Finance Research Letters 37 (2020) 101367

Contents lists available at ScienceDirect

Finance Research Letters

journal homepage: www.elsevier.com/locate/frl

# A neural approach to the value investing tool F-Score

Ruth Gimeno, Lidia Lobán, Luis Vicente*

*Department of Accounting and Finance, Facultad de Economía y Empresa. C/ Gran Vía 2, University of Zaragoza, 50005 Zaragoza, Spain*

ARTICLE INFO

*Keywords:*
Value investing
F-Score
Network data envelopment analysis
Financial statement Information

ABSTRACT

This work is the first neural approach to Piotroski's (2000) F-Score. From the same informative signals, our approach based on network data envelopment analysis allows for (1) overcoming the binary perspective of classification between companies with good/bad fundamentals, and (2) appropriately assessing the existing interaction among a company's main financial areas. The analysis of a complete sample of the largest listed companies in the Eurozone and in the U.S. market in the period 2006–2017 shows that our neural F-Score significantly improves the portfolio returns obtained by the original F-Score.

## 1. Introduction

From the seminal works of Rosenberg et al. (1984), Fama and French (1992), and Lakonishok et al. (1994), there is significant empirical evidence of the long-term success of investment strategies in worldwide companies with high book-to-market (BM) ratios (e.g., Fama and French, 2012; Asness et al., 2013). Furthermore, the use of complementary indicators improves the success of these value strategies (e.g., Novy-Marx, 2013).

Piotroski (2000) developed a fundamental score (F-Score) based on accounting signals that differentiated between companies with good and bad fundamental scores among all of those with high BM ratios. Piotroski (2000) established that official financial information was useful for the appropriate selection of these companies because (1) the companies tend to be ignored by analysts, (2) the information that companies voluntarily report to the market lacks credibility given their poor recent performance, and (3) the companies tend to be financially distressed.

Piotroski (2000) showed that an annual excess return of approximately 7.50% could be obtained in the short term with respect to the rest of the companies with high BM ratios in the U.S. stock market during the period 1976–1996. These results had a great impact on both industry and academia.[1]

However, the motivation of our paper is that the F-Score shows some methodological limitations that could affect the reliability of its valuations. The main limitations are as follows:

a) The binary approach of the accounting signals included in the F-Score is too simple to reflect the great variety of financial situations among the companies analyzed. It is not sufficient to determine whether the accounting signals have increased or decreased; rather, it is necessary to consider the magnitude of these variations for better scoring of the companies.

b) No interaction exists between the accounting signals identified in the F-Score; thus, no relationship exists among a company's three main financial areas, as defined by Piotroski (2000): (1) profitability; (2) leverage, liquidity, and source of funds; and (3) operating efficiency.

* Corresponding author.
*E-mail addresses:* rgimeno@unizar.es (R. Gimeno), lloban@unizar.es (L. Lobán), lavicent@unizar.es (L. Vicente).

[1] A large number of subsequent studies compared the F-Score for various markets and time horizons. Richardson et al. (2010) carried out an exhaustive review of applications of fundamental analysis to accounting signals.

https://doi.org/10.1016/j.frl.2019.101367
Received 14 May 2019; Received in revised form 5 November 2019; Accepted 23 November 2019
Available online 29 November 2019
1544-6123/ 

Our neural approach based on network data envelopment analysis (network DEA) allows for overcoming the two previous limitations, thus generating more reliable selections in the short term for companies with high BM ratios. Our neural F-Score keeps the rationale of all of the accounting signals included in the F-Score and additionally considers both their magnitude and the interaction between them. We empirically show that the application of our neural approach to the challenging valuation problem of large-cap companies provides more reliable estimates than the use of binary and isolated signals derived from the original F-Score.

The next section of our work summarizes the F-Score and our neural F-Score. Section 3 presents the empirical analysis. Finally, Section 4 includes the main conclusions of our study.

## 2. The F-Score and our neural F-Score proposal

### 2.1. The F-Score

Piotroski (2000) defined the F-Score as the sum of nine binary accounting signals that can assume a value of 0 or 1. These signals are defined and grouped into a company's three large financial areas. The F-Score takes a maximum value (minimum) equal to 9 (0) when a company presents positive (negative) scoring signals for all of the accounting signals included.

$$\begin{aligned} \mathrm{F-Score} &= \mathrm{ROA} + \Delta\mathrm{ROA} + \mathrm{CFO} + \mathrm{ACCRUAL} \\ &+ \Delta\mathrm{LEVER} + \Delta\mathrm{LIQUID} + \mathrm{EQ\_OFFER} \\ &+ \Delta\mathrm{MARGIN} + \Delta\mathrm{TURN} \end{aligned} \tag{1}$$

Area 1 - Profitability:

*ROA*. Net income before extraordinary items for the fiscal year preceding portfolio formation, scaled by total assets at the beginning of year *t*. This signal is 1 if *ROA* is positive and 0 otherwise.
*ΔROA*. *ROA*for year *t* less the firm's *ROA*for year *t-1*. This signal is 1 if *ΔROA* is positive and 0 otherwise.
*CFO*. Cash flow from operations scaled by total assets at the beginning of year *t*. This signal is 1 if *CFO* is positive and 0 otherwise.
*ACCRUAL*. Net income before extraordinary items less cash flow from operations, scaled by total assets at the beginning of year *t*. This signal is 1 if *ACCRUAL* is negative and 0 otherwise.

Area 2 - Leverage, Liquidity, and Source of Funds:

*ΔLEVER*. Change in the firm's debt-to-asset ratio between the end of year *t* and year *t-1*. The debt-to-asset ratio is defined as the firm's total long-term debt (including the portion of long-term debt classified as current), scaled by average total assets. This signal is 1 if *ΔLEVER* is negative and 0 otherwise.
*ΔLIQUID*. Change in the firm's current ratio between the end of year *t* and year *t-1*. The current ratio is defined as total current assets divided by total current liabilities. This signal is 1 if *ΔLIQUID* is positive and 0 otherwise.
*EQ_OFFER*. This signal is 1 if the firm did not issue common equity in the year preceding portfolio formation and 0 otherwise.

Area 3 - Operatingefficiency:

*ΔMARGIN*. Gross margin (net sales less cost of goods sold) for the year preceding portfolio formation, scaled by net sales for the year, less the firm's gross margin (scaled by net sales) from year *t-1*. This signal is 1 if *ΔMARGIN* is positive and 0 otherwise.
*ΔTURN*. Change in the firm's asset turnover ratio between the end of year *t* and year *t-1*. The asset turnover ratio is defined as net sales scaled by average total assets for the year. This signal is 1 if *ΔTURN* is positive and 0 otherwise.

### 2.2. Our neural approach to the F-Score (NF-Score)

Fig. 1 reflects the structure of our neural F-Score (NF-Score). According to Kao (2014), the NF-Score is a mixed structure, which is the combination of a parallel and a series structure. The NF-Score includes nine signals similar to those of the F-Score, but they are evaluated in terms of the output obtained versus the input consumed. That is, the NF-Score does not consider these signals in binary terms (1,0) but rather assigns them a score as a function of the existing output/input relationship of these signals. These nine signals are grouped in the same three financial areas defined by Piotroski (2000), but our neural approach interconnects them through some of the signals defined as links. These include what is considered to be an output for a signal of one financial area and what is deemed an input for a signal from another area.

After defining our neural structure in Fig. 1, we describe a suitable procedure for modeling it. We work with $n$ companies ($j = 1, \ldots, n$) consisting of nine signals ($k = =1,\ldots,9$). Let $m_k$ and $r_k$ be the numbers of the inputs and outputs to signal $k$, respectively. The link from signal $k$ to signal $h$ is denoted by $(k,h)$, and the set of links by $L$. The inputs of company $j$ at signal $k$ are $\{x_j^k \in \mathrm{R}_+^{m_k}\}$, and the outputs of company $j$ at signal $k$ are $\{y_j^k \in \mathrm{R}_+^{r_k}\}$, where ($j = 1,\ldots,n; k = =1,\ldots,9$). The link variables from signal $k$ to signal $h$ are $\{z_j^{(k,h)} \in \mathrm{R}_+^{t(k,h)}\}$ ($j = 1,\ldots,n$; $(k,h) \in \mathrm{L}$), where $t_{(k,h)}$ is the number of intermediate inputs and outputs in link $(k,h)$. $\lambda^k \in \mathrm{R}_+^n$ is the

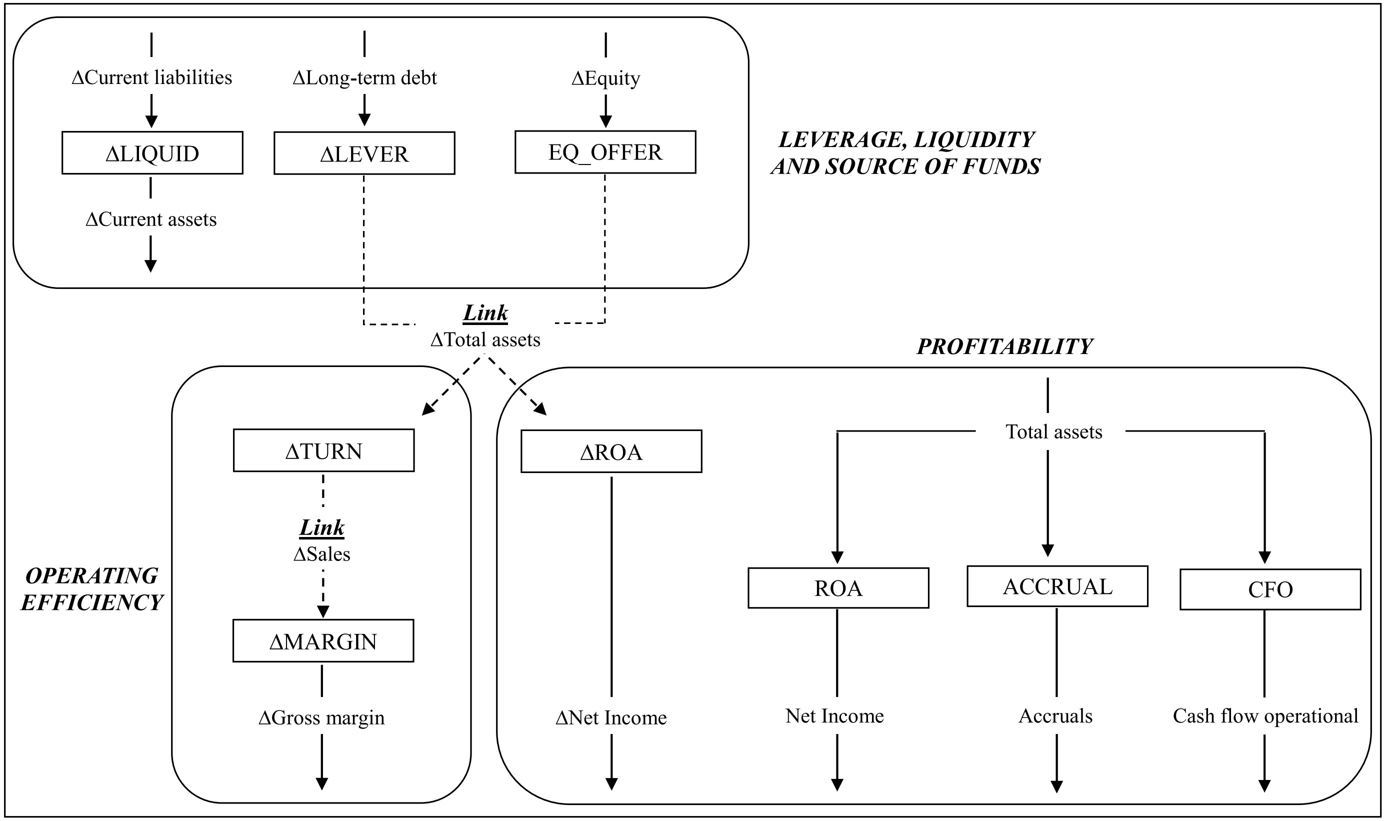


**Fig. 1.** Neural F-Score. The nine accounting signals evaluate a company's three financial areas. The continuous input arrows to each of these signals indicate the input consumed, whereas the continuous output arrows indicate the output obtained. Two links are represented by dashed lines between the various financial areas. *ΔTotal assets* is an output of the signals of *ΔLEVER* and *EQ_OFFER*, and simultaneously, it is an input of *ΔTURN* and *ΔROA*. A third link acts within the area of operational efficiency and links the signals of *ΔTURN* and *ΔMARGIN*.

intensity vector of signal $k$, and $s^{k+}$ and $s^{k-}$ are the nonnegative slack vectors of input excesses and output shortfalls, respectively.[2]

We assume a variable returns-to-scale (VRS) hypothesis because it better evaluates signals when not all of the companies operate at the optimal scale. Thus, production possibility set $P$ is spanned by the following convex hull of the existing companies.

$$P=\{(x^k, y^k, z^{(k,h)}) \mid x^k \geq X^k\lambda^k, \quad y^k \leq Y^k\lambda^k, \quad z^{(k,h)} = Z^{(k,h)}\lambda^k, \quad z^{(k,h)} = Z^{(k,h)}\lambda^h, \; e\lambda^k=1, \quad \lambda^k \geq 0\} \tag{2}$$

Based on the original slacks-based model (SBM)[3] of Tone (2001), we follow one of the network SBM extensions proposed by Tone and Tsutsui (2009) to incorporate the slack vectors of the link variables into the objective function.[4] This extension includes the slacks $s^{(f,k)-}$ of the intermediate input to signal $k$ at link $(f,k)$ and the slacks $s^{(k,h)+}$ of the intermediate output from signal $k$ at link $(k,h)$. The network SBM assumes that the overall fundamental evaluation is the weighted average of the partial fundamental evaluations of each signal, where the weights $w^k$ are set exogenously. For our research purposes, we consider that the nine signals are equally weighted. Finally, our NF-Score model is set as follows:

$$\begin{aligned}
\text{NF}-\text{Score} \;=\; & \min_{\lambda^k,\, s^{k-},\, s^{k+},\, s^{(f,k)-},\, s^{(k,h)+}} \frac{\sum_{k=1}^{K} w^k\left[1-\frac{1}{m_k+\sum_{f\in P_k} t_{(f,k)}}\left(\sum_{i=1}^{m_k}\frac{s_i^{k-}}{x_{io}^k}+\sum_{f\in P_k}\frac{s_f^{(f,k)-}}{z_{fo}^{(f,k)}}\right)\right]}{\sum_{k=1}^{K} w^k\left[1+\frac{1}{r_k+\sum_{h\in F_k} t_{(k,h)}}\left(\sum_{r=1}^{r_k}\frac{s_r^{k+}}{y_{ro}^k}+\sum_{h\in F_k}\frac{s_h^{(k,h)+}}{z_{ho}^{(k,h)}}\right)\right]} \\
& \text{subject to} \\
X^k\lambda^k + s^{k-} \;=\; & x_o^k \qquad Y^k\lambda^k - s^{k+} = y_o^k \qquad e^k = 1 \qquad (k=1,2,\cdots,K) \\
& Z^{(f,k)}\lambda^k + s^{(f,k)-} = z_o^{(f,k)} \qquad Z^{(f,k)}\lambda^f = Z^{(f,k)}\lambda^k \qquad \forall\, f, k \\
& Z^{(k,h)}\lambda^k - s^{(k,h)+} = z_o^{(k,h)} \qquad Z^{(k,h)}\lambda^h = Z^{(k,h)}\lambda^k \qquad \forall\, k, h \\
& \lambda^k, s^{k-}, s^{k+}, s^{(f,k)-}, s^{(k,h)+} \geq 0 \qquad \forall\, f, k, h
\end{aligned} \tag{3}$$

where $P_k$ is the set of signals with the link of $(f,k) \in L$ (antecessor of signal $k$), $t_{(f,k)}$ is the number of intermediate inputs and outputs in that link, $F_k$ is the set of signals with the link of $(k,h) \in L$ (successor of signal $k$), and $t_{(k,h)}$ is the number of intermediate inputs and outputs in that link.

A company will be positively evaluated overall in model [3] when optimal slacks $(s^{k-*}, s^{k+*})$ together with optimal intermediate input and output slacks $(s^{(f,k)-*}, s^{(k,h)+*})$ result in NF-Score $= 1$.

## 3. Empirical analysis

### 3.1. Data

At the end of each year between 2006 and 2017, we identified all of the companies that were listed in euros (dollars) and that were part of the FTSE Eurofirst 300 (Standard & Poor's 500) as the representative benchmarks of the Eurozone and the U.S. stock market, respectively. Following the usual practice in these studies, we excluded the financial sector. Thus, we worked with a complete sample of the largest nonfinancial companies in the world's most economically relevant stock markets. The choice of this challenging sample to validate our neural model is justified by the difficulty of finding successful value strategies in large companies with rapid information-dissemination environments (Piotroski, 2000).

Finally, our sample consisted of 1678 (5018) company observations, which means an average number of 140 (418) large-cap companies analyzed each year in the Eurozone (U.S.) market. The accounting information and the daily historical quotes of each company were obtained from Datastream-Thomson Reuters. We excluded approximately 3.5% of company observations due to the unavailability of data.

### 3.2. Empirical results

We developed an empirical analysis similar to that of Piotroski (2000). For each year, we selected companies with BM ratios higher than the median.[5] The final sample consisted of 922 (2266) observations of the large companies most undervalued in the Eurozone (U.S.) market from 2006–2017.

Next, we selected companies each year with a maximum F-Score of 9 and a minimum F-Score of 0 according to equation [1].[6] Subsequently, we applied our NF-Score model [3] to the same annual sample of companies. The NF values obtained each year were grouped into 10 clusters based on the k-medoids technique. Therefore, each year, we obtained a score for each company assigned by

[2] The inputs and outputs of the various signals that can take positive and negative values are normalized between 0 and 1 with respect to the values of the complete sample following Sánchez-González et al. (2017). This will prevent having to use negative variables in our neural model.

[3] SBM is a nonradial model for measuring efficiency when inputs and outputs may change nonproportionally. The unoriented version of SBM works with excess input and output shortfalls simultaneously.

[4] Section 6.1 of Tone and Tsutsui (2009) includes further details.

[5] Piotroski (2000) used the top quintile because his sample was much larger as a result of not limiting the size of the company.

[6] For those years in which maximum and/or minimum scores were not available for a company, the adjacent scores were taken as a reference.

**Table 1**
Buy-and-Hold Returns to a value investment strategy based on F-Scores and NF-Scores, 2006–2017 †.

Panel A: Buy-and-Hold Annual Compound Returns to a Value Investment Strategy Based on F-Scores

| | | | | | |
|---|---|---|---|---|---|
| Eurozone Market | High F-Score | Low F-Score | All ExHigh F-Score | High F_Low F | High F_All ExHigh |
| One-Year Return | 5.0169% | 5.8947% | 7.9068% | −0.8290% | −2.6788% |
| Two-Year Return | 5.1985% | 8.7524% | 9.1882% | −3.2689% | −3.6552% |
| U.S. Market | High F-Score | Low F-Score | All ExHigh F-Score | High F_Low F | High F_All ExHigh |
| One-Year Return | 5.7532% | 12.8318% | 10.5673% | −7.0786% | −4.8140% |
| Two-Year Return | 6.8250% | 16.2813% | 11.1587% | −9.4563% | −4.3338% |

Panel B: Buy-and-Hold Annual Compound Returns to a Value Investment Strategy Based on NF-Scores

| | | | | | |
|---|---|---|---|---|---|
| Eurozone Market | High NF-Score | Low NF-Score | All ExHigh NF-Score | High NF_Low NF | High NF_All ExHigh |
| One-Year Return | 8.7250% | 5.3143% | 7.9719% | 3.2392% | 0.6977% |
| Two-Year Return | 8.7683% | 7.8032% | 9.1742% | 0.8955% | −0.3719% |
| U.S. Market | High NF-Score | Low NF-Score | All ExHigh NF-Score | High NF_Low NF | High NF_All ExHigh |
| One-Year Return | 9.9174% | 9.3705% | 11.3331% | 0.5468% | −1.4157% |
| Two-Year Return | 10.0367% | 10.5756% | 11.4793% | −0.5390% | −1.4426% |

Panel C: Comparison between Buy-and-Hold Returns Based on F-Scores and NF-Scores (Annual Compound Excess Returns)

| High NF_High F (One-Year Excess Returns) | Low NF_Low F (One-Year Excess Returns) | High NF_High F (Two-Year Excess Returns) | Low NF_Low F (Two-Year Excess Returns) | |
|---|---|---|---|---|
| Eurozone Market | 3.5315%** | −0.5482% | 3.3941%** | −0.8731% |
| U.S. Market | 4.1641%** | −3.4612%** | 3.4239%** | −6.6960%** |

The annual compound excess returns provided in Panel C for the entire period of 2006–2017 have been tested using a similar bootstrapping approach from that in Piotroski (2000). Details are available upon request.

† For one-year returns (two-year returns), the analysis starts at the beginning of April 2018 ends in March 2019 (is not considered). ***significance at the 1% level.

** significance at the 5% level. *significance at the 10% level.

the original F-Score of Piotroski (2000) between 0 and 9. Likewise, we also obtained an NF-Score between 0 and 9, which refers to the 10 clusters obtained from our neural model.

Next, we obtained the equally weighted returns of the portfolios formed by the best and worst rated companies by both models. These returns were calculated as one- and two-year Buy-and-Hold Returns earned from the beginning of the second quarter of the year following the selection of the companies. Similarly, the equally weighted returns of the portfolios formed by all of the companies, excluding the highest rated companies, were computed.

Table 1 shows the main results. Panel A significantly questions the reliability of the selection made by the F-Score for both the Eurozone and U.S. markets, especially for a two-year investment horizon. For 2006–2017, companies with high F-Scores obtained negative excess returns compared with the rest of the sample. Our finding confirmed the size effect already shown by Piotroski (2000), thereby questioning again the successful application of the F-Score in large-cap companies that are extensively followed by worldwide analysts.

However, Panel B shows that the NF-Score reliably selects high-BM companies, especially in the Eurozone market, where the portfolios formed by companies with good fundamentals obtained a 3.24% annual excess return at one year compared with companies with poor fundamentals. However, these results are much less successful for the U.S. market, where the rapid information-dissemination environment seems to severely restrict the utility of the analysis of the nine accounting signals (Piotroski, 2000) with respect to the Eurozone. This finding could be explained by the higher levels of analysts' coverage of large-cap companies in the U.S. market than in the Eurozone. Bolliger (2004) finds that forecast accuracy is negatively associated with the number of countries covered by the analysts. In addition, the significant number of chartered analysts in the U.S. market (CFA, 2018) together with local analyst advantages (Bae et al., 2008) question the importance of the analysis of basic accounting signals in U.S. large-cap companies.

Finally, Panel C shows the best selection skills of the NF-Score with respect to the original F-Score. The company selection with good fundamentals from our NF-Score significantly exceeds the returns obtained by the F-Score selection in both the Eurozone and U.S. markets. This better selection is also evident, especially in the U.S. market, in which the worst fundamental companies chosen by the NF-Score obtained significantly lower returns than those selected by the F-Score. All of the results are robust for one- and two-year investment periods.

The interaction between the financial areas of a company through our NF-Score allowed us to obtain short-term annual excess returns of approximately 4% in long-short value strategies over large-cap Eurozone companies that were rated by the binary F-Score model. This improvement is even more relevant in the U.S. market, with a range of annual excess returns from 7.62% to 10.11% for one- and two-year investment periods, respectively.

## 4. Conclusion

This work is the first neural approach to the F-Score model proposed by Piotroski (2000). Using the same accounting signals, our model allows for (1) overcoming the binary classification perspective between companies with good/bad fundamentals and (2) appropriately assessing the existing interaction among a company's main financial areas.

The empirical application in a complete sample of the largest nonfinancial companies in the Eurozone and the U.S. market during the period 2006–2017 showed that our NF-Score significantly improves the short-term returns of long-short value strategies selected by the F-Score. However, the nine accounting signals proposed by Piotroski (2000) are not sufficient to get winner returns for the U.S. large-cap companies. The first implication of these results is that the interaction between the financial areas in large-cap companies should be considered in further value strategies to obtain higher levels of performance. Second, further neural valuation models should include more sophisticated accounting signals to identify winner value strategies in rapid information-dissemination environments with high levels of analysts' coverage. Both implications define the avenues for future research in other markets and investment strategies.

## 4. Funding

This work was supported by Ministerio de Ciencia, Innovación y Universidades (MCIU), la Agencia Estatal de Investigación (AEI) y el Fondo Europeo de Desarrollo Regional (FEDER) [RTI2018-093483-B-I00]; Gobierno de Aragón [CIBER S38_17R]. Ruth Gimeno thanks the specific support provided by Gobierno de Aragón [IIU/1/2017]. Lidia Lobán thanks the specific support provided by Ministerio de Ciencia, Innovación y Universidades [FPU16/03779]. We really thank all the insightful comments and suggestions of the review process.

## References

Asness, C.S., Moskowitz, T.J., Pedersen, L.H., 2013. Value and momentum everywhere. J. Financ. 68, 929–985.
Bae, K.H., Stulz, R.M., Tan, H., 2008. Do local analysts know more? A cross-country study of the performance of local analysts and foreign analysts. J. Financ. Econ. 88, 581–606.
Bolliger, G., 2004. The characteristics of individual analysts' forecasts in Europe. J. Bank. Financ. 28, 2283–2309.
CFA, 2018. Annual Report. Fiscal Year 2018. . https://www.cfainstitute.org/-/media/documents/corporate-record/annual-report-2018.ashx (accessed 4 November 2019).
Fama, E., French, K., 1992. The cross-section of expected stock returns. J. Financ. 47, 427–465.
Fama, E., French, K., 2012. Size, value and momentum in international stock returns. J. Financ. Econ. 105, 457–472.
Kao, C., 2014. Network data envelopment analysis: a review. Eur. J. Oper. Res. 239, 1–16.
Lakonishok, J., Shleifer, A., Vishny, R., 1994. Contrarian investment, extrapolation, and risk. J. Financ. 44, 1541–1578.
Novy-Marx, R., 2013. The other side of value: the gross profitability premium. J. Financ. Econ. 108, 1–28.
Piotroski, J., 2000. Value investing: the use of historical financial statement information to separate winners from losers. J. Account. Res. 38, 1–41.
Richardson, S., Tuna, I., Wysocki, P., 2010. Accounting anomalies and fundamental analysis: a review of recent research advances. J. Account. Econ. 50 (2–3), 410–454.
Rosenberg, B., Reid, K., Lanstein, R., 1984. Persuasive evidence of market inefficiency. J. Portf. Manag. 11, 9–17.
Sánchez-González, C., Sarto, J.L., Vicente, L., 2017. The efficiency of mutual fund companies: evidence from an innovative network SBM approach. Omega. Int. J. Manag. Sci. 71, 114–128.
Tone, K., 2001. A slacks-based measure of efficiency in data envelopment analysis. Eur. J. Oper. Res. 130 (3), 498–509.
Tone, K., Tsutsui, M., 2009. Network DEA: a slacks-based measure approach. Eur. J. Oper. Res. 197, 243–252.