* 14, 659–680.

Huberman, G., Regev, T., 2001. Contagious speculation and a cure for cancer: A non-event that made stock prices soar. *Journal of Finance* 56, 387–96.

Jegadeesh, N., 1990. Evidence of Predictable Behavior of Security Returns. *Journal of Finance* 45, 881–898.

Jegadeesh, N., Titman, S., 1991. Short horizon return reversals and the bid-ask spread. *Journal of Financial Intermediation* 4, 116–132.

Jegadeesh, N., Titman, S., 1993. Returns to buying winners and selling losers: Implications for stock market efficiency. *Journal of Finance* 48, 65–91.

Jegadeesh, N., Titman, S., 1995. Overreaction, delayed reaction and contrarian profits. *Review of Financial Studies* 8, 973–993.

Jegadeesh, N., Titman, S., 2001. Profitability of momentum strategies: an alternative evaluation of explanations. *Journal of Finance* 56, 699–720.

Kale, J.R., Shahrur, H., 2007. Corporate capital structure and the characteristics of suppliers and customers. *Journal of Financial Economics* 2007, 321–365.

Kaltenhaeuser, B., 2003. Country- and sector-specific spillover effects in the euro area, the United States and Japan. European Central Bank working paper no. 286.

Kini, O., Mian, S., Rebello, M., Venkateswaran, A., 2003. On the Determinants of International Analyst Research Coverage. SSRN working paper.

La Porta, R., Lopez-de-Silanes, F., Shleifer, A., Vishny, R.W., 1998. Law and finance. *Journal of Political Economy* 106, 1113–1155.

La Porta, R., Lopez-de-Silanes, F., Shleifer, A., Vishny, R.W., 2000. Agency problems and dividend policies around the world. *Journal of Finance* 55, 1–33.

Lehmann, B., 1990. Fads, martingales and market efficiency. *Quarterly Journal of Economics* 60, 1–28.

Lesmond, D.A., Schill M.J., Zhou, C., 2004. The illusionary nature of momentum profits. *Journal of Financial Economics* 71, 349-380.

Levy, H., Sarnat, M., 1970. International diversification of investment portfolios. *American Economic Review* 25, 668–675.

Lewellen, J., 2002. Momentum and autocorrelation in stock returns. *Review of Financial Studies Special* 15, 533–563.

Ling C., Chan, K., Dasgupta, S., Gao, N., 2011. When the tail wags the dog: Industry leaders, limited attention and spurious cross-industry information diffusion. SSRN working paper.

Lintner, J., 1965. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics* 47, 13–37.

Lo, A., MacKinlay, G., 1988. Stock market prices do not follow random walk: Evidence from a simple specification test. *Review of Financial Studies* 1, 41–66.

Lo, A., MacKinlay, G., 1990a. When are contrarian profits due to stock market overreaction?. *Review of Financial Studies* 3, 175–208.

Lo, A., MacKinlay, G., 1990b. An econometric analysis of nonsynchronous trading. *Journal of Econometrics* 45, 181–211.

Malloy, C.J., 2005. The Geography of equity analysis. *Journal of Finance* 60, 719–755.

Massa, M., Simonov, A., 2006. Hedging, familiarity, and portfolio choice. *Review of Financial Studies* 19, 633–685.

Mech, T.S., 1993. Portfolio return autocorrelation. *Journal of Financial Economics* 34, 307–344.

Menzly, L., Ozbas, O., 2006. Cross-industry momentum. Working paper, University of Southern California, Los Angles, CA.

Menzly, L., Ozbas, O., 2010. Market segmentation and cross-predictability of returns. *Journal of Finance* 65, 1555–1580.

Merton, R.C., 1987. A simple model of capital market equilibrium with incomplete information. *Journal of Finance* 42, 483–510.

Moskowitz, T., Grinblatt, M., 1999. Do industries explain momentum?. *Journal of Finance* 54, 1249–1290.

Petersen, M., 2009. Estimating standard errors in finance panel data sets: Comparing approaches. *Review of Financial Studies* 22, 435–480.

Richards, A.J., 1997. Winner-loser reversals in national stock market indices: can they be explained?. *Journal of Finance* 52, 2129–2144.

Roll, R., 1992. Industrial structure and the comparative behavior of international stock market indices. *Journal of Finance* 47, 3–42.

Rouwenhorst, G.K., 1998. International momentum strategies. *Journal of Finance* 53, 267–284.

Rouwenhorst, G.K., 1999a. Local return factors and turnover in emerging stock markets. *Journal of Finance* 54, 1439–1464.

Rouwenhorst, G.K., 1999b. European Equity Markets and the EMU. *Financial Analysts Journal* 55, 57–64.

Scholes, M., Williams, J., 1977. Estimating betas from nonsynchronous data. *Journal of Financial Economics* 5, 309–327.

Seasholes, M.S., Zhu, N., 2010. Individual investors and local bias. *Journal of Finance* 65, 1987–2010.

Shahrur, H., Becker Y.L., Rosenfeld, D., 2009. Return predictability along the supply chain: the international evidence. SSRN working paper.

Sharpe, W.F., 1964. Capital asset prices: A theory of market equilibrium under conditions of risk, *Journal of Finance* 19, 425–442.

Shleifer, A., Summers, L., 1990. The noise trader approach to finance. *Journal of Economic Perspectives* 4, 19–33.

Simon, H., 1955. A behavioral model of rational choice. *Quarterly Journal of Economics* 69, 99–118.

Solnik, B.H., 1974. The international pricing of risk: an empirical investigation of the world capital market structure. *Journal of Finance* 29, 365–378.

Thompson, S.B., 2011. Simple formulas for standard errors that cluster by both firm and time. *Journal of Financial Economics* 99, 1–10.

Wynne, M.A., Koo, J., 2000. Business cycles under monetary union: a comparison of the EU and US. *Economica* 67, 347–374.

Yalcin, A., 2008. Gradual information diffusion and contrarian strategies. *Quarterly Review of Economics and Finance* 48,579-604.

Zhang, F.X., 2006. Information uncertainty and stock returns. *Journal of Finance* 61, 105–137.

Zin, S.E., 2002. Are behavioral asset-pricing models structural?. *Journal of Monetary Economics* 49, 215–228.

APPENDICES

Appendix1: NACE rev 1.1 to SIC 1987 Concordance Table

This table presents the correspondences between NACE rev 1.1 and SIC 1987 codes that are used in this study to assign sample stocks into industries. The General Industrial Classification of Economic Activities within the European Communities revision 1.1 (NACE rev 1.1) industry classifications contains 59 different industry accounts. Standard Industrial Classification (SIC) codes are based on the 1987 revision.

NACE rev 1.1	Industry name (NACE rev 1.1)	SIC 1987
1	Agriculture, hunting and related service activities	111, 112, 115, 116, 119, 131, 132, 133, 134, 139, 161, 171, 172, 173, 174, 175, 179, 181, 182, 191, 211, 212, 213, 214, 219, 241, 251, 252, 253, 254, 259, 271, 272, 279, 291, 711, 721, 722, 724, 761, 782, 783, 831, 831, 971, 2084, 4971
2	Forestry, logging and related service activities	811, 851, 2411
5	Fishing, operating of fish hatcheries and fish farms; incidental service activities	273, 912, 913, 921
10	Mining of coal and lignite; extraction of peat	1221, 1222, 1231, 1311
11	Extraction of crude petroleum and natural gas; service activities incidental to oil and gas extraction excluding surveying	1321, 1381, 1389
12	Mining of uranium and thorium ores	1094, 1099
13	Mining of metal ores	1011, 1021, 1031, 1041, 1044, 1061
14	Other mining and quarrying	1411, 1422, 1423, 1429, 1442, 1446, 1455, 1459, 1474, 1475, 1479
15	Manufacture of food products and beverages	2011, 2013, 2015, 2021, 2022, 2023, 2024, 2026, 2032, 2033, 2034, 2035, 2037, 2038, 2041, 2043, 2044, 2045, 2046, 2047, 2048, 2051, 2052, 2053, 2061, 2062, 2063, 2064, 2066, 2067, 2068, 2074, 2075, 2076, 2077, 2079, 2082, 2083, 2085, 2086, 2087, 2091, 2092, 2095, 2096, 2097, 2098, 2099
16	Manufacture of tobacco products	2111, 2121, 2131, 2141
17	Manufacture of textiles	2211, 2221, 2231, 2241, 2251, 2252, 2253, 2257, 2258, 2261, 2262, 2269, 2273, 2281, 2282, 2284, 2295, 2296, 2297, 2298, 2299, 2391, 2393, 2394, 2397, 2399
18	Manufacture of wearing apparel; dressing and dyeing of fur	2311, 2321, 2322, 2323, 2325, 2326, 2329, 2331, 2335, 2337, 2339, 2341, 2342, 2353, 2361, 2369, 2371, 2381, 2384, 2385, 2386, 2387, 2389, 3151
19	Tanning and dressing of leather; manufacture of luggage, handbags, saddlery, harness and footwear	3021, 3111, 3131, 3142, 3143, 3144, 3149, 3161, 3171, 3172
20	Manufacture of wood and of products of wood and cork, except furniture; manufacture of articles of straw and plaiting materials	2421, 2426, 2429, 2431, 2435, 2436, 2439, 2441, 2448, 2449, 2452, 2491, 2493
21	Manufacture of pulp, paper and paper products	2611, 2621, 2631, 2652, 2653, 2655, 2656, 2657, 2671, 2671, 2674, 2675, 2676, 2677, 2678, 2679
22	Publishing, printing and reproduction of recorded media	2711, 2721, 2731, 2732, 2741, 2761, 2771, 2782, 2789, 2791, 2796, 3652
23	Manufacture of coke, refined petroleum products and nuclear fuels	2911, 2992
24	Manufacture of chemicals and chemical products	2812, 2813, 2816, 2821, 2822, 2823, 2824, 2833, 2834, 2835, 2836, 2841, 2842, 2843, 2844, 2851, 2861, 2865, 2873, 2874, 2875, 2879, 2891, 2892, 2893, 2895, 3087, 3695
25	Manufacture of rubber and plastic products	3011, 3052, 3061, 3081, 3082, 3083, 3084, 3085, 3086, 3088, 7534
26	Manufacture of other non-metallic mineral products	2951, 2952, 3211, 3221, 3231, 3241, 3251, 3253, 3255, 3259, 3261, 3262, 3263, 3264, 3269, 3271, 3272, 3273, 3274, 3275, 3281, 3295, 3296, 3297, 3299, 3624
27	Manufacture of basic metals	3313, 3316, 3317, 3321, 3322, 3324, 3325, 3331, 3334, 3339, 3341, 3351, 3353, 3354, 3355, 3356, 3363, 3364, 3365, 3366, 3369
28	Manufacture of fabricated metal products, except machinery and equipment	3398, 3411, 3412, 3421, 3425, 3431, 3442, 3446, 3448, 3451, 3452, 3462, 3463, 3466, 3469, 3471, 3479, 3493, 3498, 3533, 5085, 7692
29	Manufacture of machinery and equipment n.e.c.	3483, 3484, 3489, 3491, 3519, 3524, 3534, 3534, 3535, 3536, 3541, 3542, 3543, 3546, 3547, 3549, 3552, 3553, 3554, 3555, 3556, 3561, 3562, 3563, 3564, 3565, 3566, 3567, 3568, 3581, 3582, 3585, 3586, 3593, 3594, 3596, 3631, 3632, 3633, 3635, 3639, 3795
30	Manufacture of office machinery and computers	3571, 3572, 3575, 3577, 3578
31	Manufacture of electrical machinery and apparatus n.e.c.	3612, 3613, 3621, 3625, 3641, 3643, 3644, 3645, 3646, 3647, 3648, 3677, 3678, 3691, 3692, 3694, 7694
32	Manufacture of radio, television and communication equipment and apparatus	3651, 3671, 3672, 3674, 3675, 3676
33	Manufacture of medical, precision and optical instruments, watches and clocks	3812, 3822, 3823, 3824, 3826, 3827, 3844, 3845, 3851, 3873, 8072
34	Manufacture of motor vehicles, trailers and semi-trailers	2451, 3465, 3713, 3715, 3716, 3792
35	Manufacture of other transport equipment	3721, 3731, 3732, 3751

Appendix 1: Continued

This table presents the correspondences between NACE rev 1.1 and SIC 1987 codes that are used in this study to assign sample stocks into industries. The General Industrial Classification of Economic Activities within the European Communities revision 1.1 (NACE rev 1.1) industry classifications contains 59 different industry accounts. Standard Industrial Classification (SIC) codes are based on the 1987 revision.

NACE rev 1.1	Industry name (NACE rev 1.1)	SIC 1987
36	Manufacture of furniture; manufacturing n.e.c.	2434, 2511, 2512, 2514, 2515, 2517, 2519, 2521, 2522, 2531, 3931, 3942, 3944, 3949, 3951, 3953, 3991
37	Recycling	5093
40	Electricity, gas, steam and hot water supply	4911, 4923, 4924, 4925, 4931, 4932, 4961
41	Collection, purification and distribution of water	4939, 4941
45	Construction	1521, 1522, 1531, 1541, 1542, 1611, 1622, 1623, 1629, 1711, 1721, 1731, 1741, 1742, 1743, 1751, 1752, 1761, 1771, 1781, 1791, 1793, 1794, 1795, 1796
50	Sale, maintenance and repair of motor vehicles and motorcycles; retail sale services of automotive fuel	5012, 5013, 5014, 5015, 5075, 5091, 5511, 5521, 5531, 5541, 5561, 5571, 7532, 7533, 7536, 7537, 7538, 7539, 7542, 7549
51	Wholesale trade and commission trade, except of motor vehicles and motorcycles	5000, 5021, 5023, 5031, 5032, 5033, 5039, 5043, 5044, 5045, 5046, 5047, 5048, 5049, 5051, 5052, 5063, 5064, 5065, 5072, 5074, 5078, 5082, 5083, 5084, 5087, 5088, 5092, 5094, 5099, 5100, 5111, 5112, 5113, 5122, 5131, 5136, 5137, 5139, 5141, 5142, 5143, 5144, 5145, 5146, 5147, 5148, 5149, 5153, 5154, 5159, 5162, 5169, 5171, 5172, 5181, 5182, 5191, 5192, 5193, 5194, 5198, 5199
52	Retail trade, except of motor vehicles and motorcycles; repair of personal and household goods	5211, 5231, 5251, 5261, 5271, 5311, 5331, 5399, 5411, 5421, 5431, 5451, 5499, 5551, 5611, 5621, 5632, 5641, 5651, 5661, 5713, 5714, 5719, 5722, 5731, 5734, 5735, 5736, 5912, 5921, 5932, 5941, 5942, 5943, 5944, 5945, 5946, 5947, 5948, 5949, 5961, 5962, 5983, 5984, 5989, 5992, 5993, 5994, 5995, 5999, 7631, 7641
55	Hotels and restaurants	5812, 5813, 7011, 7021, 7032, 7033, 7041
60	Land transport; transport via pipelines	4011, 4111, 4121, 4131, 4141, 4142, 4151, 4213, 4214, 4612, 4613, 4619, 4922
61	Water transport	4412, 4424, 4432, 4449, 4481, 4482, 4489
62	Air transport	4512
63	Supporting and auxiliary transport activities; activities of travel agencies	4173, 4221, 4222, 4225, 4226, 4231, 4491, 4492, 4493, 4724, 4725, 4729, 4731, 4783, 4785, 7521
64	Post and telecommunications	4215, 4311, 4513, 4812, 4813, 4822, 4841
65	Financial intermediation, except insurance and pension funding	6011, 6019, 6021, 6022, 6029, 6035, 6036, 6061, 6062, 6081, 6082, 6091, 6111, 6141, 6153, 6159, 6162, 6712, 6722, 6726, 6732, 6733, 6792, 6794, 6798, 6799
66	Insurance and pension funding, except compulsory social security	6311, 6321, 6324, 6331, 6351, 6361, 6371, 6399
67	Activities auxiliary to financial intermediation	6163, 6211, 6221, 6231, 6282, 6289
70	Real estate activities	6512, 6513, 6514, 6515, 6517, 6519, 6531, 6552
71	Renting of machinery and equipment without operator and of personal and household goods	7352, 7359, 7377, 7513, 7514, 7515, 7519, 7841
72	Computer and related activities	7370, 7371, 7372, 7373, 7374, 7375, 7376, 7378, 7379
73	Research and development	8731
74	Other business activities	762, 781, 1382, 6541, 6719, 7221, 7291, 7311, 7312, 7313, 7319, 7322, 7323, 7331, 7334, 7335, 7336, 7338, 7342, 7349, 7361, 7363, 7381, 7382, 7384, 8111, 8711, 8712, 8713, 8721, 8741, 8742, 8743, 8748
75	Public administration and defence; compulsory social security	9111, 9121, 9131, 9199, 9211, 9221, 9222, 9223, 9224, 9229, 9311, 9411, 9431, 9441, 9451, 9511, 9531, 9532, 9611, 9621, 9631, 9641, 9651, 9661, 9711
80	Education	8211, 8221, 8222, 8243, 8244, 8249
85	Health and social work	741, 742, 8011, 8021, 8031, 8041, 8042, 8043, 8049, 8051, 8052, 8059, 8062, 8063, 8069, 8071, 8082, 8092, 8093, 8331, 8351, 8361
90	Sewage and refuse disposal, sanitation and similar activities	4952, 4953, 4959
91	Activities of membership organisation n.e.c.	8611, 8621, 8631, 8651, 8661
92	Recreational, cultural and sporting activities	4832, 4833, 7383, 7812, 7822, 7829, 7832, 7833, 7911, 7929, 7933, 7941, 7948, 7991, 7992, 7993, 7996, 7997, 7999, 8412, 8422
93	Other service activities	6553, 7211, 7212, 7213, 7215, 7216, 7217, 7218, 7261
95	Private households with employed persons	8811

Appendix 2: Transformed Eurostat Input-Output Table for EU27

This symmetric input-output table presents the flow of goods and services between selected industries in the EU27 area in year 2000. The table is constructed from the consolidated EU27 supply and use tables using fixed product sales structure transformation model. Industry codes are are based on the NACE rev 1.1 industry classifications. Industry sales to other industries are reported on the rows and industry purchases from other industries are reported on the columns. The figures are expressed in millions of euros.

NACE rev 1.1.	1	2	5	10	11	12	13	14	15	Total sales
1	44801	319	42	11	17	0	5	15	156354	201564
2	167	4099	1	12	5	1	4	11	193	4494
5	95	1	243	0	0	0	0	1	2795	3136
10	185	8	1	461	3	0	12	22	242	933
11	28	2	1	7	3746	0	1	10	202	3996
12	0	0	0	14	0	7	0	0	0	21
13	10	1	0	6	3	0	91	3	19	134
14	256	4	12	9	5	0	1	1595	387	2269
15	30152	36	414	28	49	0	7	64	109338	140088
Total purchases	75694	4469	715	546	3827	9	121	1721	269530	

Appendix 3: Return on Assets and Number of Companies by Country and Year

This table presents annual country-level return on assets (ROA) figures over the whole sample period ranging from year 2000 to 2009. All data items are retrieved from Worldscope database. Mean ROA is the annual country-level ROA calculated by weighting individual firm ROAs with firm assets for a given country and year. Number of firms is reported in brackets and it is the number of different firms included in the sample for a given country and year. The first figure in the brackets indicates the number of firms in the Eurozone sample and the second figure is the number of firms in the EU27 sample. These figures are slightly different for the Eurozone and EU27 samples due to industries with less than five firms being exlcuded from the samples.

Country		2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
Austria	Mean ROA	8.17 %	7.74 %	3.60 %	9.68 %	9.24 %	3.67 %	2.53 %	2.61 %	2.96 %	2.30 %
	No.of Firms	[66/66]	[65/65]	[66/66]	[63/63]	[64/64]	[67/69]	[71/74]	[72/75]	[72/75]	[69/72]
Belgium	Mean ROA	10.54 %	10.04 %	10.24 %	11.56 %	12.29 %	2.71 %	2.48 %	2.08 %	0.92 %	1.76 %
	No.of Firms	[87/87]	[84/84]	[87/88]	[88/90]	[95/98]	[98/103]	[95/100]	[103/108]	[99/104]	[98/103]
Cyprus	Mean ROA	-	-	-	-	-	1.29 %	1.68 %	2.43 %	1.50 %	0.97 %
	No.of Firms	-/-	-/-	-/-	-/-	-/-	[75/79]	[78/82]	[78/82]	[73/77]	[75/79]
Estonia	Mean ROA	-	-	11.80 %	15.28 %	13.67 %	14.57 %	10.31 %	11.69 %	5.42 %	6.03 %
	No.of Firms	-/-	-/-	[1/1]	[5/5]	[6/6]	[12/12]	[13/13]	[12/12]	[11/11]	[13/13]
Finland	Mean ROA	14.84 %	12.70 %	12.76 %	12.25 %	12.96 %	10.04 %	9.29 %	10.85 %	9.21 %	6.31 %
	No.of Firms	[89/90]	[87/88]	[90/91]	[92/93]	[94/95]	[97/100]	[97/99]	[98/100]	[99/101]	[100/102]
France	Mean ROA	6.96 %	5.56 %	6.40 %	6.40 %	6.25 %	4.60 %	4.12 %	3.93 %	2.26 %	2.35 %
	No.of Firms	[467/468]	[463/465]	[478/481]	[503/509]	[516/522]	[559/571]	[582/594]	[596/609]	[598/611]	[582/595]
Germany	Mean ROA	4.44 %	2.81 %	3.13 %	4.71 %	5.27 %	3.26 %	2.71 %	2.40 %	1.94 %	2.43 %
	No.of Firms	[512/516]	[482/489]	[470/478]	[488/496]	[517/526]	[541/556]	[558/571]	[568/580]	[563/575]	[533/545]
Greece	Mean ROA	6.15 %	6.44 %	6.57 %	6.85 %	5.98 %	3.39 %	2.79 %	2.71 %	2.04 %	1.88 %
	No.of Firms	[162/164]	[180/182]	[169/171]	[175/177]	[184/186]	[211/214]	[221/224]	[229/234]	[229/234]	[229/234]
Ireland	Mean ROA	11.65 %	9.23 %	10.25 %	4.58 %	11.66 %	2.04 %	1.92 %	2.63 %	2.16 %	0.51 %
	No.of Firms	[37/38]	[38/39]	[40/41]	[40/41]	[44/46]	[50/52]	[51/55]	[52/56]	[52/56]	[50/54]
Italy	Mean ROA	7.26 %	7.89 %	8.91 %	10.79 %	12.07 %	4.99 %	3.76 %	3.49 %	3.03 %	2.65 %
	No.of Firms	[158/158]	[163/163]	[172/172]	[175/176]	[186/187]	[201/212]	[216/227]	[219/230]	[222/233]	[222/232]
Luxembourg	Mean ROA	4.19 %	8.74 %	9.23 %	9.03 %	20.66 %	2.05 %	2.43 %	3.16 %	2.01 %	1.65 %
	No.of Firms	[14/14]	[13/13]	[16/16]	[17/17]	[19/19]	[25/26]	[26/28]	[26/28]	[29/31]	[30/32]
Malta	Mean ROA	-	-	-	-	-	2.64 %	2.47 %	2.69 %	1.63 %	1.85 %
	No.of Firms	-/-	-/-	-/-	-/-	-/-	[11/11]	[11/11]	[12/12]	[13/13]	[13/13]
Netherlands	Mean ROA	12.94 %	12.36 %	11.91 %	12.33 %	13.70 %	4.95 %	3.28 %	2.91 %	2.09 %	2.98 %
	No.of Firms	[125/126]	[115/116]	[118/120]	[116/119]	[114/117]	[112/117]	[118/124]	[122/128]	[122/129]	[120/127]
Portugal	Mean ROA	4.54 %	5.72 %	6.68 %	7.43 %	5.75 %	4.71 %	3.97 %	3.90 %	3.05 %	3.40 %
	No.of Firms	[43/44]	[43/44]	[45/46]	[46/47]	[44/44]	[45/45]	[49/49]	[50/50]	[50/50]	[48/48]
Slovakia	Mean ROA	13.84 %	13.75 %	10.73 %	4.26 %	6.48 %	4.07 %	3.21 %	3.76 %	1.81 %	1.29 %
	No.of Firms	[4/4]	[4/4]	[4/4]	[4/4]	[7/7]	[10/10]	[11/11]	[11/11]	[12/12]	[13/13]
Slovenia	Mean ROA	10.26 %	6.64 %	10.77 %	10.50 %	9.20 %	5.99 %	6.03 %	5.61 %	4.55 %	4.19 %
	No.of Firms	[1/1]	[1/1]	[8/8]	[8/8]	[9/9]	[16/17]	[29/30]	[30/32]	[30/32]	[30/32]
Spain	Mean ROA	11.13 %	6.53 %	6.64 %	6.89 %	7.06 %	5.45 %	3.67 %	3.63 %	2.97 %	3.09 %
	No.of Firms	[100/100]	[104/104]	[108/108]	[105/105]	[114/114]	[125/127]	[131/133]	[131/133]	[129/131]	[127/129]
Bulgaria	Mean ROA	-	-	-	-	-	-	-	-	-	-
	No.of Firms	-	-	-	-	-	-	-	-	-	-
Czech Republic	Mean ROA	10.92 %	10.22 %	9.74 %	10.25 %	15.43 %	9.76 %	9.73 %	10.38 %	10.21 %	9.87 %
	No.of Firms	[15]	[12]	[11]	[11]	[16]	[17]	[17]	[16]	[15]	[15]
Denmark	Mean ROA	2.06 %	2.42 %	2.42 %	2.70 %	2.66 %	2.29 %	2.29 %	1.82 %	1.95 %	1.51 %
	No.of Firms	[94]	[92]	[96]	[96]	[103]	[125]	[143]	[150]	[150]	[150]
Hungary	Mean ROA	15.54 %	14.09 %	16.10 %	14.63 %	11.54 %	10.50 %	9.18 %	8.08 %	6.86 %	5.52 %
	No.of Firms	[15]	[16]	[14]	[16]	[22]	[24]	[31]	[30]	[34]	[34]
Latvia	Mean ROA	-	-	-	-	-	9.08 %	6.26 %	8.82 %	4.55 %	3.65 %
	No.of Firms	-	-	-	-	-	[23]	[23]	[24]	[24]	[25]
Lithuania	Mean ROA	-	-	19.67 %	15.63 %	12.08 %	5.96 %	4.94 %	5.24 %	3.54 %	2.08 %
	No.of Firms	-	-	[2]	[2]	[4]	[25]	[32]	[33]	[33]	[35]
Poland	Mean ROA	11.25 %	6.59 %	7.59 %	8.22 %	8.73 %	6.74 %	6.45 %	6.23 %	4.93 %	4.71 %
	No.of Firms	[46]	[48]	[71]	[129]	[172]	[239]	[269]	[304]	[326]	[337]
Romania	Mean ROA	-	-	-	-	-	-	-	-	-	-
	No.of Firms	-	-	-	-	-	-	-	-	-	-
Sweden	Mean ROA	9.18 %	5.97 %	4.46 %	4.86 %	12.43 %	5.28 %	4.00 %	3.28 %	2.23 %	2.54 %
	No.of Firms	[184]	[186]	[191]	[200]	[218]	[283]	[306]	[335]	[341]	[332]
United Kingdom	Mean ROA	11.01 %	10.92 %	8.33 %	6.85 %	5.52 %	4.20 %	3.48 %	2.81 %	2.15 %	2.55 %
	No.of Firms	[747]	[809]	[867]	[928]	[1039]	[1151]	[1249]	[1298]	[1300]	[1256]
Eurozone	Mean ROA	6.99 %	5.39 %	5.84 %	7.01 %	7.39 %	4.17 %	3.38 %	3.18 %	2.33 %	2.49 %
	No.of Firms	[1865]	[1842]	[1872]	[1925]	[2013]	[2255]	[2357]	[2409]	[2403]	[2352]
EU27	Mean ROA	7.43 %	6.03 %	6.06 %	6.64 %	6.66 %	4.17 %	3.42 %	3.08 %	2.30 %	2.51 %
	No.of Firms	[2977]	[3020]	[3143]	[3332]	[3614]	[4208]	[4495]	[4670]	[4698]	[4607]
	Min ROA	2.06 %	2.42 %	2.42 %	2.70 %	2.66 %	1.29 %	1.68 %	1.82 %	0.92 %	0.51 %
	Max ROA	15.54 %	14.09 %	19.67 %	15.63 %	20.66 %	14.57 %	10.31 %	11.69 %	10.21 %	9.87 %

Appendix 4: Mean Return on Assets and Number of Firms by Industry and Related Industries

This table presents the average return on assets (ROA) by industry and its customer and supplier industries over the whole sample period ranging from year 2000 to 2009. The figures are reported for both Eurozone and EU27 samples separately. The industries are based on NACE rev 1.1 industry classifications and firms are allocated into industries based on their primary SIC codes. The allocation of customer and supplier industries is based on the Eurostat Input-Output tables for years 2000 and 2005. All data items are retrieved from Worldscope database. Mean ROA is the simple average of the annual industry (supplier/customer) ROAs over the the whole sample period. Industry ROA is calculated by weighting individual firm ROAs with firm assets for a given industry and year. Supplier (customer) ROA is calculated by weighting the industry ROAs of supplier (customer) industries by the inter-industry flow of goods and services reported in the Eurostat Input-Output tables. Number of firms is the total number of different firms within an industry in the sample and it is reported in brackets. Industry accounts with less than five firms for each year are exlcuded from the sample.

			Eurozone			EU27		
Industry code	Industry name		Industry	Suppliers	Customers	Industry	Suppliers	Customers
1	Agriculture, hunting and related service activities	Mean ROA	6.06 %	10.15 %	12.51 %	5.77 %	10.69 %	13.04 %
		No. of firms	[33]	-	-	[59]	-	-
2	Forestry, logging and related service activities	Mean ROA	-	-	-	7.25 %	8.63 %	9.40 %
		No. of firms	-	-	-	[11]	-	-
5	Fishing, operating of fish hatcheries and fish farms;	Mean ROA	4.49 %	9.93 %	10.59 %	4.49 %	9.72 %	11.79 %
	service activities incidental to fishing	No. of firms	[11]	-	-	[11]	-	-
10	Mining of coal and lignite; extraction of peat	Mean ROA	19.49 %	7.94 %	8.58 %	19.27 %	9.04 %	9.21 %
		No. of firms	[21]	-	-	[131]	-	-
11	Extraction of crude petroleum and natural gas; service	Mean ROA	10.16 %	7.31 %	16.63 %	11.32 %	7.53 %	14.55 %
	activities incidental to oil and gas extraction excluding	No. of firms	[15]	-	-	[29]	-	-
12	Mining of uranium and thorium ores	Mean ROA	-	-	-	-	-	-
		No. of firms	-	-	-	-	-	-
13	Mining of metal ores	Mean ROA	11.98 %	7.98 %	5.90 %	17.98 %	8.70 %	8.69 %
		No. of firms	[18]	-	-	[124]	-	-
14	Other mining and quarrying	Mean ROA	13.83 %	8.61 %	8.35 %	13.78 %	9.55 %	8.91 %
		No. of firms	[14]	-	-	[21]	-	-
15	Manufacture of food products and beverages	Mean ROA	13.41 %	6.95 %	8.36 %	13.89 %	7.49 %	8.85 %
		No. of firms	[130]	-	-	[206]	-	-
16	Manufacture of tobacco products	Mean ROA	-	-	-	16.45 %	7.73 %	9.44 %
		No. of firms	-	-	-	[5]	-	-
17	Manufacture of textiles	Mean ROA	5.56 %	8.32 %	10.47 %	5.95 %	9.53 %	11.49 %
		No. of firms	[35]	-	-	[55]	-	-
18	Manufacture of wearing apparel; dressing and dyeing	Mean ROA	11.68 %	7.05 %	8.24 %	13.38 %	7.69 %	8.91 %
		No. of firms	[37]	-	-	[51]	-	-
19	Tanning and dressing of leather; manufacture of	Mean ROA	15.71 %	8.60 %	9.27 %	15.60 %	9.36 %	10.04 %
	luggage, handbags, saddlery, harness and footwear	No. of firms	[12]	-	-	[18]	-	-
20	Manufacture of wood and of products of wood and	Mean ROA	9.25 %	7.98 %	8.11 %	9.29 %	8.82 %	8.66 %
	cork	No. of firms	[20]	-	-	[37]	-	-
21	Manufacture of pulp, paper and paper products	Mean ROA	9.21 %	8.13 %	9.30 %	9.56 %	9.10 %	10.16 %
		No. of firms	[38]	-	-	[69]	-	-
22	Publishing, printing and reproduction of recorded	Mean ROA	9.03 %	8.23 %	7.83 %	9.76 %	8.90 %	7.98 %
	media	No. of firms	[58]	-	-	[106]	-	-
23	Manufacture of coke, refined petroleum products and	Mean ROA	21.71 %	8.31 %	7.67 %	17.86 %	10.21 %	9.22 %
	nuclear fuels	No. of firms	[12]	-	-	[21]	-	-
24	Manufacture of chemicals and chemical products	Mean ROA	11.54 %	8.90 %	9.45 %	14.54 %	8.63 %	9.24 %
		No. of firms	[127]	-	-	[243]	-	-
25	Manufacture of rubber and plastic products	Mean ROA	9.33 %	8.99 %	8.75 %	9.46 %	10.28 %	9.63 %
		No. of firms	[23]	-	-	[29]	-	-
26	Manufacture of other non-metallic mineral products	Mean ROA	11.43 %	8.39 %	6.37 %	11.59 %	9.49 %	7.08 %
		No. of firms	[58]	-	-	[77]	-	-
27	Manufacture of basic metals	Mean ROA	3.40 %	8.54 %	8.49 %	4.21 %	9.77 %	9.86 %
		No. of firms	[31]	-	-	[49]	-	-
28	Manufacture of fabricated metal products, except	Mean ROA	8.02 %	6.51 %	7.85 %	9.81 %	7.34 %	8.91 %
	machinery and equipment	No. of firms	[23]	-	-	[50]	-	-
29	Manufacture of machinery and equipment n.e.c.	Mean ROA	8.90 %	7.55 %	8.24 %	11.52 %	8.25 %	8.92 %
		No. of firms	[73]	-	-	[122]	-	-
30	Manufacture of office machinery and computers	Mean ROA	5.30 %	8.11 %	8.52 %	5.54 %	8.61 %	9.06 %
		No. of firms	[38]	-	-	[59]	-	-
31	Manufacture of electrical machinery and apparatus	Mean ROA	10.80 %	7.48 %	7.66 %	10.85 %	7.72 %	8.55 %
	n.e.c.	No. of firms	[31]	-	-	[53]	-	-
32	Manufacture of radio, television and communication	Mean ROA	8.44 %	8.07 %	9.15 %	8.63 %	7.68 %	9.32 %
	equipment and apparatus	No. of firms	[53]	-	-	[82]	-	-
33	Manufacture of medical, precision and optical	Mean ROA	13.94 %	7.92 %	8.56 %	13.98 %	8.26 %	8.70 %
	instruments, watches and clocks	No. of firms	[44]	-	-	[103]	-	-
34	Manufacture of motor vehicles, trailers and semi-	Mean ROA	-	-	-	-	-	-
	trailers	No. of firms	-	-	-	-	-	-

Appendix 4: Continued

			Eurozone			EU27		
Industry code	Industry name		Industry	Suppliers	Customers	Industry	Suppliers	Customers
35	Manufacture of other transport equipment	Mean ROA	4.78 %	8.06 %	7.40 %	5.12 %	8.30 %	7.84 %
		No. of firms	[18]	-	-	[30]	-	-
36	Manufacture of furniture; manufacturing n.e.c.	Mean ROA	11.64 %	8.03 %	7.99 %	12.33 %	8.76 %	8.56 %
		No. of firms	[30]	-	-	[59]	-	-
37	Recycling	Mean ROA	-	-	-	-	-	-
		No. of firms	-	-	-	-	-	-
40	Electricity, gas, steam and hot water supply	Mean ROA	8.07 %	9.84 %	8.67 %	8.45 %	10.28 %	9.34 %
		No. of firms	[68]	-	-	[117]	-	-
41	Collection, purification and distribution of water	Mean ROA	6.36 %	7.55 %	8.35 %	7.43 %	8.19 %	8.60 %
		No. of firms	[23]	-	-	[35]	-	-
45	Construction	Mean ROA	5.12 %	8.65 %	6.36 %	5.77 %	9.25 %	6.28 %
		No. of firms	[113]	-	-	[214]	-	-
50	Sale, maintenance and repair of motor vehicles and motorcycles; retail sale services of automotive fuel	Mean ROA	6.79 %	7.31 %	7.62 %	7.04 %	7.69 %	9.10 %
		No. of firms	[17]	-	-	[45]	-	-
51	Wholesale trade and commission trade, except of motor vehicles and motorcycles	Mean ROA	8.85 %	6.97 %	8.95 %	9.41 %	7.89 %	9.72 %
		No. of firms	[160]	-	-	[243]	-	-
52	Retail trade, except of motor vehicles and motorcycles; repair of personal and household goods	Mean ROA	6.81 %	6.93 %	8.96 %	8.26 %	7.36 %	9.59 %
		No. of firms	[107]	-	-	[206]	-	-
55	Hotels and restaurants	Mean ROA	8.51 %	9.40 %	7.82 %	9.26 %	9.90 %	8.18 %
		No. of firms	[58]	-	-	[109]	-	-
60	Land transport; transport via pipelines	Mean ROA	4.00 %	8.85 %	8.96 %	9.83 %	8.55 %	9.32 %
		No. of firms	[15]	-	-	[29]	-	-
61	Water transport	Mean ROA	6.49 %	8.55 %	9.97 %	12.02 %	8.14 %	9.54 %
		No. of firms	[38]	-	-	[62]	-	-
62	Air transport	Mean ROA	6.89 %	8.93 %	7.77 %	6.47 %	8.66 %	7.77 %
		No. of firms	[16]	-	-	[23]	-	-
63	Supporting and auxiliary transport activities; activities of travel agencies	Mean ROA	6.32 %	6.99 %	7.85 %	6.35 %	8.49 %	9.46 %
		No. of firms	[28]	-	-	[46]	-	-
64	Post and telecommunications	Mean ROA	11.90 %	7.27 %	7.33 %	11.38 %	7.56 %	7.35 %
		No. of firms	[59]	-	-	[97]	-	-
65	Financial intermediation, except insurance and pension funding	Mean ROA	0.67 %	6.74 %	6.82 %	0.79 %	7.27 %	7.44 %
		No. of firms	[257]	-	-	[456]	-	-
66	Insurance and pension funding, except compulsory social security	Mean ROA	-	-	-	1.26 %	5.68 %	7.86 %
		No. of firms	-	-	-	[83]	-	-
67	Activities auxiliary to financial intermediation	Mean ROA	2.77 %	5.53 %	3.43 %	3.13 %	6.65 %	2.78 %
		No. of firms	[90]	-	-	[202]	-	-
70	Real estate activities	Mean ROA	2.73 %	4.91 %	7.78 %	2.35 %	5.56 %	8.05 %
		No. of firms	[222]	-	-	[377]	-	-
71	Renting of machinery and equipment without operator and of personal and household goods	Mean ROA	12.34 %	6.94 %	7.88 %	14.05 %	7.48 %	8.54 %
		No. of firms	[14]	-	-	[22]	-	-
72	Computer and related activities	Mean ROA	11.97 %	7.59 %	7.11 %	11.74 %	7.42 %	7.06 %
		No. of firms	[348]	-	-	[611]	-	-
73	Research and development	Mean ROA	-	-	-	-10.60 %	8.58 %	9.95 %
		No. of firms	-	-	-	[77]	-	-
74	Other business activities	Mean ROA	6.03 %	8.20 %	8.24 %	7.42 %	8.29 %	8.42 %
		No. of firms	[161]	-	-	[375]	-	-
75	Public administration and defence; compulsory social security	Mean ROA	-	-	-	-	-	-
		No. of firms	-	-	-	-	-	-
80	Education	Mean ROA	-	-	-	-	-	-
		No. of firms	-	-	-	-	-	-
85	Health and social work	Mean ROA	9.45 %	8.49 %	8.01 %	9.34 %	9.15 %	8.00 %
		No. of firms	[30]	-	-	[52]	-	-
90	Sewage and refuse disposal, sanitation and similar activities	Mean ROA	8.50 %	7.67 %	7.91 %	8.57 %	8.14 %	8.41 %
		No. of firms	[19]	-	-	[37]	-	-
91	Activities of membership organisation n.e.c.	Mean ROA	-	-	-	-	-	-
		No. of firms	-	-	-	-	-	-
92	Recreational, cultural and sporting activities	Mean ROA	13.08 %	7.60 %	7.50 %	12.86 %	8.03 %	8.08 %
		No. of firms	[100]	-	-	[208]	-	-
93	Other service activities	Mean ROA	-	-	-	-	-	-
		No. of firms	-	-	-	-	-	-
95	Private households with employed persons	Mean ROA	-	-	-	-	-	-
		No. of firms	-	-	-	-	-	-
		Min ROA	0.67 %	4.91 %	3.43 %	-10.60 %	5.56 %	2.78 %
		Max ROA	21.71 %	10.15 %	16.63 %	19.27 %	10.69 %	14.55 %

Appendix 5: Stock Market Returns and Number of Companies by Country and Year

This table presents average annual country-level stock returns for all sample countries over the sample period ranging from 2000 to 2009. Mean monthly market return is the simple average of the value-weighted monthly stock market returns for a given country and year. Market return figures are annualized. Number of firms is the number of different firms included in the sample for a given country and it is reported in brackets. The first number in the brackets indicates the number of firms in the Eurozone sample and the second number is the number of firms in the EU27 sample. These figures are slightly different for the Eurozone and EU27 samples due to industries with less than five firms being exlcuded from the samples.

Country		2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
Austria	Mean Market Return	-15.98 %	-0.99 %	-1.29 %	28.77 %	50.28 %	41.07 %	24.50 %	-0.49 %	-63.61 %	33.66 %
	No.of Firms	[56/56]	[56/56]	[57/57]	[57/57]	[59/59]	[57/57]	[59/59]	[62/62]	[69/69]	[69/69]
Belgium	Mean Market Return	-6.90 %	-11.09 %	-28.55 %	12.95 %	37.46 %	24.13 %	20.68 %	-7.59 %	-61.20 %	44.24 %
	No.of Firms	[84/85]	[87/87]	[85/86]	[85/85]	[84/85]	[83/84]	[83/85]	[87/91]	[90/95]	[94/99]
Cyprus	Mean Market Return	-72.06 %	-49.04 %	-36.59 %	-19.26 %	-7.55 %	53.97 %	86.96 %	41.15 %	-76.37 %	22.81 %
	No.of Firms	[33/33]	[44/44]	[66/66]	[73/74]	[74/75]	[74/75]	[76/77]	[76/77]	[72/73]	[76/77]
Estonia	Mean Market Return	74.05 %	-1.57 %	42.55 %	66.32 %	64.28 %	63.04 %	43.48 %	-30.39 %	-76.23 %	19.22 %
	No.of Firms	[6/6]	[6/6]	[6/6]	[6/6]	[6/6]	[6/6]	[7/7]	[9/9]	[11/11]	[12/12]
Finland	Mean Market Return	-4.66 %	-11.86 %	-5.04 %	62.65 %	25.51 %	27.94 %	20.30 %	0.99 %	-57.48 %	36.60 %
	No.of Firms	[42/42]	[44/44]	[45/45]	[46/46]	[46/46]	[47/47]	[49/49]	[53/53]	[59/59]	[62/62]
France	Mean Market Return	2.39 %	-20.41 %	-28.70 %	18.19 %	11.16 %	30.15 %	20.25 %	3.00 %	-45.54 %	24.10 %
	No.of Firms	[342/345]	[392/395]	[429/432]	[441/444]	[445/448]	[456/459]	[469/473]	[501/504]	[546/552]	[549/557]
Germany	Mean Market Return	-18.86 %	-26.47 %	-46.49 %	28.57 %	5.06 %	24.07 %	18.61 %	13.40 %	-48.57 %	22.96 %
	No.of Firms	[280/281]	[389/390]	[424/426]	[434/437]	[485/487]	[526/529]	[551/555]	[618/624]	[655/662]	[655/660]
Greece	Mean Market Return	-54.32 %	-27.00 %	-35.78 %	29.03 %	15.72 %	31.07 %	19.49 %	12.69 %	-66.70 %	10.55 %
	No.of Firms	[158/160]	[191/194]	[211/214]	[228/231]	[233/236]	[232/235]	[231/234]	[236/239]	[240/243]	[241/244]
Ireland	Mean Market Return	8.50 %	2.11 %	-34.18 %	25.10 %	26.33 %	19.88 %	17.83 %	-17.25 %	-68.01 %	27.32 %
	No.of Firms	[32/33]	[32/34]	[35/38]	[35/38]	[36/39]	[36/39]	[37/40]	[40/43]	[44/48]	[48/51]
Italy	Mean Market Return	-2.66 %	-27.60 %	-23.02 %	16.21 %	21.90 %	16.32 %	19.51 %	-3.56 %	-48.50 %	17.70 %
	No.of Firms	[140/141]	[159/160]	[183/184]	[187/188]	[182/183]	[183/184]	[184/185]	[192/194]	[204/208]	[214/219]
Luxembourg	Mean Market Return	-44.47 %	-68.66 %	-53.97 %	49.35 %	16.11 %	37.51 %	28.64 %	33.52 %	-62.04 %	66.47 %
	No.of Firms	[17/18]	[16/17]	[19/20]	[20/21]	[19/20]	[20/21]	[22/23]	[23/24]	[26/27]	[26/27]
Malta	Mean Market Return	-11.76 %	-34.99 %	-18.45 %	15.26 %	50.97 %	73.31 %	3.38 %	3.59 %	-35.47 %	13.66 %
	No.of Firms	-/0]	[6/-	[8/8]	[9/9]	[9/9]	[9/9]	[9/9]	[9/9]	[10/10]	[11/11]
Netherlands	Mean Market Return	-16.22 %	-21.22 %	-33.98 %	11.75 %	3.61 %	26.02 %	14.49 %	11.80 %	-40.06 %	29.62 %
	No.of Firms	[125/126]	[127/128]	[123/124]	[119/121]	[114/116]	[112/114]	[110/112]	[112/115]	[111/116]	[109/115]
Portugal	Mean Market Return	-11.74 %	-24.71 %	-23.19 %	18.98 %	18.59 %	16.78 %	33.92 %	15.67 %	-54.30 %	35.36 %
	No.of Firms	[48/48]	[43/43]	[42/42]	[39/39]	[41/41]	[42/42]	[42/42]	[42/42]	[42/42]	[46/46]
Slovakia	Mean Market Return	-28.27 %	39.61 %	23.33 %	39.56 %	127.96 %	26.23 %	21.13 %	26.25 %	-13.38 %	-31.97 %
	No.of Firms	[2/2]	[3/3]	[4/4]	[4/4]	[4/4]	[4/4]	[6/6]	[11/11]	[9/9]	[9/9]
Slovenia	Mean Market Return	-6.80 %	13.99 %	47.75 %	24.29 %	34.61 %	3.91 %	65.32 %	50.12 %	-68.82 %	8.04 %
	No.of Firms	[5/5]	[7/7]	[7/7]	[7/7]	[7/7]	[7/7]	[7/7]	[20/20]	[23/23]	[23/23]
Spain	Mean Market Return	-9.54 %	-2.19 %	-19.08 %	21.66 %	24.44 %	27.72 %	37.82 %	-3.50 %	-40.85 %	13.71 %
	No.of Firms	[83/83]	[85/85]	[89/89]	[90/90]	[85/85]	[84/84]	[84/84]	[86/87]	[86/87]	[85/86]
Bulgaria	Mean Market Return	-73.72 %	-17.83 %	80.87 %	318.96 %	26.11 %	53.44 %	44.48 %	45.37 %	-75.73 %	13.29 %
	No.of Firms	[2]	[4]	[9]	[10]	[11]	[13]	[13]	[107]	[124]	[125]
Czech Republic	Mean Market Return	-4.37 %	-22.08 %	6.99 %	39.49 %	61.86 %	61.81 %	5.92 %	48.46 %	-37.40 %	17.09 %
	No.of Firms	[1454/18]	[1687/18]	[1833/15]	[1880/11]	[1928/10]	[1977/17]	[2025/17]	[2177/15]	[2297/17]	[2328/15]
Denmark	Mean Market Return	8.92 %	-18.88 %	-31.16 %	22.93 %	23.63 %	40.70 %	27.08 %	10.88 %	-49.35 %	22.01 %
	No.of Firms	[126]	[121]	[120]	[118]	[120]	[119]	[119]	[122]	[134]	[146]
Hungary	Mean Market Return	-21.18 %	-14.43 %	4.40 %	18.28 %	57.22 %	38.24 %	8.72 %	12.72 %	-56.58 %	58.05 %
	No.of Firms	[25]	[24]	[23]	[21]	[22]	[23]	[26]	[28]	[30]	[31]
Latvia	Mean Market Return	56.39 %	44.18 %	-16.81 %	48.41 %	47.05 %	58.13 %	1.30 %	-18.51 %	-55.70 %	-11.41 %
	No.of Firms	[8]	[8]	[8]	[8]	[9]	[9]	[9]	[9]	[9]	[9]
Lithuania	Mean Market Return	-27.98 %	-14.53 %	-21.76 %	93.48 %	33.89 %	32.49 %	8.95 %	-5.53 %	-70.38 %	47.07 %
	No.of Firms	[6]	[9]	[10]	[15]	[19]	[23]	[25]	[25]	[27]	[28]
Poland	Mean Market Return	-9.38 %	-38.20 %	-1.29 %	27.57 %	22.24 %	29.55 %	58.95 %	-4.83 %	-50.78 %	20.50 %
	No.of Firms	[42]	[53]	[64]	[69]	[70]	[82]	[112]	[135]	[181]	[230]
Romania	Mean Market Return	-	-	-	-	-	-	-	-	-	-
	No.of Firms	-	-	-	-	-	-	-	-	-	-
Sweden	Mean Market Return	-17.37 %	-22.76 %	-43.54 %	29.66 %	22.47 %	31.24 %	22.79 %	-6.59 %	-43.00 %	45.89 %
	No.of Firms	[167]	[188]	[192]	[198]	[199]	[199]	[207]	[232]	[273]	[301]
United Kingdom	Mean Market Return	-9.49 %	-12.08 %	-23.30 %	20.01 %	11.99 %	23.55 %	12.25 %	3.53 %	-44.09 %	25.98 %
	No.of Firms	[551]	[629]	[699]	[714]	[729]	[795]	[946]	[1092]	[1186]	[1200]
Eurozone	Mean Market Return	-12.31 %	-24.47 %	-34.98 %	18.29 %	13.82 %	24.93 %	20.50 %	4.08 %	-51.10 %	20.94 %
	No.of Firms	[1454]	[1687]	[1833]	[1880]	[1928]	[1977]	[2025]	[2177]	[2297]	[2328]
EU27	Mean Market Return	-7.68 %	-15.36 %	-22.17 %	11.16 %	8.54 %	15.21 %	12.76 %	2.69 %	-36.34 %	13.29 %
	No.of Firms	[2408]	[2752]	[2988]	[3059]	[3133]	[3274]	[3519]	[3967]	[4313]	[4452]
	Min Market Return	-73.72 %	-68.66 %	-53.97 %	-19.26 %	-7.55 %	3.91 %	1.30 %	-30.39 %	-76.37 %	-31.97 %
	Max Market Return	74.05 %	44.18 %	80.87 %	318.96 %	127.96 %	73.31 %	86.96 %	50.12 %	-13.38 %	66.47 %

Appendix 6: Mean Returns and Number of Firms by Industry and Related Industries

This table presents the average monthly stock returns for all sample industries and their customer and supplier industries over the whole sample period ranging from year 2000 to 2009. The figures are reported for both Eurozone and EU27 samples separately. The industries are based on NACE rev 1.1 industry classifications and firms are allocated into industries based on their primary SIC codes. The allocation of customer and supplier industries is based on the Eurostat Input-Output tables for years 2000 and 2005. All data items are retrieved from Thomson One Banker database. Mean return is the simple average of the monthly value-weighted industry (supplier/customer) stock returns calculated over the the whole sample period. Industry returns are calculated by weighting individual firm returns with firm market values. Supplier (customer) return is calculated by weighting the industry returns of supplier (customer) industries by the inter-industry flow of goods and services reported in the Eurostata Input-Output tables. Return figures are annualized. Number of firms is the total number of different firms within an industry in the sample and it is reported in brackets. Industry accounts with less than five firms for each year are exlcuded from the sample.

			Eurozone			EU27		
Industry code	Industry name		Industry	Suppliers	Customers	Industry	Suppliers	Customers
1	Agriculture, hunting and related service	Mean Return	-1.13 %	-2.61 %	-0.07 %	-1.94 %	-2.01 %	2.01 %
	activities	No. of Firms	[31]	-	-	[49]	-	-
2	Forestry, logging and related service activities	Mean Return	-	-	-	-	-	-
		No. of Firms	-	-	-	-	-	-
5	Fishing, operating of fish hatcheries and fish	Mean Return	-26.43 %	-4.21 %	-3.25 %	-26.43 %	-3.61 %	-0.84 %
	farms; service activities incidental to fishing	No. of Firms	[9]	-	-	[9]	-	-
10	Mining of coal and lignite; extraction of peat	Mean Return	4.66 %	-2.69 %	0.14 %	6.20 %	-2.36 %	0.78 %
		No. of Firms	[19]	-	-	[123]	-	-
11	Extraction of crude petroleum and natural gas;	Mean Return	-	-	-	3.23 %	-4.24 %	0.94 %
	service activities incidental to oil and gas	No. of Firms	-	-	-	[25]	-	-
12	Mining of uranium and thorium ores	Mean Return	-	-	-	-	-	-
		No. of Firms	-	-	-	-	-	-
13	Mining of metal ores	Mean Return	-6.80 %	-3.04 %	-1.30 %	10.90 %	-2.26 %	-0.64 %
		No. of Firms	[16]	-	-	[117]	-	-
14	Other mining and quarrying	Mean Return	-	-	-	7.29 %	-3.14 %	-2.42 %
		No. of Firms	-	-	-	[18]	-	-
15	Manufacture of food products and beverages	Mean Return	1.50 %	-3.41 %	-5.47 %	3.61 %	-3.77 %	-5.29 %
		No. of Firms	[129]	-	-	[204]	-	-
16	Manufacture of tobacco products	Mean Return	-	-	-	19.36 %	-4.93 %	-4.57 %
		No. of Firms	-	-	-	[13]	-	-
17	Manufacture of textiles	Mean Return	-26.25 %	-3.97 %	-4.38 %	-20.31 %	-3.85 %	-3.97 %
		No. of Firms	[42]	-	-	[73]	-	-
18	Manufacture of wearing apparel; dressing and	Mean Return	-2.83 %	-14.75 %	-8.40 %	-2.09 %	-11.71 %	-7.39 %
	dyeing of fur	No. of Firms	[33]	-	-	[49]	-	-
19	Tanning and dressing of leather; manufacture	Mean Return	-	-	-	-	-	-
	of luggage, handbags, saddlery, harness and	No. of Firms	-	-	-	-	-	-
20	Manufacture of wood and of products of wood	Mean Return	-10.56 %	-3.86 %	-5.13 %	-10.17 %	-3.77 %	-4.16 %
	and cork	No. of Firms	[17]	-	-	[33]	-	-
21	Manufacture of pulp, paper and paper	Mean Return	-1.37 %	-4.69 %	-7.04 %	-1.29 %	-4.22 %	-5.75 %
	products	No. of Firms	[37]	-	-	[69]	-	-
22	Publishing, printing and reproduction of	Mean Return	-15.89 %	-5.49 %	-6.33 %	-13.19 %	-5.55 %	-6.44 %
	recorded media	No. of Firms	[62]	-	-	[103]	-	-
23	Manufacture of coke, refined petroleum	Mean Return	3.64 %	-4.07 %	-3.80 %	1.06 %	-1.15 %	-3.44 %
	products and nuclear fuels	No. of Firms	[12]	-	-	[21]	-	-
24	Manufacture of chemicals and chemical	Mean Return	-1.36 %	-4.30 %	-4.02 %	-1.27 %	-4.59 %	-4.40 %
	products	No. of Firms	[123]	-	-	[237]	-	-
25	Manufacture of rubber and plastic products	Mean Return	-1.56 %	-3.91 %	-3.37 %	-1.82 %	-3.78 %	-3.07 %
		No. of Firms	[20]	-	-	[28]	-	-
26	Manufacture of other non-metallic mineral	Mean Return	-3.21 %	-3.84 %	-2.23 %	-3.88 %	-2.67 %	-1.77 %
	products	No. of Firms	[62]	-	-	[79]	-	-
27	Manufacture of basic metals	Mean Return	-0.63 %	-2.57 %	1.46 %	-1.65 %	-2.62 %	-0.38 %
		No. of Firms	[31]	-	-	[50]	-	-
28	Manufacture of fabricated metal products,	Mean Return	9.10 %	-3.74 %	-2.79 %	3.03 %	-3.80 %	-2.10 %
	except machinery and equipment	No. of Firms	[23]	-	-	[50]	-	-
29	Manufacture of machinery and equipment	Mean Return	-0.50 %	-1.95 %	-3.06 %	2.03 %	-3.50 %	-3.04 %
	n.e.c.	No. of Firms	[78]	-	-	[139]	-	-
30	Manufacture of office machinery and	Mean Return	-27.17 %	-6.80 %	-5.99 %	-27.06 %	-7.25 %	-6.42 %
	computers	No. of Firms	[34]	-	-	[54]	-	-
31	Manufacture of electrical machinery and	Mean Return	-0.03 %	-3.64 %	-3.74 %	-0.97 %	-4.88 %	-4.30 %
	apparatus n.e.c.	No. of Firms	[29]	-	-	[51]	-	-
32	Manufacture of radio, television and	Mean Return	-14.43 %	-4.66 %	-6.67 %	-16.48 %	-5.46 %	-8.52 %
	communication equipment and apparatus	No. of Firms	[53]	-	-	[78]	-	-
33	Manufacture of medical, precision and optical	Mean Return	-5.00 %	-4.81 %	-5.59 %	-8.95 %	-5.78 %	-6.34 %
	instruments, watches and clocks	No. of Firms	[43]	-	-	[91]	-	-
34	Manufacture of motor vehicles, trailers and	Mean Return	-	-	-	-	-	-
	semi-trailers	No. of Firms	-	-	-	-	-	-

Appendix 6: Continued

Industry code	Industry name		Eurozone Industry	Eurozone Suppliers	Eurozone Customers	EU27 Industry	EU27 Suppliers	EU27 Customers
35	Manufacture of other transport equipment	Mean Return	-6.87 %	-3.59 %	-6.54 %	-5.80 %	-4.70 %	-6.32 %
		No. of Firms	[18]	-	-	[30]	-	-
36	Manufacture of furniture; manufacturing n.e.c.	Mean Return	-12.36 %	-5.83 %	-5.19 %	-10.46 %	-5.76 %	-4.92 %
		No. of Firms	[28]	-	-	[56]	-	-
37	Recycling	Mean Return	-	-	-	-	-	-
		No. of Firms	-	-	-	-	-	-
40	Electricity, gas, steam and hot water supply	Mean Return	0.52 %	-3.67 %	-4.45 %	1.71 %	-2.57 %	-4.05 %
		No. of Firms	[65]	-	-	[106]	-	-
41	Collection, purification and distribution of	Mean Return	-5.52 %	-3.85 %	-4.23 %	1.03 %	-3.88 %	-4.16 %
	water	No. of Firms	[20]	-	-	[30]	-	-
45	Construction	Mean Return	-1.77 %	-3.69 %	-6.52 %	-0.92 %	-4.20 %	-6.23 %
		No. of Firms	[117]	-	-	[209]	-	-
50	Sale, maintenance and repair of motor vehicles	Mean Return	-3.81 %	-5.94 %	-4.25 %	-3.08 %	-5.95 %	-3.95 %
	and motorcycles; retail sale services of	No. of Firms	[18]	-	-	[44]	-	-
51	Wholesale trade and commission trade, except	Mean Return	-5.40 %	-5.97 %	-4.21 %	-5.89 %	-5.64 %	-4.20 %
	of motor vehicles and motorcycles	No. of Firms	[150]	-	-	[232]	-	-
52	Retail trade, except of motor vehicles and	Mean Return	-7.67 %	-6.74 %	-4.47 %	-4.88 %	-6.42 %	-4.43 %
	motorcycles; repair of personal and household	No. of Firms	[98]	-	-	[183]	-	-
55	Hotels and restaurants	Mean Return	-7.00 %	-3.43 %	-5.35 %	-6.06 %	-2.62 %	-5.56 %
		No. of Firms	[57]	-	-	[118]	-	-
60	Land transport; transport via pipelines	Mean Return	-1.61 %	-4.28 %	-3.81 %	-1.72 %	-4.16 %	-3.52 %
		No. of Firms	[12]	-	-	[26]	-	-
61	Water transport	Mean Return	-8.58 %	-3.52 %	-2.72 %	-5.16 %	-2.99 %	-3.28 %
		No. of Firms	[41]	-	-	[65]	-	-
62	Air transport	Mean Return	-7.32 %	-4.11 %	-5.11 %	-8.39 %	-4.17 %	-5.19 %
		No. of Firms	[17]	-	-	[21]	-	-
63	Supporting and auxiliary transport activities;	Mean Return	-1.26 %	-5.67 %	-4.70 %	-0.73 %	-5.80 %	-4.73 %
	activities of travel agencies	No. of Firms	[28]	-	-	[45]	-	-
64	Post and telecommunications	Mean Return	-11.34 %	-7.35 %	-6.23 %	-12.59 %	-7.63 %	-6.05 %
		No. of Firms	[55]	-	-	[92]	-	-
65	Financial intermediation, except insurance and	Mean Return	-5.81 %	-7.71 %	-5.98 %	-6.34 %	-7.78 %	-5.51 %
	pension funding	No. of Firms	[261]	-	-	[460]	-	-
66	Insurance and pension funding, except	Mean Return	-10.86 %	-5.84 %	-5.39 %	-10.00 %	-6.06 %	-5.15 %
	compulsory social security	No. of Firms	[57]	-	-	[94]	-	-
67	Activities auxiliary to financial intermediation	Mean Return	-3.30 %	-8.35 %	-8.35 %	-0.68 %	-9.19 %	-7.77 %
		No. of Firms	[93]	-	-	[198]	-	-
70	Real estate activities	Mean Return	-8.69 %	-5.11 %	-6.02 %	-8.17 %	-5.01 %	-5.66 %
		No. of Firms	[221]	-	-	[373]	-	-
71	Renting of machinery and equipment without	Mean Return	-15.07 %	-5.93 %	-4.79 %	-5.91 %	-6.32 %	-4.36 %
	operator and of personal and household goods	No. of Firms	[16]	-	-	[23]	-	-
72	Computer and related activities	Mean Return	-16.14 %	-7.81 %	-6.11 %	-16.93 %	-7.74 %	-6.26 %
		No. of Firms	[346]	-	-	[607]	-	-
73	Research and development	Mean Return	-	-	-	-12.00 %	-7.76 %	-5.85 %
		No. of Firms	-	-	-	[66]	-	-
74	Other business activities	Mean Return	-7.12 %	-8.34 %	-5.30 %	-8.32 %	-7.93 %	-5.34 %
		No. of Firms	[165]	-	-	[353]	-	-
75	Public administration and defence; compulsory	Mean Return	-	-	-	-	-	-
	social security	No. of Firms	-	-	-	-	-	-
80	Education	Mean Return	-	-	-	-	-	-
		No. of Firms	-	-	-	-	-	-
85	Health and social work	Mean Return	-6.29 %	-5.30 %	-5.16 %	-8.16 %	-5.43 %	-5.44 %
		No. of Firms	[27]	-	-	[47]	-	-
90	Sewage and refuse disposal, sanitation and	Mean Return	-5.25 %	-5.74 %	-4.98 %	-7.18 %	-5.81 %	-4.84 %
	similar activities	No. of Firms	[20]	-	-	[36]	-	-
91	Activities of membership organisation n.e.c.	Mean Return	-	-	-	-	-	-
		No. of Firms	-	-	-	-	-	-
92	Recreational, cultural and sporting activities	Mean Return	-11.93 %	-7.08 %	-6.98 %	-12.00 %	-6.93 %	-7.15 %
		No. of Firms	[92]	-	-	[187]	-	-
93	Other service activities	Mean Return	-	-	-	-	-	-
		No. of Firms	-	-	-	-	-	-
95	Private households with employed persons	Mean Return	-	-	-	-	-	-
		No. of Firms	-	-	-	-	-	-
		Min Return	-27.17 %	-14.75 %	-8.40 %	-27.06 %	-11.71 %	-8.52 %
		Max Return	9.10 %	-1.95 %	1.46 %	19.36 %	-1.15 %	2.01 %
		Min Firms	9	-	-	9	-	-
		Max Firms	346	-	-	607	-	-

Appendix 7: Exports-to-Total Output –Ratio and Customer Industry Cross-Predictability Effects

This table presents Fama-MacBeth coefficient estimates calculated as time-series averages from augmented monthly cross-sectional regressions of stock returns on lagged related industry returns interacted with exports-to-total output –ratio. Panel A contains the results for the Eurozone sample and panel B contains the results for the EU27 sample. Columns 1 to 4 contain the regular regression results. Column 5 reports the average coefficients from the size sorted regressions. The industries are based on NACE rev 1.1 industry classifications. Return data is retrieved from Thomson One Banker database and the export data is from Eurostat input-output tables for years 2000 and 2005. Each month t stocks are ranked into quartiles based on their industry's exports-to-total output –ratio which measures the firms exposure to out-of-sample customers. Industries with the lowest level of out-of-sample exports are allocated into quartile 1 whereas industries with the highest level of out-of-sample exports are allocated into quartile 4. $r^{customer}$ is the return on the customer industry portfolio in month $t-1$. R^2 is calculated as the average value of the R^2s collected from the cross-sectional regressions in the first step of the Fama-MacBeth procedure. t statistics are reported in parentheses. ∗∗∗, ∗∗, or ∗ indicates that the coefficient estimate is different from zero at the 1%, 5%, or 10% level, respectively.

Panel A: Eurozone	(1)	(2)	(3)	(4)
Constant	-0.006	0.001	-0.006	0.000
	(-1.20)	(0.20)	(-1.41)	(0.07)
$r^{composite}$ × $Rank_{t-1}$ (1^{st} quartile - low)	0.081	0.071	0.088	0.076
	(1.08)	(1.01)	(1.43)	(1.35)
$r^{composite}$ × $Rank_{t-1}$ (2^{nd} quartile)	0.062	0.069	0.048	0.067
	(0.83)	(1.00)	(0.76)	(1.18)
$r^{composite}$ × $Rank_{t-1}$ (3^{rd} quartile)	0.140	0.188*	0.121	0.162*
	(1.23)	(1.81)	(1.35)	(1.97)
$r^{composite}$ × $Rank_{t-1}$ (4^{th} quartile -high)	0.057	0.079	0.071	0.091
	(0.75)	(1.15)	(1.12)	(1.61)
R^2	0.007	0.006	0.008	0.007
T	120	120	120	120
Winsorized (5^{th} and 95^{th} percentiles)	No	No	Yes	Yes
Indexed stock return	No	Yes	No	Yes

Panel B: EU27	(1)	(2)	(3)	(4)
Constant	-0.004	0.001	-0.005	0.000
	(-0.83)	(0.22)	(-1.12)	(-0.02)
$r^{composite}$ × $Rank_{t-1}$ (1^{st} quartile - low)	0.159*	0.165**	0.149**	0.156**
	(1.96)	(2.11)	(2.28)	(2.51)
$r^{composite}$ × $Rank_{t-1}$ (2^{nd} quartile)	0.330***	0.337***	0.255***	0.283***
	(2.99)	(3.14)	(2.79)	(3.18)
$r^{composite}$ × $Rank_{t-1}$ (3^{rd} quartile)	0.162**	0.191***	0.139**	0.170***
	(2.19)	(2.68)	(2.3)	(2.99)
$r^{composite}$ × $Rank_{t-1}$ (4^{th} quartile -high)	0.189**	0.189***	0.170***	0.175***
	(2.41)	(2.61)	(2.77)	(3.11)
R^2	0.061	0.005	0.007	0.006
T	120	120	120	120
Winsorized (5^{th} and 95^{th} percentiles)	No	No	Yes	Yes
Indexed stock return	No	Yes	No	Yes

Appendix 8: Descriptive Statistics for Firm Size Quintiles

This table presents descriptive statistics for the firm size quintiles. Panel A contains the statistics for the Eurozone sample and panel B contains the statistics for the EU27 sample. Each month t stocks are ranked into five quintiles based on firm size which is measured as the firm market capitalization in month $t-1$. Stocks with the lowest market capitalizations are allocated into quintile 1 and stocks with the highest market capitalizations are allocated into quintile 5. Return and market capitalization data is retrieved from Thomson One Banker database. Quintile mean return is the simple average of the monthly stock returns of firms within a quintile calculated over the whole sample period ranging from 2000 to 2009. Composite mean return is the simple average of the monthly returns on firms' composite portfolios within a quintile calculated over the whole sample period. Composite portfolio return is calculated as the simple average of monthly returns on a firm's supplier and customer industries. All return figures are annualized. Median market capitalization is the median firm market value calculated over all companies within a quintile for the whole sample period. Market values are expressed in millions of euros. Mean number of firms is the average monthly number of firms within a quintile calculated over the whole sample period. Number of industries is the number of different industries represented within a quintile. Number of countries is the number of different countries represented within a quintile. Mean quintile upper limit is the market capitalization, expressed in millions of euros, that is used as the upper limit for the quintile.

Panel A: Eurozone

Quintile	Quintile Mean Monthly Return	Composite Mean Monthly Return	Median Market Capitalization (€M)	Mean No. of Firms	No. of Industries	No. of Countries	Mean Quintile Upper Limit
1st quintile - low	-11.50 %	-5.93 %	5.96	392	45	26	14.80
2nd quintile	-13.02 %	-5.79 %	24.14	392	45	26	46.34
3rd quintile	-13.23 %	-5.64 %	72.33	392	45	26	145.11
4th quintile	-7.01 %	-5.61 %	280.00	392	45	26	708.77
5th quintile - high	-2.60 %	-5.58 %	2 536.98	392	44	26	120 013.25

Panel B: EU27

Quintile	Quintile Mean Monthly Return	Composite Mean Monthly Return	Median Market Capitalization (€M)	Mean No. of Firms	No. of Industries	No. of Countries	Mean Quintile Upper Limit
1st quintile - low	-12.63 %	-6.21 %	4.91	678	49	17	12.43
2nd quintile	-12.43 %	-6.09 %	20.82	677	49	17	39.70
3rd quintile	-12.66 %	-5.99 %	62.52	677	49	17	126.97
4th quintile	-5.56 %	-5.95 %	240.54	677	49	17	609.47
5th quintile - high	-2.18 %	-5.85 %	2 050.20	678	49	16	167 482.59

Appendix 9: Descriptive Statistics for Analyst Coverage Quintiles

This table presents descriptive statistics for the analyst coverage quintiles. Panel A contains the statistics for the Eurozone sample and panel B contains the statistics for the EU27 sample. Each month t stocks are ranked into five quintiles based on the level of analyst coverage in month $t-1$. Stocks with the lowest level of analyst coverage are allocated into quintile 1 and stocks with the highest level of analyst coverage are allocated into quintile 5. Analyst coverage measure is retrieved from I/B/E/S detail database and is defined as the numerical count of EPS estimates included in the mean EPS estimate for the stock in month $t-1$. Return and market capitalization data is retrieved from Thomson One Banker database. Quintile mean return is the simple average of the monthly stock returns of firms within a quintile calculated over the whole sample period ranging from 2000 to 2009. Composite mean return is the simple average of the monthly returns on firms' composite portfolios within a quintile calculated over the whole sample period. Composite portfolio return is calculated as the simple average of monthly returns on a firm's supplier and customer industries. All return figures are annualized. Median market capitalization is the median firm market value calculated over all companies within a quintile for the whole sample period. Market values are expressed in millions of euros. Mean number of firms is the average monthly number of firms within a quintile calculated over the whole sample period. Number of industries is the number of different industries represented within a quintile. Number of countries is the number of different countries represented within a quintile. Mean quintile upper limit is the number of analyst EPS estimates included in the mean EPS estimate for a stock that is used as the upper limit for the quintile.

Panel A: Eurozone

Quintile	Quintile Mean Monthly Return	Composite Mean Monthly Return	Median Market Capitalization (€M)	Mean No. of Firms	No. of Industries	No. of Countries	Mean Quintile Upper Limit
1st quintile - low	-13.39 %	-5.33 %	64.53	263	45	16	1
2nd quintile	-20.85 %	-8.86 %	77.76	114	45	16	2
3rd quintile	-16.39 %	-9.11 %	104.87	128	45	16	5
4th quintile	-6.97 %	-6.51 %	140.92	151	45	15	11
5th quintile - high	-7.23 %	-6.77 %	654.46	152	41	12	39

Panel B: EU27

Quintile	Quintile Mean Monthly Return	Composite Mean Monthly Return	Median Market Capitalization (€M)	Mean No. of Firms	No. of Industries	No. of Countries	Mean Quintile Upper Limit
1st quintile - low	-13.71 %	-5.91 %	55.82	425	49	25	1
2nd quintile	-14.40 %	-7.60 %	79.02	171	49	25	2
3rd quintile	-14.75 %	-8.36 %	96.21	175	49	23	4
4th quintile	-9.40 %	-6.55 %	147.87	232	49	21	10
5th quintile - high	-6.12 %	-6.86 %	500.41	234	47	18	39

Appendix 10: Industry Inclusion Frequencies for Investor Geographic Specialization Analysis

This table presents the frequencies with which industries enter a quartile based on previous-month geographic dispersion of related industries during the sample period. The figures are reported for both Eurozone and EU27 samples separately. Each month t stocks are ranked into four quartiles based on the level of geographic dispersion of their related industries as measured by the $Geodisp_{j,t}$ variable in month $t-1$. Stocks with the lowest $Geodisp_{j,t}$ value are allocated into quartile 1 and stocks with the highest $Geodisp_{j,t}$ value are allocated into quartile 4. Traded-goods industry column indicates whether the industry belongs to a traded-goods (T) or nontraded-goods industry (-) as specified by Griffin and Karolyi (1998).

		Eurozone				EU27			
Industry	Traded-goods industry (T)	1st quartile	2nd quartile	3rd quartile	4th quartile	1st quartile	2nd quartile	3rd quartile	4th quartile
1	-	0.00	0.00	0.00	1.00	0.02	0.00	0.00	0.98
5	-	0.00	0.00	0.00	1.00	0.00	0.02	0.00	0.98
10	T	0.00	0.00	0.00	1.00	0.02	0.00	0.00	0.98
11	T	-	-	-	-	0.05	0.20	0.73	0.03
13	T	0.11	0.15	0.34	0.40	0.11	0.18	0.38	0.34
14	T	-	-	-	-	0.05	0.33	0.63	0.00
15	T	0.78	0.20	0.03	0.00	0.76	0.24	0.00	0.00
16	T	-	-	-	-	0.60	0.13	0.26	0.02
17	T	0.00	0.00	0.42	0.58	0.02	0.00	0.41	0.58
18	T	0.96	0.04	0.00	0.00	0.94	0.06	0.00	0.00
20	T	0.00	0.00	0.16	0.84	0.00	0.00	0.16	0.84
21	T	0.00	0.00	0.51	0.49	0.00	0.02	0.49	0.49
22	-	0.00	0.00	0.54	0.46	0.00	0.02	0.56	0.43
23	T	1.00	0.00	0.00	0.00	0.98	0.00	0.02	0.00
24	T	0.03	0.04	0.53	0.41	0.03	0.03	0.53	0.41
25	T	0.00	0.00	0.00	1.00	0.00	0.02	0.01	0.98
26	T	0.00	0.00	0.11	0.89	0.02	0.00	0.09	0.89
27	T	0.36	0.31	0.33	0.00	0.35	0.33	0.33	0.00
28	-	0.02	0.23	0.65	0.10	0.02	0.23	0.68	0.07
29	T	0.73	0.27	0.00	0.00	0.74	0.26	0.00	0.00
30	T	0.00	0.08	0.41	0.51	0.00	0.09	0.38	0.53
31	-	0.08	0.34	0.52	0.06	0.07	0.38	0.53	0.03
32	-	0.11	0.18	0.70	0.01	0.13	0.18	0.69	0.01
33	-	1.00	0.00	0.00	0.00	0.98	0.01	0.00	0.01
35	T	0.62	0.02	0.13	0.23	0.60	0.03	0.14	0.23
36	-	0.00	0.11	0.68	0.22	0.00	0.11	0.69	0.20
40	-	0.73	0.26	0.01	0.00	0.71	0.28	0.01	0.00
41	-	0.00	0.00	0.58	0.43	0.01	0.01	0.63	0.35
45	-	0.00	0.01	0.94	0.05	0.02	0.01	0.96	0.02
50	-	1.00	0.00	0.00	0.00	0.98	0.01	0.01	0.00
51	-	0.09	0.91	0.00	0.00	0.13	0.88	0.00	0.00
52	-	0.00	0.00	0.06	0.94	0.00	0.01	0.10	0.89
55	-	0.00	0.00	0.01	0.99	0.00	0.01	0.02	0.98
60	T	0.00	0.20	0.80	0.00	0.00	0.21	0.79	0.00
61	-	0.98	0.02	0.00	0.00	0.98	0.03	0.00	0.00
62	T	0.93	0.07	0.00	0.00	0.93	0.08	0.00	0.00
63	-	0.23	0.21	0.56	0.00	0.23	0.22	0.55	0.00
64	-	1.00	0.00	0.00	0.00	0.98	0.02	0.00	0.00
65	-	0.52	0.48	0.00	0.00	0.56	0.44	0.00	0.00
66	-	0.75	0.22	0.03	0.00	0.77	0.23	0.00	0.00
67	-	0.49	0.00	0.00	0.51	0.51	0.00	0.00	0.49
70	-	0.00	0.07	0.71	0.23	0.02	0.06	0.73	0.19
71	-	0.02	0.61	0.38	0.00	0.02	0.63	0.36	0.00
72	-	0.12	0.69	0.07	0.13	0.12	0.70	0.08	0.11
73	-	-	-	-	-	0.36	0.58	0.06	0.00
74	-	0.00	0.18	0.63	0.20	0.00	0.18	0.73	0.09
85	-	0.01	0.14	0.75	0.10	0.01	0.16	0.73	0.10
90	-	0.89	0.11	0.00	0.00	0.88	0.11	0.02	0.00
92	-	1.00	0.00	0.00	0.00	0.98	0.02	0.00	0.00

Appendix 11: Descriptive Statistics for Geodisp Quartiles

This table presents descriptive statistics for the $Geodisp_{j,t}$ quartiles. Panel A contains the statistics for the Eurozone sample and panel B contains the statistics for the EU27 sample. Each month t stocks are ranked into four quartiles based on the level of geographic dispersion in their related industries as measured by the $Geodisp_{j,t}$ variable in month $t-1$. Stocks with the lowest $Geodisp_{j,t}$ value are allocated into quartile 1 and stocks with the highest $Geodisp_{j,t}$ value are allocated into quartile 4. All data is retrieved from Thomson One Banker database. Quartile mean return is the simple average of the monthly stock returns of firms within a quartile calculated over the whole sample period ranging from 2000 to 2009. Composite mean return is the simple average of the monthly returns on firms' composite portfolios within a quartile calculated over the whole sample period. Composite portfolio return is calculated as the simple average of monthly returns on a firm's supplier and customer industries. All return figures are annualized. Median market capitalization is the median firm market value calculated over all companies within a quartile for the whole sample period. Market values are expressed in millions of euros. Mean number of firms is the average monthly number of firms within a quartile calculated over the whole sample period. Number of industries is the number of different industries represented within a quartile. Number of countries is the number of different countries represented within a quartile. Mean quartile upper limit is the $Geodisp_{j,t}$ value that is used as the upper limit for the quartile.

Panel A: Eurozone							
Quartile	Quartile Mean Monthly Return	Composite Mean Monthly Return	Median Market Capitalization (€M)	Mean No. of Firms	No. of Industries	No. of Countries	Mean Quartile Upper Limit
1st quartile - low	-9.22 %	-6.17 %	115.50	572	27	26	54
2nd quartile	-8.73 %	-4.40 %	52.26	474	27	26	59
3rd quartile	-11.90 %	-6.84 %	71.90	471	28	26	70
4th quartile - high	-12.18 %	-4.49 %	63.36	446	26	26	133
Panel B: EU27							
Quartile	Quartile Mean Monthly Return	Composite Mean Monthly Return	Median Market Capitalization (€M)	Mean No. of Firms	No. of Industries	No. of Countries	Mean Quartile Upper Limit
1st quartile - low	-8.00 %	-6.05 %	86.23	993	38	17	54
2nd quartile	-8.67 %	-5.10 %	46.14	781	42	17	59
3rd quartile	-10.67 %	-5.05 %	62.67	878	34	17	70
4th quartile - high	-13.31 %	-5.22 %	57.38	743	29	17	133

Appendix 12: Trading Strategy Industry Inclusion Probabilities

This table presents the frequency with which industries enter the short and long portfolios in the self-financing strategies based on previous-month related industry returns. Panel A contains the statistics for the Eurozone sample and panel B contains the statistics for the EU27 sample. The industries are based on NACE rev 1.1 industry classifications. Each month t industries are sorted into five quintiles according to their related industry returns in month $t - 1$. Industries with the lowest previous-month related industry returns are allocated into quintile 1 and industries with the highest previous-month related industry returns are allocated into quintile 5. In the self-financing strategy an industry enters the short (long) portfolio if the industry's previous-month related industry return is in the bottom (top) quintile. The reported figure is the number of months an industry enters a portfolio-strategy combination divided by the number of trading months (120 months, January 2000 to December 2009).

Panel A: Eurozone

		$r^{supplier,t-1}$		$r^{customer,t-1}$		$r^{composite,t-1}$	
Industry code	Industry name	Low	High	Low	High	Low	High
1	Agriculture, hunting and related service activities	0.28	0.32	0.38	0.48	0.38	0.42
2	Forestry, logging and related service activities	-	-	-	-	-	-
5	Fishing, operating of fish hatcheries and fish farms	0.18	0.22	0.33	0.34	0.28	0.27
10	Mining of coal and lignite; extraction of peat	0.08	0.23	0.35	0.46	0.28	0.43
11	Extraction of crude petroleum and natural gas	-	-	-	-	-	-
12	Mining of uranium and thorium ores	-	-	-	-	-	-
13	Mining of metal ores	0.07	0.18	0.37	0.50	0.28	0.43
14	Other mining and quarrying	-	-	-	-	-	-
15	Manufacture of food products and beverages	0.32	0.35	0.37	0.33	0.31	0.33
16	Manufacture of tobacco products	-	-	-	-	-	-
17	Manufacture of textiles	0.16	0.15	0.39	0.36	0.28	0.28
18	Manufacture of wearing apparel; dressing and dyeing of fur	0.54	0.15	0.24	0.06	0.53	0.08
19	Tanning and dressing of leather	-	-	-	-	-	-
20	Manufacture of wood and of products of wood and cork	0.05	0.08	0.29	0.29	0.18	0.17
21	Manufacture of pulp, paper and paper products	0.13	0.13	0.31	0.18	0.20	0.13
22	Publishing, printing and reproduction of recorded media	0.21	0.10	0.20	0.07	0.18	0.13
23	Manufacture of coke, refined petroleum products and nuclear fuels	0.11	0.13	0.11	0.19	0.13	0.18
24	Manufacture of chemicals and chemical products	0.06	0.04	0.18	0.22	0.12	0.17
25	Manufacture of rubber and plastic products	0.26	0.25	0.06	0.09	0.15	0.17
26	Manufacture of other non-metallic mineral products	0.03	0.05	0.28	0.40	0.19	0.32
27	Manufacture of basic metals	0.09	0.19	0.28	0.48	0.22	0.43
28	Manufacture of fabricated metal products, except machinery and equipment	0.28	0.38	0.20	0.36	0.21	0.33
29	Manufacture of machinery and equipment n.e.c.	0.18	0.43	0.07	0.17	0.13	0.34
30	Manufacture of office machinery and computers	0.25	0.19	0.18	0.15	0.27	0.16
31	Manufacture of electrical machinery and apparatus n.e.c.	0.15	0.28	0.23	0.24	0.15	0.27
32	Manufacture of radio, television and communication equipment and apparatus	0.12	0.12	0.27	0.19	0.18	0.16
33	Manufacture of medical, precision and optical instruments, watches and clocks	0.17	0.19	0.36	0.36	0.26	0.28
34	Manufacture of motor vehicles, trailers and semi-trailers	-	-	-	-	-	-
35	Manufacture of other transport equipment	0.18	0.31	0.43	0.35	0.28	0.28
36	Manufacture of furniture; manufacturing n.e.c.	0.26	0.29	0.03	0.01	0.14	0.16
37	Recycling	-	-	-	-	-	-
40	Electricity, gas, steam and hot water supply	0.11	0.19	0.04	0.05	0.07	0.06
41	Collection, purification and distribution of water	0.08	0.21	0.05	0.08	0.03	0.11
45	Construction	0.18	0.31	0.33	0.21	0.20	0.18
50	Sale, maintenance and repair of motor vehicles and motorcycles	0.01	0.02	0.16	0.21	0.07	0.07
51	Wholesale trade and commission trade, excl. motor vehicles	0.15	0.12	0.03	0.03	0.02	0.03
52	Retail trade, except of motor vehicles and motorcycles; repair of personal and household goods	0.22	0.09	0.01	0.02	0.05	0.02
55	Hotels and restaurants	0.30	0.32	0.04	0.07	0.16	0.17
60	Land transport; transport via pipelines	0.20	0.21	0.08	0.17	0.13	0.20
61	Water transport	0.26	0.29	0.18	0.27	0.22	0.29
62	Air transport	0.32	0.33	0.18	0.15	0.25	0.25
63	Supporting and auxiliary transport activities; activities of travel agencies	0.25	0.20	0.23	0.26	0.24	0.18
64	Post and telecommunications	0.27	0.18	0.12	0.08	0.22	0.10
65	Financial intermediation, except insurance and pension funding	0.42	0.32	0.17	0.11	0.31	0.23
66	Insurance and pension funding, except compulsory social security	0.34	0.38	0.13	0.10	0.29	0.33
67	Activities auxiliary to financial intermediation	0.38	0.18	0.43	0.39	0.39	0.34
70	Real estate activities	0.25	0.31	0.15	0.10	0.18	0.13
71	Renting of machinery and equipment without operator and of personal and household goods	0.19	0.13	0.03	0.02	0.07	0.08
72	Computer and related activities	0.26	0.10	0.17	0.08	0.21	0.07
73	Research and development	-	-	-	-	-	-
74	Other business activities	0.34	0.18	0.05	0.02	0.22	0.09
75	Public administration and defence; compulsory social security	-	-	-	-	-	-
80	Education	-	-	-	-	-	-
85	Health and social work	0.13	0.09	0.19	0.12	0.10	0.10
90	Sewage and refuse disposal, sanitation and similar activities	0.08	0.07	0.10	0.09	0.03	0.02
91	Activities of membership organisation n.e.c.	-	-	-	-	-	-
92	Recreational, cultural and sporting activities	0.15	0.06	0.26	0.14	0.23	0.10
93	Other service activities	-	-	-	-	-	-
95	Private households with employed persons	-	-	-	-	-	-
	Mean	0.20	0.20	0.20	0.20	0.20	0.20
	Min	0.01	0.02	0.01	0.01	0.02	0.02
	Max	0.54	0.43	0.43	0.50	0.53	0.43

Appendix 12: Continued

Panel B: EU27							
		$r^{supplier,t-1}$		$r^{customer,t-1}$		$r^{composite,t-1}$	
Industry code	Industry name	Low	High	Low	High	Low	High
1	Agriculture, hunting and related service activities	0.26	0.32	0.37	0.47	0.35	0.43
2	Forestry, logging and related service activities	-	-	-	-	-	-
5	Fishing, operating of fish hatcheries and fish farms	0.23	0.25	0.34	0.39	0.31	0.36
10	Mining of coal and lignite; extraction of peat	0.06	0.22	0.33	0.45	0.25	0.43
11	Extraction of crude petroleum and natural gas	0.07	0.11	0.39	0.45	0.32	0.40
12	Mining of uranium and thorium ores	-	-	-	-	-	-
13	Mining of metal ores	0.11	0.28	0.31	0.43	0.23	0.39
14	Other mining and quarrying	0.11	0.22	0.24	0.35	0.17	0.29
15	Manufacture of food products and beverages	0.28	0.34	0.31	0.32	0.30	0.33
16	Manufacture of tobacco products	0.26	0.23	0.03	0.02	0.13	0.13
17	Manufacture of textiles	0.16	0.18	0.32	0.33	0.25	0.28
18	Manufacture of wearing apparel; dressing and dyeing of fur	0.50	0.23	0.19	0.07	0.45	0.16
19	Tanning and dressing of leather	-	-	-	-	-	-
20	Manufacture of wood and of products of wood and cork	0.07	0.11	0.31	0.34	0.21	0.21
21	Manufacture of pulp, paper and paper products	0.12	0.13	0.34	0.18	0.19	0.13
22	Publishing, printing and reproduction of recorded media	0.19	0.11	0.28	0.08	0.19	0.08
23	Manufacture of coke, refined petroleum products and nuclear fuels	0.31	0.44	0.17	0.25	0.26	0.36
24	Manufacture of chemicals and chemical products	0.06	0.03	0.22	0.25	0.11	0.13
25	Manufacture of rubber and plastic products	0.33	0.28	0.05	0.11	0.15	0.18
26	Manufacture of other non-metallic mineral products	0.11	0.24	0.29	0.42	0.17	0.33
27	Manufacture of basic metals	0.07	0.17	0.22	0.43	0.14	0.35
28	Manufacture of fabricated metal products, except machinery and equipment	0.26	0.36	0.19	0.37	0.18	0.30
29	Manufacture of machinery and equipment n.e.c.	0.16	0.29	0.05	0.13	0.10	0.19
30	Manufacture of office machinery and computers	0.33	0.22	0.25	0.16	0.31	0.23
31	Manufacture of electrical machinery and apparatus n.e.c.	0.23	0.24	0.22	0.28	0.19	0.24
32	Manufacture of radio, television and communication equipment and apparatus	0.20	0.21	0.35	0.23	0.28	0.22
33	Manufacture of medical, precision and optical instruments, watches and clocks	0.21	0.19	0.36	0.35	0.31	0.25
34	Manufacture of motor vehicles, trailers and semi-trailers	-	-	-	-	-	-
35	Manufacture of other transport equipment	0.19	0.22	0.43	0.33	0.34	0.29
36	Manufacture of furniture; manufacturing n.e.c.	0.23	0.29	0.01	0.00	0.13	0.14
37	Recycling	-	-	-	-	-	-
40	Electricity, gas, steam and hot water supply	0.26	0.38	0.03	0.08	0.14	0.27
41	Collection, purification and distribution of water	0.08	0.15	0.13	0.10	0.05	0.10
45	Construction	0.15	0.21	0.30	0.21	0.20	0.16
50	Sale, maintenance and repair of motor vehicles and motorcycles	0.03	0.03	0.06	0.17	0.06	0.07
51	Wholesale trade and commission trade, excl. motor vehicles	0.10	0.08	0.03	0.03	0.01	0.00
52	Retail trade, except of motor vehicles and motorcycles; repair of personal and household goods	0.16	0.08	0.02	0.01	0.03	0.05
55	Hotels and restaurants	0.28	0.35	0.08	0.06	0.18	0.17
60	Land transport; transport via pipelines	0.17	0.18	0.10	0.12	0.15	0.17
61	Water transport	0.26	0.35	0.13	0.23	0.22	0.30
62	Air transport	0.23	0.27	0.18	0.15	0.18	0.27
63	Supporting and auxiliary transport activities; activities of travel agencies	0.23	0.23	0.24	0.28	0.27	0.27
64	Post and telecommunications	0.33	0.17	0.15	0.11	0.28	0.13
65	Financial intermediation, except insurance and pension funding	0.35	0.26	0.11	0.05	0.28	0.17
66	Insurance and pension funding, except compulsory social security	0.31	0.29	0.10	0.06	0.24	0.21
67	Activities auxiliary to financial intermediation	0.39	0.18	0.43	0.36	0.39	0.27
70	Real estate activities	0.24	0.28	0.18	0.08	0.15	0.12
71	Renting of machinery and equipment without operator and of personal and household goods	0.18	0.09	0.03	0.05	0.07	0.05
72	Computer and related activities	0.29	0.10	0.23	0.08	0.27	0.05
73	Research and development	0.20	0.13	0.29	0.23	0.27	0.15
74	Other business activities	0.33	0.10	0.03	0.03	0.19	0.05
75	Public administration and defence; compulsory social security	-	-	-	-	-	-
80	Education	-	-	-	-	-	-
85	Health and social work	0.16	0.13	0.16	0.11	0.10	0.04
90	Sewage and refuse disposal, sanitation and similar activities	0.07	0.03	0.14	0.12	0.05	0.03
91	Activities of membership organisation n.e.c.	-	-	-	-	-	-
92	Recreational, cultural and sporting activities	0.13	0.04	0.32	0.14	0.25	0.09
93	Other service activities	-	-	-	-	-	-
95	Private households with employed persons	-	-	-	-	-	-
	Mean	0.20	0.20	0.20	0.20	0.20	0.20
	Min	0.03	0.03	0.01	0.00	0.01	0.00
	Max	0.50	0.44	0.43	0.47	0.45	0.43