JEL classification: G12, G14, G32, M41
Keywords: fundamental analysis, stock valuation, mispricing, market efficiency, limits to arbitrage

Separating Winners from Losers: Composite Indicators Based on Fundamentals in the European Context*

Borja AMOR-TAPIA—(borja.amor@unileon.es), *corresponding author*
María T. TASCÓN—(m.tascon@unileon.es)
both authors: University of Leon (Spain)

Abstract
Our work tests the usefulness of score measures based on fundamental signals in an out-of-sample study during the past two decades. While previous research, primarily focused on the US, demonstrates that fundamental signals derived from financial statements allow for future abnormal stock returns, our European sample documents that Xue-Zhang's FSCORE2 and Wahlen-Wieland's PEIS do not serve as a source of market anomalies. By contrast, other fundamental signals—Piotroski's FSCORE and Mohanram's GSCORE—still allow for abnormal returns in our sample and period. We also contribute to the market efficiency debate by documenting the role of idiosyncratic volatility, transaction costs and noise trader risk in the persistence of market anomalies, supporting the existence of limits to arbitrage for these investment strategies.

1. Introduction

In an efficient market, prices should incorporate available information in a timely manner. From this point of view, a market anomaly is a pattern of stock returns that appears to contradict traditional asset pricing models. Although stock returns may be affected by multiple pieces of information, financial statements are the primary source of information because they summarize firm performance. This information can be employed to forecast cash flows, estimate risk and obtain the intrinsic value of the given firm, which can be compared to market prices. If this information is not incorporated in a timely fashion by stock returns, an anomaly may arise and arbitrage opportunities may emerge. However, once everyone is aware of the anomaly, it should disappear. The analysis of an anomaly previously discovered raises the question of whether profit opportunities survive.

One way of summarizing the information contained in financial statements is to build measures that bring together a set of positive/negative signals. In broad terms, we can formulate the following equation:

$$Composite_{i,t} = \sum_{1}^{j} Signal_j \tag{1}$$

where a "composite" or "score" to the firm i in the year t is formed by adding j

* We acknowledge the helpful comments of the Editor and the anonymous referees of this journal and the participants of the 36th Annual Congress of the European Accounting Association (Paris) and the IX Workshop on Empirical Research in Financial Accounting (Las Palmas de Gran Canaria University, Spain). All mistakes are ours.
We thank the support of the Spanish Ministry of Science and Innovation [ECO2011-29144-C03-01] and the University of Leon [ULE2012-1].

signals (positive/negative news about the firm) drawn from the accounting information. This paper explores whether investors are able to exploit the documented abnormal returns to fundamental signals reflected in financial statement information. In particular, we revisit some previously documented anomalies. FSCORE1 (Piotroski, 2000) summarizes an array of nine binary signals from financial statements; FSCORE2 (Xue and Zhang, 2011) summarizes eleven binary signals formed on a relative basis by comparing a firm's financials with industry averages; GSCORE (Mohanram, 2005) summarizes eight binary signals formed by comparisons between a firm's financials and industry medians; and PEIS (Wahlen and Wieland, 2011) employs six signals (positive, negative or no news) formed by a firm's quintile position of six accounting ratios.

We examine whether anomalies based on fundamental scores exist several years after the anomaly has been identified and, in the affirmative case, we test whether there are frictions such as transaction costs or other risks allowing the existence of these patterns. It is important to note that the samples are mainly extracted from US markets, while we perform out-of-sample tests in fourteen European markets.

This paper makes several important contributions. First, we provide evidence of the persistence of the fundamental signals and contribute to the debate on the information impounded in prices.In an efficient market, information should be incorporated into prices once it becomes known. Therefore, previously published anomalies should have disappeared. Our out-of-sample evidence finds that utilizing Xue and Zhang's FSCORE2 and Wahlen and Wieland's PEIS in a hedge strategy that goes long in strong firms—high fundamentals—and short in weak firms—low fundamentals—does not generate one-year-ahead buy-and-hold abnormal returns. These results suggest that markets are efficient. Second, we demonstrate that two surviving anomalies, Piotroski's and Mohanram's, allow investors to construct hedge portfolios that earn one-year buy-and-hold abnormal returns. These results suggest that markets may not be semi-strong efficient or, alternatively, scores could capture some underlying risk. Third, our findings provide evidence of the reasons behind the persistence of these surviving anomalies. Although the efficient markets theory establishes that anomalies would be exploited and eliminated by rational arbitrageurs, frictions could prevent anomalies from disappearing completely. Specifically, the argument of the limits of arbitrage asserts that idiosyncratic risk, noise trader momentum risk and implementation costs make arbitrage difficult. Our results support the relevant role of these arguments. We find that both FSCORE1 and GSCORE exhibit greater idiosyncratic volatility concentrated in firms with weak fundamentals that protect the existence of greater levels of mispricing. Along with idiosyncratic volatility, we find support for the presence of noise trader momentum risk. Finally, employing a set of proxies for direct transaction costs, indirect transaction costs, short selling risk and investor sophistication, our results suggest that firms with weak fundamentals have high implementation costs that limit the action of arbitrageurs.

Our overall conclusion is that investors should pay attention to the limits of arbitrage before implementing previously documented anomalies, and out-of-sample tests should be performed before employing previously documented anomalies in markets other than those where the anomalies were identified. We find evidence that some anomalies previously documented in the US do not work in the European markets included in our sample, while surviving anomalies are found to be dominated by some limits of arbitrage that protect the existence of semi-strong inefficiencies. Hence, hedge portfolio strategies based on financial statement analysis should be taken with caution.

The remainder of the paper is organized as follows: Section 2 reviews the previous evidence pertaining to fundamental analysis and the limits of arbitrage. Section 3 describes the data to be employed in the empirical tests and presents descriptive evidence. Section 4 tests whether the predictive ability of a number of fundamental signals documented in the previous literature could be exploited in European markets. Section 5 analyzes the limits of arbitrage for those scores that we find useful in Europe. Finally, we present the conclusions in Section 6.

2. Previous Evidence

Efficient markets have two characteristic features (Brav and Heaton, 2002): investors are assumed to have essentially complete knowledge of the fundamental structure of their economy—*information*—and are assumed to be completely rational information processors who make optimal decisions—*rationality*. If one of these two assumptions fails, abnormal stock returns may appear. We examine the possibility that by employing fundamental analysis, abnormal stock returns are obtained years after the anomaly was discovered.

2.1 Fundamental Analysis Research

Fundamental analysis focuses on the translation of the information contained in financial statements into estimates of values to distinguish "winners" (undervalued firms) from "losers" (overvalued firms). One approach to carrying out this task is to obtain the firm intrinsic value and the systematic errors in market expectations (see, for example, Frankel and Lee, 1998, among others). Another approach is to trade on signals of financial performance. The abnormal returns generated by the signals could be due to the market's inability to fully understand a particular piece of information or due to failures in the rational decision-making process. In the literature, there are many examples of individual signals such as accruals and post-earnings-announcement drift and composite signals built upon various pieces of information, such as FSCORE (Piotroski, 2000), GSCORE (Mohanram, 2005), FSCORE (Xue and Zhang, 2011) and PEIS (Wahlen and Wieland, 2011). These composite signals aggregate the information contained in an array of performance measures or screens from financial statements and form portfolios on the basis of a firm's overall signal. Previous research has shown that these investment strategies earn abnormal buy-and-hold returns (ABHR).

Piotroski (2000) builds FSCORE based on nine individual binary signals derived from accounting data (profitability, financial leverage/liquidity and operating efficiency). He finds that strong (high FSCORE) value firms—or low book-to-market (BM)—experience improved future firm performance and stock returns relative to weak (low FSCORE) value firms, suggesting that the market does not impound financial statement information into prices in a timely manner.

FSCORE is employed in related papers. Fama and French (2006) find that FSCORE employs proxies for expected net cash flows, but it seems to incorporate no economically important information on expected returns beyond the information in lagged profitability, asset growth and accruals. They emphasize that relationships among average returns and BM, profitability, asset growth, accruals and FSCORE may be due to rational or irrational pricing.

Mohanram (2005) complements Piotroski's (2000) study by introducing measures of the quality of by-firm growth-related financial signals to identify firms likely to grow continuously. Thus, Mohanram (2005) focuses on a set of eight variables, combined in the GSCORE index, that are able to strongly separate future winners from losers in glamour firms (low BM). Mohanram attributes these results to a mispricing-based explanation for the BM effect on low-BM firms. Despite the robustness of the reported results, the subsequent discussion by Piotroski (2005) raises some doubts about the inefficient markets story.

In this stream of literature, a set of questions emerges on the sources of those documented abnormal returns. If prices are not efficient and fundamental signals present profitable arbitrage opportunities, it is determinant to investigate whether sophisticated investors trade on these signals. Xue and Zhang (2011) examine institutional investors' trading on fundamental signals and its implications for stock valuation. They employ a modified Piotroski FSCORE with eleven signals (we have renamed it as FSCORE2). Several signals are the same as in Piotroski (2000), but Xue and Zhang choose fundamentals to describe the financial conditions of ordinary listed firms and therefore include the financial ratios that are most visible to investors. Instead of focusing on high-BM firms as in the Piotroski measure, tests of FSCORE2 are applied to all firms in US markets with available data. Their evidence demonstrates that institutional investors contribute to reducing abnormal returns based on financial statements information, although fundamentals-based abnormal returns point to the presence of limits to arbitrage for this investment strategy.

Finally, Wahlen and Wieland (2011) compute another score, called PEIS (predicted earnings increase score), to determine whether financial statement information could be exploited for identifying firms with more likely future earnings increases. The findings demonstrate that high-score stocks are more likely to increase future earnings and abnormal returns generated by a hedge portfolio traded on these signals exceed the consensus recommendations of trading analysts.

Because this evidence is based on US samples, we might ask whether it is present in other markets and, if so, whether the anomalies persist some years after the publication of the papers mentioned above.

2.2 Limits of Arbitrage

A financial anomaly is a pattern of prices found inconsistent with the rational expectations of traditional efficient markets.[1] Two competing explanations for these patterns exist. First, the return to portfolio strategies (anomalies) represents compensation for risk, as suggested by Fama and French (1992, 1993, 1996). This rational story views anomalies as by chance results. Apparently, overreaction should be as common as underreaction. In addition, anomalies should tend to disappear after being learned of or once they have been published and the market incorporates the publicly

[1] For example, the disconnection between credit and market risks (Choi *et al.,* 2010).

available information (learning vs. arbitrage in the words of Brav and Heaton, 2002). According to the discount-rate theories (Cochrane, 2011), anomalies would be discount-rate variations that we do not understand yet.[2]

Second, "anomalous" returns result from systematic mispricing, as people's expectations are wrong or at least some agents are not fully rational. If anomalies represent mispricing due to systematic bias in the expectations of noise traders, then we can examine why smart traders do not eliminate mispricing in a timely manner. Shleifer and Vishny (1997) argue that arbitrage is costly and any systematic mispricing would not be quickly and completely traded away in situations where arbitrage costs exceed arbitrage benefits. In other words, strategies designed to correct mispricing could be both risky and costly. Barberis and Thaler (2003) identify three risks of mispricing that limit the possibility to obtain abnormal returns: fundamental (or idiosyncratic) risk, noise trader (momentum) risk and implementation costs.

2.2.1 Idiosyncratic Risk

In Shleifer and Vishny's (1997) model, stocks are not rationally priced and idiosyncratic risk deters arbitrage—volatile securities will exhibit greater mispricing and a higher average return to arbitrage. In particular, some stocks with high idiosyncratic variance may be overpriced, and that overpricing is not eliminated by arbitrage because shorting them is risky. As a result, these volatile, overpriced stocks earn a lower expected return, making larger pricing errors possible. Idiosyncratic risk is proposed as a limit to arbitrage in several papers (Ali *et al.,* 2003; Mashruwala *et al.*, 2006; Pontiff 2006; Brav *et al.,* 2010). According to previous evidence and the limits of the arbitrage view, we can test the following hypothesis: *Idiosyncratic volatility protects the existence of mispricing.*

2.2.2 Noise Trader Momentum Risk

In a world with noise and smart traders, it could be possible that pessimistic investors causing a stock to be undervalued in the first place become even more pessimistic, lowering the price even further, and vice versa. As Shleifer and Vishny (1997) note, arbitrage is conducted by relatively few professional smart traders. The main feature of such arbitrage is that brains (for example, mutual fund managers) and resources (the money of millions of minor traders—noise traders) are linked by an agency relationship.[3] In this setting, losses can induce the withdrawal of some capital. Therefore, noise trader momentum risk is the risk of irrational beliefs getting worse in the same direction of mispricing. We can test the following hypothesis: *Momentum in the same direction protects the existence of mispricing.*

2.2.3 Implementation Costs

When securities are mispriced, transaction costs can make it less attractive to exploit arbitrage. We consider three types of implementation costs (Ali *et al.,* 2003): direct transaction costs, indirect transaction costs and costs associated with short selling. Direct transaction costs include bid-ask spreads and brokerage commissions.

[2] For example, Garcia-Blandon *et al.* (2011) find no evidence of ex-dividend day anomalies.

[3] If "smart" arbitrageurs manage the money of uninformed investors, those investors will only observe whether they lose or win money in the short run as signals of the managers' intelligence.

Bhardwaj and Brooks (1992) and Ali *et al.* (2003) suggest that quoted bid-ask spreads and commissions per share as percentages of share prices are inversely related to share prices. Thus, we first employ share prices as a measure of direct transaction costs, but we also utilize bid-ask spreads as an additional measure.

Because markets provide liquidity and price discovery, indirect transaction costs impound the adverse price effects of the trade and the delay in processing the transaction. Trading volume in terms of monetary units is a relevant determinant of these indirect costs (see, for example, Kyle, 1985; Bhushan, 1991). Price discovery requires the inclusion of new information into asset prices and liquidity therefore becomes a relevant factor. If stocks have low volumes of trading, transactions are less likely to be completed quickly and are more likely to cause adverse price effects.

Additionally, short selling provides an adjustment of prices. However, this activity is costly because short sellers must borrow the shorted securities and must return the securities on demand. The risk of a short squeeze is likely to be lower for stocks with substantial institutional ownership because it is easier to find alternative lenders of such stocks (Dechow *et al.,* 2001; Ali *et al.,* 2003; Brav *et al.,* 2010). We employ the percentage of institutional ownership as a proxy for the costs of short selling. According to the previous literature, we test the following hypothesis: *Implementation costs protect the existence of mispricing.*

3. Research Design

3.1 Composite Definitions

To form the scores, it is necessary to aggregate the information contained in an array of measures from the financial statements. The general idea is to transform some accounting ratios into binary signals—usually, 0 for bad news, 1 for good news—which can be aggregated into a single measure. Thus, Piotroski's FSCORE1 is formed by aggregating nine binary signals related to profitability—return on assets, cash flow from operations, change in return on assets and accruals, leverage, liquidity, source of funds—issues of common equity—and operating efficiency—change in margins and in turnovers. The resulting score equals zero (nine) when the firm shows the least (most) favorable set of financial signals. Xue and Zhang's FSCORE2 equals the sum of eleven individual binary signals from financial statements. The resulting score equals zero (eleven) when the firm shows the least (most) favorable set of financial signals. Mohanram's GSCORE1 equals the sum of eight individual binary signals related to earnings and cash flow profitability, naive extrapolation and accounting conservatism. A score of zero (eight) represents the least (most) favorable set of financial signals. Wahlen and Wieland's PEIS equals the sum of six individual binary signals. Each signal equals +1 (–1) if the underlying realization is a good (bad) signal about future firm performance. PEIS equals –6 (6) when the firm shows the least (most) favorable set of fundamental signals. This measure is quite complex because each individual signal is obtained after ranking firms in quintiles each year based on a particular accounting ratio.

3.2 Sample Selection and Composite Calculation

We collect data from the ThomsonOne database for the years 1981–2011. Accounting data come from Worldscope; monthly returns, prices and volumes come from DataStream; and data on analysts come from the IBES database. We consider all firms from 14 European countries[4] with the required data for the years $t-1$, t, and $t+1$ and with the fiscal year ending in December. We compute FSCORE1, FSCORE2, GSCORE and PEIS in the same way as in the original papers, but instead of choosing a sample of high- or low-BM firms, we apply scores utilizing the entire sample. In the case of FSCORE1, we compute FSCORE1a as in Piotroski (2000) and FSCORE1b with two modifications: accruals are computed following Sloan (1996) and cash flows are computed as net income less accruals. Similarly, we compute FSCORE2 twice, obtaining FSCORE2a and FSCORE2b. In the case of GSCORE, we compute GSCORE1a in the same way as in Mohanram (2005), but we replace the missing values of R&D and advertising expenses with 0 due to data restrictions in the European context. The resulting GSCORE1a a ranges from 1 to 8. Additionally, we compute GSCORE1b, a composite of G1 to G5 as in Mohanram, but with G6 indicating a signal of "intangible intensity", defined as the ratio of *other assets*$_t$ to *total assets*$_{t\text{-}1}$. G6 equals 1 if a firm's intangible intensity is higher than the contemporaneous median for all the firms in the same industry and 0 otherwise. Thus, GSCORE1b ranges from 0 to 6, where 0 (6) indicates a negative (positive) outlook of the firm's future performance.

Firm-years with any fundamental signal missing are excluded from the sample (this is the case of financial firms). When the signal is benchmarked against the average value by industry and year, we employ the 49 Fama-French classifications of industries. To avoid industry averages being driven by a small group of firms, we drop industry-years with fewer than four observations. As a result, in our paper we test seven signals—FSCORE1a, FSCORE1b, FSCORE2a, FSCORE2b, GSCORE1a, GSCORE1b and PEIS—based on the four scores documented in the previous literature. The time period covered by each composite varies from one measure to another due to the different data requirements associated with building the scores, but in broad terms the period covers 1989 to 2011.

3.3 Calculation of Returns

As the previous research commonly focused one year ahead, we examine the firms' market-adjusted buy-and-hold returns over a twelve-month window beginning on March 31 after the end of the previous fiscal year. An abnormal return is defined as the firm-specific buy-and-hold raw return less the market (benchmark) return over the same time period. To control by the measurement of the abnormal return, we calculate four measures of market return: abnormal returns with equally weighted returns, with value weighted returns, with Fama and French's size and BM quintiles (25 intersections of five size portfolios and five BM portfolios), and size-decile adjusted returns. The buy-and-hold raw return is based on monthly total returns (adjusted by splits, and including dividends and capital distributions). To avoid data errors, we winsorize returns at the 3% level. Our empirical analyses proceed in two steps. The first step examines whether abnormal returns driven by fundamental signals are present in the main European markets and the second step examines whether the surviving abnormal returns present higher levels of limits to arbitrage.

[4] Austria, Belgium, Germany, Denmark, Finland, France, Greece, Italy, the Netherlands, Norway, Portugal, Spain, Sweden and the United Kingdom.

Table 1 Sample Distribution

Country	FSCORE1a	FSCORE1b	FSCORE2a	FSCORE2b	GSCORE1a	GSCORE1b	PEIS
AUT	481	366	225	201	604	603	398
BEL	873	659	235	191	1,062	1,071	363
DEU	5,264	3,642	2,390	2,352	5,987	6,007	3,833
DNK	1,074	708	439	438	1,499	1,346	655
ESP	705	549	140	111	874	880	188
FIN	1,338	612	135	110	1,434	1,440	539
FRA	4,863	2,225	1,688	1,473	5,361	5,386	1,034
GBR	5,981	4,088	708	603	6,135	6,146	5,064
GRC	1,695	906	362	263	1,671	1,675	1,545
ITA	2,134	1,835	966	901	2,349	2,363	2,160
NLD	1,493	1,172	507	477	1,678	1,679	971
NOR	1,481	863	494	455	1,634	1,548	739
PRT	556	428	133	119	607	611	36
SWE	2,815	1,218	589	508	2,831	2,876	1,402
Total Obs.	**30,753**	**19,271**	**9,011**	**8,202**	**33,726**	**33,631**	**18,927**
Firms	3,456	2,274	1,516	1,413	3,983	3,965	2,632
Score range	[0, 9]	[0, 9]	[0, 11]	[0, 11]	[1, 8]	[0, 6]	[-6, +6]
Sample years	1989–2011	1990–2011	1990–2011	1991–2011	1989–2011	1989–2011	1989–2011

Notes: FSCORE1a takes the same definition as in Piotroski (2000), and FSCORE1b is built with accruals computed as in Sloan (1996); cash flows are computed as net income less accruals. FSCORE2a takes the same definition as in Xue and Zhang (2011), and FSCORE2b is built with accruals computed as in Sloan (1996); cash flows are computed as net income less accruals. GSCORE1a takes the same definition as Mohanram (2005) but with two signals set to zero; GSCORE1b gathers the original three signals G6–G8 into one signal of "intangible intensity". As a result, GSCORE1a takes values from 1 to 8, and GSCORE1b from 0 to 6. PEIS is Wahlen and Wieland's (2011) measure.

4. Empirical Results: Future Returns in European Markets

4.1 Descriptives

Table 1 provides the distribution of the scores by country, number of observations and firms, value ranges and sample years covered. Some measures are easy to compute, but others require more complex calculations such as industry-level figures or quintiles. As a result, we end up with a range of 8,202 observations—1,413 firms in the period 1991–2011 with FSCORE2b—to 33,726 observations—3,983 firms in the period 1989–2011 with GSCORE1a.

To find evidence of whether the measures provide complementary or substitutive information, we perform a correlation analysis. *Table 2* presents correlations between the individual fundamental signals. Pairwise correlations are displayed below the diagonal and Spearman correlations above it. As expected, FSCORE1a and FSCORE1b present a high significant positive correlation. The same occurs with FSCORE2a-FSCORE2b, and GSCORE1a-GSCORE1b. The rest of the correlations are low, indicating that they probably capture different aspects of firm performance.

Table 2 Analysis of Correlation between Scores

	FSCORE1a	FSCORE1b	FSCORE2a	FSCORE2b	GSCORE1a	GSCORE1b	PEIS
FSCORE1a	1	0.9378*	0.2929*	0.2755*	0.3500*	0.3397*	0.2514*
FSCORE1b	0.9314*	1	0.2857*	0.2741*	0.3053*	0.2900*	0.2358*
FSCORE2a	0.3058*	0.2969*	1	0.9399*	0.1460*	0.1275*	0.1440*
FSCORE2b	0.2830*	0.2806*	0.9431*	1	0.1254*	0.1088*	0.1263*
GSCORE1a	0.4292*	0.3348*	0.1910*	0.1707*	1	0.8760*	-0.0307*
GSCORE1b	0.4167*	0.3137*	0.1590*	0.1386*	0.8712*	1	-0.0592*
PEIS	0.2018*	0.2426*	0.1443*	0.1291*	-0.0199*	-0.0661*	1

Note: Pairwise correlations are displayed below the diagonal and Spearman correlations above it.

To test whether scores documented to be useful in the construction of successful strategies in the US are implementable in Europe, we form portfolios to compute one-year-ahead buy-and-hold returns.

4.2 Evidence of Returns by Sorts of Scores

This subsection shows the usefulness of each composite in the identification of winners and losers. To this end, sorts supply a simple picture of how returns vary across the composite levels.

Table 3 presents one-year-ahead buy-and-hold returns by score levels to each fundamental investment strategy for our full sample of European firms. In addition, we compute a hedge strategy that takes a long position in firms with the highest score and a short position in firms with the lowest score on a yearly basis. As in the previous research, a small number of observations obtained extreme scores. Therefore, we form two groups—high and low scores—with a similar number of firm-years and compute the H-L difference between both, presenting the *t*-statistics.

Consistent with US evidence, FSCORE1 discriminates between firms with future strong and weak return performance. Except for the lowest score—note the very high standard deviation—the most striking result is the positive relationship between FSCORE1 and the subsequent returns more than ten years after the publication of Piotroski's paper. Firms with a high FSCORE1 significantly outperform those with a low FSCORE1 in the year following formation of the portfolio (mean ABHR of 0.05 versus –0.22, respectively, in FSCORE1a). The mean return difference of 0.28 is significant at the 1% level. This pattern of returns also extends beyond the mean performance of portfolios. This investment approach shifts the entire distribution of returns, not only in ABHR but also in raw BHR. In the untabulated results, we find that the returns in the 10th percentile, the 25th percentile, the median, the 75th percentile and the 90th percentile of the high FSCORE1a portfolio are higher than the corresponding returns of the low FSCORE1a portfolio. We perform the same type of analysis with FSCORE1b and we reach similar conclusions. Overall, FSCORE1 discriminates between winners and losers.

Different patterns can be observed in Xue and Zhang's (2011) FSCORE2 with no significant H-L portfolio and negative returns in hedge portfolios, although low levels of the score are associated with negative returns and high levels with positive returns. In addition, firms with available data decrease considerably due to the higher

Table 3 Buy-and-Hold Returns (end of March) from year t to year t + 1 based on each Score with 5 Size x 5 BM portfolios

Panel A FSCORE1a, FSCORE1b, FSCORE2a, FSCORE2b

	FSCORE1a					FSCORE1b					FSCORE2a					FSCORE2b				
	Raw	Abnormal BHR				Raw	Abnormal BHR				Raw	Abnormal BHR				Raw	Abnormal BHR			
Levels	Mean	Mean	StD.	Med.	Obs.	Mean	Mean	StD.	Med.	Obs.	Mean	Mean	StD.	Med.	Obs.	Mean	Mean	StD.	Med.	Obs.
0	**0.31**	**0.07**	0.91	-0.08	44	**0.32**	**0.10**	0.74	-0.04	13	**0.31**	**0.17**	0.84	-0.07	89	**0.38**	**0.28**	0.82	0.02	71
1	**-0.03**	**-0.26**	0.75	-0.44	365	**0.06**	**-0.16**	0.79	-0.37	192	**0.12**	**-0.07**	0.66	-0.19	363	**0.16**	**-0.04**	0.71	-0.19	356
2	**0.02**	**-0.22**	0.74	-0.39	1411	**0.12**	**-0.13**	0.76	-0.27	700	0.14	-0.08	0.63	-0.19	825	0.18	-0.04	0.66	-0.17	720
3	0.09	-0.15	0.72	-0.29	3186	0.11	-0.12	0.71	-0.23	1729	0.17	-0.05	0.64	-0.14	1293	0.22	-0.01	0.68	-0.13	1201
4	0.14	-0.10	0.68	-0.21	5300	0.15	-0.08	0.66	-0.19	3155	0.19	-0.04	0.67	-0.15	1759	0.20	-0.03	0.65	-0.13	1561
5	0.18	-0.05	0.65	-0.16	6633	0.20	-0.04	0.63	-0.13	4139	0.21	-0.03	0.66	-0.14	1671	0.23	-0.01	0.68	-0.14	1514
6	0.20	-0.03	0.63	-0.14	6218	0.21	-0.02	0.63	-0.13	4113	0.22	-0.02	0.66	-0.11	1342	0.23	-0.01	0.66	-0.11	1251
7	0.23	-0.01	0.62	-0.12	4657	0.20	-0.03	0.59	-0.12	3241	0.23	-0.01	0.65	-0.12	888	0.26	0.02	0.67	-0.09	827
8	**0.29**	**0.04**	0.63	-0.08	2398	**0.27**	**0.03**	0.62	-0.09	1632	0.30	0.05	0.63	-0.03	465	0.36	0.12	0.69	-0.03	423
9	**0.35**	**0.10**	0.68	-0.05	541	**0.35**	**0.11**	0.68	-0.05	357	**0.30**	**0.06**	0.67	-0.09	216	**0.32**	**0.08**	0.66	-0.02	188
10	**n.d.**					**n.d.**					**0.29**	**0.07**	0.64	-0.08	84	**0.34**	**0.09**	0.60	-0.03	78
11	**n.d.**					**n.d.**					**0.15**	**-0.14**	0.34	-0.17	16	**0.21**	**0.02**	0.25	-0.02	12
Total	0.18	-0.06	0.66	-0.17	30753	0.19	-0.04	0.64	-0.15	19271	0.20	-0.03	0.65	-0.13	9011	0.23	0.00	0.67	-0.12	8202
Hedge		0.14					0.02					-0.02					-0.14			
High		0.05					0.04					0.05					0.08			
Low		-0.22					-0.13					-0.02					0.01			
H-L		**0.28**					**0.17**					**0.07**					**0.07**			
t		13.46***					6.41***					1.46					1.32			

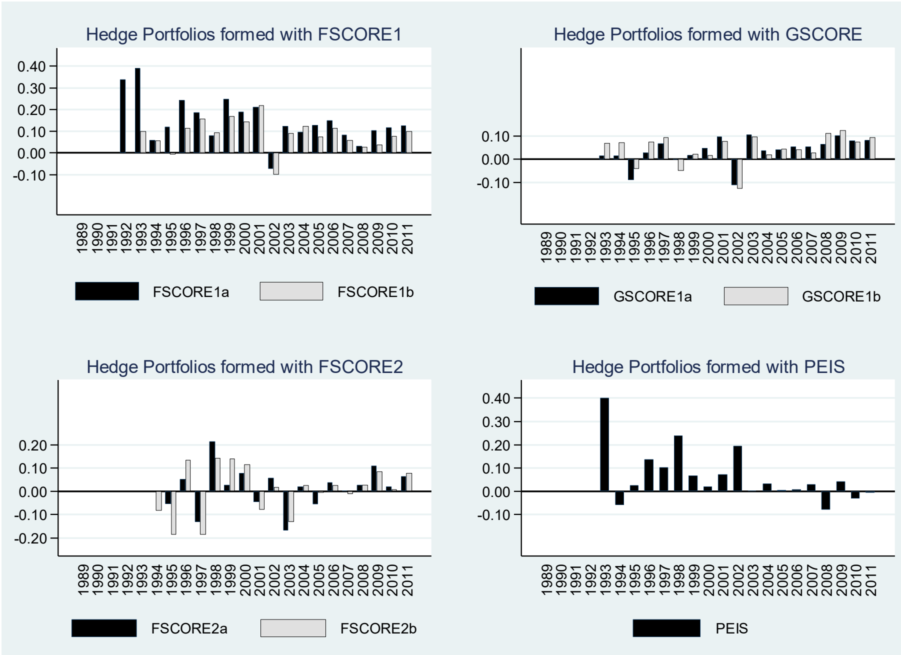
Notes: n.d. means score not defined (out of range). ***, ** and * indicate significance at the 1%, 5% and 10% levels, respectively. Means in bold represent the groups of high and low scores. The hedge portfolio is formed by a long position in the highest score and a short position in the lowest score. H-L represents the difference between the groups of high and low scores.

Panel B GSCORE1a, GSCORE1b, PEIS

	GSCORE1a					GSCORE1b					PEIS				
	Raw	Abnormal BHR				Raw	Abnormal BHR				Raw	Abnormal BHR			
Levels	Mean	Mean	StD.	Med.	Obs.	Mean	Mean	StD.	Med.	Obs.	Mean	Mean	StD.	Med.	Obs.
-6	**n.d.**					**n.d.**									0
-5	**n.d.**					**n.d.**					**0.13**	**-0.12**	0.70	-0.18	26
-4	**n.d.**					**n.d.**					**0.03**	**-0.20**	0.57	-0.21	134
-3	**n.d.**					**n.d.**					**0.14**	**-0.10**	0.70	-0.22	564
-2	**n.d.**					**n.d.**					**0.16**	**-0.09**	0.66	-0.19	1632
-1	**n.d.**					**n.d.**					**0.14**	**-0.09**	0.62	-0.17	3841
0	**n.d.**					**0.00**	**-0.24**	0.67	-0.37	642	0.15	-0.08	0.61	-0.15	5432
1	**-0.04**	**-0.26**	0.74	-0.48	33	**0.09**	**-0.14**	0.70	-0.26	3574	**0.16**	**-0.07**	0.65	-0.17	4092
2	**0.02**	**-0.18**	0.76	-0.33	674	**0.10**	**-0.13**	0.67	-0.22	6316	**0.18**	**-0.05**	0.69	-0.17	2057
3	0.06	-0.16	0.71	-0.29	3036	0.14	-0.08	0.62	-0.16	7432	**0.15**	**-0.11**	0.74	-0.26	806
4	0.11	-0.12	0.66	-0.22	6884	**0.17**	**-0.05**	0.58	-0.13	7157	**0.11**	**-0.16**	0.76	-0.32	261
5	0.14	-0.08	0.62	-0.16	8069	**0.20**	**-0.05**	0.55	-0.11	5769	**0.14**	**-0.15**	0.83	-0.37	69
6	0.18	-0.05	0.57	-0.12	7895	**0.18**	**-0.06**	0.50	-0.10	2741	**0.38**	**0.05**	0.82	-0.10	13
7	0.19	-0.05	0.53	-0.11	5573	**n.d.**					**n.d.**				
8	**0.19**	**-0.05**	0.52	-0.10	1562	**n.d.**					**n.d.**				
	0.15	-0.08	0.61	-0.16	33726	0.14	-0.09	0.61	-0.16	33631	0.15	-0.08	0.65	-0.17	18927
Hedge		0.07					0.17					0.17			
High		-0.05					-0.05					-0.07			
Low		-0.19					-0.14					-0.1			
H - L		**0.14**					**0.09**					**0.03**			
t		4.9***					11***					2.28*			

Notes: ***, ** and * indicate significance at the 1%, 5% and 10% levels, respectively. Means in bold represent the groups of high and low scores. The hedge portfolio is formed by a long position in the highest score and a short position in the lowest score. H-L represents the difference between the groups of high and low scores..

Figure 1 Time Series of ABHR by Hedge Portfolios in Each Score



Note: One-year-ahead abnormal buy-and-hold returns to hedge portfolios based on each score by calendar year taking a long position in firms with high scores and a short position in firms with low scores.

computation requirements associated with building the measure. These results are robust to various specifications of abnormal returns. Better results can be observed in Mohanram's (2005) GSCORE, with the mean and medians shifting the distribution of abnormal returns to the left for lower score portfolios and to the right for higher score portfolios, in line with the US evidence. However, except for the H-L portfolio of approximately 3%, Wahlen and Wieland's (2011) PEIS is unable to discriminate between firms with future strong and weak return performance in our sample of European markets.

4.3 Time-Series Evidence

We examine the robustness of the fundamental strategies based on each score over time. In line with the previous literature and due to the small size of each yearly score subsamples, we classify all firms in two groups according to scores: each year, firms where the majority of the signals are good are assigned to the portfolio of high scores, and firms where the majority of the signals are bad news are assigned to the portfolio of low scores. Then, a hedge portfolio is formed as a long position in high scores and a short position in low scores. *Figure 1* depicts the time series of the abnormal buy-and-hold returns generated by the hedge portfolio of each score. In line with our previous evidence, the strategy is robust across time only in the case of FSCORE1 (19 out of 20 years with positive ABHR) and GSCORE (16 out of 19 years with positive ABHR). FSCORE2 presents positive and negative abnormal returns resulting in yearly an average ABHR close to zero. Finally, PEIS displays a diminishing pattern with remarkably high ABHRs in the 1990s but with ABHRs of approximately zero after 2002.

4.4 Cross-Section Evidence

Fundamental strategies could (i) potentially be influenced by the countries of the sample, (ii) be correlated with other risk factors or anomalies, or (iii) be biased by some years or some time periods. To ensure the robustnes of our findings, following Piotroski (2000) and Mohanram (2005), we estimate Fama-MacBeth (1973) regressions of abnormal buy-and-hold returns in the next year:

$$ABHR_{i,t+1} = \alpha_{0,it} + \alpha_{1,it}SCORE_{it} + \alpha_{2,it}SIZE_{it} + \alpha_{3,it}BM_{it} + \alpha_{4,it}MOM_{it} + \\ + \alpha_{5,it}ACCRD_{it} + \alpha_{6,it}EQOFF_{it} + \sum_{j=1}^{14}\gamma_j C_j + \varepsilon_{it} \quad (2)$$

Controls are similar to those employed by Piotroski and Mohanram. SIZE is the log of market value; BM is the book-to-market; MOM is *momentum* defined as the firm's 12-months-previous market-adjusted return prior to portfolio formation; ACCRD, or *accruals*, equals the firm's total accruals scaled by total assets; EQOFF equals one if the firm raised equity in the previous fiscal year; and C_j is a set of dummies that equals 1 if the firm belongs to the country j and 0 otherwise. Following Piotroski, MOM and ACCRD were replaced with their portfolio decile ranking (1 through 10) based on yearly cutoffs.

Table 4 presents the summary from the Fama-MacBeth regressions with country fixed effects. The first group of regressions includes all years for each score. Except for PEIS, all the scores are significant in some way. If we look at FSCORE1, for each point of increase ABHR increases by aproximately 3%. In the case of GSCORE, each point of increase results in an increase of aproximately 1% of ABHR. In addition, FSCORE2a exhibits an increase of aporximately 0.6% of ABHR with each point of the score increase.

To control by time periods, the remainder of the regressions includes only a particular period of time, demonstrating that FSCORE1 and GSCORE were significant over the past twelve years, whereas FSCORE2 and PEIS were not significant in the subperiods considered, with the exception of PEIS in 2000–2007. These results are in line with our evidence found by employing sorts of scores and time series.

Overall, the findings in this section suggest that part of the previously documented successful strategies to earn abnormal returns in US markets are not implementable in our European sample (PEIS and FSCORE2). Apart from data restrictions on building the measures, our results suggest that previous research conclusions could be addressed by the sample or, alternatively, our sample shows a learning effect of the anomalies. After reviewing four fundamental scores, we find that only the FSCORE1 and GSCORE1 strategies could be profitable.

Table 4 Sources of ABHR Predictability. Fama MacBeth Regressions

MEASURE	SCORE	SIZE	BM	MOM	ACCRD	EQOFF	Intercept	Obs.	R.Sq.
FSCORE1a	0.0355***	-0.00112	0.00295	0.0114	-0.00203	0.0374	-0.390***	28,672	0.257
FSCORE1b	0.0291**	-0.00203	0.0116	0.0394	0.00738	0.0358**	-0.503**	18,082	0.247
FSCORE2a	0.00671*	0.0167***	0.0407**	0.00288	-0.0333	0.0526***	-0.403***	8,294	0.235
FSCORE2b	0.00446	0.0117**	0.0479**	0.00394	-0.0325	0.0429**	-0.330***	7,414	0.195
GSCORE1a	0.00998*	0.0266	0.0214	0.00957	-0.00818	0.0312	-0.651*	32,068	0.253
GSCORE1b	0.0132*	0.0220	0.0269	0.0111	0.00411	0.0334	-0.541	31,970	0.254
PEIS	-0.0382	0.0272	0.00241	0.0317	0.0332	0.170	-0.662**	17,734	0.278
				Subsamples of FSCORE1a					
1990–1999	0.0440	-0.00237	0.0196	0.0265	-0.00796	0.0426	-0.552**	3,353	0.381
2000–2007	0.0296***	-0.01000*	0.00177	0.0115	-0.00609	0.0395*	-0.172	14,529	0.108
2008–2011	0.0371***	0.0165	0.000615	-0.0199	0.0190	0.0311	-0.435**	10,787	0.092
				Subsamples of FSCORE1b					
1990–1999	0.0369	-0.00540	-0.000451	0.0898	0.0624	0.0391	-0.807*	2,542	0.422
2000–2007	0.0235***	-0.00732	0.0297	0.0118	-0.0352	0.0408**	-0.255	8,835	0.120
2008–2011	0.0229*	0.0161	0.00253	-0.0189	-0.0312	0.0184	-0.315	6,705	0.108
				Subsamples of FSCORE2a					
1990–1999	0.00760	0.0388***	0.0467	-0.00489	0.00862	0.0658*	-0.671***	919	0.432
2000–2007	0.00441	-0.000276	0.0457	0.0150	-0.0543	0.0509**	-0.257	4,210	0.124
2008–2011	0.00976	0.0120	0.0201	-0.00783	-0.0645	0.0330	-0.229	3,165	0.112
				Subsamples of FSCORE2b					
1990–1999	-0.00295	0.0259**	0.0618	-0.000692	0.0157	0.0373	-0.468**	991	0.323
2000–2007	0.00778	-0.00160	0.0473	0.0135	-0.0602	0.0530**	-0.231	3,584	0.123
2008–2011	0.0108	0.0135	0.0245	-0.00714	-0.0613	0.0324	-0.285	2,839	0.117
				Subsamples of GSCORE1a					
1990–1999	-0.00550	0.0648	0.0618	0.0159	-0.0268	-0.00656	-1.196	3,455	0.368
2000–2007	0.0148**	-0.00825	0.00519	0.0181	-0.0101	0.0617**	-0.235	16,480	0.107
2008–2011	0.0377***	0.0145	0.000431	-0.0194	0.0357	0.0631**	-0.292*	12,130	0.1
				Subsamples of GSCORE1b					
1990–1999	0.0109	0.0595	0.0597	0.0197	0.00175	-0.00570	-1.139	3,456	0.37
2000–2007	0.0123*	-0.00836*	0.00519	0.0180	-0.0118	0.0647***	-0.0610	16,388	0.105
2008–2011	0.0398***	0.0128	0.00317**	-0.0193	0.0423	0.0672**	-0.364	12,123	0.105
				Subsamples of PEIS					
1990–1999	-0.0924	0.0649	0.00114	0.0602	0.116	0.322	-1.184*	1,924	0.425
2000–2007	0.0119*	-0.00890	0.00465	0.0255	-0.0292*	0.0738**	-0.250	8,449	0.114
2008–2011	-0.00508	0.0214	0.00138	-0.0122	-0.0194	0.0623*	-0.362	7,359	0.095

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Fama-MacBeth regressions of ABHR in each score with controls and country fixed effects.

5. Returns Conditional to the Limits of Arbitrage

In this section, we examine the ABHR of both FSCORE1 and GSCORE conditional to three types of limits of arbitrage. The rational-pricing story (or the "discount-rate" effect) poses that strategies implemented to earn apparent abnormal returns track time-variation in discount rates: they predict returns because they capture information about the risk premium. By contrast, according to the mis-pricing view (inefficient markets), scores predict future returns as prices return to fundamentals. It is assumed that prices do not impound all available information in a timely manner. These explanations encounter the objection that irrationality-induced anomalies would be eliminated by rational arbitrageurs. The argument on the limits of arbitrage counters this objection by asserting that in some circumstances arbitrage is difficult. The argument is testable because it implies stronger financial anomalies when the limits to arbitrage condition diverse aspects of trading.

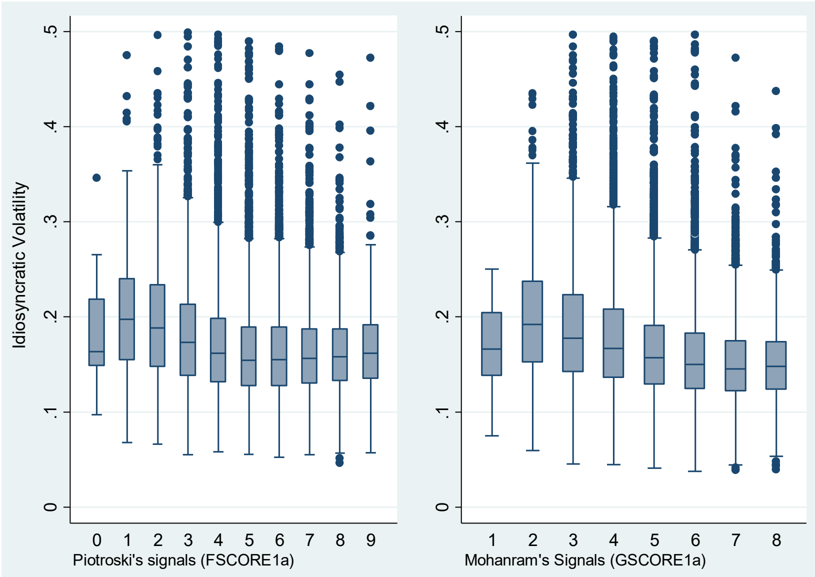
To test the hypothesis that idiosyncratic volatility promotes mispricing, we employ the Fama and French four-factor model as a measure of idiosyncratic risk, and then we sort securities by the size of their residual variability. In this way, we can check whether this financial anomaly is more prominent for higher values of residual variability. To test the hypothesis of mispricing caused by momentum, we form portfolios by utilizing FSCORE1 and GSCORE to explore the possibility that noise traders extrapolate the previous sign of performance into the future (Lakonishok *et al.*, 1994). Under the hypothesis of implementation costs, we explore the evidence with respect to the stock price level, institutional ownership, number of analysts following liquidity and market capitalization.

5.1 Returns Conditional to Idiosyncratic Risk

Our measure of idiosyncratic risk is idiosyncratic volatility. Because it is unobservable and model dependent, we take the approach common in the financial literature. For each surviving anomaly, we estimate a by-firm four-factor regression utilizing monthly return data from the preceding five years, beginning in March 1996. We utilize 60 monthly returns to estimate the regressions. The magnitude of the estimated residual variability after employing a four-factor asset-pricing model in which we include Fama and French's RMRF, SMB and HML factors and a momentum factor, WML, serves as a proxy for idiosyncratic risk. Under this hypothesis, we explore whether financial anomalies increase in absolute terms with the amount of residual variability. *Figure 2* presents our results for FSCORE1a and GSCORE1a, excluding outliers greater than 0.5. An interesting initial result is the negative relationship between the scores and idiosyncratic volatility, with the exception of the lowest score, which comprises a small group of volatility firms.[5] In untabulated results, we form yearly quartiles of idiosyncratic volatility and then we combine each score level with the yearly quartiles. We find that weak scores are dominated by high idiosyncratic volatility firms (quartiles 3 and 4), whereas strong scores are dominated by low idiosyncratic volatility firms (quartiles 1 and 2). In previous sections, we have shown that abnormal returns are dominated by a negative ABHR, while the results in this section indicate that abnormal returns are dominated by weak

[5] In *Table 3*, it can be seen that the number of firms is very low, average and median returns are relatively high, and standard deviations of returns are very high.

Figure 2 Idiosyncratic Volatility of the Firms in each Score of Surviving Anomalies



Notes: For each firm-year, we estimate a four-factor regression employing monthly return data from the preceding 60 monthly returns. The four factors are the Fama and French RMRF, SMB, and HML factors including a momentum factor, WML. The four factors are obtained from Ken French's website for European markets. The data commence in 1991, so the first regression window ends in 1996.

firms with high idiosyncratic risk. The same behavior can be observed in medians as well as in percentiles.

We can conclude that the surviving strategies, FSCORE1a and GSCORE1a, display greater idiosyncratic volatility that makes the levels of mispricing persistent.

5.2 Returns Conditional to Noise Trader Momentum Risk

In this subsection, we test the role of the limits of arbitrage in both undervaluation and overvaluation anomalies. The noise trader momentum risk would induce mispricing persistence, i.e. mispriced firms would have momentum in the direction of the mispricing and arbitrage actions would thus be riskier. If firms are mispriced due to the extrapolation of recent performance, those firms that are weak (strong) and have had recent good (bad) performance have momentum in the direction of the mispricing, and selling these firms would therefore be riskier. According to the fundamental principle of risk and return, riskier firms should be more profitable. Consequently, weak (strong) firms with good (bad) recent performance, or winner-weak (loser-strong), should have higher returns than weak (strong) firms with bad (good) recent performance, or loser-weak (winner-strong).

Thus, we check whether this is true for weak (or low-score) firms when their recent previous results have been bad and for strong (or high-score) firms when recent returns have been good. To test this hypothesis, we form four portfolios that are rebalanced yearly. First, firms are sorted into quintiles based on their pre-formation momentum characteristics, beginning with their buy-and-hold returns

Table 5 Tests of Noise Trader Risk as a Limit to Arbitrage: Intersection of Momentum and Fundamental Strategies based on FSCORE1a and GSCORE1b

	FSCORE1a			GSCORE1a		
Portfolio	**Avg. ABHR**	**Std. Dev.**	**Obs.**	**Avg. ABHR**	**Std. Dev.**	**Obs.**
Loser-Weak	***-0.2760***	***0.7520***	***804***	***-0.3420***	***0.6870***	***209***
Winner-Weak	-0.2790	0.6270	160	-0.1360	0.7640	112
Loser-Strong	0.0449	0.6430	140	-0.0127	0.7100	74
Winner-Strong	***-0.0376***	***0.5490***	***889***	***-0.1200***	***0.4520***	***349***
Total	-0.1470	0.6620	1993	-0.1740	0.6140	744

Notes: Firms are sorted into portfolios based on their pre-formation fundamental score and their momentum characteristics. We begin with the universe of all firms traded on four European markets. We form four portfolios rebalanced yearly. The first portfolio, labeled “Loser-Weak”, holds firms in the bottom momentum quintile and with the lowest fundamental score. The second portfolio, “Winner-Weak”, holds firms in the highest momentum quintile and with the lowest fundamental score. The third portfolio, “Loser-Strong”, holds stocks in the bottom momentum quintile and with the highest fundamental score. The fourth portfolio, “Winner-Strong”, holds stocks in the highest momentum quintile and with the highest fundamental score. We report the resulting portfolio abnormal buy-and-hold returns, standard deviations and number of firm-year observations.

in the previous twelve months. Then we form the four intersections between the two extreme momentum quintiles and the two groups of fundamental scores (lowest and highest) for each signal. Therefore, the first portfolio (loser-weak) holds firms in the bottom momentum quintile and the lowest fundamental score. The second portfolio (winner-weak) holds firms in the highest momentum quintile and the lowest fundamental score. The third portfolio (loser-strong) holds stocks in the bottom momentum quintile and the highest fundamental score. The fourth portfolio (winner-strong) holds stocks in the highest momentum quintile and the highest fundamental score.

Table 5 reports the abnormal buy-and-hold returns, standard deviations and number of firm-year observations for the resulting portfolio. For Mohanram’s (2005) score, we demonstrate that weak firms that are recent winners have higher returns than weak firms that are recent losers and strong firms that are recent losers do better than strong firms that are recent winners; in addition, more profitable portfolios display higher standard deviations, which is consistent with the noise trader momentum risk. Similar patterns can be found for Piotriski’s (2000) FSCORE, with the exception of the weak firms. In broad terms, our results are consistent with the noise trader momentum risk version of the limits of arbitrage argument.

5.3 Returns Conditional to Implementation Costs

Implementation costs limit the capacity of investors to take advantage of mispricing and thus to trade in order to eliminate them. Direct transaction costs include bid-ask spreads and brokerage commissions. To test the limits imposed by direct transaction costs, we calculate two measures. The first one, the share price, is taken because the financial literature finds that quoted bid-ask spreads and commissions are inversely related to share prices. Our second measure is the bid-ask spread, calculated in the manner that is common in the financial literature: first, we compute the monthly (ask – bid) / (0.5 * (bid + ask)), and then we obtain the average value for the previous twelve months, from April of year $t-1$ to March of year t.

Table 6 Implementation Costs as a Limit to Arbitrage

Piotroski's (2000) FSCORE1a									
FSCORE1a	**ABHR**	**Price**	**Bid-Ask**	**Volume**	**Amihud**	**Analysts**	**Insiders**	**Size**	**BM**
0	0.07	2.98	0.0242	69.09	0.0004	1	0.39	10.69	0.427
1	-0.26	2.77	0.0250	88.05	0.0007	2	0.39	10.54	0.415
2	-0.22	3.20	0.0272	71.42	0.0008	2	0.40	10.59	0.496
3	-0.15	3.78	0.0257	66.07	0.0006	2	0.43	10.74	0.537
4	-0.10	4.50	0.0207	100.70	0.0004	4	0.46	11.28	0.599
5	-0.05	5.94	0.0173	139.80	0.0002	4	0.48	11.74	0.610
6	-0.03	6.41	0.0171	138.30	0.0002	4	0.49	11.90	0.595
7	-0.01	6.57	0.0170	148.60	0.0002	5	0.50	12.02	0.604
8	0.04	6.40	0.0188	108.80	0.0003	4	0.52	11.90	0.617
9	0.10	6.99	0.0233	48.91	0.0005	3	0.56	11.57	0.651
Total	**-0.06**	**5.50**	**0.0196**	**112.10**	**0.0003**	**4**	**0.48**	**11.55**	**0.592**
Mohanram's (2005) GSCORE1a									
GSCORE1a	**ABHR**	**Price**	**Bid-Ask**	**Volume**	**Amihud**	**Analysts**	**Insiders**	**Size**	**BM**
1	-0.26	4.64	0.0239	59.00	0.0012	2	0.46	10.82	0.513
2	-0.18	3.37	0.0302	49.06	0.0009	2	0.44	10.65	0.572
3	-0.16	3.23	0.0297	56.02	0.0007	2	0.44	10.69	0.612
4	-0.12	3.70	0.0255	61.99	0.0006	2	0.47	10.91	0.701
5	-0.08	5.27	0.0196	108.80	0.0003	3	0.49	11.60	0.707
6	-0.05	7.61	0.0153	179.40	0.0002	5	0.48	12.22	0.646
7	-0.05	10.73	0.0124	260.50	0.0001	6	0.49	12.68	0.566
8	-0.05	10.39	0.0110	370.30	0.0001	7	0.45	12.84	0.543
Total	-0.08	**6.09**	**0.0185**	**116.50**	**0.0003**	**4**	**0.47**	**11.70**	**0.643**

Notes: This table presents a characterization of each fundamental signal by average ABHR and the medians of transaction costs and investor sophistication. *Price* is the median closing price of a common stock at the end of March of year *t*. *Bid-Ask* is the median percentage bid-ask spread, defined as (ask – – bid)/(0.5*(bid + ask)) averaged over the last trading day of each of the 12 months, beginning in April of year $t-1$ and ending in March of year *t*. *Volume* is the median average volume of trade during the 52 previous weeks in the firm's shares ending in March of year t in thousands of euros. *Amihud* is an illiquidity measure calculated as the median ratio of absolute value of monthly returns scaled by *Volume*. *Analysts* are the median number of analysts' estimates included in the IBES database in March of year *t*. *Insiders* are the median percentage of common stock owned by insiders and investors with more than 5% of the firms' shares at the end of year t – 1. *Size* is the median log of market capitalization in March of year *t*.

Indirect transaction costs imply a delay in the processing of transactions, which prevents news from being priced in a timely manner. To test the limits imposed by indirect transaction costs, we calculate two illiquidity measures. The first, "volume", is the average value of volume in monetary units (thousands of euros) over the previous 52 weeks. Our second measure, inspired by Amihud (2002), is the ratio of the absolute value of the monthly return scaled by "volume". This measure can be interpreted as the price response associated with one euro of trading volume. Amihud proposes that the expected stock excess returns also reflect compensation for the expected market illiquidity, and the returns would thus be an increasing function of the expected market illiquidity. As a consequence, we expect incremental mispricing with the Amihud measure. Additionally, we utilize the number of analysts following the company at the beginning of the portfolio formation as a measure of investor sophistication (Walter, 1997). If there are a large number of analysts, it is likely that investors have access to more information about the company.

Finally, we consider a proxy for short selling. We employ the percentage of "closely held shares" that represent those shares held by insiders because it would be more difficult to do short selling in firms with concentrated ownership (Ali *et al.*, 2003).

Our results (*Table 6*) strongly support the hypothesis that abnormal returns are affected by implementation costs that limit arbitrage. FSCORE1a in Panel A indicates that median prices rise with the score (from less than 3 for signal 0 to 6.99 for signal 9). This is consistent with the intuition that transaction costs make arbitrage difficult and foster some mispricing in weak firms. Our second measure suggests the same direction: low and high FSCORE1a have wider bid-ask spreads than scores in the middle levels. In addition, firms at extreme levels have low volume, are less liquid—larger values of the Amihud measure and are followed by few analysts: with 0 FSCORE1a, the median number of analysts is 1 with a progressive increase until the final scores. However, for score 9, the median value is only 3. Additionally, low score levels have low insider holdings and this variable increases with the levels of the fundamental signal. We expect that firms with low levels of insider holdings may be harder to arbitrage because short selling would be riskier. This is because firms with concentrated ownership are easier to arbitrage because higher proportions of these shares are available to be borrowed and less likely to be subject to a "short squeeze" (Dechow *et al.*, 2001; Ali *et al.*, 2003). Furthermore, as ABHR of FSCORE1a is dominated by large negative returns, hedging these stocks appears to be complicated because the low concentration of ownership limits this possibility. Panel B also reports the presence of transaction costs that limit arbitrage for Mohanram's (2005) GSCORE1. The lowest scores are concentrated in firms with low prices, wide bid-ask spreads, low volume and few analysts.

Our results support the transaction cost explanation as a limit to arbitrage and indicate that mispricing should be greater in this context. In other words, abnormal buy-and-hold returns could be possible because arbitrage is riskier.

6. Conclusions

Piotroski (2000), Xue and Zhang (2011), Mohanram (2005) and Wahlen and Wieland (2011), among others, develop scores based on financial statement analysis that allow the investor to earn abnormal returns. This apparent anomaly was initially documented in US markets. However, if markets are efficient, anomalies should tend to disappear once they have been discovered, either by learning or arbitrage. We present new, out-of-sample evidence related to these anomalies. First, we demonstrate that in four European markets, a hedge strategy that goes long in strong firms (high fundamentals) and short in weak firms (low fundamentals) does not reward investors with one-year-ahead buy-and-hold abnormal returns in two measures, Xue and Zhang's FSCORE2 and Wahlen and Wieland's PEIS, which is consistent with the efficient markets view.

Second, we demonstrate that two surviving anomalies, Piotroski's FSCORE1 and Mohanram's GSCORE, produce hedge portfolios that earn one-year-ahead buy-and-hold abnormal returns. This finding is consistent with the inefficient markets view. Without entering into the rational vs. irrational prices debate, we should note that returns are dominated by few abnormal losses in low scores, i.e. the short strategies of the hedge portfolio.

Third, we try to answer why hedge strategies based on FSCORE1 or GSCORE, which allow one-year-ahead buy-and-hold abnormal returns, are still persistent. Although irrational explanations of anomalies encounter the objection that anomalies would be exploited and eliminated by rational arbitrageurs, the limits of arbitrage argument asserts that idiosyncratic risk, noise trader momentum risk and implementation costs make arbitrage difficult. Under the hypothesis of idiosyncratic risk, we test whether financial anomalies increase in absolute terms with the amount of residual variability. Furthermore, we find that both FSCORE1 and GSCORE exhibit greater idiosyncratic volatility concentrated in firms with weak fundamentals that protect the existence of higher levels of mispricing. Under the hypothesis of noise trader momentum risk, we ask whether momentum might allow more mispricing of weak and strong firms (according to their fundamental scores) when recent returns have been bad (for weak, or low-score, firms) or good (for strong, or high-score, firms), finding support for this limit of arbitrage. Under the hypothesis of implementation costs, we examine whether high costs protect greater levels of mispricing. Employing a set of proxies of direct transaction costs, indirect tranaction costs, short selling risk and investor sophistication, we find that firms labeled with weak fundamentals exhibit greater costs such as low prices, wide bid-ask spreads, low volumes, few analysts and low levels of insiders. It appears that firms with weak fundamentals have high implementation costs that limit the actuation of arbitrageurs.

Taken together, these results support the presence of some anomalies due to the limits of arbitrage, but present a serious challenge for some other anomalies documented in the previous literature.

APPENDIX: Measurement of Scores

Measurement of Piotrosky's (2000) Signals

Signals	Measurement	Indicator Variable
Financial performance signals: Profitability	$ROA_t = \frac{Net\ Income\ Bef\ Ext\ Items_t}{Total\ Assets_{t-1}}$	*F_ROA* = 1 if *ROA* > 0 *F_ROA* = 0 if *ROA* ≤ 0
	$CFO_t = \frac{Cash\ Flow\ From\ Operations_t}{Total\ Assets_{t-1}}$	*F_CFO* = 1 if *CFO* > 0 *F_CFO* = 0 if *CFO* ≤ 0
	$DROA_t = ROA_t - ROA_{t-1}$	*F_DROA* = 1 if *DROA* > 0 *F_DROA* = 0 if *DROA* ≤ 0
	$ACCRUAL_t$ = (*Net Income Before Extraordinary Items*$_t$ – – *Cash Flow from Operations*$_t$) / *Total Assets*$_{t-1}$	*F_ACCRUAL* = 1 if *CFO* > *ROA* *F_ACCRUAL* = 0 if *CFO* ≤ *ROA*
Financial performance signals: Leverage, liquidity, and source of funds	$DLEVER_t = \frac{LT\ Debt_t - LT\ Debt_{t-1}}{\frac{Total\ Assets_t + Total\ Assets_{t-1}}{2}}$	*F_ DLEVER* = 1 if *DLEVER* < 0 *F_ DLEVER* = 0 if *DLEVER* ≥ 0
	$DLIQUID_t = \frac{Current\ Assets_t}{Current\ Liabil_t} - \frac{Current\ Assets_{t-1}}{Current\ Liabil_{t-1}}$	*F_DLIQUID* = 1 if *DLIQUID* > 0 *F_ DLIQUID* = 0 if *DLIQUID* ≤ 0
	Issues of common equity	*EQ_OFFER* = 1 if the firm did not issue common equity in the year preceding portfolio formation, zero otherwise
Financial performance signals: Operating efficiency	$DMARGIN_t = \frac{Sales_t - COGS_t}{Sales_t} - \frac{Sales_{t-1} - COGS_{t-1}}{Sales_{t-1}}$	*F_DMARGIN* = 1 if *DMARGIN* > 0 *F_DMARGIN* = 0 if *DMARGIN* ≤ 0
	$\Delta TURN_t = \frac{Sales_t}{Total\ Assets_{t-1}} - \frac{Sales_{t-1}}{Total\ Assets_{t-2}}$	*F_DTURN* = 1 if *DTURN* > 0 *F_DTURN* = 0 if *DTURN* ≤ 0
Composite score	*FSCORE* = *F_ ROA* + *F_CFO* + *F_ΔROA* + *F_ACCRUAL* + *F_ΔLEVER*+ *F_ΔLIQUID* + *EQ_OFFER* + *F_ΔMARGIN* + *F_ΔTURN*	

Measurement of Xue and Zhang's (2011) Signals

Signals	Measurement	Indicator variable
Financial performance signals: Profitability	Δ*Profit Margin*$_t$—Percentage change in profit margin (defined as income from continuing operations divided by sales) from year t − 1 to year t.	*F_PM* = 1 if Δ*PM* > Industry average of Δ*PM* *F_PM* = 0 otherwise
	ROA$_t$—Return on assets (defined as income from continuing operations divided by the average of the fiscal-year-beginning and fiscal-year-end total assets) in year t.	*F_ROA* = 1 if *ROA* > Industry average of *ROA* *F_ROA* = 0 otherwise
	Δ*ROA*$_t$—Percentage change in ROA from year t − 1 to year t.	*F_DROA* = 1 if Δ*ROA* > Industry average of Δ*ROA* *F_DROA* = 0 otherwise
	Δ*Cash Flows–to–Assets*$_t$—Percentage change in operating cash flows scaled by fiscal year-beginning total assets from year t − 1 to year t. *Note*: For FSCORE2b, we derive cash flows as net income less accruals.	*F_CFO* = 1 if Δ*CFO* > Industry average of Δ*CFO* *F_CFO* = 0 otherwise
	Operating Accruals$_t$ (ACCR)—Total operating accruals in year t. Total operating accruals = ΔAccounts receivable + + ΔInventories + ΔPrepaid expenses − – ΔAccounts payable − ΔTax payable. *Note*: For FSCORE2b, we derive Accruals as Sloan (1996).	*F_ACCR* = 1 if *ACCR* < 0 *F_ACCR* = 0 otherwise
Financial performance signals: Operating efficiency	Δ*Accounts Receivable Turnover*$_t$—Percentage change in accounts receivable turnover ratio (defined as total sales divided by the average of the fiscal-year-beginning and fiscal-year-end accounts receivables) from year t − 1 to year t).	*F_ARTN* = 1 if Δ*ARTN* > Industry average of Δ*ARTN* *F_ARTN* = 0 otherwise
	Δ*Inventory Turnover*$_t$ – Percentage change in inventory turnover ratio (defined as total cost of goods sold divided by the average of the fiscal-year-beginning and fiscal-year-end inventories) from year t − 1 to year t.	*F_INVTN* = 1 if Δ*INVTN* > > Industry average of Δ*INVTN* *F_INVTN* = 0 otherwise
	Δ*Asset Turnover*$_t$—Percentage change in asset turnover ratio (defined as total sales scaled by the average of the fiscal-year-beginning and fiscal-year-end total assets) from year t − 1 to year t.	*F_ASTN* = 1 if Δ*ASTN* > Industry average of Δ*ASTN* *F_ASTN* = 0 otherwise
Financial performance signals: Liquidity	Δ*Current Ratio*$_t$—Percentage change in current ratio (defined as current assets divided by current liability at fiscal year end) from year t − 1 to year t.	*F_CR* = 1 if Δ*CR* > Industry average of Δ*CR* *F_CR* = 0 otherwise
	Δ*Quick Ratio*$_t$—Percentage change in quick ratio (current assets net of inventory and prepaid items divided by current liability at fiscal year end) from year t − 1 to year t.	*F_QKR* = 1 if Δ*QKR* > Industry average of Δ*QKR* *F_ QKR* = 0 otherwise
	Δ*Working Capital*$_t$—Percentage change in net working capital (defined as current assets minus current liabilities at fiscal year end) from year t − 1 to year t.	*F_WC* = 1 if Δ*WC* > Industry average of Δ*WC* *F_WC* = 0 otherwise
Composite score	*FSCORE2* = *F_PM* + *F_ROA* + *F_DROA* + *F_CFO* + *F_ACCR* + *F_ARNT* + + *F_INVTN* + *F_ASTN* + *F_CR* + *F_QKR* + *F_WC*	

Measurement of Mohanram's(2005) Signals

Signals	Measurement	Indicator variable
Category 1: Signals based on Earnings and Cash Flow Profitability	$ROA_t = \dfrac{Net\ Income\ Bef\ Ext\ Items_t}{\dfrac{Total\ Assets_{t-1} + Total\ Assets_{t-1}}{2}}$	$G1$ = 1 if ROA ≥ Ind. Median ROA 0 otherwise
	$CFROA_t = \dfrac{Cash\ Flow\ from\ Operations_t}{\dfrac{Total\ Assets_{t-1} + Total\ Assets_{t-1}}{2}}$	$G2$ = 1 if $CFROA$ ≥ Ind. Median $CFROA$ 0 otherwise
	ACCRUALS	$G3$ = 1 if $CFROA$ ≥ ROA 0 otherwise
Category 2: Signals Related to Naive Extrapolation	$VARROA_t = \sigma^2_{ROA_t} = \dfrac{\sum_{t=0}^{-3}\left(ROA_t - \overline{ROA_t}\right)^2}{4}$	$G4$ = 1 if $VARROA$ ≤ Ind. Median $VARROA$ 0 otherwise
	$SGR_t = \dfrac{Sales_t - Sales_{t-1}}{Sales_{t-1}}$ $VARSGR_t = \sigma^2_{ROA_t} = \dfrac{\sum_{t=0}^{-3}\left(SGR_t - \overline{SGR_t}\right)^2}{4}$	$G5$ = 1 if $VARSGR$ ≤ Ind. Median $VARSGR$ 0 otherwise
Category 3: Signals Related to Accounting Conservatism	$RDINT_t = \dfrac{R\&D_t}{Total\ Assets_{t-1}}$	$G6$ = 1 if $RDINT$ ≥ Ind. Median $RDINT$ 0 otherwise
	$CAPINT_t = \dfrac{CAPEX_t}{Total\ Assets_{t-1}}$	$G7$ = 1 if $CAPINT$ ≥ Ind. Median $CAPINT$ 0 otherwise
	$ADINT_t = \dfrac{Advertising\ Expenses_t}{Total\ Assets_{t-1}}$	$G8$ = 1 if $ADINT$ ≥ Ind. Median $ADINT$ 0 otherwise
Composite score	$GSCORE = G1 + G2 + G3 + G4 + G5 + G6 + G7 + G8$	

Measurement of Wahlen and Wieland's (2011) Signals

Signal	Measure	Information content		Quintile scoring		
				+1	0	-1
RNOA	= Operating Income / / [($NOA_t - NOA_{t-1}$) / 2] *Note*: *NOA*: Shareholder's equity + Short Term Debt + + Long Term Debt + + Preferred Stock – Cash & Short Term Investments + + Other	Mean reversion in earnings when RNOA is extreme		Bottom	Middle	Top
ΔGM	= Δ%Sales – Δ%Cost of goods sold *Note*:Δ% is the rate of change Gross Margin is: GM = Sales – Cost of goods sold	Firm's changing position in input markets relative to output markets		Top	Middle	Bottom
ΔSGA	= (SGA_t / $Sales_t$) – – (SGA_{t-1} / $Sales_{t-1}$) *Note*: SGA: Selling, General and Administrative expense	Changes in operating costs relative to sales	Sales growth	Bottom	Middle	Top
			Sales decline	Top	Middle	Bottom
ΔATO	= ($Sales_t$ / $Total\ Assets_{t-1}$) – – ($Sales_{t-1}$ / $Total\ Assets_{t-2}$)	Changes in the efficiency of the firm's total assets		Top	Middle	Bottom
G^{NOA}	= ($NOA_t - NOA_{t-1}$) / NOA_{t-1}	Changes in operating asset efficiency, conditional on the level of RNOA	Within RNOA quintiles	Bottom	Middle	Top
ACC	= [Operating Income – Cash Flow From Operations] / / Average NOA	Persistence of the accruals component in earnings, controlling for the level of RNOA	Within RNOA quintiles	Bottom	Middle	Top
Composite SCORE	*PEIS = SRNOA + SGM + SSGA + SATO + SGNOA + SACC*					

REFERENCES

Ali A, Hwang L, Trombley MA (2003): Arbitrage risk and the book-to-market anomaly. *Journal of Financial Economics*, 69:355–373.

Barberies N, Thaler R (2003): A Survey of Behavioral Finance. In: *Handbook of the Economics of Finance,* Volume 1B, Financial Markets and Asset Pricing (Eds. Constantinides GM, Harris M, Stulz RM, Chapter 18, pp. 1053–1128.

Bhardwaj RK, Brooks LD (1992): The January Anomaly: Effects of Low Share Price, Transaction Costs, and Bid-Ask Bias. *Journal of Finance*, 47:553–575.

Bhushan R (1991): Trading Costs, Liquidity, and Asset Holdings. *Review of Financial Studies*, 4:343–360.

Brav A, Heaton JB (2002): Competing Theories of Financial Anomalies. *Review of Financial Studies*, 15:575–606.

Brav A, Heaton JB, Li S (2010): The Limits of the Limits of Arbitrage. *Review of Finance*, 14: 157–187.

Choi J-G, Sil S (2010): A Look Back at the 2008 Financial Crisis: The Disconnect between Credit and Market Risks. *Finance a úvěr-Czech Journal of Economics and Finance*, 60(5):400–413.

Cochrane JH (2011): Presidential Address: Discount Rates. *Journal of Finance*, 66:1047–1108.

Dechow PM, Hutton AP, Meulbroek L, Sloan RG (2001): Short-Sellers, Fundamental Analysis, and Stock Returns. *Journal of Financial Economics*, 61:77–106.

Fama EF, French KR (1992): The Cross-Section of Expected Stock Returns. *Journal of Finance*, 47:427–465.

Fama EF, French KR (1993): Common Risk Factors in the Returns on Stocks and Bonds. *Journal of Financial Economics*, 33:3–56.

Fama EF, French KR (1996): Multifactor Explanations of Asset Pricing Anomalies. *Journal of Finance*, 51:55–84.

Frankel R, Lee C (1998): Accounting valuation, market expectation, and cross-sectional stock returns. *Journal of Accounting and Economics*, 25(3):283–319.

Garcia-Blandon J, Martinez-Blasco M, Argiles-Bosch J (2011): Ex-Dividend Day Returns when Dividend and Capital Gains are Taxed at the Same Rate. *Finance a úvěr-Czech Journal of Economics and Finance*, 61(2):140–152.

Kyle AS (1985): Continuous Auctions and Insider Trading. *Econometrica*, 53:1315–1335.

Lakonishok J, Shleifer A, Vishny RW (1994): Contrarian Investment, Extrapolation and Risk. *Journal of Finance*, 49:1541–1578.

Mashruwala C, Rajgopal S, Shevlin T (2006): Why is the Accrual Anomaly not Arbitraged Away? The Role of Idiosyncratic Risk and Transaction Costs. *Journal of Accounting and Economics*, 42: 3–33.

Mohanram PS (2005): Separating Winners from Losers among Low Book-to-Market Stocks using Financial Statement Analysis. *Review of Accounting Studies*, 10:133–170.

Piotroski JD (2000): Value Investing: The Use of Historical Financial Statement Information to Separate Winners from Losers. *Journal of Accounting Research*, 38:1–41.

Piotroski JD (2005): Discussion of 'Separating Winners from Losers among Low Book-to-Market Stocks Using Financial Statement Analysis'. *Review of Accounting Studies*, 10:171–184.

Pontiff J (2006): Costly Arbitrage and the Myth of Idiosyncratic Risk. *Journal of Accounting and Economics*, 42:35–52.

Shleifer A, Vishny RW (1997): The Limits of Arbitrage. *Journal of Finance*, 52:35–55.

Sloan RG (1996): Do Stock Prices Fully Reflect Information in Accruals and Cash Flows About Future Earnings? *Accounting Review*, 71:289–315.

Wahlen J, Wieland M (2011): Can Financial Statement Analysis Beat Consensus Analysts' Recommendations? *Review of Accounting Studies*, 16:89–115.

Walther B (1997): Investor Sophistication and Market Earnings Expectations. *Journal of Accounting Research*, 35(2):157–179.

Wieland M (2011): Identifying Consensus Analysts' Earnings Forecasts that Correctly and Incorrectly Predict an Earnings Increase. *Journal of Business Finance & Accounting*, 38:574–600.

Xue Y, Zhang MH (2011): Fundamental Analysis, Institutional Investment, and Limits to Arbitrage. *Journal of Business Finance & Accounting*, 38:1156–1183.