Communities and regularities in the behavior of investment fund managers

Andrea Flori,a Fabio Pammolli,a,b Sergey V. Buldyrev,a,c,t Luca Regis,d and H. Eugene Stanley*b,e,f

*aDepartment of Management, Economics and Industrial Engineering, Polytechnic University of Milan, 20156 Milan, Italy; bCenter for Analysis, Decisions, and Society, Human Technopole, 20156 Milan, Italy; cDepartment of Physics, Yeshiva University, New York, NY 10033; dDepartment of Economics and Statistics, University of Siena, 53100 Siena, Italy; tCenter for Polymer Studies, Boston University, Boston, MA 02115; and fDepartment of Physics, Boston University, Boston, MA 02115

Contributed by H. Eugene Stanley, November 7, 2018 (sent for review March 14, 2018; reviewed by Xavier Gabaix and Alessandro Vespignani)

We analyze a large microlevel dataset on the full daily portfolio holdings and exposures of 22 complex investment funds to shed light on the behavior of professional investment fund managers. We introduce a set of quantitative attributes that capture essential distinctive features of manager allocation strategies and behaviors. These characteristics include turnover, attitude toward hedging, portfolio concentration, and reaction to external events, such as changes in market conditions and flows of funds. We find the existence and stability of three main investment attitude profiles: conservative, reactive, and proactive. The conservative profile shows low turnover and resilience against external shocks; the reactive one is more prone to respond to market condition changes; and members of the proactive profile frequently adjust their portfolio allocations, but their behavior is less affected by market conditions. We find that exogenous shocks temporarily alter this configuration, but communities return to their original state once these external shocks have been absorbed and their effects vanish.

Against that background, we examine a microlevel dataset of complex portfolios, use metrics overlooked in previous studies, and construct a vector of behavioral attributes to describe manager investment decisions. These attributes are synthetic measures derived from portfolio holdings and their dynamic adjustments. Differently from the literature that focuses on performance determinants (21–24), this paper studies the co-occurrences of behavioral traits to determine whether professional investors differ are similar, in terms of trading intensity, derivative exposures, response to changes in market conditions, and position concentration, as well as sectoral, asset-type–based, geographical, and market-based portfolio compositions. We focus on managers’ attitudes toward risk and uncertainty by examining the role of derivatives when hedging, the use of liquidity as a buffer when calibrating asset allocation, the response to market instability, and the net variation of assets under management due to the issuance or redemption of fund shares. Our detailed microlevel dataset constitutes an ideal setting to unravel how professional investors behave and react to macroevents.

To identify communities with homogeneous behavioral features, we apply a hierarchical clustering algorithm and find that community membership is stable. Our analysis detects the existence of four persistent communities shaped by three main behavioral profiles. The analysis of performances across and within communities reveals no particular pattern and does show an orthogonality between different investment attitudes and performances. We also find that community formation is not related to the size distribution of funds.

Our analysis of community stability uncovers two aspects that confirm that mental models, beliefs, and routines shape expert investors’ decision making. First, the composition of the

Significance

This paper relies on a unique database of fund managers’ holdings to map their behavior across asset classes. We unravel the existence of stable and persistent communities. This paper characterizes three different main behavioral attitudes: conservative, reactive, and proactive. Macroeconomic shocks temporarily perturb the configuration of the system, altering the differences between communities. This paper represents a significant step forward in understanding how heuristics, attitudes, and routines shape the behavior of expert investors. It opens a research trajectory in the analysis of behavioral interdependencies in financial markets.

Author contributions: A.F., F.P., S.V.B., L.R., and H.E.S. designed research, performed research, contributed new analytic tools, analyzed data, and wrote the paper.

Reviewers: X.G., Harvard University; and A.V., Northeastern University.

The authors declare no conflict of interest.

Published under the PNAS license.

1To whom correspondence may be addressed. Email: buldyrev@yu.edu or hes@bu.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1802976116/-/DCSupplemental.
communities tends to be stable, thus indicating that community membership is characterized by distinctive and persistent traits. Second, although communities temporarily dissolve when facing an exogenous shock, they then return to their original configuration.

Data and Methodology
Portfolio composition indicators used by prior literature on investment strategies (25–27) include standard fund classification characteristics constructed from publicly available data. Here, we use additional information on fund manager behavior, usually not publicly available on a daily basis.

Our dataset provides the portfolio allocations of 22 flexible open-ended funds from an asset management company, for which we were able to gather rich and reliable daily data. Fund identities are kept anonymous and denoted as idxx, where xx ranges from 1 to 22. Data are from 2015. The funds have different sizes, with assets under management ranging from a few million to more than 2 billion Euros. Portfolios include over 4,000 constituents with issuers belonging to ~70 geographical regions. Because these data are available at a daily frequency, they allow a closer scrutiny of management actions relative to publicly disclosed data. For each day, data include the full list of end-of-day portfolio constituents, their market values, prices, quantities, exposures, and registry information. The constituents are stocks, bonds, and derivatives. Each position is classified according to asset class, market, sector, and geographical location of the issuer. Data also include daily fund returns and the total values of the assets under management. Funds invest in a wide range of instruments, geographical areas, and sectors and are flexible in their allocation strategies. Thus, the dataset allows us to investigate a comprehensive set of different investment choices.

For each fund $i$, we construct a daily vector $x_i(t)$ of synthetic indicators that characterize the investment choices of a fund manager. We use these attributes to map trading intensity, exposure to derivative positions, approach to stock selection and asset diversification, and response to exogenous factors such as market instability and liquidity injections. For each fund, the vector $x$ is thus formed by measures of both portfolio composition and manager response to external signals, as follows: 10 attributes related to portfolio composition (we obtain these 10 indicators by applying a principal component analysis to 33 categories that indicate market, geographical, sectoral, and asset class); the turnover index (TI), which is the ratio of the market value of trades in one day to the value of fund assets under management (it measures therefore the manager’s intensity of trading); the hedging coefficient (HC), which indicates whether equity derivatives are used for hedging purposes or not; the Herfindahl–Hirschman index (HHI), which quantifies the investment concentration or diversification among equity, corporate bond, and government bond markets; the correlation between the TI and the (lagged) Chicago Board Options Exchange volatility index (VIX), which measures the manager response to changes in the market volatility level through variations in the trading intensity; and the correlation between the TI and (lagged) net flows, which measures manager reaction to changes in liquidity when retail investors decide to invest in or redeem fund shares.

We estimate correlations using 6-mo rolling windows. Thus, we use the last 6 mo of our dataset to detect communities and analyze their stability. To smooth the estimates and limit potential noise in daily observations, all measures at time $t$ are averages of their values across the preceding 10 d. Results are robust across different averaging window levels, ranging from 5 d to 15 d. (SI Appendix, Tables S1–S5 summarize the descriptive statistics of the indicators for the funds in our sample.)

For each date, from July 1, 2015 to December 30, 2015, we construct a network, whose nodes are the funds. Our objective is to detect the partition of the nodes that best represents the network structure, i.e., to properly identify funds that behave similarly in a given day. In total, we have 128 dates, corresponding to 128 network configurations. An alternative approach would be to treat our network as a multilayer network, as in ref. 28. The vectors $x_i(t)$ provide information that allows us to identify commonalities in fund managers’ behaviors and to cluster the funds accordingly. To measure the degree of similarity between funds, we compute the cosine similarities between their vectors of attributes. Although clustering methods for signed networks have been extensively used (for instance, refs. 29 and 30), we decided to apply a preserving transformation that turns the cosine similarity into a metric, which both assigns more weight to more similar nodes and avoids negative edges. We denote the similarity matrix as $SM(t)$, whose elements $SM_{ij}(t)$ are

$$SM_{ij}(t) = 1 - \sqrt{0.5(1 - CS(x_i(t), x_j(t)))}, \quad [1]$$

where $CS(x_i(t), x_j(t))$ indicates the cosine similarity between the vectors $x_i$ and $x_j$ at time $t$. $SM_{ij}(t) \in [0, 1]$ measures, indeed, the degree of similarity between the two funds. Then, we apply the Louvain clustering algorithm (31) to the daily matrices $SM(t)$ and obtain, for each of the 128 dates, the clusters of similar funds. The detection of the partitions is thus performed by maximizing the modularity, a measure that quantifies the strength of a partition in a system (32). The higher the modularity, the denser are the connections between members belonging to the same community, and the sparser are the links between members of different communities. We follow ref. 33 to remove redundant links; we refer the interested reader to SI Appendix for further details.

Each date constitutes a different network, since the nodes, i.e., the funds, are always the same, while weights change. Daily configurations embed high-frequency information and are thus informative, but they can be affected by market noise that influences investment behavior. It would have been possible to aggregate some information and compute the similarity matrix at a lower frequency or averaging the weights connecting each node over some dates to reduce the number of network configurations. However, longer time windows would have generated an over-smoothing effect. As a consequence, we decided to focus on daily networks. Given our choice, the stability of the different daily communities detected across the entire period of observation is an issue. We therefore identify communities that can be considered as persistent across the 128 d and that identify the groups of funds that behave similarly throughout the whole sample period. In practice, we examine the daily configurations, find co-occurrences in time among funds’ community members, and select communities that $(i)$ have a higher number of persistent memberships over the sample period and $(ii)$ are stable. Other papers dealt with the problem of identifying persistent (robust) communities and analyze their stability and properties through time. Ref. 34, for instance, identifies clusters of exchange rates and discusses their persistence across time from 1991 to 2008. More precisely, we adopt the following procedure.

We calculate the matrix of intersections $M_{ij}$, which quantifies the number of funds in daily community $i$ present in persistent community $j$. We next arrange the elements $M_{ij}$ in descending order $M_{i1,j1} \geq M_{i2,j2} \geq M_{i3,j3} \geq \ldots$ and identify $i_1$ with $j_1$. When $i_1 = i_k$, we skip element $k$ in the list until we find $i_{k} \neq i_1$ and $j_k \neq j_1$ and we identify them with each other. We continue to scan the list, ignoring communities that have been already identified, until we find list $i_1(t), j_1(t); i_2(t), j_2(t); \ldots i_n(t), j_n(t)$, which identifies persistent communities with daily communities that exist on day $t$. For each day we
define the size $S_i(t)$ of persistent community $i$ to be the number of funds in the daily community identified with it. We define the daily core of persistent community $i$ to be the number of funds held by the persistent community that are also present in the daily community identified with it.

**Results**

**Identification of Communities of Experts.** We introduce an indicator that measures how often funds are assigned to the same community in time. The level of cohesiveness of a certain community $g_i$, i.e., $\Gamma_{g_i}$, is

$$\Gamma_{g_i} = \frac{\sum_{i,j} F_{ij}}{n^2 - n},$$

where $F_{ij}$ is the frequency co-occurrence percentage of funds $i$ and $j$ in the same fund community, and $n$ is the number of funds in that community. Thus, a homogeneous community will have a cohesiveness indicator that approaches 1.

Fig. 1 shows co-occurrences among fund pairs in the second half of 2015. The dark green cells are fund pairs more frequently belonging to the same community, and lighter green cells are fund pairs less frequently belonging to the same community. Using the analysis of the more frequent co-occurrences, we identify four persistent communities. The largest ($C_3$) consists of seven funds, community $C_4$ consists of six funds, community $C_1$ consists of five funds, and community $C_2$ consists of three funds. One fund is a separate singleton community ($C_0$) for the entire period. Note that these four communities collapse into two larger aggregates when we observe the system is less granular. The identified persistent communities are consistent across time windows, and our daily network snapshots allow us to capture behavioral signals otherwise over-smoothed in wider intervals. Communities $C_1$ and $C_4$ are stable in time and extremely cohesive (with values above 0.85). Communities $C_2$ and $C_3$ are slightly more volatile, with cohesiveness values of $\sim 0.60$ and 0.70, respectively, although, on average, their core members are stable. (SI Appendix, Table S8 reports the cohesiveness values for each community averaged over the entire sample period. Discarding id13 and/or id15 in $C_2$ or id2 and/or id10 in $C_3$ significantly increases their cohesiveness levels.)

**Stability Analysis.** Fig. 2 shows the evolution of the size and core of each community as a function of time. In the first 75 d the communities remain relatively stable. At day 75 (corresponding to September 8, 2015) their sizes and cores begin to fluctuate, indicating a change in manager behavior. On days 90 and 91 (November 3, 2015 and November 4, 2015) community $C_1$ merges with $C_5$ and community $C_3$ merges with $C_4$. In Discussion, we connect these substantial changes in communities’ configuration to exogenous shocks, caused by major financial and political events that occurred in the second half of 2015. While these exogenous shocks in the autumn of 2015 may have pushed some managers to temporarily adopt a different behavior, the original set of communities returns at the end of 2015.

Over time, some funds never change their persistent community, while others switch from one community to another. We define the loyalty of a fund to a persistent community as the percentage of observations in which the fund belongs to the community. Note that the sum of the loyalties of a fund is not always 1 because on some days it may be assigned to a daily community not identified with any persistent community. The loyalty of funds to their persistent communities is always greater than 0.5, while the average stability of a persistent community, defined as the average loyalties of its constituent funds (SI Appendix, Table S9), is greater than 0.7. The only exception is id15, which is 0.44. Fund id5 switches from $C_1$ to $C_2$, spending 30% of its time in $C_1$ and 54% of its time in $C_2$.

**Behavioral Communities’ Features.** We summarize the characteristics of the four persistent communities we have identified by examining the average daily values of the vector components. SI Appendix, Table S6 lists these averages and their SDs. Often, the attributes linked to portfolio composition alone, although important, do not clearly characterize a community. Marked differences between communities emerge instead when we consider the whole set of indicators.

Funds in communities $C_1$ and $C_4$ adjust their allocations less frequently and display lower TI values, but those in $C_2$ and $C_3$ display a more volatile portfolio allocation behavior. Funds in $C_1$ and $C_4$ are less sensitive to net flow dynamics and rely less on liquidity as a buffer to stabilize portfolios. In contrast, funds in $C_2$ and $C_3$ trade more frequently when faced with additional liquidity. Other indicators point to marked differences among the members of pairs $C_1 - C_4$ and $C_2 - C_3$. The HC is relatively high in $C_4$, although its members have on average a low equity exposure, but funds in $C_3$ with a similar level of equity exposure have a very low average HC. Funds in $C_1$ have minimal HC despite a consistent equity position. In contrast, $C_4$ has an average portfolio composition similar to $C_1$ but very high HC. This is due more to manager investment attitude than to sector type or geographical market. Similarly, the HHIs indicate diversified or concentrated investments in similar portfolio compositions, dependent on the asset class composition. Finally, funds in $C_4$ respond to changes in market volatility by adjusting their positions, while investments in the other communities seem less sensitive to market dynamics.

Note that funds belonging to different behavioral classes differ in some ways and not in others. This confirms the importance of our identification strategy, which builds upon granular, multidimensional, data. Although portfolio composition is an important feature that characterizes funds, our analysis highlights that more weight should be given to fund manager behavior. To find whether the communities we identified are distinct, we apply nonparametric tests to the distributions of behavioral indicators. The results are presented in SI Appendix, Table S7. We use the Kruskal–Wallis nonparametric equality-of-medians test to verify whether at least two communities have differing median values for each feature. Results indicate that this is the case.
for the majority of the medians. The Dunn posthoc multiple-pairwise comparison test also supports the presence of distinct communities.

Discussion

Fig. 3 shows the normalized average values of the attributes of each community. While portfolio composition values display few notable differences among the four communities, the other attributes indicate peculiar patterns among them. This confirms that our approach is able to better capture heterogeneity in investment manager behavior and provide richer information about the allocation decision process.

We do not find statistically relevant differences in the performances of our communities, which emerge independently of market results. Our technique departs from previous analyses that identify similarities by examining the relation between funds’ extra returns and the performance of specific portfolios. We propose a taxonomy for the communities we detect. Community $C_1$ has low values for portfolio TI and high levels of resilience against external signals. Thus, its fund members are assigned a “conservative” profile. In contrast, funds in $C_4$ are more prone to respond against changes to market conditions, thus have high values of correlation between TI and VIX, more often use derivatives for hedging purposes, thus have high HC values, and therefore are assigned a “reactive” profile. Finally, communities $C_2$ and $C_3$ often change their portfolio allocations, showing high TI values and positive correlation between trading intensity and net flows. We assign a “proactive” profile to both of them, even if their concentration/diversification strategies differ.

SI Appendix, Fig. S3 shows that all communities, apart from the reactive one, have an inelastic relation between daily stock trades and returns. In other words, on average, they do not react differently to positive or negative price swings. This suggests that they are playing a somewhat “stabilizing” role for the stocks they hold. $C_4$ shows instead a slightly negative relation between stock returns and holding changes, highlighting that members of this community tend to buy (sell) when prices go down (up), behaving as negative feedback traders in this particular period.

Interestingly enough, our analysis enables us to show that behavioral attitudes are influenced by exogenous shocks. Fig. 2 shows that around October 13 communities became less stable, and some funds suddenly changed community membership. Turbulence became more intense on November 3 and November 4, when community $C_1$ merged with $C_2$ and community $C_3$ merged with $C_4$. Afterward, the original configuration of communities emerged again. Note that this transitory shift happened in correspondence with a series of relevant macroevents that occurred during the second part of our sample period. The Greek legislative election took place on September 20 and Syriza won by 7.5 points over New Democracy. The new austerity package was enacted on November 19 by the Greek government. Monetary policy actions by both the European Central Bank (ECB) and the Federal Reserve (FED) took place at the end of this sample period and hit markets that had been experiencing a long period of stability. Then on December 9, the ECB reduced the deposit facility rate to $-0.30\%$ (the previous
modification occurred in September 2014 when it was fixed at 0.25% and the FED raised the target range for the federal funds rate to [0.25; 0.50]% on December 17 (the previous modification occurred in December 2008 when it was fixed at [0; 0.25]%). The Eurozone debt crisis and the Greek instability resulted in high credit spreads on government bonds throughout 2015, and during the summer and autumn the effect was especially severe. All these events heavily affected the decisions of managers and may have concurred with the reduced heterogeneity in manager behaviors we observe when the communities merge. When the effects of the exogenous shocks vanished, the original configuration returned. This finding, apart from supporting the presence of persistent commonalities in the way fund managers allocate their portfolios, opens a perspective in the analysis of the interdependence between observed behaviors and the emergence and resolution of phases of systemic instability.

Conclusions

Agents tend to apply complex decision-making mechanisms, but formal rules of rational choice can be overturned by subjective views, beliefs, and habits, which generate personal mental references, are conditioned by a limited set of opportunities, suffer from bounded rationality, and rely on routines, all of which we understand to be investment behavioral features.

Behavioral attitudes contribute to induce manager allocation strategies that go beyond traditional classifications based on portfolio compositions. The goal of our approach is to quantify financial indicators that may be related to well-known patterns detected in behavioral finance, e.g., anchoring, belief perseverance, overconfidence, and conservatism, that influence how financial indicators may be related to well-known patterns detected in behavioral finance, e.g., anchoring, belief perseverance, overconfidence, and conservatism, that influence how portfolios are allocated and managed. By looking at the attitudes shown by fund managers toward trading intensity, the use of the derivatives, position concentration, the reaction to liquidity injections, and market condition changes, we managed to identify behavioral patterns and shed light on the emergence of commonalities in fund managers’ behaviors. Further complementary research may couple our financial attributes with survey-based, self-reported fund managers’ attitudes or characteristics (complementing previous analysis on the effect of managers’ characteristics and behaviors, for instance refs. 35 and 36) to specifically assess which behavioral biases may drive community formation the most.

We believe that the behavioral commonalities we have detected in our analysis of professional fund managers’ allocation strategies are relevant for several reasons. First, our evidence suggests that mental models, personal preferences, and routines play an important role in expert decision making. Although our analysis highlights their importance, further research is needed to disentangle the effects of behavioral traits and other fund characteristics that we could not observe, such as management fees, the structure of transaction costs, or business constraints, on the adoption of specific strategies. Also, similar analysis on higher-frequency, intraday data may contribute to shed further light on managers’ behavior. Second, we show that exogenous shocks temporarily alter the configurations of communities. During the out-of-equilibrium phase, expert investors seem to converge toward more similar strategies, and then the system returns to its precrisis configuration. If confirmed by future investigations on other datasets, this pattern might contribute to our understanding of the emergence and evolution of market instabilities. Finally, at the microlevel, our approach allows us to characterize investment manager behavior and may pave the way to (i) better understand how beliefs, managers’ personal preferences, and mental models affect the risk profile of a fund and (ii) distinguish between the contributions of standard and nonstandard behavioral drivers to performance.

ACKNOWLEDGMENTS.

We are very grateful to Sara Zaltron, Francesco De Matties, and the team at Azimut Analytics for insightful discussions and for providing us with the dataset. We acknowledge funding from the National Research Project CRISIS Laboratory. Financial support from the Competitività, Regole, Mercati (CERM) Foundation is gratefully acknowledged. The Boston University Center for Polymer Studies is supported by NSF Grants PHY-1505000, CMMI-1125290, and CHE-1213217; by Defense Threat Reduction Agency (DTRA) Grant HDTRA1-14-1-0017; and by Department of Energy Contract DE-AC07-05ID14517.