Replicant Investment Platforms

Mauricio Ferraresi Junior*†

April 5, 2025

Abstract

This research examines how investment platforms determine asset exposure, distinguishing authentic from replicant behavior. I develop an exposure-based demand model to study how retailer demand shocks move through fund networks and influence platform decisions. The model introduces replicant risk, a new concept capturing how competition and imitation among platforms affect asset pricing via exposure elasticity. Using Brazilian fund data, I estimate the model through a Bayesian dynamic factor approach to quantify this elasticity. Between 2016 and 2021, a replicant risk-based strategy earned an annual premium of 7.6% with 15.47% volatility. The strategy is nearly market-neutral and shows no momentum. The model highlights stocks particularly sensitive to replicant behavior, deepening our understanding of asset pricing and investor dynamics. This framework links financial intermediation and platform competition, offering insight into how platforms shape prices. It underscores the growing importance of demand-side forces in financial markets and the strategic behavior of investment intermediaries.

Keywords: investment platforms, asset allocation, fund networks, exposure elasticity, Bayesian factor models, asset pricing, financial intermediation.

[1]I am indebted to Marcelo Fernandes for his valuable comments and suggestions on this paper.

[2]São Paulo School of Economics, FGV.

Financial support by FAPESP grant 2023/05812-5.

Contact: mauricio.ferraresi@fgv.br.

"I've seen things you people wouldn't believe. Attack ships on fire off the shoulder of Orion. I watched C-beams glitter in the dark near the Tannhaüser Gate. All those moments will be lost in time, like tears in rain. Time to die."

Replicant Roy Batty in Blade Runner

1 Introduction

Distinguishing authentic behavior from mere imitation is a classic problem in fields ranging from artificial intelligence to evolutionary biology. The Turing test, for example, aims to determine whether a machine can reproduce human communication patterns so effectively that it is indistinguishable from a real person (Turing, 1950). Similarly, in the fictional universe of Blade Runner, the Voight-Kampff test assesses whether an individual responds emotionally like a human or merely simulates such responses.[1] This challenge leads to a question in the study of herding dynamics in the financial context: Is it possible to explain the asset exposure decisions of investment platforms on the basis of their own principles, or do they merely imitate prevailing patterns? I examine whether investment platforms' equity exposure decisions reflect independent choices or whether they emerge from a replicant behavior. This replicant behavior is explained by the role of investment platforms in the intermediation of retail investor resources.

These platforms play a central role in the allocation of retail investor resources, directing them to various assets. These platforms are made up of a range of mutual funds offered to retail investors by both large financial institutions and small asset managers. The way these platforms organize the funds is similar to a "spoke-hub" network structure: A fund is responsible for trading directly in the asset market (hub), and a fundraising network in the form of quota funds is exposed to the same assets managed by the hub fund, even without participating directly in the market. This network structure of mutual funds available to the retail market provides managers with economies of scale. The fragmentation of fundraising across funds prevents managers that buy assets directly from the market from being exposed to flows concentrated in their own liabilities. Thus, instead of studying supply and demand for equities solely from the perspective of the direct holdings of a few funds (hubs), an exposure-based model will propose to study supply and demand from the perspective of the direct and indirect holdings of investment funds available on investment platforms.

The exposure of investment platforms to stocks helps to introduce a new type of risk into the financial economy, which I call replicant risk. By assessing the exposure elasticity of IBRX-100 stocks on different investment platforms, it becomes feasible to create an investable replicant portfolio. The portfolio's long position includes stocks displaying non-replicant behavior on these platforms, while the short position comprises stocks with replicant behavior. The replicant portfolio achieved an annual return of 7.6% with a volatility of 15.46% between 2016 and 2021. This portfolio offers key insights for platform risk management. In particular, the strategy is nearly market neutral: the correlation between the market betas of the long and short positions is -0.5 during the studied period. In addition, the yearly returns of the long and short positions are closely correlated and varies above and below

[1] In the book "Do Androids Dream of Electric Sheep?" by Philip K. Dick, published in 1968.

zero in diverse intervals, indicating that there is no momentum driver in this strategy. For platforms, being able to identify replicant behavior is crucial, as participating in either position of this strategy can have significant implications depending on the month. Specifically, during the COVID-19 outbreak (February to March 2020), the short position comprising replicant-influenced stocks outperformed the long position. For platforms like Santander, which did not exhibit replicant behavior during this time, this knowledge is highly valuable for managing risks across various funds tied to the institution.

Recent literature in finance has shown that demand in a financial market has a significant impact on asset price formation. Demand-based asset pricing seeks to explain why shocks associated with investors' resources flows are able to generate significant price fluctuations independently of asset fundamentals (Koijen and Yogo, 2019; Gabaix and Koijen, 2021; Koijen and Yogo, 2020; Koijen, Richmond, and Yogo, 2024; Fuchs, Fukuda, and Neuhann, 2023; Van der Beck, 2022; Kim, 2025; Camanho, Hau, and Rey, 2022; Jiang, Richmond, and T. Zhang, 2024; Ben-David et al., 2022; Jansen, Wenhao Li, and Schmid, 2024; Benetton and Compiani, 2024; Huebner, 2023; Albuquerque, Cardoso-Costa, and Faias, 2024). Specifically, Koijen and Yogo (2019) proposes a demand system based on stock characteristics to study the impact of demand on stock prices, focusing on the role of institutional investors. In a complementary way, Gabaix and Koijen (2021) proposes the hypothesis of inelastic markets, arguing that the price elasticity of aggregate demand for equities is low enough to amplify the effects of inter-asset class flows (flows between bonds and equities) on the prices of these assets.[2]

Currently, there are no models in the financial literature that adequately explain how investors choose their exposure to stocks through mutual fund networks. Retail investors typically allocate their funds via financial services platforms, which offer a variety of mutual funds. However, there is a lack of an exposure-based demand model that takes into account the network of these funds. The key to understanding this issue lies in how data are collected through direct holdings of outstanding shares. By examining consolidated share positions among mutual funds in the Brazilian market, we can propose a new model that better captures retailer investor demand.

In contrast to this research, there are proposals in finance that relate the concept of networks to portfolio optimization, treating assets and their correlations as a network. An example is the work of Peralta and Zareei (2016), where the authors treat the stocks themselves as nodes in a network whose links are market returns. In the context of asset allocation, identifying the position of each stock in a financial network can significantly improve the composition of the portfolio. In general, centrality measures[3] capture the influence of each asset in the network, allowing investors to adjust the weights according to the connectivity between securities. In stable markets, there is evidence that the most central assets can provide a higher average return. However, during turbulent periods, the overweight of assets far from the network nodes tends to reduce risk and improve diversification (Ren et al., 2017; Peralta and Zareei, 2016; Ioannidis, Sarikeisoglou, and Angelidis, 2023).

Some authors point out that the preference for allocations to assets with lower centrality occurs

[2]The inelastic market hypothesis belongs to an extensive recent finance literature that explores themes like the rise of passive indexing, the role of institutional investors, and imperfect competition in financial economics (see Pavlova and Sikorskaya, 2023; Bond and Garcia, 2022; Berk and Van Binsbergen, 2025; Chinco and Sammon, 2024; Heath et al., 2022; Davis, Kargar, and J. Li, 2023; Haddad and Muir, 2021; Rey et al., 2024; Coimbra and Rey, 2024, Greenwood and Sammon, 2022, Chodorow-Reich, Ghent, and Haddad, 2021, Haddad and Muir, 2025; Azar, Schmalz, and Tecu, 2018; Haddad, Huebner, and Loualiche, 2025; Rostek and Yoon, 2023, Coles, Heath, and Ringgenberg, 2022, Charles, Frydman, and Kilic, 2024; Brown et al., 2023; Neuhann and Sockin, 2024; Loseto and Mainardi, 2023; Yu An, Benetton, and Song, 2023; Sammon and Shim, 2024; Clark, Houde, and Kastl, 2021).

[3]The central nodes of the network tend to coincide with older, larger capitalization, cheaper and more financially risky assets.

because these assets have a lower correlation with the assets considered central in the network. This preference provides greater protection during crises or sudden drops in liquidity. For example, the study by Ren et al. (2017) shows that in drawdown scenarios, allocating more to peripheral assets led to better performance than traditional portfolio selection strategies. The study by Peralta and Zareei (2016) discusses, based on an extensive data set, that weighting inversely proportional to centrality not only reduces portfolio variance but also improves performance metrics such as Carhart's alpha (Carhart, 1997). However, during periods of stability, some research suggests that the core assets in a network may be more effective in capturing upswings (C. He et al., 2022; Ioannidis, Sarikeisoglou, and Angelidis, 2023). In these scenarios, the higher internal correlation between core assets does not seem to hinder diversification as much, as the market tends to move more homogeneously, diluting some of the risk. In addition, the dynamic monitoring of connections through temporal measures of centrality can help in the transition of weights to take advantage of uptrends or minimize sudden losses in crises (Zhao et al., 2018).

The robustness of these portfolio optimization methods depends on their ability to capture changes in the structure of the network over time, as well as the way in which transaction costs are incorporated, a factor that is little addressed in these studies. Work such as Ioannidis, Sarikeisoglou, and Angelidis (2023) shows that the dynamic adaptation of the portfolio to changes in the market regime can indeed outperform traditional Markowitz techniques (Markowitz, 1952). In this research, I adopt a different approach. The exposure portfolio is made up of a network of mutual funds, which I refer to as an investment platform. A concrete example of such a platform is the selection of funds the financial Brazilian institutions offer to clients, e.g., the Itaú investment platform. To effectively address the question of replicant behavior among these funds, it is essential to estimate an exposure-based demand model of each investment platform on the market. Additionally, this approach allows an analysis of the impact that different sets of funds have on asset pricing in financial markets, as well as the risks associated with a market structure composed of platforms.

Thus, a pricing model based on the exposure of investors with funds on investment platforms seeks to explain the reasons for the variability of asset prices, which do not seem to be explained by fundamentals but rather by replicant behavior. One of the explanations for price variability is the inelastic markets hypothesis (Gabaix and Koijen, 2021). An exposure-based asset pricing model does not need to assume that investors are inelastic, providing an alternative perspective on asset market dynamics. Regardless of the price elasticity of investor demand on investment platforms, the market for exposures is concentrated on a limited number of platforms. This high concentration means that these platforms respond to changes in investor demand by adjusting the supply of exposures simultaneously. In such a concentrated environment, the price elasticity of demand is just one factor to consider when evaluating which platform holds market power, especially if investors are price inelastic. Market power can be indicated by the compensation of fund managers, which also takes into account the transaction costs that are passed on to investors.

Therefore, the purpose of an exposure-based demand model is to assess the consequences of price variability as investment platforms react to shocks in investor demand. This alternative view of the investor demand model assumes that the financial market can be interpreted as an oligopoly of investment platforms competing for investors' resources. It is a Cournot oligopoly, by exposure, not by price. In this context, the interaction between platforms can be described as an exposure oligopoly, in which each platform adjusts its allocation strategies in response to the decisions of its competitors. In this context, the most important decision variable for understanding the demand for exposure on a platform would not be the asset price but rather demand shocks from competing platforms. This

new way of interpreting market participants makes it possible to study risk through inter-platform spillovers and their consequences for the variability of asset prices. I define inter-platform spillovers in terms of an exposure elasticity of demand. This new definition of elasticity makes it possible to determine whether a platform can be considered a replicant or not.

Replicant behavior in finance has been widely studied in the literature. For example, digital investment platforms have been linked to speculative trading patterns. Barber et al. (2022) show that retailer investors trading through easily accessible apps such as Robinhood exhibit coordinated behavior, generating temporary price spikes followed by negative corrections. In turn, Apesteguia, Oechssler, and Weidenholzer (2020) analyze copy trading platforms and show that the ability to automatically replicate the decisions of experienced traders induces investors to take excessive risks, thus amplifying the replicative dynamics in financial markets. In Section 2, I systematically review part of the vast literature on replicant behavior.

The data used in this research are described in Section 3. The dataset includes consolidated positions in mutual fund portfolios. It will be essential to distinguish between direct holdings in outstanding shares and exposure to stocks. This distinction allows me to propose an exposure-based demand model that may not necessarily be suitable for optimizing the portfolios of a single manager. The sample consists of 7 million observations of exposures in FI (hub) and FIC (network fundraising) mutual funds from 41 investment platforms in the Brazilian market. These data were obtained from Quantum's company.[4] The exposures involve 166 stocks that were part of the IBRX-100 index at any time between January 2016 and December 2021. During this period, there were approximately 1,300 funds with assets under management exceeding R$15 million, more than 10 shareholders, and exposures equal to or greater than zero. The funds were classified by ANBIMA in categories related to free multi-market funds (often associated with US market hedge funds) and included eight different classifications of shares (which can be defined as mutual funds).

In Section 4, I explore the theoretical foundations of an exposure-based demand model. I propose representing exposure-based demand through a dynamic factor model, where each investment platform has its own specific model. The dynamics of these factors are designed to capture the temporal structure of the correlations between the aggregate exposures of each platform. In Section 4.3, I elaborate on the analogy between a factor model as described by Ross (1976) and my proposal for exposure-based factors. This analogy aims to establish connections between stock pricing models based on returns and my model, which relies solely on exposure. Additionally, I discuss the potential of treating the collection of exposure-based factors and heterogeneous beliefs about the expected values of stock returns as a single aggregate factor. This topic will be further examined in another paper (Ferraresi, 2025c). A part of this discussion is presented in the Appendix A and B.

To quantify replicant behavior, I use the exposure elasticity of demand for each investment platform, which I will denote by the Greek letter δ. This parameter is related to a concept known as a granular variable (Gabaix, 2011; Gabaix and Koijen, 2024). In Section 4.2, I will elaborate not only on the idea of a granular variable but also on my proposal of a granular network variable. The granular network variable is the linear combination of demand shocks in other platforms. This granular network variable provides crucial insights into an investor's demand for exposure on an investment platform. The only reference to this type of granular network variable comes from the preliminary work of Chodorow-Reich, Gabaix, et al. (2024). Because a granular network variable considers the demand shocks of other platforms, it can be used as an instrument to identify the slope of residual demand curves.

[4] https://quantumfinance.com.br

It is important to note a key distinction between the exposure-based demand model proposed in this paper and the demand models for equities found in the finance literature: the role of the demand agent differs. In the financial literature's demand models, the interests of those seeking shares are mediated by mutual fund managers. These managers decide on the direct holdings of their funds, which in turn affects the holdings of financial institutions. The main takeaway is that focusing solely on direct ownership of outstanding shares in financial institutions offers a limited perspective on the market.

How can exposure elasticity explain the existence of a replicant risk? Replicant risk is a premium associated with uncertainty about the behavior of investment platforms. This uncertainty exists because, for any level of market risk, there will be uncertainty associated with the context in which the platforms are inserted. In other words, whether it is a coordination context or not. I will assess uncertainty by forming an investment portfolio that is based on the exposure elasticities of individual stocks.

Each investment platform will have a different exposure elasticity for each stock. I define this exposure elasticity by $\delta_{m,n}$, where m is the platform and n is the stock. Each 1% variation in the granular network variable of a platform will be associated with $\delta_{m,n}$% variation in the platform's exposure. Thus, for each stock, there is a proportion p of investment platforms that show a replicating behavior $\delta_n > 0$, and a proportion $1 - p$ that does not have a replicating behavior $\delta_n \leq 0$. The proportions p or $1 - p$ define two possible market regimes for the dynamics of stocks. A simplified example helps to better understand these regimes and the new type of replicant risk in share premiums. The Figure 1 represents the possibilities of premiums on two stocks given the behavior of the set of platforms in coordination and anti-coordination contexts.

		stock 1	
		$\delta_1 > 0$	$\delta_1 \leq 0$
stock 2	$\delta_2 > 0$	**(2%;2%)**	(2%;0%)
	$\delta_2 \leq 0$	(0%;2%)	**(3%;3%)**

coordination

		stock 1	
		$\delta_1 > 0$	$\delta_1 \leq 0$
stock 2	$\delta_2 > 0$	(0%;0%)	**(4%;2%)**
	$\delta_2 \leq 0$	**(2%,4%)**	(3%,3%)

anti-coordination

Figure 1: Exposure-based game in two stocks given the set of platforms. In bold the possibles expected returns of equilibrium.

Examine the scenario of anti-coordination: when neither stock exhibits replicant behavior, the risk premium aligns with the market premium. For instance, both stocks show a market premium of 3%, indicating that exposure-based shocks are generally offset by exposure supply from platforms. Anti-coordination involves two costs. First, there is a distrust cost tied to replicant behavior, which can potentially cancel out the market premium if both stocks exhibit such behavior. In this scenario, the expected returns will be 0%, and this market setup will not be an equilibrium. The second cost is related to non-replicant behavior, with a stock facing costs associated with the potential shift from non-replicant to replicant (outside option). In an anti-coordination environment, the distrust cost exceeds the outside option cost, resulting in the anti-diagonal payoff illustrated in Figure 1. This scenario achieves equilibrium as a Nash equilibrium with mixed strategies, where replicant behavior proportions in stocks are exactly $p = 0.5$.

In a coordination context, consider again the market premium of 3% for stocks not showing replicant behavior of the platforms, serving as the baseline market condition. Here, the outside option cost related to replicant behavior is 1%, akin to the anti-coordination case, thus leading to 2% expected

returns when all platforms are replicant. Evaluating a coordination scenario, distrust costs may reduce premiums, as platforms gain no advantage from synchronized actions. Again, distrust costs for non-replicant stocks surpass the outside option costs. The outcomes in this coordination scenario are shown in Figure 1, with a Nash equilibrium in mixed strategies existing between two market regimes, where replicant behavior in stocks is $p < 0.5$.

It is crucial to understand the motivations behind the coordination of the platform. For platforms that manage funds exposed to the same assets, there is always a counterparty risk inherent in the exposures offered to the retail market. In finance, diversification plays a central role, not only in individual managers' decisions but also between portfolios managed by different managers. Exposure elasticity captures exactly this diversification between platforms. When all platforms have a positive elasticity of exposure in a stock, there are necessarily trading opportunities that make it possible to reduce exposure through other assets, transforming counterparty risk into replicant risk due to the coordination context. In the specific case of the fictitious values used in this example, the replicant premium is the outside option of 1%, and given this asymmetry, the Nash equilibrium is for the proportion of replicant platforms in each of the two stocks to be different from 0.5.

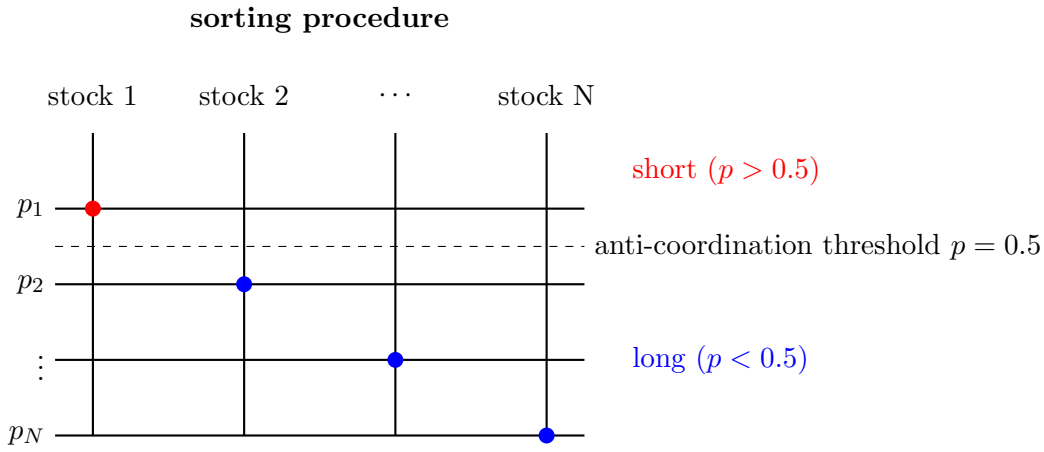


Figure 2: Example of constructing a long-short portfolio using the platforms proportions p as criterion.

The replicant premium in a stock is the difference in premium between non-replicant and replicant behavior, taking into account the uncertainty of whether the platforms are coordinated or not. This premium can be quantified using the Fama and MacBeth (1973) procedure. I illustrate this procedure in Figure 2. For a set of stocks in a panel of exposure data each month, it is possible to construct a portfolio long the stocks with exposure elasticity $\delta > 0$ and short the stocks with exposure elasticity $\delta \leq 0$. If the proportions of platforms with replicant behavior approach $p = 0.5$, the replicant risk would turn into a counterparty risk in a context in which the platforms do not coordinate.

Therefore, it is crucial to infer the demand parameters for each investment platform, which will be discussed in detail in Section 5. I employ a Bayesian approach to infer these parameters because the demand models for each platform have high-dimensional parametric space. In contrast, the frequentist approach only provides aggregate estimates of price elasticity due to convergence issues in the generalized method of moments (GMM). For instance, Koijen and Yogo (2019) classify retail investment providers into categories (such as banks and pension funds) and estimate a low-dimensional portfolio choice model for each category on a quarterly basis. Instead of referring to them as investment providers or financial institutions, I prefer the term investment platforms, considering the interconnectedness of funds. The Bayesian approach allows for the estimation of panels of high-dimensional portfolio choice models for each platform using monthly data. Additionally, Bayesian modeling captures the uncertainty

in the exposure-based demand model by utilizing the posterior distribution of each parameter.

Bayesian approaches to dynamic factor models have become increasingly popular due to their ability to capture parameter uncertainty through full posterior distributions (Frühwirth-Schnatter, Hosszejni, and Lopes, 2024; Lopes and M. West, 2004; Bai and P. Wang, 2015; Aguilar and M. West, 2000; Koop, Korobilis, et al., 2010). By treating parameters as random variables, Bayesian methods incorporate prior information and provide a more comprehensive representation of uncertainty, which can be especially important in volatile financial situations (Bauder et al., 2018; Nakajima and M. A. West, 2013). In contrast, the frequentist perspective focuses more on point estimates and confidence intervals. Although this approach offers simpler computational methods, it often overlooks the full extent of the estimation risk. Some comparative analyses highlight the limited uncertainty quantification of the frequentist framework but also note that frequentist inference is generally easier to apply in large-scale situations where computational efficiency is crucial (Sigauke, 2016). Ultimately, the choice between these two perspectives largely depends on the specific context, considering the complexity of the data, the dimensionality of the model, and the amount of prior knowledge available. I will discuss these points in Section 5.1.

In Section 6, I examine the empirical findings. To measure the degree of replicant behavior in the market, Bayesian estimates are used to determine parameters for each platform using the same approach of Ward et al. (2019).[5] A comprehensive list of estimated parameters for each platform is presented in Appendix D. With the exposure elasticity of the demand parameters available, it is feasible to gauge the proportion of platforms that feature positive or negative exposure elasticity across stocks in the IBRX-100 index. Only positive exposures are considered for these estimates. The analysis does not include evaluations of decisions regarding the platforms' network of fund exposures on the extensive margin. The entire data set is used, to ensure that the parameters do not change over time. The study identified replicant proportions in 42 stocks from 15 platforms and highlights the significance of this topic for these platforms' risk management, as the risk ties to the fund selection available to retail investors. The platform exhibiting the greatest replicant behavior was the financial institution Itau, showing such behavior in 21 of the 42 stocks examined. Itau stands as Brazil's largest private bank by market value,[6] implying that a risk metric applicable to a platform like Itau's fund set could enhance risk management in the context of imperfect competition for retail funds, encompassing both Itau and other platforms.

Finally, this research seeks to improve the literature on demand-driven asset pricing by incorporating insights into the financial intermediation of investor resources, competition between platforms, and the possibility of replicant behavior when choosing exposures in mutual fund network portfolios.

2 Related Literature

Replicant behavior in financial markets has received significant academic attention (Hirshleifer, 2020). Researchers have studied why investors, both individual and institutional, often make similar trading decisions and how these aggregated actions influence asset prices and overall market dynamics. Early theoretical models primarily focused on information asymmetries, suggesting that rational investors can derive valuable insights from observing the trades of others. In these "information cascades," market participants, even when equipped with private signals, may choose to follow the crowd simply because

[5] Ward et al. (2019) utilized the STAN package in R, which was explained in Carpenter et al. (2017)

[6] The most valuable brand in Latin America, valued at US$8.7 billion, according to Brand Finance's 2023 Global 500 Ranking (https://brandfinance.com/)

they believe that the consensus incorporates aggregated information (Bikhchandani, Hirshleifer, and Welch, 1992). Such models highlighted how replicant behavior can arise from rational considerations when the cost of collecting or processing personal information is high (Cipriani and Guarino, 2014). However, literature soon showed that concerns about reputation can amplify these effects: analysts, fund managers, and other financial professionals risk harming their careers by deviating from the majority, thus reinforcing a tendency toward conformity (Dasgupta, Prat, and Verardo, 2005; Nirei, Stamatiou, and Sushko, 2012).

Over time, empirical studies have provided substantial evidence that replicant behavior is not just a theoretical construct, but a measurable phenomenon with material consequences (Tan, Xiaoyan Zhang, and Xinran Zhang, 2023; Eaton et al., 2022; Van der Beck and Jaunin, 2021; Boehmer et al., 2021; Pedersen, 2022). In US equity markets, for example, researchers have documented institutional herding in both cross-sectional and time-series data, showing that large money managers often buy or sell the same stocks in overlapping periods (Puckett and Yan, 2008; Teh, De Bondt, et al., 1997). This clustering of trades can temporarily push prices away from fundamentals, only for valuations to revert once contrarian investors step in or momentum dissipates. Various studies estimate that these short-term distortions can amount to a few percentage points of price deviation before reversal (Cipriani and Guarino, 2005; Gutierrez and Kelley, 2009). In addition, some scholars find that buy-side herding can accelerate the price discovery process in certain contexts when the group is collectively responding to new, accurate information; however, sell-side herding tends to coincide with faster and more dramatic downward price pressures, especially in times of crisis (Cai et al., 2019; Stephanie Kremer and Nautz, 2013).

Evidence of replicant behavior can be observed in various regions and asset classes. In the US stock market, researchers have found that 38 percent of large institutional investors tend to intentionally herd toward technology stocks (Uwilingiye et al., 2019). Furthermore, studies on digital currency exchanges conducted by Song (2023) have introduced a psychological and behavioral metric called the herd behavior index, which measures the herd instinct of investors in cryptocurrency markets and identifies anomalies in cryptocurrency returns. Similarly, research on emerging markets—from Malaysia (Lai and Lau, 2004) to Taiwan (Shyu and Sun, 2010), from China (Wei Li, Rhee, and S. S. Wang, 2017) to India (Garg and Gulati, 2013)—indicates that both institutional and individual investors often align their trading strategies in response to macroeconomic announcements or changes in regulatory conditions. In particular, herding behavior in the Chinese A-share market is observed to be more pronounced among individual investors than among institutional investors. This discrepancy may be due to differences in experience levels, risk preferences, or access to reliable information (Wei Li, Rhee, and S. S. Wang, 2017).

Several studies use laboratory experiments to unravel the micro-level motivations for replicant behavior, supporting the view that herding can serve as a strategy to reduce uncertainty in highly ambiguous environments (Cipriani and Guarino, 2005). When investors face complex signals or volatile market conditions, following the majority provides psychological reassurance and a form of social validation (Petersen and Spickers, 2023). In addition, periods of optimism, characterized by bullish sentiment, can lead to buy-side convergence as rising prices reinforce positive feedback loops. In contrast, when negative sentiment dominates, investors can herd to sell, accelerating downward price spirals (Guo et al., 2024; Bastías and Ruiz, 2022). Although both buy-side and sell-side herding lead to temporary price distortions, the magnitude and duration can differ, with some scholars noting that sell-side waves tend to occur more quickly and can be more destabilizing (Cai et al., 2019; Boortz et al., 2013).

A particularly rich line of research focuses on how institutional frameworks influence the intensity and impact of replicant behavior. Incentive structures that tie fund managers' compensation to short-term performance or relative benchmarks can exacerbate herding, as few managers want to underperform their peers or deviate sharply from the prevailing consensus (Pedraza and Pulga, 2019; Bootrz and Kremer, 2013). Peer benchmarking increases the perceived risks of going against the group; If the entire cohort goes in one direction and loses, blame is diffused, but if a single contrarian underperforms, reputational damage is more severe (Dasgupta, Prat, and Verardo, 2005; Nirei, Stamatiou, and Sushko, 2012). Regulatory factors also play a role. Studies in emerging markets, such as the Chilean context, suggest that changes in capital controls or pension fund guidelines can trigger large-scale asset reallocations that look very much like herding (Bastías and Ruiz, 2022). Similar patterns appear in developed markets as well; Fiechter and Mangeney (2020) documents that acquisitions of brokerage firms and consolidation of research analysts reduce the diversity of opinion in the market, potentially increasing herding as fewer independent research signals circulate among large investors.

The intensity and impact of replicant behavior on asset pricing also occur in the short term. Herding can distort valuations, pushing certain securities above or below their fundamental value in the short term. When these movements coincide with market stress or liquidity shortages, volatility can spill over into other assets and even trigger cross-asset correlations. For example, Ben-Horin and Kedar-Levy (2013) finds that equity herding among large institutional investors can also affect corporate bonds, likely due to rebalancing decisions or changes in perceived risk across a portfolio. As correlated trades accumulate, seemingly disparate assets can begin to move in tandem, reducing the potential benefits of diversification. In the medium term, some segments of the literature suggest that markets correct these mispricings, especially once contrarian participants enter the fray (Clarke, Ornthanalai, and Tang, 2015). However, other studies warn that when herding episodes coincide with cheap credit or excessive leverage, distortions can persist and fuel bubble-like conditions that increase systemic fragility (Petersen and Spickers, 2023; Krokida, Makrychoriti, and Spyrou, 2020).

Another important issue is the distinction between informed and uninformed herding. The former can improve informational efficiency when the crowd actually converges on superior insights or fundamental news that has not yet been fully digested by the market. In such cases, herding accelerates the incorporation of relevant information into prices (Clarke, Ornthanalai, and Tang, 2015; Christoffersen and Tang, 2010). However, uninformed or sentiment-driven herding can obscure true value signals, potentially undermining market efficiency and exacerbating volatility. Identifying which type of herding predominates is an empirical challenge. Many econometric studies attempt to determine whether abnormal trading patterns correlate with meaningful events such as earnings announcements or macroeconomic data releases (Jurkatis, Stephanie Kremer, and Nautz, 2012). When trades concentrate on the absence of clear informational catalysts, it is suggested that herding is largely driven by bias or imitation rather than rational inference.

Market stress, such as financial crises, often intensifies replicant behavior among investors. During turbulent times, risk aversion increases, making the perceived safety of following the crowd more appealing. For example, research conducted during the global financial crisis indicates that pension funds and other long-term institutional investors resorted to collective selling as credit risks became more apparent (Bastías and Ruiz, 2022). Similarly, Boortz et al. (2013) argues that the risk of information - where the reliability or availability of signals decreases - increases the tendency to herd. In such situations, decision-makers may distrust their own analyses and give undue importance to observable trades made by well-known institutions. Consequently, these behavioral shifts can lead to liquidity shortages and exacerbate downward price spirals.

Despite the disruption it can cause, replicant behavior can stabilize markets in certain contexts. When a fundamentally justified signal is strong - for example, a clear indication that a company's future earnings have improved - uniform buying by a large number of institutional players can quickly bring the security back to its fair value. In normal, non-crisis environments, many herding episodes appear to be short-lived and self-correcting, especially in large, liquid markets where contrarian traders can profit from temporary mispricings (Gutierrez and Kelley, 2009; Clarke, Ornthanalai, and Tang, 2015). However, there is disagreement among scholars on whether the long-term net effect of herding is beneficial or detrimental to price discovery and market efficiency. Some emphasize the negative externalities of correlated trading, such as increased systemic risk and potential contagion between asset classes (Krokida, Makrychoriti, and Spyrou, 2020; Ben-Horin and Kedar-Levy, 2013). Others note that herding driven by legitimate fundamental signals could ultimately improve overall market quality.

From a policy point of view, the literature suggests a range of measures that could mitigate the most destructive aspects of replicant behavior. Enhanced transparency is widely advocated, and timely disclosure of large trades or short positions can potentially diminish the incentive to cluster. Policymakers have also considered adjusting performance benchmarking practices to encourage longer investment horizons. If fund managers are no longer strictly judged on quarterly results or relative ranking, they may feel freer to deviate from group behavior, improving the diversity of strategies in the market (Pedraza and Pulga, 2019). Macroprudential regulations, such as countercyclical capital buffers or circuit breakers, could help dampen the impact of large-scale, simultaneous sell-offs that are triggered by herd instincts. At the same time, regulatory bodies must tread carefully to avoid unintended consequences, such as stifling beneficial information flows or reducing the depth of secondary markets.

Recent work has also begun to examine the role of technology and alternative data in replicant behavior (Yang and Loang, 2024). As high-frequency traders and algorithmic strategies proliferate, some researchers suggest that automated models may "learn" from each other's trades and inadvertently create new forms of herding, possibly on very short time scales (K. Li, 2014). In addition, a greater reliance on machine learning for investment decisions could reduce the diversity of analytical approaches, particularly if many firms use similar algorithms trained on overlapping data sets. As a result, the next generation of research may focus on how artificial intelligence amplifies or attenuates replicant behavior, and whether advanced methods can be developed to identify emerging clusters before they destabilize markets (Loang, 2025; Dou, Goldstein, and Ji, 2024; Kaniel et al., 2023; Feng, J. He, et al., 2024; Axtell and Farmer, 2025; Dyer et al., 2024; LeBaron, 2006).

References indicate that replicant behavior is influenced by a combination of rational, behavioral, and institutional factors. Cognitive biases, such as uncertainty aversion, shifts in sentiment, and peer benchmarking, tend to lead investors toward imitation behavior, particularly in times of increased ambiguity or crisis. In addition, regulatory frameworks, incentive structures, and industry norms can either strengthen or mitigate these behavioral tendencies. Empirical evidence shows that herding can cause short-term price distortions of various magnitudes, with potential long-term effects that are still a topic of debate. While some aspects of herding can facilitate the market's integration of genuinely useful information, others contribute to mispricing and increase systemic vulnerabilities. By integrating findings from behavioral finance and institutional economics, researchers aim to clarify the conditions under which aggregated investor behavior acts as a stabilizing force rather than a threat to financial stability.

My contribution to the literature focuses on understanding how the behavior of investment platforms, which often resembles that of a replicant, can significantly affect the risk management

of these platforms. This article examines the equity exposure of investment platforms and analyzes whether their asset allocation decisions are independent strategies or simply imitations of existing trends. It also assesses how the structure of fund networks affects asset pricing, taking into account the presence of imperfect competition among platforms. The effect of replicant investment platforms on asset pricing is defined by replicant risk. The general contribution of my research is to the use and interpretation of the exposure data of the fund network in the form of an investment platform, as is in various unpublished works (Ferraresi, 2025b, 2025c; Ferraresi and Urbano, 2025; Ferraresi and Leal, 2025; Ferraresi and Di Pietra, 2025; Ferraresi, Castro, and Yoshinaga, 2025).

3 Data

Retailer investors demand assets in exchange for their financial resources. Since this demand is delegated to mutual fund managers, it is natural to think of the demand-based model from the manager's perspective. However, the demand for assets from each fund or manager is an incomplete view of the structure of the financial market. The market is made up of investment platforms. These platforms can be large financial institutions or small asset managers that have a network of funds in which their investors can invest. The demand for assets intermediated by multiple funds differs from the demand for assets intermediated by a single fund. This difference is due to the fact that raising funds through a network of investment funds provides exposure to assets acquired by any mutual fund on the same platform. Therefore, the structure of investment funds available to retail investors through investment platforms requires special care in the description and interpretation of data.

For example, the demand for a network of funds is not only due to the direct holdings of its funds. It is very important to differentiate between direct participation in outstanding shares and exposure to shares. An investment fund's exposure must consider the direct and indirect holdings due to the fund's participation in other funds. The exposure in an asset corresponds to the consolidated holdings of the funds. Consolidation takes into account the indirect holding in each asset due to purchases of stakes in other funds.

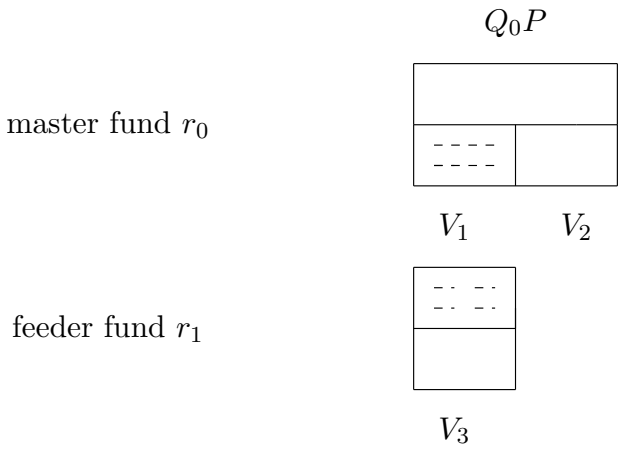


Figure 3: The assets are the upper rectangles, and the liabilities are the lower rectangles. Only the master fund r_0 directly participates in the outstanding shares, but both funds have exposure to the stock.

Consider a simple example of an investment platform with exposure to a stock with Q outstanding shares. This platform, represented in Figure 3, makes only two funds available to retailers: A master fund r_0 and a feeder fund r_1. Retail funding is represented in each fund's liabilities as V_2 and V_3. Only

the master fund r_0 acquires a fraction Q_0 in the outstanding shares worth $Q_0 \times P$, using the total resource $V_1 + V_2$. Both funds are exposed to the stock due to the V_1 of the feeder r_1 invested in master r_0.

To better understand the difference between exposure and direct participation, note that there are two ways of calculating a stock's relative weight in a fund's portfolio. The first calculation is associated with the relative weight due to the direct holding $\omega^{(d)}$, and the second ω, to the indirect + direct holding associated with an exposure E. These two forms coincide for the master r_0, since in this case, the exposure is the direct participation itself $E_0 = Q_0$, with

$$\omega_0^{(d)} = \frac{Q_0 P}{V_1 + V_2} = 1 \quad \text{and} \quad \omega_0 = \frac{E_0 P}{V_1 + V_2} = \frac{Q_0 P}{V_1 + V_2} = 1.$$

These two ways of calculating relative weight do not coincide in r_1. Although there is no direct participation in the outstanding shares $Q_1 = 0$, there is an indirect participation due to the amount V_1 invested in r_0. The r_0 uses the r_1 to raise funds, with $V_1 = V_3$. So, although $w_1^d = Q_1 P / V_3 = 0$, it is possible to obtain an indirect participation:

$$w_1 = \frac{E_1 P}{V_3} \quad \Rightarrow \quad \frac{V_1}{Q_0 P} = \frac{E_1 P}{V_3} \quad \Rightarrow \quad E_1 = \frac{V_1^2}{Q_0 P^2},$$

therefore, an investment platform, the fund network, can have a greater exposure to shares ($Q_0 + V_1^2 / Q_0 P^2$) than direct participation of its funds Q_0. An exposure-based demand should consider the fundraising structure through a network of investment funds available to retailers. This fundraising structure is a form of a multiplier because the institution's exposure is always greater than direct participation.

The investment platform organizes the network of funds for retailers in a spoke-hub structure. In this representation, the platform establishes master funds to manage the shares that circulate in the market through direct participation. The master fund serves as the central hub in the network, while all other nodes act as feeders, investing in the master fund's shares. The end nodes collect resources from retailers and transfer them to the central hub by exchanging exposure for quota shares.

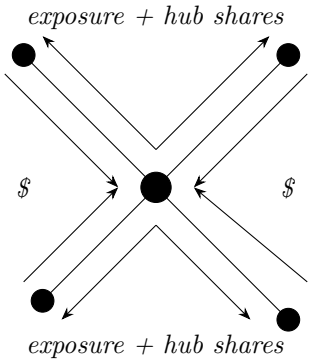


Figure 4: A retailer investment platform in the form of a spoke-hub network. Only the master fund in the network's center (hub) has a direct participation in the outstanding shares. All the other four nodes are responsible for retail fundraising.

The structure of a network of investment funds available by investment platforms is more complex than the representation in Figure 4. Funds identified as masters often hold shares in other funds and feeders may invest directly in stocks. In addition, a fund on one platform can acquire shares of other platforms, and funds managed by one platform can be distributed to retailer investors on different platforms. In any case, Figure 4 represents a simplified platform. If other platforms were considered,

the fund networks would have intersections, given the fund distribution network of the fundraising industry.

The most relevant data for the exposure-based demand model presented in Section 4 are the stock exposures of their fund networks made available to the retail market. The sample included around 7 million observations of stock exposures in investment funds classified as master and feeder funds qualified to buy stakes in other funds from 41 platforms in the Brazilian market. The company Quantum[7] collected the data. The exposures covered 166 stocks, some of which were part of the IBRX-100 index between January 2016 and December 2021. Throughout this period, around 1300 investment funds were available to retail investors with assets under management exceeding R$15 million; it has more than five shareholders and only includes exposures with values greater than or equal to zero.

Some minimum criteria are required to include or exclude funds from the sample. A minimum threshold for the funds' assets under management and the number of shareholders limit the importance of small, young, and exclusive investment funds. To better understand any criteria, it is necessary to understand that the fund sector in the Brazilian market has segmented the different responsibilities inherent in its financial intermediation services. The functions associated with investment funds can be divided into custody, management, administration, and distribution. Different platforms can perform these different functions in a single investment fund.

For example, custody is the safekeeping of asset in an investment fund, including the physical and financial settlement of assets, that is, the buying and selling of shares, bonds, and other assets. The institution that carries out custody is called a depository institution and is responsible for holding the assets on behalf of the investors' funds. The administration is responsible for ensuring that the fund functions properly. The institutional administrator has the power to carry out the necessary acts to manage the fund. One of these acts is to calculate the value of a fund's liquid assets, also called assets under management.

In addition, management and distribution are the two other most important functions in understanding asset demand. Management is responsible for making investment decisions, such as selecting the assets the fund will buy and sell. Distribution is responsible for raising funds from investors, that is, selling the investment fund's shares. The distributor can be the administrator himself or a third party he hired. Over time, most investment funds begin with a main fund that manages the resources and some feeder funds distributed by one platform or different platforms.

Thus, small investment funds in the data sample (with few assets under management and few shareholders) can significantly affect the results of a platform's exposure. Platform's exposure would be related to the pressure caused by strong retail outflows due to the high mortality of funds in their early stages of implementation. In this case, the demand for exposure would be strongly related to the outflow of funds and not to the risk allocation decisions when funds decide to expose themselves to a stock. As a criterion, I chose funds with at least 5 shareholders and R$15 million in assets under management in the months between 2016 and 2021.

The Brazilian Financial and Capital Markets Association (ANBIMA) classified these investment funds in the sample into nine classifications related to equity management strategies. These funds related to equity management could be interpreted as mutual funds. The nomenclature of hedge funds and mutual funds in the Brazilian market is fuzzy. The ANBIMA did not use this nomenclature. Since July 2015, ANBIMA has classified investment funds on three levels. This hierarchy of levels seeks to

[7] https://quantumfinance.com.br

reflect logic in the investment process. The first classification level describes the asset classes defined by the Brazilian Securities Commission (CVM) regulations. Assets at this first level are divided into fixed income, equities, multi-market, and foreign exchange. The other two levels are associated with self-regulation of the funds. The intermediate level is related to types of risk management and is differentiated into indexed, active, and foreign investment. The last level varies according to strategy.

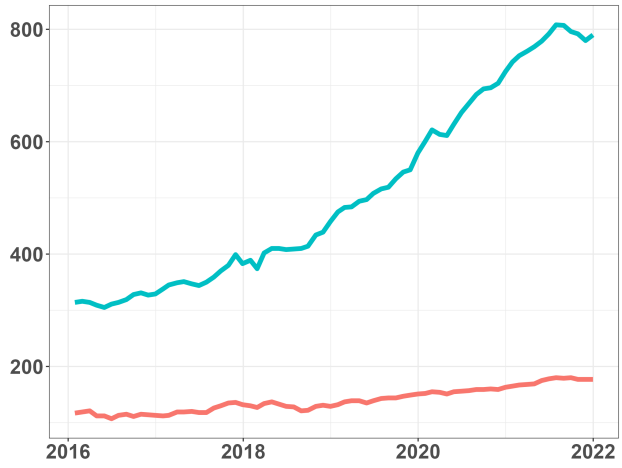
The database contains the investment funds classified as multi-market with a free strategy that allows leverage. These funds do not engage in specific strategies. These funds are associated with hedge funds in the US market. The sample also contains all the equity funds associated with mutual funds. These funds are classified by indexes and with active management based on strategies related to value/growth, sector, dividends, small caps, sustainability/governance, active index and free management. In the Brazilian market, equity funds must hold at least 67% of their portfolio in shares, warrants or subscription receipts, depositary receipts, shares in equity funds, and shares in equity index funds. Figure 5 offers a more thorough investigation of the selected sample.

Figure 5 describes (a) the evolution of the number of funds, (b) the evolution of the median assets under management each month of these funds, (c) the evolution of the median shareholders each month, and (d) the median of the net redemption/applications flows of these shareholders each month of the sample. The funds classified as network feeders are turquoise, and the Brazilian market has the acronym FIC (quota investment fund) in the name of the fund. Funds classified as masters are red and have the acronym FI (investment funds) in their names. Disclosure of exposure data (consolidated form) and direct holdings in outstanding shares (non-consolidated form) occurs monthly. By regulation of the CVM, these data are released with a delay of one quarter at the request of financial institutions.

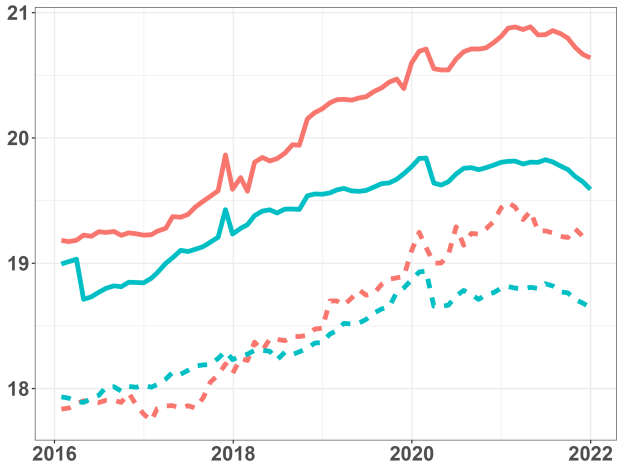
Although portfolio data are only released at the end of each month, the prices of the fund's shares, their assets under management, the number of shareholders, and the values of redemption and applications flows occur daily. The option was to smooth the higher frequency information and synchronize the daily and monthly information. All the information on the investment funds that is disclosed daily is carried out on average in the last fifteen business days. The idea is that the monthly disclosure of positions in portfolios is not necessarily the closing positions of the month. Each fund has its portfolio disclosure policy, and the regulator does not standardize the disclosure of these numbers.

Figure 5 suggests gains in the scale of the fund industry in raising funds in the form of a network of investment funds. The number of investment funds that behave as feeders of the funding network and the number of shareholders of these funds indicate a faster evolution in this period than the funds that play a role in the direct management of outstanding shares. On the other hand, the evolution of median wealth in the form of assets under management has the opposite behavior. Despite the suggestion of probable gains of scale in building a fundraising network in the form of a network, the evolution of investor flows suggests a more significant variability in the flows of funds associated with the direct management of outstanding shares, indicating the possibility of a dilemma between gains of scale and risks associated with liquidity. Investor flows in network feeder funds are the funds with distributions between institutions. As these flows are primarily the origin of the resources that master funds use to manage portfolios with direct participation, there is a risk associated with the liquidity that retail investors provide to the fund industry.

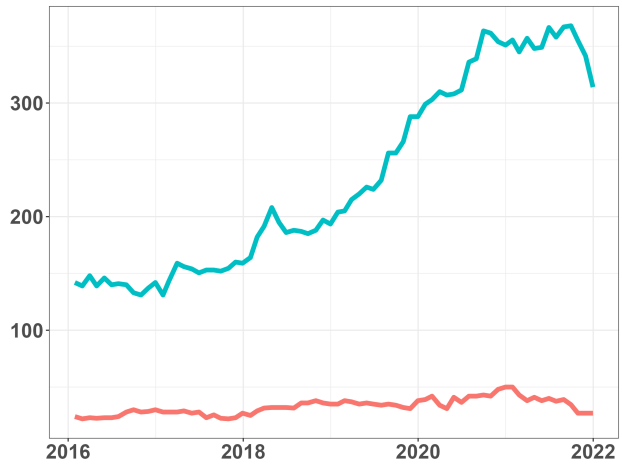
The exposure data collected by Quantum was consolidated, considering up to fifty levels of funds on funds. In most cases, consolidation down to three or four levels is sufficient. This type of data is challenging to collect and contributes significantly to the finance literature. The data currently used in the finance literature is only on direct participation. The exposure observations $E(m, a, r, n, t)$ in



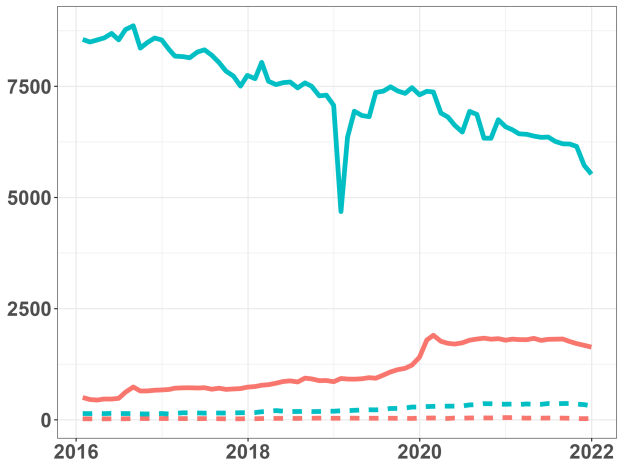
(a) number of funds



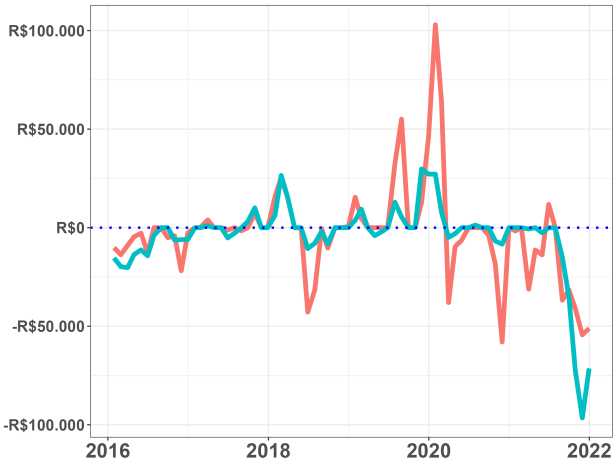
(b) mean and median in dashed (AUM in logs)



(c) median number of retailers

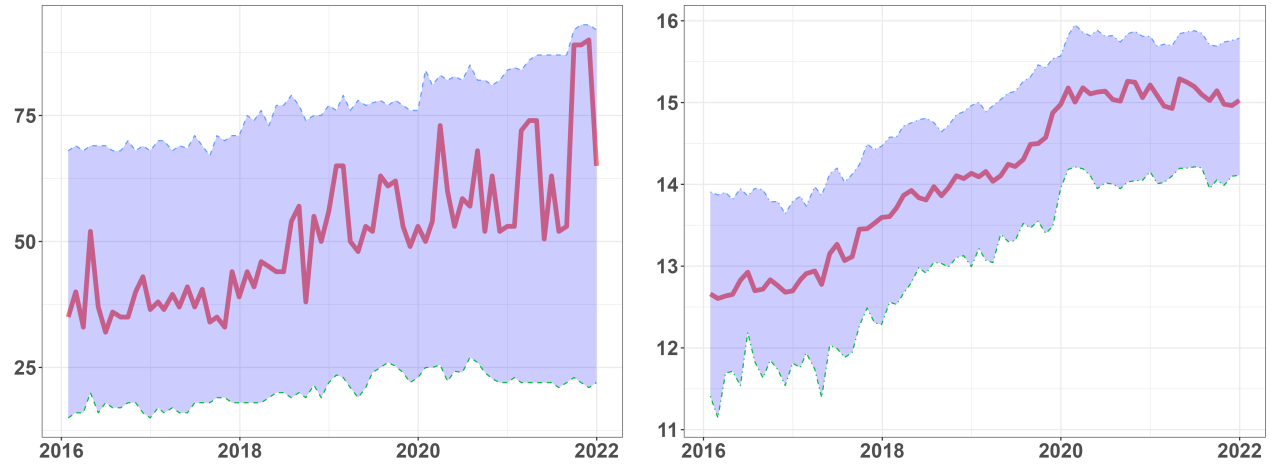


(d) mean and median in dashed of retailers



(e) median net flows

Figure 5: The network of investment funds (IFs) among sampled institutions has significantly expanded. Feeders funds (FF) are turquoise, and masters funds (MFs) are red. The number of feeders has increased over time above masters (a), and the assets under management (AUM) for masters have risen quickly (b). The number of retailers in feeders exceeds the number of feeders in masters (c). The flow pressure is more volatile in masters (d).



(a) number of stocks

(b) Bradesco stocks exposure (in logs)

Figure 6: IBRX-100 stocks in platform portfolios. Median in red. On the left, the distance between max and min in purple. On the right, the distance between the quintiles in purple.

the sample are indexed by $m = 1, \ldots, 41$ investment platforms, which made available $r = 1, \ldots, 1134$ investment funds to investors with $a = 1, \ldots, 9$ different ANBIMA classifications. The ANBIMA classification will be omitted in the notation. Each observation refers to stocks $n = 1, \ldots, 166$ that are part of the IBRX-100 index between January 2016 and December 2021 $t = 1, \ldots, 72$.

The number of stocks in the IBRX-100 index never reaches 100 in any given month (Figure 6a). Exposure refers to investment funds; if a fund has no exposure to any index stocks, I assign zero to its portfolio. The exposure data are derived from the portfolio weights $\omega_{m,r,n,t} = E_{m,r,n,t}P_{n,t} \,/\, \text{AUM}_{m,r,t}$, where the assets under management are defined as $\text{AUM}_{m,r,t} = \sum_n P_{n,t}E_{m,r,n,t} + H_{m,r,t} + O_{m,r,t}$. This represents the net worth of the fund, including the aggregate exposure to stock $\sum_n P_{n,t}E_{m,n,t}$ and the value of the inflows/outflows of retail investors $H_{m,t}$ along with the market value of external asset exposures $O_{m,t}$. Thus, the exposure for each fund is $E_{m,r,n,t}$ and is calculated using the observables $\omega_{m,r,n,t}$, $\text{AUM}_{m,r,t}$, and $P_{n,t}$. The exposure-based demand model looks at the total exposure of the fund network, exemplified in Figure 6b, which displays the Bradesco platform's exposure timeline.

Consequently, the platform exposure is an aggregate measure that totals the investments of the platform's funds,

$$E_{m,n,t} = \sum_r E_{m,r,n,t}. \tag{1}$$

The exposure variable can be consolidated across all platforms to evaluate asset exposure, as shown by:

$$E_{n,t} = \sum_m E_{m,n,t}. \tag{2}$$

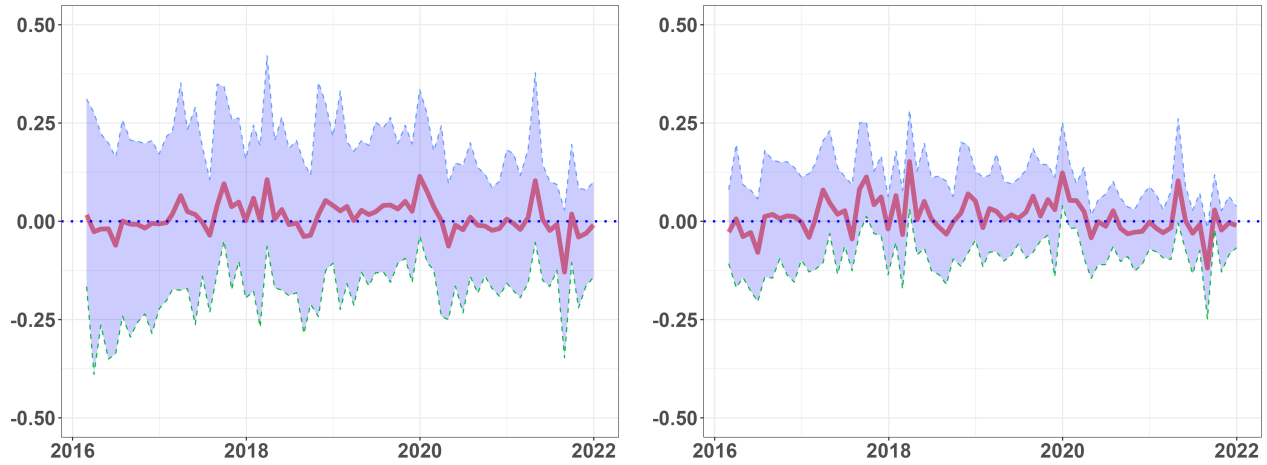
Thus, alterations in platform or asset exposure lead to correlated exposure patterns. The exposure change by platform is given by:

$$e_{m,n,t} = \log \frac{E_{m,n,t}}{E_{m,n,t-1}}, \tag{3}$$

while the exposure change by asset is defined as:

$$e_{n,t} = \log \frac{E_{n,t}}{E_{n,t-1}}. \tag{4}$$

The Figure 7 illustrate these two types of changes, demonstrating that changes are more pronounced at the platform level, suggesting correlated behaviors. The variable of interest seeks to examine the variation in volatility. The core modeling strategy employs the detailed hypothesis of aggregated fluctuations introduced by Gabaix and Koijen (2011), which posits that most of the aggregate fluctuation originates from idiosyncratic shocks to individual agents. In the case of investment platforms, the retailer exposure-based demand shocks play the role of idiosyncratic shocks.



(a) change exposure on platforms

(b) change exposure in assets

Figure 7: The red line represents the median evolution, while the blue region indicates the inter-quintile range.

Two points of view can be taken when examining the data. One is the concept of granular instrumental variables as discussed by Gabaix and Koijen (2024). This instrument leverages the notion that variations in exposure levels with different averages can lead to latent shocks if the exposure model on each platform can be depicted as an exposure-based factor model. These variations have been suggested for identifying parameters linked to aggregate variables, such as stock prices. To utilize similar approaches for analyzing platform-specific exposures, it is feasible to consider only the shocks affecting other platforms. The granular network variable is essentially a linear combination of exposures from various platforms,

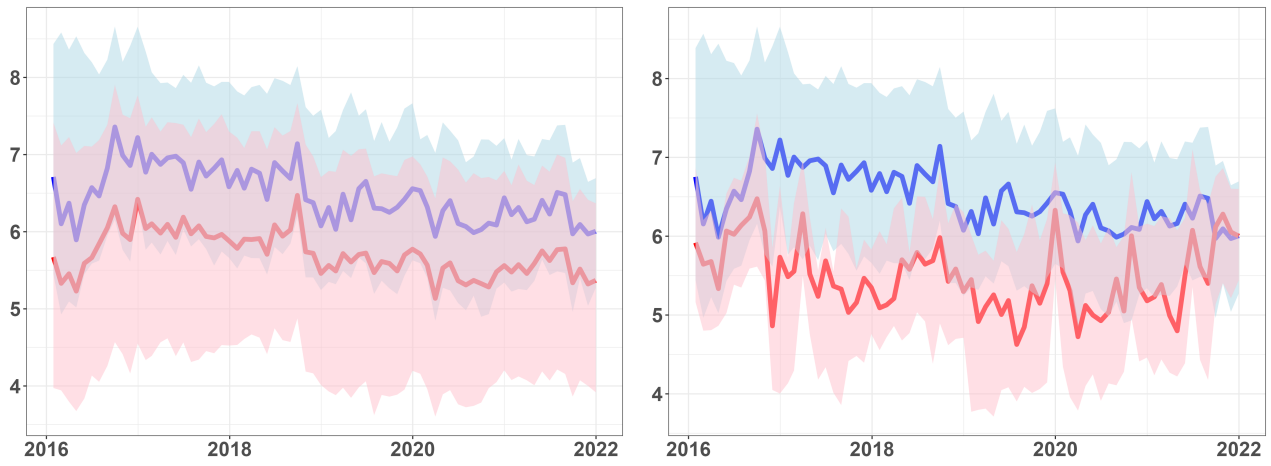
$$G_{m,t} \quad = \quad \sum_{j \neq m} \left(\tilde{S}_{j,t} - \frac{1}{M-1} \right) E_{j,t}, \tag{5}$$

where, the summing results in idiosyncratic shocks from other platforms if it assuming specific factor model to explain the exposures over time.[8] The variable $G_{m,t}$ is for one stock. In the case of the portfolio of stocks, the granular variable is a vector $\boldsymbol{G}_{m,t} \in \mathbb{R}^N$.

This concept resembles the research by Chodorow-Reich, Gabaix, et al. (2024), who introduce network granular instrumental variables, or network GIV. They extended the concept of granular instruments to account for network economies. For example, they suggest using a network GIV to determine and quantify the influence of a TFP shock in the steel industry on overall economic prices and output. In the case of investment platforms, the granular network variable is a persistent stochastic process over time, for example, in three different platforms in Figure 8.

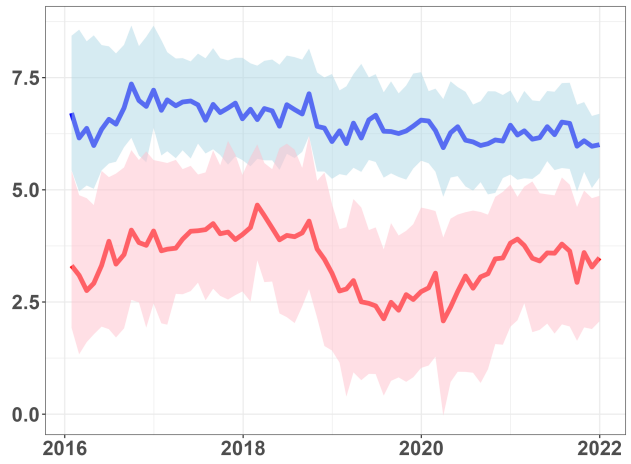
The progression of the granular variable is influenced by both the dimensions and exposure strategies of the platforms. Figure 8 depicts the temporal series of the granular network variable in

[8] More details in Section 4.



(a) Bradesco

(b) Verde



(c) Itau

Figure 8: Times series of granular network variable and granular variable. Median granular network in red and the total granular in blue (in logs).

red and the granular variable in blue. The granular variable is expressed as the sum of the values described in Equation (5), excluding the platform itself, m. As for Itau, its platform's size results in a significantly larger discrepancy compared to others. Conversely, although the Verde asset management platform exhibits granularity similar to that of the Bradesco platform, its higher variability suggests that fluctuations in Bradesco impact it more significantly, while its lower variability has a lesser effect on Bradesco.

In the next section, I will develop the exposure-based demand model using the granular network variable to infer the parameter of exposure elasticity of demand. This parameter will be essential for the empirical analysis of replicant risk in Section 6.

4 An Exposure-Based Demand

In this section, I will analyze the theoretical foundations of exposure-based demand. Section 4.1 presents the fundamentals of an exposure market, while Section 4.2 discusses stock pricing using the granular instrumental variable proposed by Gabaix and Koijen (2024). Moreover, I investigate the potential to generalize this instrument into a granular network variable, as proposed in the initial research by Chodorow-Reich, Gabaix, et al. (2024). This variable could be used as an explanatory variable in examining how platforms respond to exposure choices, with its associated impact between platforms defined as the exposure elasticity of demand. Platform decisions, regarding aggregate exposures to retail investors among available mutual funds, are explored in Section 4.3, where I investigate the interplay between equity portfolio models and the demand of retail investors for exposure. I explore a key comparison between asset returns from non-arbitrage pricing models (Ross, 1976) and a dynamic factor model for platform exposures. Lastly, Section 4.4 introduces a dynamic factor model to access investors' equity exposure demand, facilitating the monitoring of investment platform dynamics over time.

4.1 Exposure Market

The market for equity exposure consists of the supply of exposure from investment platforms and the demand for exposure from retailer investors. The demand for exposure is always a demand for market participation. The key point is that market participation, in this case, exposure, is not limited to direct participation in the outstanding shares. The exposure market is an intermediary market that exposes investors to stocks through a network of funds. Only a few specific funds buy shares directly in the market, but because funds buy stakes in each other's funds, all funds are exposed to fluctuations in share prices. In this way, an investment platform is always a network of funds that, when made available to the retail market, can create exposure to the stock market.

Consider a set $\mathcal{M} = \{1, \ldots, M\}$ of investment platforms. The exposure of a representative retailer investor of a platform in a stock is the amount $E_{m,t}$. Thus, the total exposure of the set $\mathcal{M}$ is $E_t = \sum_m E_{m,t}$. The demand for exposure from investors on each platform $E^{(d)}_{m,t}$ is represented by a demand for market share

$$\frac{E^{(d)}_{m,t}}{\overline{E}_t} \quad = \quad S_{m,t}(1 + e_{m,t}), \tag{6}$$

with the market share given by $S_{m,t} = E_{m,t} \,/\, \sum_m E_{m,t} = E_{m,t} \,/\, E_t$, and defining the total exposure market for this set of platforms by $\sum_m S_{m,t} = 1$. The term $(1 + e_{m,t})$ is a deviation factor from a trend

$\overline{E}_t$ and therefore represents a percentage change in exposure $e_{m,t} = (E_{m,t} - \overline{E}_t) / \overline{E}_t$.[9] The trend of the random variable total exposure, denoted as $\overline{E}_t$, captures the dynamics of its average, as exposures are not constant random variables over time. The trends at each time point are estimated using the information available up to the previous period. If the deviation in Equation (6) is $e_{m,t} \neq 0$, this indicates that platform m has investors who demand a change in their market share: $E^{(d)}_{m,t} / \overline{E}_t \neq S_{m,t-1}$ at time t.

Thus, a demand for exposure could also be represented by Equation (7), in which percentage price changes p_t, platform-specific controls $\boldsymbol{X}_{m,t}$, and latent demand of an idiosyncratic nature $\tau_{m,t}$ would be able to induce changes in market share through an assumed linear relationship

$$e_{m,t} \quad = \quad \mathcal{E}_m p_t + \boldsymbol{\delta}'_m \boldsymbol{X}_{m,t} + \tau_{m,t}, \tag{7}$$

where $\mathcal{E}_m$ is the price elasticity of demand for platform m. This means that variations in 1% in price cause variations of $\mathcal{E}_m$% in the exposure of the platforms.

It should be noted that the lowercase notation for price p_t represents a percentage change in price relative to a trend: $p_t = (P_t - \overline{P}_t) / \overline{P}_t$. So p_t is not a price, but a price variation.[10] Since any trend is measurable with the information present in the data up to the previous period, if the best estimate of the price trend is P_{t-1}, then p_t could be interpreted as the return of the stock. In this way, the demand for exposure is described by a model of changes in investor exposure on each platform. While in the market for direct stock ownership, the price adjusts due to the quantities traded, the exposure market studies the variability of prices.

In principle, in addition to the set of platforms, there are the general equilibrium effects of the rest of the market. Since the exposures supplied by the set of platforms have risk premiums associated with the multiplier effects,[11] it is always possible to assume that the demand shocks from the rest of the market are negligible. I assume that the rest of the market has exposure shocks $e_t = 0$ and is therefore a price taker.

The model allows for controls for each platform. An example of a control for investor demand would be the percentage change in the value of the platform's assets under management $\text{aum}_{m,t}$ caused by investor flows.[12] Other examples of signals would be expected stock returns and the correlation between those returns. Finally, latent demand would include the vector of common demand shocks across platforms.

Therefore, the total demand for exposure of the set of platforms $E^{(d)}_t$ is the total demand for exposure that changes over time only due to changes in exposure, weighted by the market shares of each platform,

$$E^{(d)}_t \quad = \quad \sum_m E^{(d)}_{m,t} = \overline{E}_t \left(1 + \sum_m S_{m,t} E_{m,t}\right) = \overline{E}_t \left(1 + e^{(s)}_t\right), \tag{8}$$

and so the aggregate demand curve, minus the specific controls of each platform, is the relationship between percentage changes in stock prices and percentage changes in total market exposure, represented

[9] When logarithms of the variables are used, lowercase letters represent random variables that can also be interpreted as proportional changes.

[10] Which may not be the best possible notation. Since I will use lowercase to represent any percentage change, I prefer to keep this choice for p_t.

[11] I discuss the multiplier effects in detail in Ferraresi and Urbano (2025)

[12] Note that this term contains the prices of all assets in the platform's funds.

by Equation (9)

$$\begin{aligned}
\frac{E_t^{(d)}}{\overline{E}_t} &= 1 + \sum_m S_{m,t} \left(\mathcal{E}_m p_t + \tau_{m,t} \right) \\
\frac{E_t^{(d)}}{\overline{E}_t} &= 1 + \sum_m S_{m,t} \mathcal{E}_m p_t + \sum_m S_{m,t} \tau_{m,t} \\
\frac{E_t^{(d)}}{\overline{E}_t} &= 1 + \mathcal{E} p_t + \tau_t^{(s)} \\
e_t^{(s)} &= \mathcal{E} p_t + \tau_t^{(s)}, \qquad (9)
\end{aligned}$$

where the price elasticity of aggregate demand $\mathcal{E}$ and the latent shocks to aggregate demand $\tau_t^{(s)}$ are $\mathcal{E} = \sum_m S_{m,t} \mathcal{E}_m$ and $\tau_t^{(s)} = \sum_m S_{m,t} \tau_{m,t}$.

Finally, the other side of the exposure market is the supply of exposure. Supply is the dilemma between the marginal benefit of the network offering exposure, given by percentage price changes, and the marginal cost to platforms of expanding or contracting their network of funds, given by negative supply shocks. Thus, the total exposure supplied $E_t^{(o)}$ is given by

$$E_t^{(o)} = \overline{E}_t \left(1 + \frac{p_t - u_t}{\psi} \right), \qquad (10)$$

where the aggregate parameters $\mathcal{E}$ and ψ are not identified due to the correlation between the supply and price shock $\mathrm{E}[u_t p_t] \neq 0$, which would not allow the parameter ψ to be identified. Note that even if $\mathrm{E}[p_t \tau_t^{(s)}] = 0$ in Equation (9), the parameter $\mathcal{E}$ of the price elasticity of aggregate demand is not identifiable due to the simultaneity between supply and aggregate demand.

Given the demand, Equation (8), and supply, Equation (10), the next step is to study the equilibrium price and aggregate exposure. To discuss the problems related to the non-identification of the aggregate elasticity parameters $\mathcal{E}$ and ψ, I will present the idea of a granular instrumental variable proposed by Gabaix and Koijen (2024), already adapted to the context of an equity exposure market.

4.2 Granular-Based Pricing

The price equation for a stock in the exposure market is derived from the equilibrium. The idea is to explain how the change in the stock price is affected by changes in the exposure of the platforms. To determine a price change equation, consider the equilibrium in the exposure market with the pricing equation (11)

$$\begin{aligned}
\overline{E}_t \left(1 + e_t^{(s)} \right) &= \overline{E}_t \left(1 + \frac{p_t - u_t}{\psi} \right) \\
p_t &= \psi e_t^{(s)} + u_t. \qquad (11)
\end{aligned}$$

The change in stock price p_t depends on both the change in aggregate exposure $e_t^{(s)}$ and the supply shock u_t. A fundamental problem is that it is not possible to estimate the parameter ψ in Equation (11) by OLS, because

$$\mathrm{E}[u_t e_t^{(s)}] \neq 0. \qquad (12)$$

The estimation of the aggregate parameters $\mathcal{E}$ and ψ is based on the idea of a granular instrumental variable in Gabaix and Koijen (2024). The main assumption of a granular instrument is that idiosyncratic shocks are independent of aggregate shocks. In principle, a linear combination of idiosyncratic

shocks affecting the platforms $\tau_t^{(s)} = \sum_m S_{m,t}\tau_{m,t}$ could be used as an instrumental variable, since by hypothesis $\mathrm{E}[\tau_t^{(s)} u_t] = 0$.[13].

The example discussed by Gabaix and Koijen (2024) in Appendix $D.2$ shows how to recover the parameters of aggregate elasticities in this simpler case, assuming the granular variable $\tau_t^{(s)}$. The expressions for aggregate demand and supply are rewritten in terms of Equations (13) and (14),

$$\begin{aligned} \frac{E_t^{(d)}}{\overline{E}_t} - 1 &= \mathcal{E} p_t + \tau_t^{(s)} \\ e_t^{(s)} &= \mathcal{E} p_t + \tau_t^{(s)}, \end{aligned} \tag{13}$$

$$\begin{aligned} \frac{E_t^{(o)}}{\overline{E}_t} - 1 &= \frac{1}{\psi} p_t - \frac{1}{\psi}\epsilon_t \\ e_t^{(o)} &= \frac{1}{\psi} p_t - \frac{1}{\psi}\epsilon_t, \end{aligned} \tag{14}$$

therefore, it is possible to obtain "pass-through" expressions between linear combinations of idiosyncratic platform shocks and prices in (15)) and total equilibrium exposure in (16). Starting from the market equilibrium,

$$\begin{aligned} e_t^{(s)} &= e_t^{(o)} \\ \mathcal{E} p_t + \tau_t^{(s)} &= \frac{1}{\psi} p_t - \frac{1}{\psi}\epsilon_t \\ \tau_t^{(s)} + \frac{1}{\psi}\epsilon_t &= \left(\frac{1}{\psi} - \mathcal{E}\right) p_t \\ p_t &= \mu^{(p)} \tau_t^{(s)} + \epsilon_t^{(p)}, \qquad (15) \\ e_t^{*} &= \mu^{(e)} \tau_t^{(s)} + \epsilon_t^{(e)}, \qquad (16) \end{aligned}$$

where $\mu^{(p)} = \psi \,/\, (1 - \psi\mathcal{E})$ is the effect of shocks on p_t, and $\mu^{(e)} = 1 \,/\, (1 - \psi\mathcal{E})$ is the effect of shocks on equilibrium total exposure changes e_t^*. Note that the errors $\epsilon_t^{(p)} = \left(\mu^{(p)} \,/\, \psi\right)\epsilon_t$ and $\epsilon_t^{(e)} = (\mu^{(p)} \,/\, \psi^2 - 1 \,/\, \psi)\epsilon_t$ depend only on the aggregate supply shock ϵ_t and the parameters, and thus the parameters $\mu^{(p)}$ and $\mu^{(e)}$ are identified under the assumption that $\tau_t^{(s)} \perp\!\!\!\perp \epsilon_t$. The recovery of the original aggregate parameters of the supply and demand curves is possible because Equation (15) is the first stage of a two-stage instrumental variable estimation to recover the parameter ψ in (11), and thus the parameter $\mathcal{E}$ is also identified.[14]

More caution is needed when using a linear combination of idiosyncratic platform shocks as an instrument. The price change is not the only aggregate shock that can explain the demand for exposure. If other possibilities of aggregate shocks are considered, it is necessary to evaluate the effect of these other shocks on the linear combinations made with the observations of exposures in order to obtain only idiosyncratic shocks in the variable to be used as an instrument.

Gabaix and Koijen (2024) present in detail how to construct a granular variable that is a linear combination of idiosyncratic shocks, assuming other common aggregate shocks across countries in the oil market. I will briefly discuss the idea of the article, which has already been adapted to the context of the exposure market. Consider the exposure of a platform to a single stock without taking into account other controls, written as

$$e_{m,t} = \mathcal{E}_m p_t + \boldsymbol{\lambda}' \boldsymbol{\eta}_t + \tau_{m,t},$$

[13]Note that a bit more care is needed when making such a statement because the responses to aggregate shocks $\boldsymbol{\lambda}_n' \boldsymbol{\eta}_t$ are included in the idiosyncratic shocks of each platform. In addition, a single combination of idiosyncratic errors would not be sufficient, as the instrument must depend on the observed variables

[14]under the condition that $\psi\mathcal{E} \neq 1$

with $\boldsymbol{\lambda}' \in \mathbb{R}^{1\times P}$ and $\boldsymbol{\eta}_t \in \mathbb{R}^P$ representing the P possibilities of joint demand shocks. The idea behind the construction of an instrumental variable is the existence of a vector of weights $\boldsymbol{\Gamma}_t \in \mathbb{R}^M$, which would make it possible to construct a random variable z_t from idiosyncratic shocks only

$$z_t = \boldsymbol{\Gamma}_t' \boldsymbol{e}_t, \tag{17}$$

with vector $\boldsymbol{e}_t \in \mathbb{R}^M$, which is composed of the percentage changes in exposure to a stock on each platform in the set.[15]. Thus, consider the factor loading matrix $\boldsymbol{\Lambda} \in \mathbb{R}^{M\times P}$, where each loading is the sensitivity of each platform to common demand shocks in the stock $\boldsymbol{\eta}_t \in \mathbb{R}^P$. The two hypotheses for z_t to be a valid instrument are

$$\begin{aligned} \boldsymbol{\Gamma}_t' \boldsymbol{\Lambda} &= \mathbf{0}, \\ \boldsymbol{\Gamma}_t' \boldsymbol{\iota}_M &= 0, \end{aligned}$$

because if the price elasticity of demand is the same for each platform, it is possible to verify that

$$\begin{aligned} z_t &= \boldsymbol{\Gamma}_t' \left(\boldsymbol{\iota}_M \mathcal{E} p_t + \boldsymbol{\Lambda} \boldsymbol{\eta}_t + \boldsymbol{\tau}_t \right) \\ &= \boldsymbol{\Gamma}_t' \boldsymbol{\tau}_t, \end{aligned}$$

with $\boldsymbol{\tau}_t \in \mathbb{R}^M$ are the idiosyncratic shocks of each platform in the stock n. Therefore, the variable z_t can be used to identify $\mu^{(p)}$ and $\mu^{(e)}$ in Equations (15) and (16). Gabaix and Koijen (2024) discuss in more detail how to obtain a vector of optimal weights Γ_t^* based on two-stage statistical procedures since it will be essential to take into account estimates of loadings and factors.

Assuming only a single common factor η_t and a load vector $\boldsymbol{\Lambda} = \boldsymbol{\iota}_M$, i.e. a time-fixed effect has absorbed the aggregate shocks, it is possible to use a known form for the vector of weights $\boldsymbol{\Gamma}_t$

$$\boldsymbol{\Gamma}_t = \boldsymbol{S}_t - \frac{1}{M} \boldsymbol{\iota}_M, \tag{18}$$

with $\boldsymbol{S}_t \in \mathbb{R}^M$. The weights are the difference between the arithmetic and share-weighted averages of the platforms. Gabix and Koijen (2024) show that assuming idiosyncratic shocks are independent across platforms, $\forall m \neq j, \tau_{m,t} \perp\!\!\!\perp \tau_{j,t}$ and homoscedastic $\sigma_\tau^2 > 0$, the vector of weights given by Equation (18) is optimal.[16]

The granular variable z_t is built to identify the aggregate parameters $\mathcal{E}$ and ψ. How can we identify specific parameters for each platform using the same idea? Assuming that the idiosyncratic shocks of different platforms are independent of each other, each platform can be reinterpreted as an "asset insulator" (Chodorow-Reich, Ghent, and Haddad, 2021). In an environment of competition for exposure, where each platform responds strategically to the others, platforms compete for quantity, and the most relevant variable for understanding their decisions would be the spillovers between them.

Using the granular variable z_t to identify $\mu^{(p)}$ and $\mu^{(e)}$ in Equations (15) and (16) opens up the possibility of using only part of the variability of z_t to identify the effects of other platforms $\mu_{\neq m}^{(p)}$ and $\mu_{\neq m}^{(e)}$ on price changes and exposure changes. In this way, it is always possible to subtract a platform's own idiosyncratic shock and thus obtain a new granular variable, which I call the granular network variable,

$$\begin{aligned} g_{m,t} &= \sum_{j\neq m} \left(\tilde{S}_{j,t} - \frac{1}{M-1} \right) e_{j,t} \\ g_{m,t} &= \boldsymbol{\Gamma}_{\neq m,t}' \boldsymbol{e}_{\neq m,t}, \end{aligned} \tag{19}$$

[15]The above expression could be identified by $z_{n,t} = \boldsymbol{\Gamma}_{n,t}' \boldsymbol{e}_{n,t}$, since it depends on which stock n we are evaluating. To be consistent with my notation, I prefer not to make the dependence on variables or matrices explicit so as not to overwhelm the reader

[16]Lowest asymptotic variance for the aggregate parameter estimates $\mathcal{E}$ and ψ among any other vector of weights.

where $\mathbf{\Gamma}_{\neq m,t} = \tilde{\boldsymbol{S}}_t - (1 \,/\, (M-1))\boldsymbol{\iota}_{M-1} \in \mathbb{R}^{M-1}$, $\boldsymbol{e}_{\neq m,t} \in \mathbb{R}^{M-1}$ and $\sum_{j \neq m} \tilde{S}_{j,t} = 1$. In this way, the granular network variable $g_{m,t}$ makes it possible to study the effects on price changes and changes in aggregate exposures caused by other platforms from the perspective of the platform m.

Before presenting the model of exposure on the part of the investment platforms, in Section 4.4, I address the fact that the exposures are of equities, and, therefore, it is necessary to include peculiarities related to risk and return in the discussion.

4.3 Factor-Based Exposure Portfolio

Is it possible to relate the demand for exposure given by Equation (7) to stock portfolio models? It is common knowledge in finance that any model that explains the percentages of stocks that investors choose to hold in their investment portfolios is a demand model. The first example in finance literature was Markowitz (1952), in which the demand of a representative risk-averse investor with quadratic utility would choose an optimal portfolio that depends on the correlations between stock returns and the expected values of returns at each point in time.

Portfolio models are dynamic. It is unreasonable to assume an independent choice process over time. Therefore, rewriting a static portfolio problem in terms of percentage changes is the simplest way to represent it dynamically. To arrive at a portfolio model associated with percentage changes in exposures, prices, and assets under management, we start with Markowitz's solution of an investment platform with constant risk aversion and quadratic preferences,

$$\boldsymbol{\omega}_{m,t} = \frac{1}{\gamma_m} \mathbf{\Sigma}_{r,t}^{-1} \boldsymbol{\mu}_{m,t}, \tag{20}$$

where the portfolio $\boldsymbol{\omega}_{m,t} \in \mathbb{R}^N$ is a vector of exposure weights. The weights are explained by the risk aversion γ_m, the expected value of the stock returns $\mu_m \in \mathbb{R}^N$ of each platform, and the covariance matrix of returns common to all platforms $\mathbf{\Sigma}_{r,t} \in \mathbb{R}^{N \times N}$.

To obtain expressions for the expected returns and an analytical solution for the inverse of the covariance matrix, I use the APT risk model proposed by Ross (1976). In that model, stock returns $\boldsymbol{r}_t \in \mathbb{R}^N$ are explained by a vector of risk factors $\boldsymbol{\zeta}_t \in \mathbb{R}^K$ with premiums given by the vector $\mathrm{E}[\boldsymbol{\zeta}_t] \in \mathbb{R}^K$. The relationship between returns and these factors is given by a pricing error vector, $\boldsymbol{\alpha} \in \mathbb{R}^N$, and a risk factor loading matrix, $\boldsymbol{\beta}' \in \mathbb{R}^{N \times K}$. Therefore, the expressions representing the APT model are

$$\begin{aligned}
\boldsymbol{r}_{t+1} &= \boldsymbol{\alpha} + \boldsymbol{\beta}' \boldsymbol{\zeta}_{t+1} + \boldsymbol{\epsilon}_{t+1}, \\
\boldsymbol{r}_{t+1} | \boldsymbol{\zeta}_{t+1} &\sim \mathcal{N}\left(\boldsymbol{\alpha} + \boldsymbol{\beta}' \boldsymbol{\zeta}_{t+1}; \sigma_\epsilon^2 \boldsymbol{I}_N\right), \\
\boldsymbol{\zeta}_t &\sim \mathcal{N}\left(\mathrm{E}[\boldsymbol{\zeta}_t]; \mathbf{\Sigma}_{\zeta,t}\right), \\
\boldsymbol{\epsilon}_t &\sim \mathcal{N}\left(\mathbf{0}; \sigma_\epsilon^2 \boldsymbol{I}_N\right), \\
\boldsymbol{\zeta}_t &\perp\!\!\!\perp \boldsymbol{\epsilon}_t,
\end{aligned}$$

and thus the expected value of the returns $\mathrm{E}_t^{(m)}[\boldsymbol{r}_{t+1}] = \boldsymbol{\mu}_m$, and the covariance matrix of the returns $\mathbf{\Sigma}_{r,t}$, respectively,

$$\boldsymbol{\mu}_m = \boldsymbol{\alpha}_m + \boldsymbol{\beta}' \,\mathrm{E}[\boldsymbol{\zeta}_t], \tag{21}$$

$$\mathbf{\Sigma}_{r,t} = \boldsymbol{\beta}' \mathbf{\Sigma}_{\zeta,t} \boldsymbol{\beta} + \sigma_\epsilon^2 \boldsymbol{I}_N, \tag{22}$$

where the vector of pricing errors $\boldsymbol{\alpha}$ is given by the weights of the average platform shares of each

pricing error $\boldsymbol{\alpha} = \boldsymbol{S}'\boldsymbol{\alpha}_m$.[17] Note that if the vector of risk factors is correctly specified, the value of the pricing error vector is $\boldsymbol{\alpha} = \mathbf{0}$, and a zero-sum game can represent the investment platforms interactions.

To obtain an analytical form for the inverse of the covariance of returns, I use Woodbury's identity

$$(\boldsymbol{A}+\boldsymbol{UCV})^{-1} = \boldsymbol{A}^{-1} - \boldsymbol{A}^{-1}\boldsymbol{U}\left(\boldsymbol{C}^{-1}+\boldsymbol{V}\boldsymbol{A}^{-1}\boldsymbol{U}\right)^{-1}\boldsymbol{V}\boldsymbol{A}^{-1},$$

with the following matrix definitions,

$$\begin{aligned}
\boldsymbol{A} &= \sigma_\epsilon^2 \boldsymbol{I}_N \in \mathbb{R}^{N\times N},\\
\boldsymbol{C}_t &= \boldsymbol{\Sigma}_{\zeta,t} \in \mathbb{R}^{K\times K},\\
\boldsymbol{U} &= \boldsymbol{\beta} \in \mathbb{R}^{N\times K},\\
\boldsymbol{V} &= \boldsymbol{\beta}' \in \mathbb{R}^{K\times N},
\end{aligned}$$

so, I rewrite Markowitz's solution in the form of a dynamic factor model, using a factor loading matrix per platform $\boldsymbol{\Lambda}_m \in \mathbb{R}^{N\times K}$, a vector of factors $\boldsymbol{\kappa}_{m,t} \in \mathbb{R}^K$, and a platform fixed effect $\boldsymbol{\xi}_m \in \mathbb{R}^N$,

$$\begin{aligned}
\frac{1}{\gamma_m}\boldsymbol{\Sigma}_{r,t}^{-1}\boldsymbol{\mu}_m &= \frac{1}{\gamma_m}\left[\sigma_\epsilon^2\boldsymbol{I}_N + \boldsymbol{\beta}\boldsymbol{\Sigma}_{\zeta,t}\boldsymbol{\beta}'\right]^{-1}\boldsymbol{\mu}_m\\
&= \frac{1}{\gamma_m}\left\{\sigma_\epsilon^{-2}\boldsymbol{I}_N - \sigma_\epsilon^{-2}\boldsymbol{I}_N\boldsymbol{\beta}\left[\boldsymbol{\Sigma}_{\zeta,t}^{-1}+\boldsymbol{\beta}'\sigma_\epsilon^{-2}\boldsymbol{I}_N\boldsymbol{\beta}\right]^{-1}\boldsymbol{\beta}'\sigma_\epsilon^{-2}\boldsymbol{I}_N\right\}\boldsymbol{\mu}_m\\
&= \frac{1}{\gamma_m\sigma_\epsilon^2}\boldsymbol{\mu}_m - \frac{1}{\gamma_m\sigma_\epsilon^2}\boldsymbol{\beta}\left[\sigma_\epsilon^2\boldsymbol{\Sigma}_{\zeta,t}^{-1}+\boldsymbol{\beta}'\boldsymbol{\beta}\right]^{-1}\boldsymbol{\beta}'\boldsymbol{\mu}_m\\
&= \frac{1}{\gamma_m\sigma_\epsilon^2}\boldsymbol{\mu}_m - \frac{1}{\gamma_m\sigma_\epsilon^2}\boldsymbol{\beta}\boldsymbol{\kappa}_{m,t}\\
&= \boldsymbol{\xi}_m + \boldsymbol{\lambda}_m\boldsymbol{\kappa}_{m,t},
\end{aligned}$$

with the platform fixed effect, factor loadings, and dynamic factors being written, respectively, by

$$\boldsymbol{\xi}_m = \frac{1}{\gamma_m\sigma_\epsilon^2}\boldsymbol{\mu}_m = \boldsymbol{A}_m\boldsymbol{\mu}_m, \tag{23}$$

$$\boldsymbol{\lambda}_m = -\frac{1}{\gamma_m\sigma_\epsilon^2}\boldsymbol{\beta} = -\boldsymbol{A}_m\boldsymbol{\beta}, \tag{24}$$

$$\boldsymbol{\kappa}_{m,t} = \left[\sigma_\epsilon^2\boldsymbol{\Sigma}_{\zeta,t}^{-1}+\boldsymbol{\beta}'\boldsymbol{\beta}\right]^{-1}\boldsymbol{\beta}'\boldsymbol{\mu}_m = \boldsymbol{B}_t\boldsymbol{\mu}_m, \tag{25}$$

with the linear transformation matrices $\boldsymbol{A}_m \in \mathbb{R}^{N\times N}$ and $\boldsymbol{B}_t \in \mathbb{R}^{K\times N}$ of the above expressions written as

$$\boldsymbol{A}_m = \frac{1}{\gamma_m\sigma_\epsilon^2}\boldsymbol{I}_N, \tag{26}$$

$$\boldsymbol{B}_t = \left[\sigma_\epsilon^2\boldsymbol{\Sigma}_{\zeta,t}^{-1}+\boldsymbol{\beta}'\boldsymbol{\beta}\right]^{-1}\boldsymbol{\beta}'. \tag{27}$$

Therefore, based on a risk model and beliefs about stock pricing errors, it is possible to define two linear transformation matrices, $\boldsymbol{A}_m$ and $\boldsymbol{B}_t$, which summarize the expected returns, taking into account a fixed effect per platform (average latent demand for exposure) and a dynamic factor model (due to the volatility of risk factors). Thus, Markowitz's equation (20) can be written in the form of a dynamic factor model:

$$\boldsymbol{\omega}_{m,t} = \boldsymbol{\xi}_{m,t} + \boldsymbol{\lambda}_m\boldsymbol{\kappa}_{m,t}.$$

[17] Although the shares are time-dependent, given an order of platforms by size that is constant over time, considering the average shares of platforms is the simplest way to arrive at a constant $\boldsymbol{\alpha}$ parameter

For didactic purposes, consider only the weight of a single exposure in the portfolio and a single-factor dynamic factor model. To rewrite the weight vector as a number

$$\omega_{m,t} \quad = \quad \xi_{m,t} + \lambda_m \kappa_{m,t},$$

and, rewriting the left side of the above expression in terms of percentage changes in exposure, price, and AUM of the platform, it is possible to obtain an expression for the demand for factor-based exposure

$$\begin{aligned}
\frac{E_{m,t} P_{m,t}}{\text{AUM}_{m,t}} &= \xi_m + \lambda_m \kappa_{m,t} \\
\frac{\overline{E}_{m,t}\,(1+e_{m,t})\,\overline{P}_{m,t}\,(1+p_{m,t})}{\overline{\text{AUM}}_{m,t}\,(1+\text{aum}_{m,t})} &= \xi_m + \lambda_m \kappa_{m,t} \\
\overline{\omega}_m\,(1+e_{m,t}+p_{m,t}-\text{aum}_{m,t}) &\simeq \xi_m + \lambda_m \kappa_{m,t} \\
e_{m,t}+p_{m,t} &\simeq \frac{\xi_m + \lambda_m \kappa_{m,t}}{\overline{\omega}_m} - 1 + \text{aum}_{m,t} \\
e_{m,t} &= \xi_m - p_{m,t} + \lambda_m \kappa_{m,t} + \tau_{m,t}. \qquad (28)
\end{aligned}$$

where $\overline{\omega}_{m,t} = \overline{\omega}_m$ represents a constant base scenario, e.g., diversification with equally weighted $1/N$ weights.[18]

Note that the term $(\xi_m + \lambda_m \kappa_{m,t}) \,/\, \overline{\omega}_m - 1$ represents the percentage difference between the weights chosen by a factor model and the weights chosen by a base scenario $\overline{\omega}_m$. This weight of the base scenario can be included in the fixed effect of the platform and factor loading simply by substituting a new coefficient of risk aversion of the platform $\gamma_m = \overline{\omega}\gamma_m^*$ in Equations (23) and (24). In addition, because the weights chosen by the factor model are very extreme,[19] I include a latent demand term $\tau_{m,t}$ that incorporates changes in assets under management so that the resulting percentage changes on the left-hand side of Equation (28) do not become so extreme.

Finally, let us return to the case of a platform with a portfolio exposed to N stocks. The demand (in percentage terms) for this platform's exposure can be written as

$$\boldsymbol{e}_{m,t} \quad = \quad \boldsymbol{\xi}_m - \boldsymbol{p}_t + \boldsymbol{\lambda}_m \boldsymbol{\kappa}_{m,t} + \boldsymbol{\tau}_{m,t}, \qquad (29)$$

where $\boldsymbol{\tau}_{m,t} \sim \mathcal{N}\left(\mathbf{0}; \sigma_m^2 \boldsymbol{I}_N\right)$ and $\boldsymbol{\kappa}_{m,t} \perp\!\!\!\perp \boldsymbol{\tau}_{m,t}$, since I assume an exact decomposition of latent demand. Platform exposure changes $\boldsymbol{e}_{m,t}$, stock price changes $\boldsymbol{p}_t$, and latent demand shocks $\boldsymbol{\tau}_{m,t}$ are decided simultaneously. Equation (29) reflects the fact that these simultaneous decisions can be written by a model of dynamic factors that refer only to percentage changes.

In the last subsection, I discuss which exposure model would be most appropriate for investment platforms, given that exposure changes and their effect on prices are the result of equilibrium in the exposure market. The main idea to be discussed is the importance of spillover effects between platforms, both for demanders and providers of exposure.

4.4 Exposure-Based Demand

Retail investors seek exposure through investment platforms that form an exposure market. Here, investment platforms behave as an oligopoly of exposure providers. The choices these platforms make

[18] The fact that the three trends are time-dependent while the base scenario is not is consistent with the fact that the three stochastic processes are cointegrated.

[19] A feature inherited from the Markowitz model

regarding exposure influence stock prices and are determined simultaneously across platforms. Because price adjustments happen more rapidly than exposure changes, these investment platforms engage in competition for exposures that resemble a Cournot oligopoly.

Suppose that the set of platforms provides a stock exposure that is considered homogeneous by investors. The platforms simultaneously choose which exposures to produce. Thus, the stock market adjusts prices so that the aggregate demand for exposures by investors equals the aggregate supply of exposures by platforms. In this way, a residual demand curve becomes of great interest for the platforms. The residual demand corresponds to all possible combinations of exposure and equilibrium price that a platform can observe, taking into account the exposures of the other platforms. Assuming the existence of marginal cost and revenue curves, it is possible to obtain reaction curves between platforms that represent the optimal supply decisions of each platform depending on changes in the exposures of the other platforms. The granular network variable $g_{m,t}$ constructed to assess the interactions of the platform would identify both the slopes of these response curves and the slopes of the residual demand of each platform.

I propose a dynamic model of demand for exposure across various platforms, where investors do not focus on market prices but rather on the impact of latent demand shocks between these platforms. If the stock were chosen by a mutual fund manager, it would be logical to treat stock prices as their primary decision variable. The key point is that, from the perspective of exposure choices, the opportunity cost of resource allocation decisions depends on how each platform responds to shocks from others. Thus, exposure-based demand is a reduced form, as my goal is not to estimate the price elasticity of demand in this exposure market. The allocation of resources to investment platforms should be viewed as a form of passive investment, where the exposure portfolio of the fund network can be interpreted as an investment strategy. This concept has been elaborated in Ferraresi and Urbano (2025) by comparing investment platforms with ETFs. It is possible to develop strategies that focus on exposure to these platforms. These strategies aim to exploit the risk premium associated with the multiplier network of the platforms (Ferraresi and Urbano, 2025).

4.4.1 State Space Model

Exposure-based demand is the state space model representing retail investor demand from a specific platform

$$\boldsymbol{e}_{m,t} = \boldsymbol{\xi}_m + \boldsymbol{\delta}_m \boldsymbol{g}_{m,t} + \boldsymbol{\lambda}_m \boldsymbol{\kappa}_{m,t} + \boldsymbol{\tau}_{m,t}, \tag{30}$$

$$\boldsymbol{\kappa}_{m,t+1} = \boldsymbol{\Phi}_m \boldsymbol{\kappa}_{m,t} + \boldsymbol{\nu}_{m,t}, \tag{31}$$

where $\boldsymbol{\xi}_m \in \mathbb{R}^N$, $\boldsymbol{\lambda}_m \in \mathbb{R}^{N \times K}$, and $\boldsymbol{\kappa}_{m,t} \in \mathbb{R}^K$ given by Equations (23), (24), and (25), respectively. The transition matrix in the dynamics of the factors is $\boldsymbol{\Phi} \in \mathbb{R}^{K \times K}$. I define the exposure elasticity of demand as the diagonal elements of the matrix $\boldsymbol{\delta}_m \in \mathbb{R}^{N \times N}$. The cross-exposure elasticity is zero as an assumption. This exposure elasticity parameter is related to the vector of inter-platform spillovers, also known as the granular network variable $\boldsymbol{g}_{m,t} \in \mathbb{R}^N$ represented by,

$$\boldsymbol{g}_{m,t} = \begin{pmatrix} g_{m,t}^{(1)} \\ \vdots \\ g_{m,t}^{(N)} \end{pmatrix} = \begin{pmatrix} \sum_{j \neq m} \left(\tilde{S}_{j,t}^{(1)} - \frac{1}{M-1} \right) e_{j,t}^{(1)} \\ \vdots \\ \sum_{j \neq m} \left(\tilde{S}_{j,t}^{(N)} - \frac{1}{M-1} \right) e_{j,t}^{(N)} \end{pmatrix},$$

and, finally, the following hypotheses about errors,

$$\begin{aligned}\boldsymbol{\tau}_{m,t} &\sim \mathcal{N}\left(\mathbf{0}; \sigma_m^2 \boldsymbol{I}_N\right), \\ \boldsymbol{\nu}_{m,t} &\sim \mathcal{N}\left(\mathbf{0}; \boldsymbol{I}_K\right), \\ \boldsymbol{\tau}_{m,t} &\perp\!\!\!\perp \boldsymbol{\nu}_{m,t}.\end{aligned}$$

The model to be estimated, using a methodology discussed in the next chapter and the results discussed in the empirical analysis chapter, is an exposure demand in the form of a state-space model with only a single factor, $K = 1$. I represent, for one stock, the observation and state transition equations, respectively, by:

$$e_{m,t} = \xi_m + \delta_m g_{m,t} + \lambda_m \kappa_{m,t} + \tau_{m,t}, \tag{32}$$

$$\kappa_{m,t} = \phi_m \kappa_{m,t-1} + \nu_{m,t}, \tag{33}$$

with the following assumptions about the errors

$$\begin{aligned}\tau_{m,t} &\sim \mathcal{N}\left(0, \sigma_m^2\right), \\ \nu_{m,t} &\sim \mathcal{N}\left(0, 1\right), \\ \tau_{m,t} &\perp\!\!\!\perp \nu_{m,t}.\end{aligned}$$

4.4.2 Exposure Elasticity

The impact of spillovers between platforms is assessed from the point of view of investors in platform m through the use of the parameter δ_m. This parameter represents the gradient of the response curve of the platform m' to unique demand fluctuations originating from other platforms. Consequently, a 1% variation in the granular network metric $g_{m,t}$ results in a δ_m% variation in the demand for exposure on the platform m. Since idiosyncratic shocks across platform demands are presumed to be independent, the parameter δ_m is a causal parameter linked to the granular network variable $g_{m,t}$.

I define the parameter δ_m as the exposure elasticity of demand. The concept of inter-platform exposure elasticity, which refers to how sensitive positions on one platform are to changes in exposures on other platforms, serves as a measure of the interconnectedness and fragility of the network of market-participating platforms. Despite this measure does not directly represent price elasticity, it can be related to broader market variables as price, returns, volatility, and systemic risk. This paper concentrates on replicant risk, a type of systematic risk that I discuss more thoroughly in the empirical analysis section, where I explain how the results for the vectors $\boldsymbol{\delta}_m \in \mathbb{R}^N$ for each platform in the study can be used to uncover replicant behavior.

Some conceptual approaches can already establish links between exposure elasticity and the aggregate market variables mentioned in the previous paragraph.[20] First, there is a relationship between exposure elasticity and equity prices/returns. If the elasticity of exposure across platforms is high, changes in one platform's portfolio can lead to significant adjustments in the portfolios of others. In a market context, this can amplify price movements. For example, if one large investment platform reduces its exposure to a stock, others affected by that move may also reduce their exposure, creating a herding effect.

This aggregated behavior can increase the selling pressure, leading to higher price declines or more volatile returns. Thus, estimates of inter-platform exposure elasticities can reveal an important

[20] price, return, volatility, and systemic risk

transmission mechanism for the propagation of platform behavior to prices and returns. High-exposure elasticities may be associated with replicant behavior. This type of behavior is often associated with more pronounced up/down cycles, as everyone tends to enter and exit the asset at the same time, affecting returns and potentially creating price distortions relative to fundamentals. The relationship between elasticity-exposure and volatility is important because of the possibility of stock price shocks being amplified. If the network of funds is highly sensitive internally (i.e., if there is high elasticity exposure), an initial shock (e.g., a price drop due to an idiosyncratic event) may be amplified as the other platforms adjust their positions, increasing volatility. The volatility of the stock would depend not only on news or external shocks but also on the internal dynamics of the fund network. The greater the sensitivity, the greater the likelihood of endogenous fluctuations, i.e. volatility generated by the platforms' own architecture and reallocation decisions.

Elasticity exposure can also be interpreted as a measure of the fragility of the financial network. By relating it to traditional risk measures such as Value at Risk (VaR), Expected shortfall, or stress tests, it is possible to see whether platforms with greater internal sensitivity would need to hold larger capital buffers, or whether a given shock would have a greater aggregate impact, leading to higher levels of systemic risk. In this way, the elasticity exposure acts as a link between the microstructural dimension (fund portfolios and their mutual sensitivity) and the macro dynamics of the market (prices, volatility, and systemic risk).

There exists a link between exposure elasticity and systemic risk. Exposure elasticity serves as a measure of the contagiousness of positions among participants. Networks characterized by high exposure elasticity are more prone to contagion, where a disturbance on one platform can significantly impact the exposures of other platforms, potentially leading to systemic risk. Consequently, assessing exposure elasticity allows us to gauge the financial system's susceptibility to localized shocks. This paper examines a single aspect of exposure elasticity: when exposure elasticity is positive within an investment platform, this procyclical tendency is labeled as a replicant behavior.

In situations where platforms are unable to synchronize their exposure choices over time, copying the behavior of others is disadvantageous, as it subjects the platform to counterpart risk. The key insight from this study is that in scenarios where platforms engage in coordination, the risk previously termed as counterparty risk is now referred to as replicant risk. Replicant risk refers to the unpredictability associated with the herding of other platforms within a coordinated environment. To be able to measure this risk, it is necessary to infer the parameters of this model for each platform in the sample. This inference problem will be discussed in the next section.

5 High-Dimensional Inference of Exposure Based Demand

The exposure-based demand model, as formulated by Equations (30) and (31), is assessed for each investment platform using Bayesian inference techniques. Given that the exposure portfolio $\boldsymbol{e}_{m,t} \in \mathbb{R}^N$ of a platform is N-dimensional, the parametric space can be viewed as being high-dimensional, leading to potentially hundreds of parameters in the inference process. The parameters that require estimation are the elements of Θ_m-set

$$\begin{aligned} \Theta_m &= \mathcal{P}_m \cup \{\boldsymbol{\kappa}_{m,1}; \ldots; \boldsymbol{\kappa}_{m,T}\} \\ \Theta_m &= \{\boldsymbol{\xi}_m, \boldsymbol{\delta}_m; \sigma_m^2; \boldsymbol{\Phi}_m; \boldsymbol{\lambda}_m; \boldsymbol{\kappa}_{m,1}; \ldots; \boldsymbol{\kappa}_{m,T}\} \end{aligned}$$

where the elements of the parameters set $\mathcal{P}$ are $\boldsymbol{\xi}_m; \boldsymbol{\delta}_m \in \mathbb{R}^N$, $\boldsymbol{\Lambda}_m \in \mathbb{R}^{N\times K}$, $\boldsymbol{\Phi}_m \in \mathbb{R}^{K\times K}$, and σ_m^2. The factors $\boldsymbol{\kappa}_{m,1}, \ldots, \boldsymbol{\kappa}_{m,T} \in \mathbb{R}^K$. So $\#\Theta_m = 2N + K(T+N) + K^2 + 1$.[21] Take, for example, the scenario where the average exposure across platforms within the IBRX-100 index investment universe is approximately $N \sim 50$, and the data spans around $T \sim 72$ months. Considering a single-factor model ($K = 1$), with the transition parameter for the factor dynamics and the observational variance error, the parametric space for each platform would encompass ~ 225 dimensions. I will elaborate on how to conduct Bayesian inference on this parameter set for each platform to derive posterior distributions for the model parameters using the framework proposed in Ward et al. (2019) and discussed in Section 5.2.[22] The parameter estimates for 15 platforms are presented in Appendix D.

5.1 Beliefs

In Bayesian approach, there is an analogy between parameter inference and updating prior beliefs in the form of posterior beliefs. This article is the first to deal with investors' demand for exposure to a set of funds in the form of a 'spoke-hub' network. These investment platforms are competing to offer exposure to investors and thus formulate beliefs about the parameters of the model proposed in the previous section. The a prior probability of the set of parameters is considered to be the same for all platforms and is written as

$$\pi\big(\mathcal{P}\big) \quad = \quad \pi(\boldsymbol{\xi})\pi(\boldsymbol{\delta})\pi(\sigma^2)\pi(\boldsymbol{\Phi})\pi(\boldsymbol{\lambda}) \tag{34}$$

the without prior beliefs, with $\boldsymbol{\xi}, \boldsymbol{\delta} \propto 1$. Polson and Scott (2011) argues that the half-Cauchy distribution should replace the inverse-Gamma distribution as a default prior for a top-level scale parameter in Bayesian hierarchical models; at least for cases where a proper prior is necessary, I adopt a similar idea for the error variance $\pi(\sigma^2) \sim \text{Half-}t_3(0, 1)$. For the transition parameter in factor dynamics and loadings, I adopt, respectively, $\pi(\boldsymbol{\Phi}) \sim \mathcal{N}(0, \boldsymbol{I}_K)$, $\pi(\boldsymbol{\lambda}) \sim \mathcal{N}(0, \boldsymbol{I}_K)$ as in traditional Gaussian dynamic factor models.

How different prior specifications can markedly influence the accuracy, computational efficiency, and predictive performance of these models? A central takeaway from the literature is that thoughtfully chosen priors help stabilize parameter estimates and improve forecasts in settings where a multitude of unobserved factors drive financial or macroeconomic time series. For instance, work by Amir-Ahmadi, Matthes, and M.-C. Wang (2020) demonstrates how Inverse Wishart and Inverse Gamma priors reduce forecast errors, whereas Bai and P. Wang (2015) highlight the robustness of Jeffrey's diffuse priors in sampling latent factors for dynamic factor models. Both studies underscore that switching from diffuse or fixed hyperparameter approaches to more carefully tailored distributions can reduce root mean squared errors and increase the reliability of inference.

Various authors investigate the versatility of alternative prior distributions to address different modeling challenges. Kastner, Frühwirth-Schnatter, and Lopes (2016) use normal priors for factor loadings in a multivariate factor stochastic volatility model and show that the resulting parameter estimates converge substantially faster in the Monte Carlo sampling of the Markov chain, sometimes by several orders of magnitude relative to the baseline approaches. Other contributions emphasize how the incorporation of threshold structures or latent processes can further refine these models. Nakajima and M. A. West (2013) propose a latent threshold dynamic factor model that combines normal, uniform, beta, and gamma priors, illustrating that such a rich specification delivers sharper estimates of factor loadings and improved predictions over multiple classes of financial assets.

[21] Factors can also represent uncertainty similar to parameters, so they might be considered parameters as well.

[22] The inference methodology using the Hamiltonian Monte Carlo algorithm is also discussed in Appendix C.

In addition to improvements in estimation accuracy, the choice of priors significantly affects computational considerations. Researchers like Bai and P. Wang (2015) demonstrate how carefully tuned priors can stabilize the Gibbs sampling procedure, enabling tens of thousands of iterations within a matter of minutes. Similarly, Amir-Ahmadi, Matthes, and M.-C. Wang (2020) present a tractable approach to estimating hyperparameters in time-varying parameter models, showing that well-chosen priors can mitigate the inherent computational burden. In contrast, studies that do not specify their priors or rely solely on diffuse assumptions often face slower convergence and occasional identification pitfalls, hinting that thoughtful prior design is central to efficiency gains.

Empirical performance in diverse financial settings provides further support for using carefully selected priors in dynamic factor models. Young et al. (2015) compare weakly informative Bayesian methods with Kalman filters and principal component-based approaches, concluding that weakly informative priors achieve estimation results on par with standard filters when evaluating factor correlations, coverage intervals, or variance decompositions. Meanwhile, research on large-dimensional data, such as exchange rates, commodity prices, or equity returns, shows that normal priors for factor loadings can shorten estimation times even in high dimensions (Kastner, Frühwirth-Schnatter, and Lopes, 2016). Additionally, non-parametric priors, such as the truncated Dirichlet process investigated by Chow et al. (2011), provide flexibility in modeling heterogeneous or shifting market conditions, further demonstrating the breadth of prior choices.

The unifying conclusion is that priors can be tailored to address structural complexities, data peculiarities, and computational burdens, thus improving inference and forecasting. From inverse distributions for covariance parameters to weakly informative or diffuse forms for latent factors, these approaches underscore the essential role of priors in shaping model behavior, ensuring robust inference, and promoting efficient estimation across a wide spectrum of financial applications.

5.2 Updating Beliefs

As all investment platforms have the same a prior beliefs about the parameters, updating these beliefs will be based on each likelihood function. The results of this belief updating process will be done in a single step, resulting in the posterior distributions of the parameters. These posterior distributions of the parameters will be considered hypotheses of common knowledge between the platforms. Therefore, to compute the posterior of the parameters, it is necessary to assume a likelihood function. The log-likelihood function is the Equation (35),

$$\begin{aligned}
\ell\big(\mathbf{e}_{1:T} \mid \mathbf{g}_{1:T}; \boldsymbol{\kappa}_{1:T}; \mathcal{P}\big) &= \log\left\{ \prod_{t=1}^{T} \big(2\pi\big)^{-\frac{N}{2}} \big|\sigma^2 \boldsymbol{I}_N\big|^{-\frac{1}{2}} \exp\left[-\frac{1}{2\sigma^2} \big(\mathbf{e}_t - \mathrm{E}_t[\boldsymbol{e}_t \mid \mathbf{g}_{1:T}; \boldsymbol{\kappa}_{1:T}; \mathcal{P}]\big)' \big(\mathbf{e}_t - \mathrm{E}_t[\boldsymbol{e}_t \mid \mathbf{g}_{1:T}; \boldsymbol{\kappa}_{1:T}; \mathcal{P}]\big) \right] \right\} \\
&= \sum_{t=1}^{T} \left\{ -\frac{N}{2} \log\big(2\pi\big) - \frac{1}{2} \log\big|\sigma^2 \boldsymbol{I}_N\big| - \frac{1}{2\sigma^2} \Big[\big(\mathbf{e}_t - \mathrm{E}_t[\boldsymbol{e}_t \mid \mathbf{g}_{1:T}; \boldsymbol{\kappa}_{1:T}; \mathcal{P}]\big)' \big(\mathbf{e}_t - \mathrm{E}_t[\boldsymbol{e}_t \mid \mathbf{g}_{1:T}; \boldsymbol{\kappa}_{1:T}; \mathcal{P}]\big) \Big] \right\} \\
&= -\frac{NT}{2} \log\big(2\pi\big) - \frac{NT}{2} \log(\sigma^2) - \frac{1}{2\sigma^2} \sum_{t=1}^{T} \big\|\mathbf{e}_t - \mathrm{E}_t[\boldsymbol{e}_t \mid \mathbf{g}_{1:T}; \boldsymbol{\kappa}_{1:T}; \mathcal{P}]\big\|^2, \qquad (35)
\end{aligned}$$

where

$$\begin{aligned}
\mathrm{E}_t[\boldsymbol{e}_t \mid \mathbf{g}_{1:T}; \boldsymbol{\kappa}_{1:T}; \mathcal{P}] &= \boldsymbol{\xi} + \boldsymbol{\delta}\, \mathrm{E}_t[\boldsymbol{g}_t \mid \boldsymbol{\kappa}_t; \mathcal{P}] + \boldsymbol{\lambda}\, \mathrm{E}_t[\boldsymbol{\kappa}_t \mid \mathcal{P}] \qquad (36) \\
\mathrm{E}_t[\boldsymbol{\kappa}_t \mid \mathcal{P}] &= \boldsymbol{\Phi} \boldsymbol{\kappa}_{t-1} \qquad (37) \\
\hat{\mathbb{E}}_t[\boldsymbol{g}_t \mid \boldsymbol{\kappa}_t; \mathcal{P}] &= \frac{1}{T} \sum_{t=1}^{T} \boldsymbol{g}_t. \qquad (38)
\end{aligned}$$

Implicit in Equations (36), (37) and (38) is the assumption of weak exogeneity (Engle, Hendry, and Richard, 1983) of the granular network variable $\boldsymbol{g}_t$ relative of parameters of interest.[23] It is not necessary to specify a complete joint distribution between all random variables $\boldsymbol{e}_{1:T}$, $\boldsymbol{g}_{1:T}$, and $\boldsymbol{\kappa}_{1:T}$, thus excluding likelihood $\ell\big(\boldsymbol{g}_{1:T}|\boldsymbol{\kappa}_{1:T};\mathcal{P}\big)$ the purposes of estimating the parameters of interest it entails no loss of information to confine one ones attention to the conditional distribution.

Adopting the assumption of Gaussian errors $\boldsymbol{\nu}_{m,t} \sim \mathcal{N}\left(\mathbf{0};\boldsymbol{I}_K\right)$ also for the factor dynamics error allows a second log-likelihood function to be written for the factor vector

$$\begin{aligned}
\ell\big(\boldsymbol{\kappa}_{1:T} \mid \mathcal{P}\big) &= \log \prod_{t=1}^{T} \left(2\pi\right)^{-\frac{K}{2}} \exp\left\{ -\frac{1}{2}\big(\boldsymbol{\kappa}_t - \mathrm{E}_t[\boldsymbol{\kappa}_t \mid \mathcal{P}]\big)'\big(\boldsymbol{\kappa}_t - \mathrm{E}_t[\boldsymbol{\kappa}_t \mid \mathcal{P}]\big)\right\} \\
&= \sum_{t=1}^{T}\left\{ -\frac{K}{2}\log\left(2\pi\right) - \frac{1}{2}\big(\boldsymbol{\kappa}_t - \mathrm{E}_t[\boldsymbol{\kappa}_t \mid \mathcal{P}]\big)'\big(\boldsymbol{\kappa}_t - \mathrm{E}_t[\boldsymbol{\kappa}_t \mid \mathcal{P}]\big)\right\} \\
&= -\frac{KT}{2}\log\left(2\pi\right) - \frac{1}{2}\sum_{t=1}^{T}\|\boldsymbol{\kappa}_t - \mathrm{E}_t[\boldsymbol{\kappa}_t \mid \mathcal{P}]\|^2. \qquad (39)
\end{aligned}$$

Given the likelihood functions (35) and (39), the posterior distributions of the parameters are

$$\log \pi\big(\boldsymbol{\kappa}_{1:T} \mid \mathbf{e}_{1:T};\mathbf{g}_{1:T};\mathcal{P}\big) \propto \ell\big(\mathbf{e}_{1:T} \mid \mathbf{g}_{1:T};\boldsymbol{\kappa}_{1:T};\mathcal{P}\big) + \ell\big(\boldsymbol{\kappa}_{1:T} \mid \mathcal{P}\big) \qquad (40)$$

$$\log \pi\big(\mathcal{P} \mid \mathbf{e}_{1:T};\mathbf{g}_{1:T};\boldsymbol{\kappa}_{1:T}\big) \propto \ell\big(\mathbf{e}_{1:T} \mid \mathbf{g}_{1:T};\boldsymbol{\kappa}_{1:T};\mathcal{P}\big) + \ell\big(\boldsymbol{\kappa}_{1:T} \mid \mathcal{P}\big) + \log \pi\big(\mathcal{P}\big). \qquad (41)$$

This study implements an algorithm that diverges from conventional forward filtering backward sampling by avoiding the generation of the sample in sequential blocks.[24] To obtain samples from the non-normalized posteriors in Equations (41) and (40), I use the algorithm implemented in STAN, a software for Bayesian data analysis (Carpenter et al., 2017).

The initial point contains $\boldsymbol{\kappa}_0 \sim \mathcal{N}(\mathbf{0},\boldsymbol{I}_K)$ and for each element in the set Θ, create another set Θ^*, where the elements are $\boldsymbol{\theta}^* \sim \mathcal{N}(\mathbf{0};\boldsymbol{\Sigma}^*)$.

$$\log \mathcal{H}(\Theta;\Theta^*) = -\log \pi\big(\boldsymbol{\kappa}_{1:T} \mid \mathbf{e}_{1:T};\mathbf{g}_{1:T};\mathcal{P}\big) - \log \pi\big(\mathcal{P} \mid \mathbf{e}_{1:T};\mathbf{g}_{1:T};\boldsymbol{\kappa}_{1:T}\big) + \frac{1}{2}\boldsymbol{\theta}^{*\prime}\left(\boldsymbol{\Sigma}^*\right)^{-1}\boldsymbol{\theta}^*, \qquad (42)$$

which allow to write the following Hamiltonian equations

$$\frac{d\boldsymbol{\theta}^*}{dt} = \nabla_{\boldsymbol{\kappa}} \log \pi\big(\boldsymbol{\kappa}_{1:T} \mid \mathbf{e}_{1:T};\mathbf{g}_{1:T};\mathcal{P}\big) + \nabla_{\mathcal{P}} \log \pi\big(\mathcal{P} \mid \mathbf{e}_{1:T};\mathbf{g}_{1:T};\boldsymbol{\kappa}_{1:T}\big), \qquad (43)$$

$$\frac{d\boldsymbol{\theta}}{dt} = \left(\boldsymbol{\Sigma}^*\right)^{-1}\boldsymbol{\theta}^*. \qquad (44)$$

It is possible to adopt the Leapfrog algorithm, which modifies Euler's discretization method by using a discrete step size ϵ individually for $\boldsymbol{\theta}^*$ and $\boldsymbol{\theta}$, with a full step ϵ in $\boldsymbol{\theta}$ sandwiched between two haf steps ϵ / 2 for $\boldsymbol{\theta}^*$ (Ruth, 1983). The algorithm performs L sequential steps, Equations (45), (46), and (47), before reaching a new point in the parametric space.

$$\boldsymbol{\theta}^*\left(t+\frac{\epsilon}{2}\right) = \boldsymbol{\theta}^*(t) + \frac{\epsilon}{2}\left[\nabla_{\boldsymbol{\kappa}} \log \pi\big(\boldsymbol{\kappa}_{1:T} \mid \mathbf{e}_{1:T};\mathbf{g}_{1:T};\mathcal{P}\big) + \nabla_{\mathcal{P}} \log \pi\big(\mathcal{P} \mid \mathbf{e}_{1:T};\mathbf{g}_{1:T};\boldsymbol{\kappa}_{1:T}\big)\right], \qquad (45)$$

$$\boldsymbol{\theta}(t+\epsilon) = \boldsymbol{\theta} + \epsilon\left(\boldsymbol{\Sigma}^*\right)^{-1}\boldsymbol{\theta}^*\left(t+\frac{\epsilon}{2}\right), \qquad (46)$$

$$\boldsymbol{\theta}^*(t+\epsilon) = \boldsymbol{\theta}^*\left(t+\frac{\epsilon}{2}\right) + \frac{\epsilon}{2}\left[\nabla_{\boldsymbol{\kappa}} \log \pi\big(\boldsymbol{\kappa}_{1:T} \mid \mathbf{e}_{1:T};\mathbf{g}_{1:T};\mathcal{P}\big) + \nabla_{\mathcal{P}} \log \pi\big(\mathcal{P} \mid \mathbf{e}_{1:T};\mathbf{g}_{1:T};\boldsymbol{\kappa}_{1:T}\big)\right] \qquad (47)$$

where the step size ϵ, the number of leapfrog steps L and the covariance matrix $\boldsymbol{\Sigma}^*$. At the end of each L step, the acceptance of the new point uses the same idea as the Metropolis Hastings algorithm[25]

[23] more than weak exogeneity I also assume $\boldsymbol{g}_t \perp \boldsymbol{\tau}_t, \forall t$

[24] Carter and Kohn (1994), Frühwirth-Schnatter (1994), or De Jong and Shephard (1995)

[25] Metropolis-Hastings, developed in the early 1950s, is an algorithm well-explained in Chib and Greenberg's work (1995).

5.3 Convergence of Beliefs

This subsection addresses the convergence of four Hamiltonian Monte Carlo (HMC) chains used to estimate 240 parameters in the Itau investment platform model. Convergence is verified through multiple diagnostics, such as the potential scale reduction factor $\hat{R}$, principal component projections, effective sample sizes, integrated autocorrelation times, correlation structures, and trace plots. These methods jointly provide a detailed view of the sampling behavior in a high-dimensional setting.

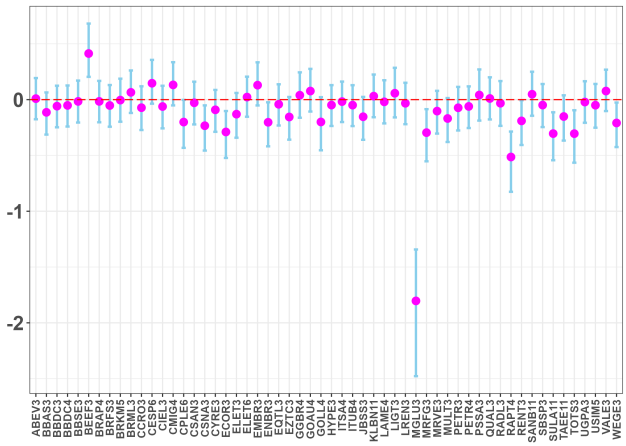
For example, in Figure 9 I present 240 parameter estimations from Itau's platform. 56 stocks with intensive margin during March 2016 and December 2021. The Figure 10 illustrates the $\hat{R}$ values for each parameter by integrating data from all four chains. When $\hat{R}$ is near 1, it indicates that the variance between chains is equivalent to the variance within chains, implying that all chains are sampling from the same posterior distribution. The majority of parameters displayed in this figure have $\hat{R}$ values beneath the usual thresholds, confirming that the sampler has reached a stable state. Figure 11 examines the chain trajectories using principal component (PC) projections. This technique reduces the high-dimensional parameter space to its most critical variability directions, enabling a visual assessment of how thoroughly each chain explores the posterior. Overlapping trajectories along these principal components suggest that the chains do not become isolated in separate regions, which is generally an indicator of effective mixing.

The following figures categorize parameters based on their conceptual roles within the model. These categories are analyzed regarding effective sample size (Neff), the Neff to integrated autocorrelation time (IACT) ratio, and the correlation structure. A large Neff value suggests that each chain produces numerous uncorrelated samples, thereby enhancing the precision of posterior estimates. Additionally, the Neff/IACT ratio serves as another measure to evaluate sample independence, with higher values indicating more efficient sampling.

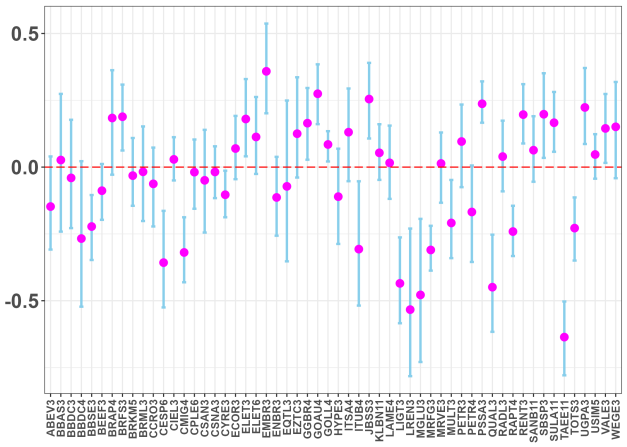
Figure 12 focuses on the exposure elasticities parameters, which may encode interactions or cross effects within the investment platform. In the left panel, most of the parameters show robust Neff values and favorable Neff/IACT ratios, indicating minimal autocorrelation. In the right panel, the correlation among these parameters does not exhibit strong off-diagonal elements, suggesting that the sampler moves freely in the associated subspace.

In Figure 13, the factor parameters are depicted as indicators of the drivers behind investment trends. By comprehensively sampling these parameters, the latent structure they represent becomes clearer. The left panel emphasizes that high effective sample sizes help reduce bias in estimations, while the right panel shows that, although correlations may exist, these latent factors are not excessively clustered, which could otherwise impede their identification. Figure 14 illustrates how loadings connect latent factors with observed data dimensions, crucial for the empirical manifestation of factor-specific effects. A thorough exploration by the sampler is essential to ascertain reliable relationships in factor loadings. On the left panel, the significance of maintaining large effective sample sizes is reiterated, and the right panel exhibits correlations that are not excessively strong, maintaining the uniqueness of each loading parameter.

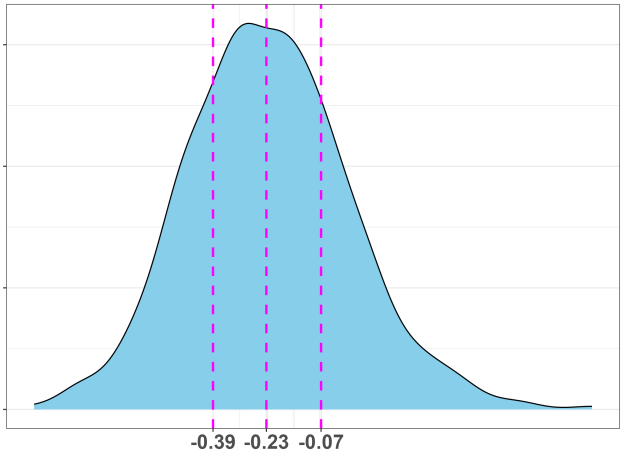
Figure 15 illustrates the intercept parameters involved in accounting for boundary effects or additional structural elements within the Itau model. Much like prior groups, the combination of robust Neff values, favorable Neff/IACT ratios, and a moderate correlation structure suggests these parameters achieve efficient mixing. Furthermore, Figure 16 presents trace plots for a chosen set of 15 parameters, offering a visual inspection of the progression of each of the four chains across iterations. Trace plots assist in determining whether a chain adequately traverses the posterior or remains confined



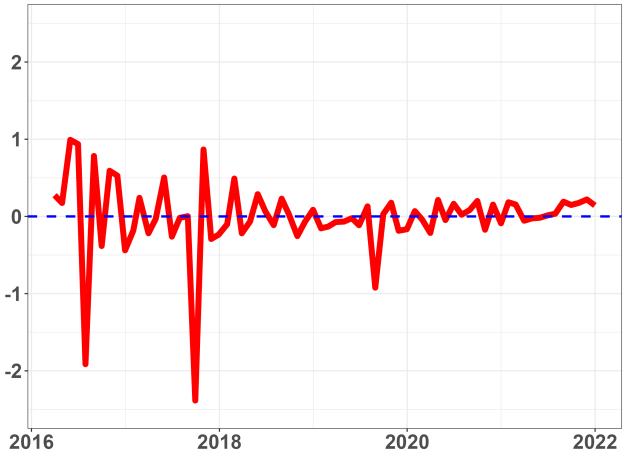
(a) loading



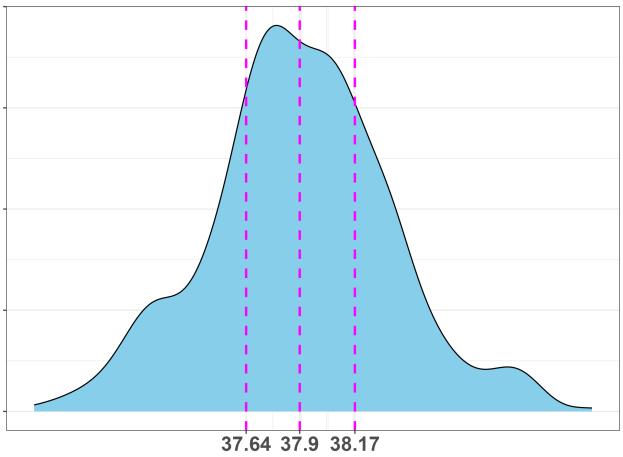
(b) exposure elasticity



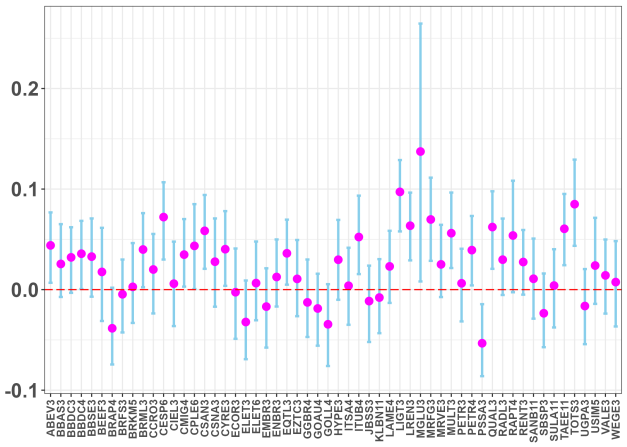
(c) ϕ -0.23 CI[-0.39,-0.07]



(d) factor



(e) σ_τ 37.9% CI[37.64%,38.17%]



(f) intercept

Figure 9: Itau

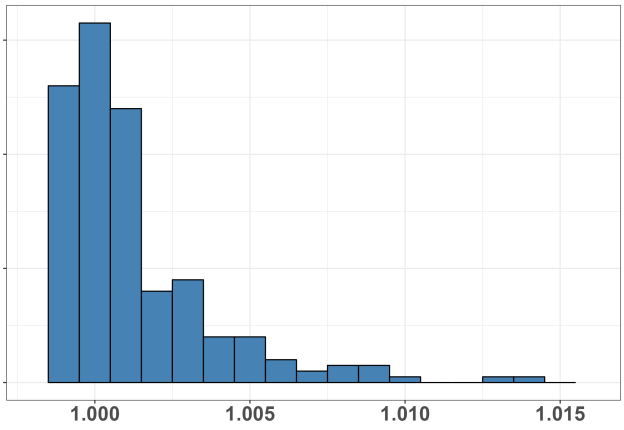


Figure 10: The values of $\hat{R}$ for each parameter using four chains

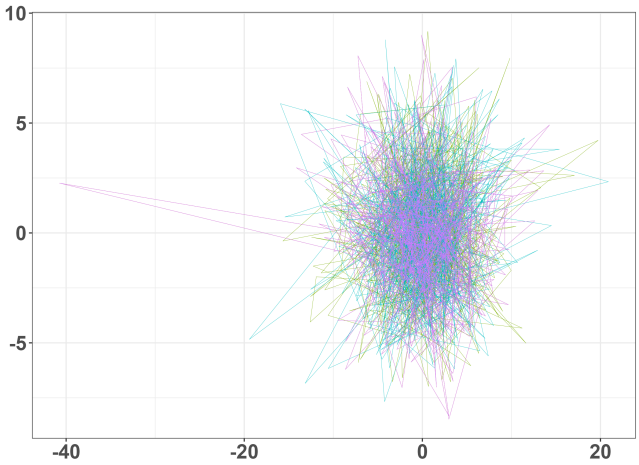
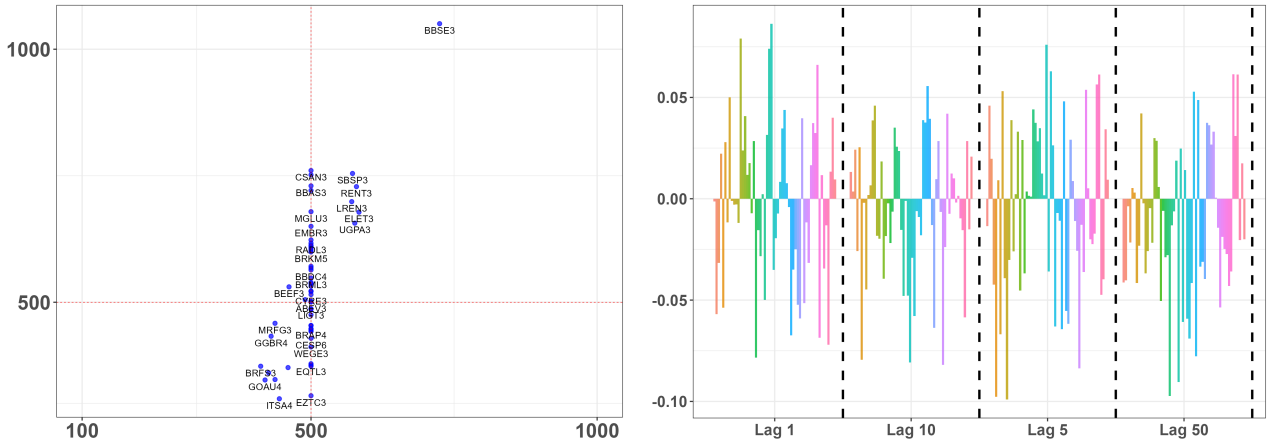


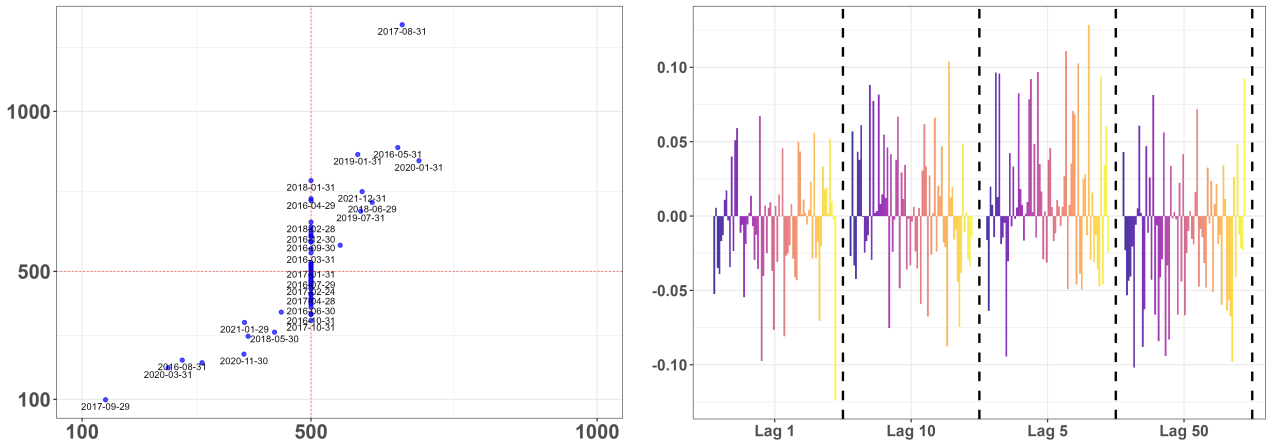
Figure 11: The chains trajectories through principal component (PC) projections



(a) Neff × Neff/IACT

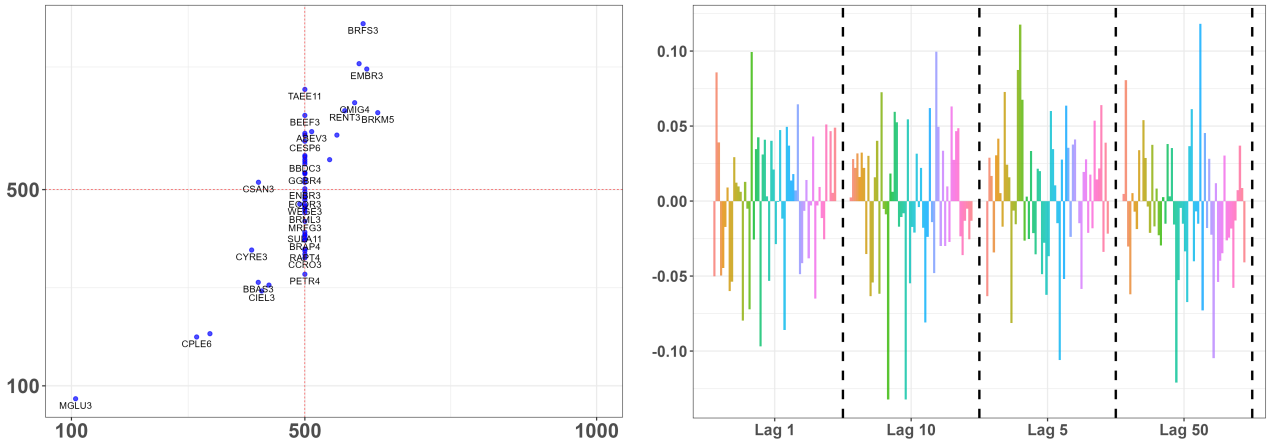
(b) autocorrelation

Figure 12: The exposure elasticity



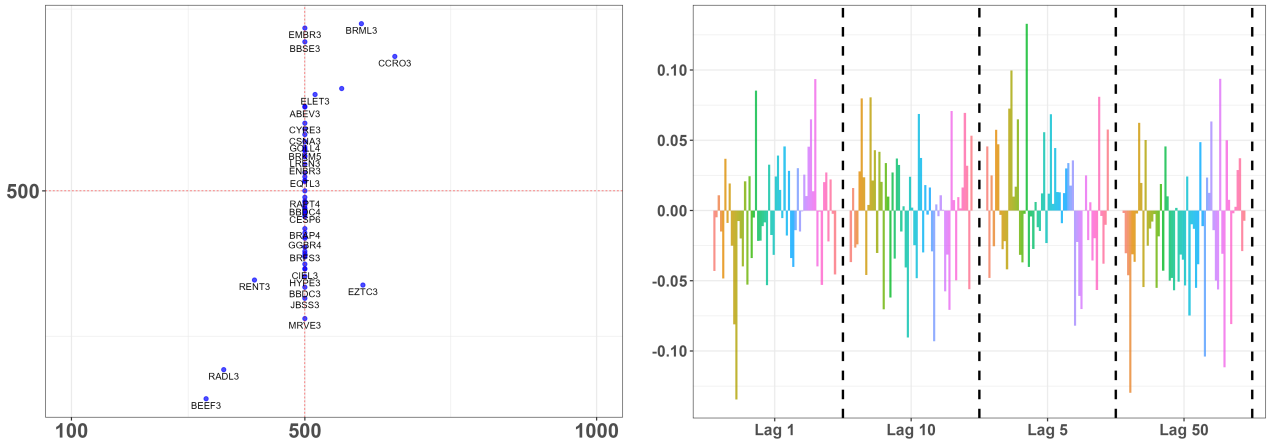
(a) Neff × Neff/IACT (b) autocorrelation

Figure 13: Factor-based exposure



(a) Neff × Neff/IACT (b) autocorrelation

Figure 14: The factors loadings



(a) Neff × Neff/IACT (b) autocorrelation

Figure 15: The intercept

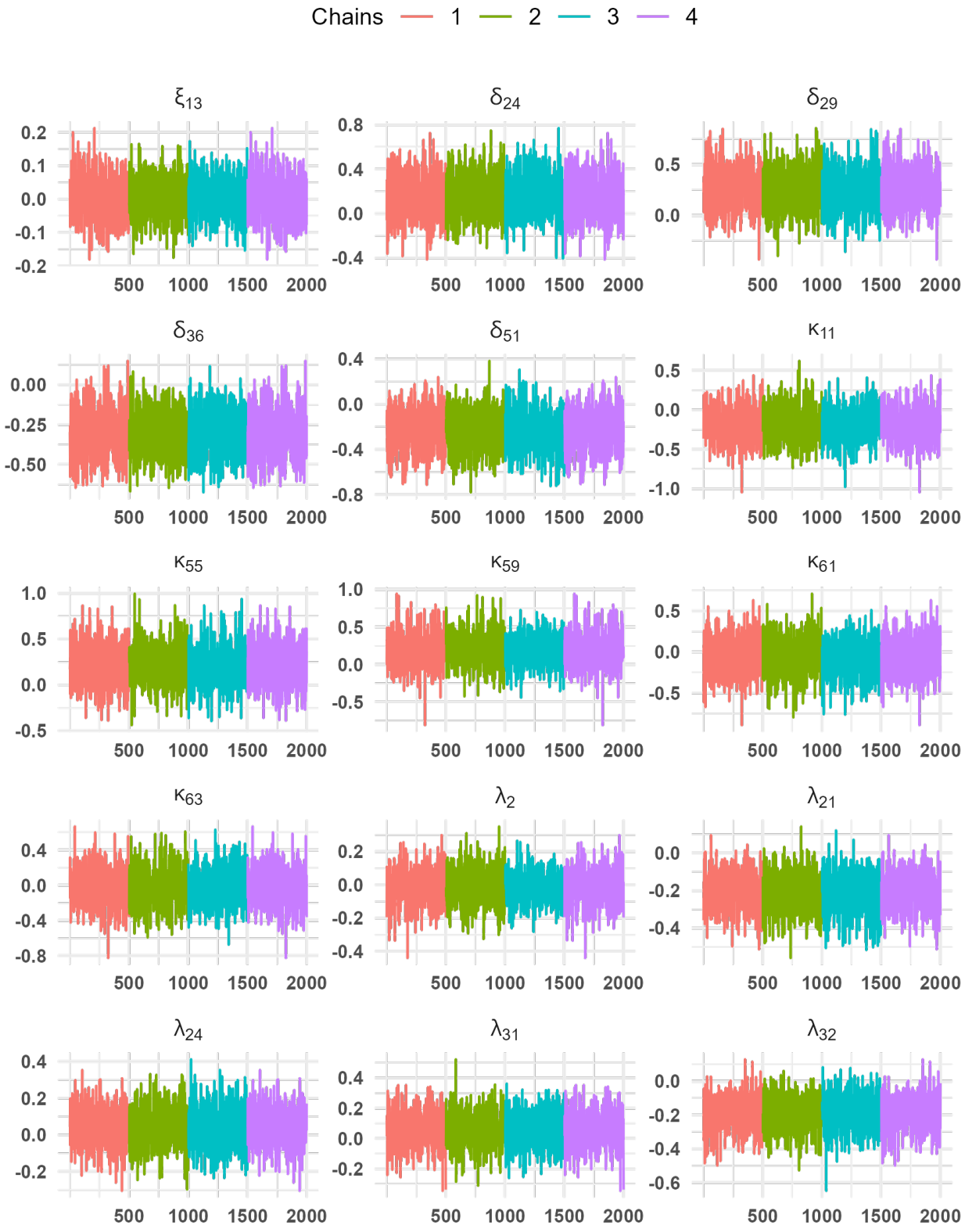


Figure 16: Trace plots

within a specific range. The chains are color-coded to differentiate their trajectories, generally showing overlapping variations around steady central tendencies. This observation reinforces the notion that the sampling algorithm is effectively navigating the posterior landscape.

Across all these figures and diagnostics, the HMC sampler shows stable convergent behavior, evidenced by nearly uniform $\hat{R}$ values below the usual thresholds, sufficient effective sample sizes, moderate correlation patterns, and well-overlapped principal component and trace plots. These observations reinforce confidence in the posterior estimates, indicating that the model's set of high-dimensional parameter parameters has been thoroughly explored by the four chains.

6 Replicant Behavior

Estimates of the exposure elasticities of the investment platforms for the stocks that made up the IBRX-100 index during the years 2016 and 2021 are presented in Appendix D. The elasticities are constant values for the entire 72 months of the sample. The aggregate exposures of investment funds from 14 investment platforms were evaluated. The criterion used to choose these platforms was to evaluate only decisions at the intensive margin[26] and thus allow the evaluation of the exposure elasticity of at least dozens of stocks. A minimum number of stocks is required to be able to sort stocks given the elasticity value. The number of stocks evaluated was 42, which made up the IBRX-100 index. The investment platforms are listed in Table 1 in descending order of replicant behavior.

Platform	**Qtd**	%
Itau	21	50
Caixa	16	38
Julius Baer	15	36
BNP Paribas	13	31
Opportunity	13	31
TNA Wealth Management	13	31
BTG Pactual	12	29
XP	12	29
BB	11	26
Safra	11	26
Vinci Partners	10	24
Bradesco	4	10
Credit Suisse	4	10

Table 1: Frequency of Replicant Platforms in 42 stocks belongs to IBRX-100 index.

Replicant behavior is described by exposure elasticity $\delta > 0$. In this case, the platform reacts to demand shocks originating from other platforms in a pro-cyclical manner, proportionally replicating these shocks. The estimates show a low elasticity of exposure in most cases, with δ varying between -0.5% and 0.5%. Itau's investment platform has replicant behavior in 21 of the 42 stocks, making it the platform with the highest number of replicant behavior. In comparison, Santander's investment platform has no replicant behavior in any of the 42 stocks.

[26]Exposures greater than zero in all months of the sample

Replicant behavior is a characteristic based on quantities and not prices, since exposure is a form of quantity. Therefore, given that platforms compete for demanders' resources for exposure, this new information on replicant behavior reveals a new perspective for risk assessment that is not based on stock returns. This new perspective for risk assessment depends on whether the aggregate exposure decisions of the platforms' funds are in a context of coordination, as this means that there is a market premium associated with similar behavior among the platforms, whether replicant or not. The definition of a replicant risk is the uncertainty that platforms have about the replicant behavior of competing platforms, i.e. the difference between the risk premiums of stocks with non-replicant behavior and stocks with replicant behavior.

That is why it is important to list in Table 2, for each of the stocks, the proportion of platforms that have replicating behavior. The lowest proportion of replicant behavior is that of National Steel Company (CSNA3),[27] this stock has no platform with exposure elasticity $\delta > 0$. The stock that is subject to the highest proportion of replicant behavior is Localiza (RENT3),[28] 10 of the 14 platforms have exposure elasticity $\delta > 0$. These differences make it possible to order the shares due to the replicant behavior of the platforms.

Consider that the stocks listed in Table 2 are in equilibrium, with prices adjusted to platform exposures. Instead of representing the exposure-based market equilibrium by the set of stock prices and platform exposures, I propose to represent the market equilibrium by the proportions of replicant behavior listed in the column "p" of Table 2. This vector of proportions represents the Nash equilibrium in mixed strategies in a context of coordination between platforms, given that positive premiums on shares would only be related to coordinated behavior - both replicant and non-replicant (see the example discussed in the introduction, in Figure 1). Furthermore, this equilibrium will have different proportions of 0.5 in each share, as the premiums related to replicant behavior are lower than the premiums related to non-replicant behavior. Share premiums in an environment of coordination between platforms.

Only two stocks have exactly 0.5 proportions of replicant platforms. This is the ratio that would represent the Nash equilibrium in a mixed strategy in the context of uncoordinated platforms. Coordination or non-coordination of platforms is a conjecture at first. As a first approach, I prefer to remove the shares of Sabesp (SBSP3)[29] and JBS(JBSS3)[30] from the replicating risk analysis.

One of the most relevant aspects in Table 2 is the generalized non-replicating behavior among investment platforms. In this case, the expected returns are not explained by demand shocks due to investors' exposure to the platforms. The explanation lies in the negative exposure elasticity: for any percentage of demand shocks on competing platforms, the platforms would offset these shocks by adjusting their intensive exposure margin decisions. In a scenario where all platforms had a negative exposure elasticity for all stocks, this would mean that the expected values of returns would not depend on shocks to investors' demand for exposure. It is precisely the fact that some platforms have a positive exposure elasticity for some stocks that allows for the emergence of replicating risk and its importance for stock pricing.

[27]National Steel Company is the largest fully integrated steel producer in Brazil and one of the largest in Latin America in terms of crude steel production.

[28]Localiza is a Brazilian car rental company founded in 1973 and is the largest car rental in Latin America and one of the largest in the world by size of the fleet and market capitalization.

[29]Sabesp is a Brazilian water and waste management company headquartered in Sao Paulo. It is the largest water and waste management company in Latin America

[30]JBS S.A. is a Brazilian multinational company that is the largest meat processing enterprise in the world, producing factory processed beef, chicken, salmon, sheep, pork, and also selling by-products from the processing of these meats

Stock	p	1-p	Replicant	Non-replicant
CSNA3	0.00	1.00		BB, BNP Paribas, Bradesco, BTG Pactual, Caixa, Credit Suisse, Itaú, Julius Baer, Opportunity, Safra, Santander, TNA Wealth Management, Vinci Partners, XP
BBDC4	0.07	0.93	XP	BB, BNP Paribas, Bradesco, BTG Pactual, Caixa, Credit Suisse, Itaú, Julius Baer, Opportunity, Safra, Santander, TNA Wealth Management, Vinci Partners
HYPE3	0.07	0.93	XP	BB, BNP Paribas, Bradesco, BTG Pactual, Caixa, Credit Suisse, Itaú, Julius Baer, Opportunity, Safra, Santander, TNA Wealth Management, Vinci Partners
LAME4	0.07	0.93	Itaú	BB, BNP Paribas, Bradesco, BTG Pactual, Caixa, Credit Suisse, Julius Baer, Opportunity, Safra, Santander, TNA Wealth Management, Vinci Partners, XP
PETR3	0.07	0.93	Itaú	BB, BNP Paribas, Bradesco, BTG Pactual, Caixa, Credit Suisse, Julius Baer, Opportunity, Safra, Santander, TNA Wealth Management, Vinci Partners, XP
QUAL3	0.07	0.93	Caixa	BB, BNP Paribas, Bradesco, BTG Pactual, Credit Suisse, Itaú, Julius Baer, Opportunity, Safra, Santander, TNA Wealth Management, Vinci Partners, XP
USIM5	0.07	0.93	Itaú	BB, BNP Paribas, Bradesco, BTG Pactual, Caixa, Credit Suisse, Julius Baer, Opportunity, Safra, Santander, TNA Wealth Management, Vinci Partners, XP
LREN3	0.14	0.86	BNP Paribas, Caixa	BB, Bradesco, BTG Pactual, Credit Suisse, Itaú, Julius Baer, Opportunity, Safra, Santander, TNA Wealth Management, Vinci Partners, XP
MRFG3	0.14	0.86	Safra, TNA Wealth Management	BB, BNP Paribas, Bradesco, BTG Pactual, Caixa, Credit Suisse, Itaú, Julius Baer, Opportunity, Santander, Vinci Partners, XP
MRVE3	0.14	0.86	Itaú, Vinci Partners	BB, BNP Paribas, Bradesco, BTG Pactual, Caixa, Credit Suisse, Julius Baer, Opportunity, Safra, Santander, TNA Wealth Management, XP
MULT3	0.14	0.86	Opportunity, Safra	BB, BNP Paribas, Bradesco, BTG Pactual, Caixa, Credit Suisse, Itaú, Julius Baer, Santander, TNA Wealth Management, Vinci Partners, XP
SANB11	0.14	0.86	Itaú, Julius Baer	BB, BNP Paribas, Bradesco, BTG Pactual, Caixa, Credit Suisse, Opportunity, Safra, Santander, TNA Wealth Management, Vinci Partners, XP
ABEV3	0.21	0.79	BNP Paribas, Opportunity, Vinci Partners	BB, Bradesco, BTG Pactual, Caixa, Credit Suisse, Itaú, Julius Baer, Safra, Santander, TNA Wealth Management, XP
BRML3	0.21	0.79	BNP Paribas, Safra, XP	BB, Bradesco, BTG Pactual, Caixa, Credit Suisse, Itaú, Julius Baer, Opportunity, Santander, TNA Wealth Management, Vinci Partners
BBAS3	0.21	0.79	BTG Pactual, Itaú, Julius Baer	BB, BNP Paribas, Bradesco, Caixa, Credit Suisse, Opportunity, Safra, Santander, TNA Wealth Management, Vinci Partners, XP
BRFS3	0.21	0.79	Itaú, Opportunity, Vinci Partners	BB, BNP Paribas, Bradesco, BTG Pactual, Caixa, Credit Suisse, Julius Baer, Safra, Santander, TNA Wealth Management, XP
CCRO3	0.21	0.79	BB, BNP Paribas, BTG Pactual	Bradesco, Caixa, Credit Suisse, Itaú, Julius Baer, Opportunity, Safra, Santander, TNA Wealth Management, Vinci Partners, XP
CSAN3	0.21	0.79	BNP Paribas, Julius Baer, XP	BB, Bradesco, BTG Pactual, Caixa, Credit Suisse, Itaú, Opportunity, Safra, Santander, TNA Wealth Management, Vinci Partners
EMBR3	0.21	0.79	Caixa, Itaú, Safra	BB, BNP Paribas, Bradesco, BTG Pactual, Credit Suisse, Julius Baer, Opportunity, Santander, TNA Wealth Management, Vinci Partners, XP
EQTL3	0.21	0.79	BNP Paribas, Caixa, Vinci Partners	BB, Bradesco, BTG Pactual, Credit Suisse, Itaú, Julius Baer, Opportunity, Safra, Santander, TNA Wealth Management, XP
GOAU4	0.21	0.79	Caixa, Credit Suisse, Itaú	BB, BNP Paribas, Bradesco, BTG Pactual, Julius Baer, Opportunity, Safra, Santander, TNA Wealth Management, Vinci Partners, XP
ITUB4	0.21	0.79	BNP Paribas, Julius Baer, Safra	BB, Bradesco, BTG Pactual, Caixa, Credit Suisse, Itaú, Opportunity, Santander, TNA Wealth Management, Vinci Partners, XP
KLBN11	0.21	0.79	BB, Itaú, Opportunity	BNP Paribas, Bradesco, BTG Pactual, Caixa, Credit Suisse, Julius Baer, Safra, Santander, TNA Wealth Management, Vinci Partners, XP
RADL3	0.21	0.79	Caixa, Itaú, TNA Wealth Management	BB, BNP Paribas, Bradesco, BTG Pactual, Credit Suisse, Julius Baer, Opportunity, Safra, Santander, Vinci Partners, XP
WEGE3	0.21	0.79	Itaú, Safra, XP	BB, BNP Paribas, Bradesco, BTG Pactual, Caixa, Credit Suisse, Julius Baer, Opportunity, Santander, TNA Wealth Management, Vinci Partners
BBSE3	0.29	0.71	BTG Pactual, Julius Baer, Opportunity, TNA Wealth Management	BB, BNP Paribas, Bradesco, Caixa, Credit Suisse, Itaú, Safra, Santander, Vinci Partners, XP
BBDC3	0.29	0.71	Credit Suisse, Opportunity, Safra, TNA Wealth Management	BB, BNP Paribas, Bradesco, BTG Pactual, Caixa, Itaú, Julius Baer, Santander, Vinci Partners, XP
CMIG4	0.29	0.71	BB, Caixa, TNA Wealth Management, Vinci Partners	BNP Paribas, Bradesco, BTG Pactual, Credit Suisse, Itaú, Julius Baer, Opportunity, Safra, Santander, XP
CYRE3	0.29	0.71	BB, BTG Pactual, Caixa, XP	BNP Paribas, Bradesco, Credit Suisse, Itaú, Julius Baer, Opportunity, Safra, Santander, TNA Wealth Management, Vinci Partners
GGBR4	0.29	0.71	BNP Paribas, Caixa, Itaú, Julius Baer	BB, Bradesco, BTG Pactual, Credit Suisse, Opportunity, Safra, Santander, TNA Wealth Management, Vinci Partners, XP
PETR4	0.29	0.71	BB, Bradesco, Julius Baer, TNA Wealth Management	BNP Paribas, BTG Pactual, Caixa, Credit Suisse, Itaú, Opportunity, Safra, Santander, Vinci Partners, XP
BRKM5	0.36	0.64	BNP Paribas, BTG Pactual, Caixa, Julius Baer, Vinci Partners	BB, Bradesco, Credit Suisse, Itaú, Opportunity, Safra, Santander, TNA Wealth Management, XP
ENBR3	0.36	0.64	BNP Paribas, BTG Pactual, Credit Suisse, Opportunity, Safra	BB, Bradesco, Caixa, Itaú, Julius Baer, Santander, TNA Wealth Management, Vinci Partners, XP
UGPA3	0.36	0.64	BB, Caixa, Itaú, Opportunity, TNA Wealth Management	BNP Paribas, Bradesco, BTG Pactual, Credit Suisse, Julius Baer, Safra, Santander, Vinci Partners, XP
BRAP4	0.43	0.57	BB, Bradesco, Itaú, Julius Baer, Opportunity, XP	BNP Paribas, BTG Pactual, Caixa, Credit Suisse, Safra, Santander, TNA Wealth Management, Vinci Partners
ITSA4	0.43	0.57	BB, Caixa, Itaú, Opportunity, TNA Wealth Management, XP	BNP Paribas, Bradesco, BTG Pactual, Credit Suisse, Julius Baer, Safra, Santander, Vinci Partners
JBSS3	0.50	0.50	Bradesco, BTG Pactual, Caixa, Itaú, Julius Baer, Safra, XP	BB, BNP Paribas, Credit Suisse, Opportunity, Santander, TNA Wealth Management, Vinci Partners
SBSP3	0.50	0.50	BNP Paribas, BTG Pactual, Caixa, Itaú, Julius Baer, Opportunity, TNA Wealth Management	BB, Bradesco, Credit Suisse, Safra, Santander, Vinci Partners, XP
CESP6	0.57	0.43	BTG Pactual, Credit Suisse, Julius Baer, Opportunity, Safra, TNA Wealth Management, Vinci Partners, XP	BB, BNP Paribas, Bradesco, Caixa, Itaú, Santander
CIEL3	0.57	0.43	BB, BNP Paribas, BTG Pactual, Caixa, Itaú, Julius Baer, TNA Wealth Management, Vinci Partners	Bradesco, Credit Suisse, Opportunity, Safra, Santander, XP
VALE3	0.64	0.36	BB, Bradesco, BTG Pactual, Itaú, Julius Baer, Opportunity, TNA Wealth Management, Vinci Partners, XP	BNP Paribas, Caixa, Credit Suisse, Safra, Santander
RENT3	0.71	0.29	BB, BNP Paribas, BTG Pactual, Caixa, Itaú, Julius Baer, Safra, TNA Wealth Management, Vinci Partners, XP	Bradesco, Credit Suisse, Opportunity, Santander

Table 2: Stocks sorted by non-replicant platform's behavior.

6.1 Replicant Risk

The ambiguity surrounding the coordination of intensive margin decisions across investment platforms introduces a novel risk, which I define replicant risk. This risk represents the premium between stocks exhibiting non-replicant behavior and those subject to replicant behavior, irrespective of the level of stock market risk. By proposing a risk grounded in exposure rather than returns, this concept offers a fresh point of view for the literature on risk factors.

Long	Short
BRADESCO PN N1 - BBDC4	BRADESPAR PN N1 - BRAP4
HYPERMARCAS ON NM - HYPE3	CESP PNB N1 - CESP6
LOJAS AMERIC PN - LAME4	CIELO ON NM - CIEL3
LOJAS RENNER ON NM - LREN3	ENERGIAS BR ON NM - ENBR3
PETROBRAS ON - PETR3	ITAUSA PN N1 - ITSA4
QUALICORP ON NM - QUAL3	LOCALIZA ON NM - RENT3
SID NACIONAL ON - CSNA3	ULTRAPAR ON NM - UGPA3
USIMINAS PNA N1 - USIM5	VALE ON N1 - VALE3

Table 3: Replicant portfolio, long (Quintile 1) vs. short (Quintile 5).

Table 3 shows the composition of the replicant portfolio formed by stocks in the extremes of the distribution of a given characteristic, with stocks in Quintile 1 on the long side and those in Quintile 5 on the short side. This portfolio is constructed to capture the premium associated with the characteristic in question. The allocation includes major Brazilian companies such as BRADESCO (BBDC4), PETROBRAS (PETR3), and USIMINAS (USIM5) on the long side, and VALE (VALE3), ITAÚSA (ITSA4), and CESP (CESP6) on the short side. The strategy aims to isolate the return differential between the two tails of the distribution, representing a market-neutral position.

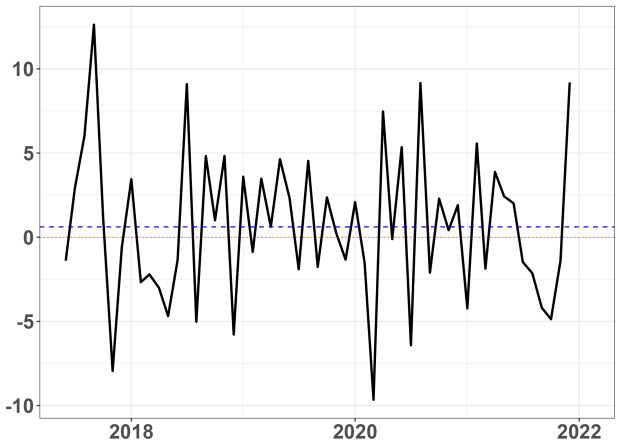


Figure 17: Replicant Risk. Blue dashed line is the premium of 0.61% per month, or 7.6% per year with annualized standard deviation 15.47%.

Figure 17 illustrates the monthly return timeline for the replicant portfolio. The average premium of the strategy, represented by the blue dashed line, is estimated to be 0.61% monthly or 7.6% annually. The portfolio's risk is signified by an annualized standard deviation of 15.47%. Despite the positive

average return, there are significant fluctuations, with periods of substantial gains followed by marked losses. This indicates that the strategy's profitability is heavily influenced by market trends and the stability of the signal used for sorting stocks into quintiles.

The Figure 18 presents the evolution of the market beta for the replicant portfolio, calculated separately for the long and short legs using a rolling window of twelve months of returns. The beta of each stock is estimated relative to the excess return of the market portfolio, using the average monthly CDI rate as the risk-free benchmark.[31] This dynamic approach captures time-varying sensitivities to market movements and provides a clearer assessment of whether the portfolio maintains its intended market neutrality over time. The beta of the long leg (red) and the short leg (turquoise) tends to move in opposite directions, with a negative correlation of -0.50. This inverse relationship is a key feature of a market-neutral strategy, as it helps ensure that market-wide shocks affect both sides of the portfolio in offsetting ways. The dashed black line represents the net beta of the combined portfolio, which remains close to zero for most of the sample period, providing evidence that the strategy effectively neutralizes market exposure. Monitoring beta in this way is crucial to verifying that the return of the replicant portfolio is driven by the targeted characteristic, rather than by systematic risk factors.

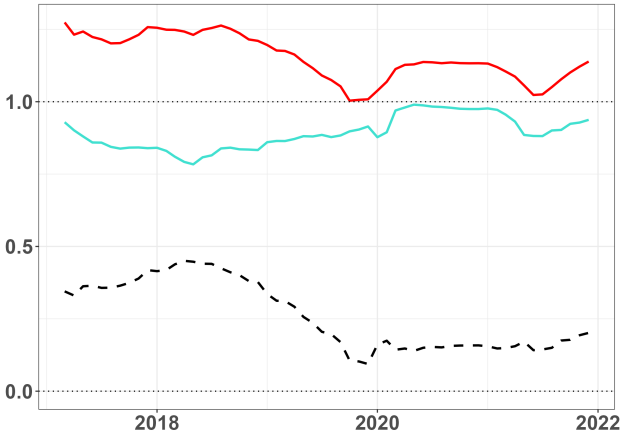
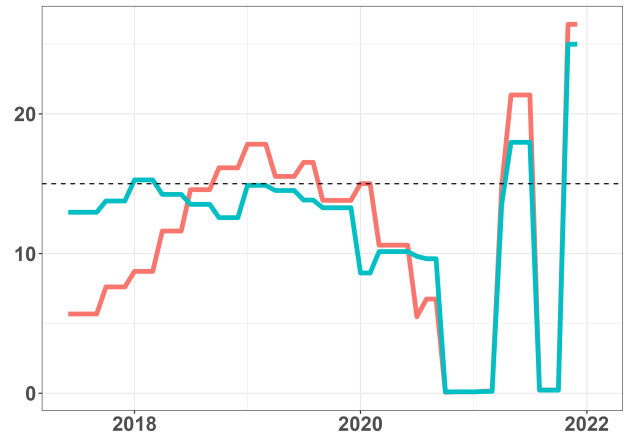


Figure 18: Evidence of Market Neutral Strategy. Long portfolio in red and short portfolio in turquoise. Correlation between long e short portfolio is -0.50

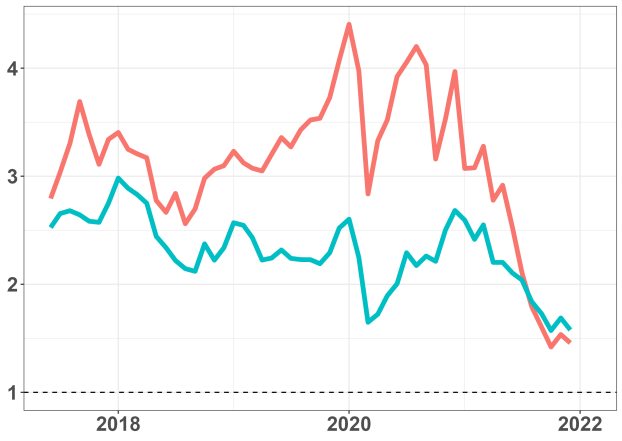
In Figure 19 the characteristics of the long (red) and short (turquoise) legs of the replicant portfolio are displayed over time, (a) the Return on Equity (ROE), (b) the price-to-book (P/B) ratio, (c) the price–earnings (P/E) ratio, (d) the logarithm of market capitalization, (e) the dividend yield, and (f) the annual return.

The ROE, calculated as the net income of the firms over the past twelve months divided by their total shareholder equity, serves as a measure of the financial performance and efficiency in generating returns from shareholders' investments. In particular, both legs exhibit a highly synchronized pattern, as indicated by the high positive correlation (0.82). Interestingly, while the dashed line at 15% ROE serves as a reference point, both legs of the portfolio frequently fluctuate around this threshold, suggesting similar profitability dynamics despite their opposing portfolio positions. The observed pattern, particularly simultaneous peaks and sharp downturns, reinforces the notion that the replicant portfolio not only captures differential characteristics between quintiles, but also highlights parallel fundamental shifts in corporate profitability across market segments.

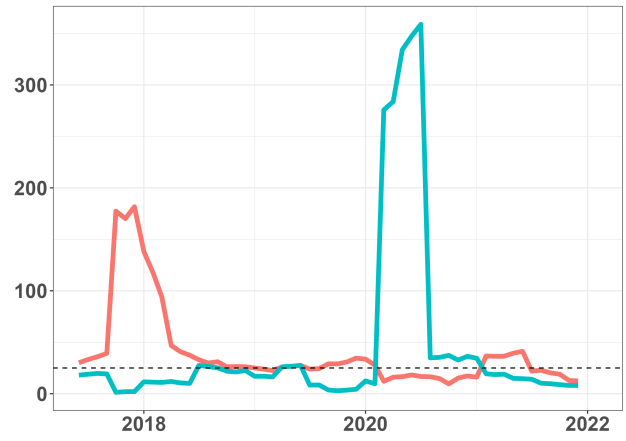
[31]The CDI is a very short-term security issued by banks in the Brazilian market to allow them to lend and borrow money from each other overnight, ensuring that the cash register ends the day with a positive balance.



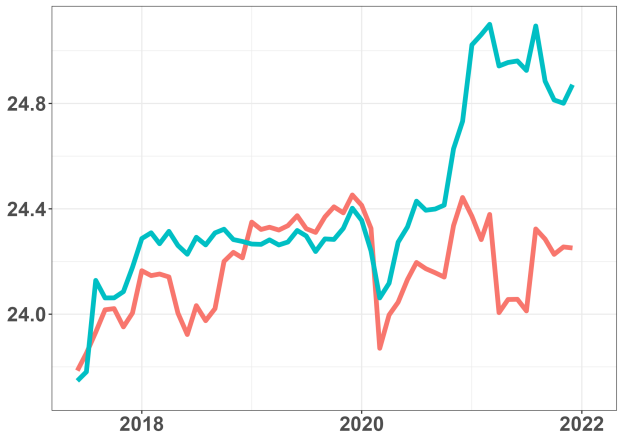
(a) ROE. Dashed line $ROE = 15\%$. Correlation 0.82



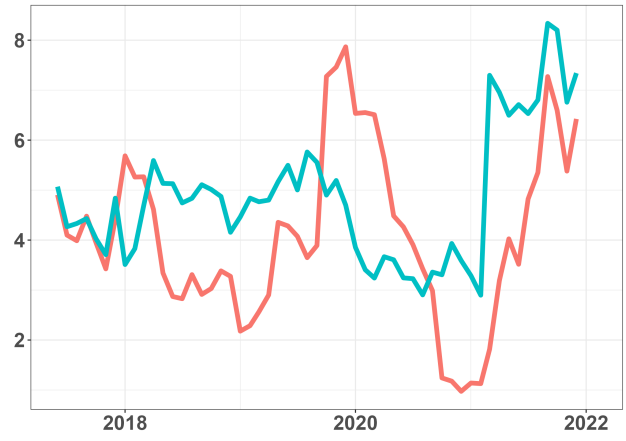
(b) P/B ratio. Dashed line $P/B = 1$. Correlation 0.51



(c) P/E ratio. Dashed line $P/E = 25$. Correlation -0.23



(d) log market cap. Correlation 0.38



(e) dividend yield. Correlation 0.26

(f) yearly return. Correlation 0.87

Figure 19: The replicant portfolio characteristics. The correlations are between long portfolio in red and short portfolio in turquoise

The P/B ratio, computed as a company's market price per share divided by its book value per share, reflects market expectations regarding firm value relative to its underlying assets. Both legs of the portfolio display a moderate positive correlation of 0.51, suggesting a general comovement in market valuation trends. Interestingly, the P/B ratio of both legs decreased significantly over the period, converging closer to the reference line ($P/B = 1$). This pattern indicates a reduction in market optimism or perceived growth opportunities in both segments of the portfolio, with the long leg consistently trading at a premium relative to the short leg. The narrowing gap towards the end of the sample period suggests a decreasing dispersion in valuation, potentially impacting the portfolio's future replicant premium.

The P/E ratio, calculated as the company's share price divided by its earnings per share over the past 12 months, provides insight into market valuation relative to recent profitability. In particular, two distinct spikes occur during the analyzed period, each affecting predominantly one leg of the portfolio: first, a peak appears around early 2018 primarily impacting the long leg, followed by an even more pronounced peak in early 2020 concentrated in the short leg. These abrupt increases suggest periods of very low or negative earnings, which temporarily drive valuations excessively high due to diminishing denominators. Outside of these anomalies, both legs generally trade below the benchmark line (P/E = 25), exhibiting modest volatility and low correlation (0.23), underscoring that the replicant strategy experiences episodic valuation extremes driven by unique profitability shocks rather than persistent common market forces.

Market capitalization, defined as the total market value of a firm's outstanding shares, offers a measure of firm size and investor valuation of future profitability and growth. Noticeably, after the onset of the COVID-19 pandemic in early 2020, the short leg exhibits a pronounced upward trend, significantly outpacing the growth observed in the long leg. This divergence could reflect flight-to-quality behavior among investors, favoring larger and presumably more stable firms within the short portfolio. Alternatively, the sectors represented more heavily in the short leg may have experienced a positive reassessment due to changing economic expectations post-pandemic. In a counterfactual scenario where the COVID-19 crisis did not occur, one could expect a more parallel trajectory between the two legs, potentially reducing the performance divergence and influencing the replicant strategy's risk and returns substantially.

The dividend yield, calculated as the total dividends paid per share divided by the initial unadjusted price within a given period, reflects the income-generating potential relative to the share price. Although both legs exhibit substantial variability, their correlation is relatively low (0.26). This weak correlation arises because dividend policies vary significantly between firms and sectors, often independently of broader market conditions. Furthermore, corporate decisions about dividend payments are based on factors such as cash flow stability, growth opportunities, and specific strategic considerations, rather than systematic market factors alone. Consequently, while both legs might occasionally respond similarly to macroeconomic conditions, such as shifts in interest rates or market risk sentiment, firm-specific dividend strategies reduce the overall coherence between the two portfolios' yields.

The yearly return are defined as the accumulated return over the last 12 months. Unlike strictly annual returns aligned with calendar years, these continuous yearly returns capture ongoing trends and changes in investor sentiment over any continuous 12-month period. The high correlation (0.87) indicates strong co-movement between the two components, signaling that both portfolios react similarly to broader market conditions or common economic shocks. Importantly, the pattern shows evidence of momentum and reversal effects: periods of persistent upward trends in returns (momentum) are often followed by substantial corrections or negative reversals. But this cyclical pattern demonstrates the

irrelevance of momentum strategies - buying recent winners and selling recent losers - and reversal strategies, which anticipate reverting to the mean after prolonged price movements for understanding replicant risk. Regardless of whether it doesn't clearly follow a factor such as momentum or reversal, understanding the dynamics of replicant risk is crucial, as it allows investment platforms to better manage timing, risk and profitability when implementing replicant portfolio strategies.

7 Concluding Remarks

This article may appear to be just one of many other articles in the literature on risk factors, which Cochrane (2011) called a "zoo of factors" in an annual speech to the American Finance Association. At that point in the discussion,[32] a recent literature that catalogs and verifies the capacity of new factors in pricing shares has emerged (Harvey, Liu, and Zhu, 2016; Harvey, 2017; Fama and French, 2018; 2020, 2020; Jensen, Kelly, and Pedersen, 2023; Kozak, Nagel, and Santosh, 2018). This article does not claim to be just a new risk factor capable of pricing stocks. This statement may seem contradictory, since the main implication of the existence of an exposure-based demand model was to allow access to a replicant factor through the inference of investment platforms' exposure elasticities. The main message of this article is that the existence of a replicant factor in the Brazilian market is the consequence of proposing a *quantitative approach* to the asset price problem.

A quantitative approach to understanding asset pricing is not clearly established in the literature. Examples of promising proposals in this direction are the recent works by Yi An, Su, and C. Wang (2024), Lopez-Lira and Roussanov (2020) and Rostek and Yoon (2023). Any study of asset pricing is based on market equilibrium. Equilibrium means both prices and quantities chosen by market participants consistent with the sharing of risk between them. This article has taken a different approach to market equilibrium: The pricing problem does not depend solely on the supply and demand for outstanding shares because market participants are pools of investment funds. Asset pricing depends on the balance between supply and demand for exposure between investment platforms and retail investors. In this way, asset prices would depend on decisions about exposures by investment platforms.

In this way, the presence of investment platforms as economic agents in market transactions influences asset pricing due to the existence of a market equilibrium based on exposure. Quantities in circulation and exposure are both information on quantities, but exposure is mostly indirect due to the participation of numerous funds in the quotas of master funds considered to be hubs in the network and which participate directly in the market. Observing a sample of exposures between investment funds was only possible due to the availability of consolidated data on fund portfolios in the Brazilian market. This consolidation is available in both the regulator's database (CVM) and in private information systems, such as the database used in this article by Quantum Axis. I consider it an important contribution of this article to highlight the importance of collecting consolidated data for researchers to investigate asset demand. Once again, I would like to point out that the valuation of the outstanding shares is not enough to capture the impact of demand shocks on share prices.

Therefore, this article proposed a model that allows access to replicant behavior between platforms in a context of imperfect competition for exposure. The replicant behavior of competing platforms is new information for risk management. Rather than assessing the risk of portfolio positions in

[32]The literature on risk factors is much older than the article by Cochrane (2011), the literature on this topic dates back to the 1960s with the works that proposed the CAPM equilibrium model (Sharpe, 1964)

investment funds, this article offers a perspective on assessing the risk of a set of funds that have exposure to shares. When investment platforms do not have information on their own behavior in relation to others in a context of imperfect competition for exposure, it means that the decisions of their groups of funds could be better coordinated. This intra-investment fund coordination does not seem to be a concern for large financial institutions, but it seems a concern for the stock exchange because it is the regulator of counterparty relationships between participants. This article offers a model that can be used by financial institutions and regulators by exploring the replicant behavior of investment platforms to contribute to the demand-based asset pricing literature by integrating insights into financial intermediation and imperfect competition for exposure.

References

Aguilar, Omar and Mike West (2000). "Bayesian dynamic factor models and portfolio allocation". In: *Journal of Business & Economic Statistics* 18.3, pp. 338–357.

Albuquerque, Rui, José Miguel Cardoso-Costa, and José Afonso Faias (2024). "Price elasticity of demand and risk-bearing capacity in sovereign bond auctions". In: *The Review of Financial Studies* 37.10, pp. 3149–3187.

Amir-Ahmadi, Pooyan, Christian Matthes, and Mu-Chun Wang (2020). "Choosing prior hyperparameters: With applications to time-varying parameter models". In: *Journal of Business & Economic Statistics* 38.1, pp. 124–136.

An, Yi, Yifan Su, and Chunyang Wang (Oct. 2024). *Quantity, risk, and return.* Disponível em: `https://papers.ssrn.com/abstract=4098609`. Acesso em: 29 out. 2024. Rochester, NY.

An, Yu, Matteo Benetton, and Yang Song (2023). "Index providers: Whales behind the scenes of etfs". In: *Journal of Financial Economics* 149.3, pp. 407–433.

Apesteguia, Jose, Jörg Oechssler, and Simon Weidenholzer (2020). "Copy trading". In: *Management Science* 66.12, pp. 5608–5622.

Axtell, Robert L and J Doyne Farmer (2025). "Agent-based modeling in economics and finance: Past, present, and future". In: *Journal of Economic Literature* 63.1, pp. 197–287.

Azar, José, Martin C Schmalz, and Isabel Tecu (2018). "Anticompetitive effects of common ownership". In: *The Journal of Finance* 73.4, pp. 1513–1565.

Bai, Jushan and Peng Wang (2015). "Identification and Bayesian estimation of dynamic factor models". In: *Journal of Business & Economic Statistics* 33.2, pp. 221–240.

Barber, Brad M et al. (2022). "Attention-induced trading and returns: Evidence from Robinhood users". In: *The Journal of Finance* 77.6, pp. 3141–3190.

Bastías, Jaime and José L Ruiz (2022). "Equity fire sales and herding behavior in pension funds". In: *Research in International Business and Finance* 62, p. 101708.

Bauder, D. Mark et al. (2018). "Bayesian mean–variance analysis: Optimal portfolio selection under parameter uncertainty". In: *Quantitative Finance* 21, pp. 221–242.

Ben-David, Itzhak et al. (2022). "Ratings-driven demand and systematic price fluctuations". In: *The Review of Financial Studies* 35.6, pp. 2790–2838.

Ben-Horin, Moshe and Haim Kedar-Levy (2013). "Herding, heterogeneity, and momentum trading of institutional investors across asset classes". In: *Journal of Reviews on Global Economics* 2, p. 455.

Benetton, Matteo and Giovanni Compiani (2024). "Investors' beliefs and cryptocurrency prices". In: *The Review of Asset Pricing Studies* 14.2, pp. 197–236.

Berk, Jonathan B and Jules H Van Binsbergen (2025). "The impact of impact investing". In: *Journal of Financial Economics* 164, p. 103972.

Betancourt, Michael (2017). "A conceptual introduction to Hamiltonian Monte Carlo". In: *arXiv preprint arXiv:1701.02434*.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch (1992). "A theory of fads, fashion, custom, and cultural change as informational cascades". In: *Journal of political Economy* 100.5, pp. 992–1026.

Boehmer, Ekkehart et al. (2021). "Tracking retail investor activity". In: *The Journal of Finance* 76.5, pp. 2249–2305.

Bond, Philip and Diego Garcia (2022). "The equilibrium consequences of indexing". In: *The Review of Financial Studies* 35.7, pp. 3175–3230.

Boortz, Christopher et al. (2013). *Herding in financial markets: Bridging the gap between theory and evidence*. Tech. rep. SFB 649 Discussion Paper.

Bootrz, C and S Kremer (2013). "The impact of information risk and market stress on institutional herding". In: *Freie Universitat Berlin, Working Paper*.

Brown, Zach Y et al. (2023). *Why do index funds have market power? Quantifying frictions in the index fund market*. Tech. rep. National Bureau of Economic Research.

Cai, Fang et al. (2019). "Institutional herding and its price impact: Evidence from the corporate bond market". In: *Journal of Financial economics* 131.1, pp. 139–167.

Camanho, Nelson, Harald Hau, and Hélene Rey (2022). "Global portfolio rebalancing and exchange rates". In: *The Review of Financial Studies* 35.11, pp. 5228–5274.

Carhart, Mark M (1997). "On persistence in mutual fund performance". In: *The Journal of finance* 52.1, pp. 57–82.

Carpenter, Bob et al. (2017). "Stan: A probabilistic programming language". In: *Journal of statistical software* 76, pp. 1–32.

Carter, Chris K and Robert Kohn (1994). "On Gibbs sampling for state space models". In: *Biometrika* 81.3, pp. 541–553.

Charles, Constantin, Cary Frydman, and Mete Kilic (2024). "Insensitive investors". In: *The Journal of Finance* 79.4, pp. 2473–2503.

Chib, Siddhartha and Edward Greenberg (1995). "Understanding the metropolis-hastings algorithm". In: *The american statistician* 49.4, pp. 327–335.

Chinco, Alex and Marco Sammon (2024). "The passive ownership share is double what you think it is". In: *Journal of Financial Economics* 157, p. 103860.

Chodorow-Reich, Gabriel, Xavier Gabaix, et al. (2024). "Propagation of shocks in networks: Identification and applications". In: *Available at SSRN*.

Chodorow-Reich, Gabriel, Andra Ghent, and Valentin Haddad (2021). "Asset insulators". In: *The Review of Financial Studies* 34.3, pp. 1509–1539.

Chow, Sy-Miin et al. (2011). "Bayesian estimation of semiparametric nonlinear dynamic factor analysis models using the Dirichlet process prior." In: *The British journal of mathematical and statistical psychology* 64 Pt 1, pp. 69–106.

Christoffersen, Susan Kerr and Ya Tang (2010). "Institutional herding and information cascades: Evidence from daily trades". In: *Available at SSRN 1572726*.

Cipriani, Marco and Antonio Guarino (2005). "Herd behavior in a laboratory financial market". In: *American Economic Review* 95.5, pp. 1427–1443.

— (2014). "Estimating a structural model of herd behavior in financial markets". In: *American Economic Review* 104.1, pp. 224–251.

Clark, Robert, Jean-François Houde, and Jakub Kastl (2021). "The industrial organization of financial markets". In: *Handbook of Industrial Organization*. Vol. 5. 1. Elsevier, pp. 427–520.

Clarke, Jonathan, Chayawat Ornthanalai, and Ya Tang (2015). *Institutional herding and asset price: The role of information*. Citeseer.

Cochrane, John H (2011). "Presidential address: Discount rates". In: *The Journal of finance* 66.4, pp. 1047–1108.

Coimbra, Nuno and Hélene Rey (2024). "Financial cycles with heterogeneous intermediaries". In: *Review of Economic Studies* 91.2, pp. 817–857.

Coles, Jeffrey L, Davidson Heath, and Matthew C Ringgenberg (2022). "On index investing". In: *Journal of Financial Economics* 145.3, pp. 665–683.

Dasgupta, Amil, Andrea Prat, and Michela Verardo (2005). "The price of conformism". In: *London School of Economics Working Paper*.

Davis, Carter, Mahyar Kargar, and Jiacui Li (2023). "Why is asset demand inelastic?" In: *Why is Asset Demand Inelastic?: Davis, Carter— uKargar, Mahyar— uLi, Jiacui*.

De Jong, Piet and Neil Shephard (1995). "The simulation smoother for time series models". In: *Biometrika* 82.2, pp. 339–350.

Dou, Winston Wei, Itay Goldstein, and Yan Ji (2024). "Ai-powered trading, algorithmic collusion, and price efficiency". In: *Jacobs Levy Equity Management Center for Quantitative Financial Research Paper*.

Duane, Simon et al. (1987). "Hybrid monte carlo". In: *Physics letters B* 195.2, pp. 216–222.

Dyer, Joel et al. (2024). "Black-box Bayesian inference for agent-based models". In: *Journal of Economic Dynamics and Control* 161, p. 104827.

Eaton, Gregory W et al. (2022). "Retail trader sophistication and stock market quality: Evidence from brokerage outages". In: *Journal of Financial Economics* 146.2, pp. 502–528.

Engle, Robert F, David F Hendry, and Jean-Francois Richard (1983). "Exogeneity". In: *Econometrica: Journal of the Econometric Society*, pp. 277–304.

Fama, Eugene F and Kenneth R French (2018). "Choosing factors". In: *Journal of financial economics* 128.2, pp. 234–252.

Fama, Eugene F and James D MacBeth (1973). "Risk, return, and equilibrium: Empirical tests". In: *Journal of political economy* 81.3, pp. 607–636.

Feng, Guanhao, Stefano Giglio, and Dacheng Xiu (2020). "Taming the factor zoo: A test of new factors". In: *The Journal of Finance* 75.3, pp. 1327–1370.

Feng, Guanhao, Jingyu He, et al. (2024). "Deep learning in characteristics-sorted factor models". In: *Journal of Financial and Quantitative Analysis* 59.7, pp. 3001–3036.

Ferraresi, Mauricio (2025b). "Beliefs of investment platforms". Unpublished manuscript, São Paulo.

— (2025c). "Exposure market with granular investment platforms: Equilibrium and causality". Unpublished manuscript, São Paulo.

Ferraresi, Mauricio, Henrique Castro, and Claudia Yoshinaga (2025). "Performance of sustainable investment platforms with social responsability during COVID-19". Unpublished manuscript, São Paulo.

Ferraresi, Mauricio and Giovanni Di Pietra (2025). "Opportunistic investment platforms behavior". Unpublished manuscript, São Paulo.

Ferraresi, Mauricio and Laura Leal (2025). "Forecasting exposure on investment platforms". Unpublished manuscript, São Paulo.

Ferraresi, Mauricio and Fernando Urbano (2025). "Analogy between investment platforms and ETFs: Risk and return of exposure-based strategies". Unpublished manuscript, São Paulo.

Fiechter, Peter and Louis Mangeney (2020). "The impact of the institutional environment on analysts' herding behavior: Evidence from broker acquisitions". In: *Available at SSRN 3741879*.

Frühwirth-Schnatter, Sylvia (1994). "Data augmentation and dynamic linear models". In: *Journal of time series analysis* 15.2, pp. 183–202.

Frühwirth-Schnatter, Sylvia, Darjus Hosszejni, and Hedibert Freitas Lopes (2024). "Sparse Bayesian factor analysis when the number of factors is unknown". In: *Bayesian Analysis* 1.1, pp. 1–44.

Fuchs, William, Satoshi Fukuda, and Daniel Neuhann (2023). "Demand-system asset pricing: Theoretical foundations". In: *Available at SSRN 4672473*.

Gabaix, Xavier (2011). "The granular origins of aggregate fluctuations". In: *Econometrica* 79.3, pp. 733–772.

Gabaix, Xavier and Ralph SJ Koijen (2021). *In search of the origins of financial fluctuations: The inelastic markets hypothesis*. Tech. rep. National Bureau of Economic Research: National Bureau of Economic Research.

— (2024). "Granular instrumental variables". In: *Journal of Political Economy* 132.7, pp. 2274–2303.

Garg, Ashish and Rachita Gulati (2013). "Do investors herd in Indian market". In: *Decision* 40, pp. 181–196.

Girolami, Mark and Ben Calderhead (2011). "Riemann manifold langevin and hamiltonian monte carlo methods". In: *Journal of the Royal Statistical Society Series B: Statistical Methodology* 73.2, pp. 123–214.

Greenwood, Robin and Marco Sammon (2022). "The disappearing index effect". In: *The Journal of Finance*.

Guo, Xu et al. (2024). "Institutional herding and investor sentiment". In: *Journal of Financial Markets* 68, p. 100891.

Gutierrez, Roberto C and Eric K Kelley (2009). "Institutional herding and future stock returns". In: *Available at SSRN* 1107523.

Haddad, Valentin, Paul Huebner, and Erik Loualiche (2025). "How competitive is the stock market? Theory, evidence from portfolios, and implications for the rise of passive investing". In: *American Economic Review* 115.3, pp. 975–1018.

Haddad, Valentin and Tyler Muir (2021). "Do intermediaries matter for aggregate asset prices?" In: *The Journal of Finance* 76.6, pp. 2719–2761.

— (2025). *Market macrostructure: Institutions and asset prices*. Tech. rep. National Bureau of Economic Research.

Harvey, Campbell R (2017). "Presidential address: The scientific outlook in financial economics". In: *The Journal of Finance* 72.4, pp. 1399–1440.

Harvey, Campbell R, Yan Liu, and Heqing Zhu (2016). "... and the cross-section of expected returns". In: *The Review of Financial Studies* 29.1, pp. 5–68.

He, Chengying et al. (2022). "Volatility correlation structure, dynamic network and portfolio implications of Chinese stock market". In: *Procedia Computer Science* 202, pp. 122–127.

Heath, Davidson et al. (2022). "Do index funds monitor?" In: *The Review of Financial Studies* 35.1, pp. 91–131.

Hirshleifer, David (2020). "Presidential address: Social transmission bias in economics and finance". In: *The Journal of Finance* 75.4, pp. 1779–1831.

Hoffman, Matthew D, Andrew Gelman, et al. (2014). "The No-U-Turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo." In: *J. Mach. Learn. Res.* 15.1, pp. 1593–1623.

Huebner, Paul (2023). "The making of momentum: A demand-system perspective". In: *Proceedings of the EUROFIDAI-ESSEC Paris December Finance Meeting*.

Ioannidis, Evangelos, Iordanis Sarikeisoglou, and Georgios Angelidis (2023). "Portfolio construction: A network approach". In: *Mathematics* 11.22, p. 4670.

Jansen, Kristy AE, Wenhao Li, and Lukas Schmid (2024). *Granular treasury demand with arbitrageurs*. Tech. rep. National Bureau of Economic Research.

Jensen, Theis Ingerslev, Bryan Kelly, and Lasse Heje Pedersen (2023). "Is there a replication crisis in finance?" In: *The Journal of Finance* 78.5, pp. 2465–2518.

Jiang, Zhengyang, Robert J Richmond, and Tony Zhang (2024). "A portfolio approach to global imbalances". In: *The Journal of Finance* 79.3, pp. 2025–2076.

Jurkatis, Simon, Stephanie Kremer, and Dieter Nautz (2012). *Correlated trades and herd behavior in the stock market*. Tech. rep. SFB 649 Discussion Paper.

Kaniel, Ron et al. (2023). "Machine-learning the skill of mutual fund managers". In: *Journal of Financial Economics* 150.1, pp. 94–138.

Kastner, Gregor, Sylvia Frühwirth-Schnatter, and Hedibert Freitas Lopes (2016). "Efficient Bayesian inference for multivariate factor stochastic volatility models". In: *Journal of Computational and Graphical Statistics* 26, pp. 905–917.

Kim, Byungwook (2025). "Correlated demand shocks and asset pricing". In: *Available at SSRN 4705500*.

Koijen, Ralph SJ, Robert J Richmond, and Motohiro Yogo (2024). "Which investors matter for equity valuations and expected returns?" In: *Review of Economic Studies* 91.4, pp. 2387–2424.

Koijen, Ralph SJ and Motohiro Yogo (2019). "A demand system approach to asset pricing". In: *Journal of Political Economy* 127.4, pp. 1475–1515.

— (2020). *Exchange rates and asset prices in a global demand system*. Tech. rep. National Bureau of Economic Research: National Bureau of Economic Research.

Koop, Gary, Dimitris Korobilis, et al. (2010). "Bayesian multivariate time series methods for empirical macroeconomics". In: *Foundations and Trends® in Econometrics* 3.4, pp. 267–358.

Kozak, Serhiy, Stefan Nagel, and Shrihari Santosh (2018). "Interpreting factor models". In: *The Journal of Finance* 73.3, pp. 1183–1223.

Kremer, Stephanie and Dieter Nautz (2013). "Short-term herding of institutional traders: New evidence from the German stock market". In: *European Financial Management* 19.4, pp. 730–746.

Krokida, Styliani-Iris, Panagiota Makrychoriti, and Spyros Spyrou (2020). "Monetary policy and herd behavior: International evidence". In: *Journal of Economic Behavior & Organization* 170, pp. 386–417.

Lai, Ming-Ming and Siok-Hwa Lau (2004). "Herd behavior and market stress: The case of Malaysia". In: *Academy of Accounting and Financial Studies Journal* 8.3, p. 85.

LeBaron, Blake (2006). "Agent-based computational finance". In: *Handbook of computational economics* 2, pp. 1187–1233.

Li, Kai (2014). *Asset price dynamics with heterogeneous beliefs and time delays*. University of Technology Sydney (Australia).

Li, Wei, Ghon Rhee, and Steven Shuye Wang (2017). "Differences in herding: Individual vs. institutional investors". In: *Pacific-Basin Finance Journal* 45, pp. 174–185.

Loang, Ooi Kok (2025). "Can machine learning surpass human investors? Evidence from adaptive herding behaviour in US, China and India". In: *Journal of Applied Economics* 28.1, p. 2435796.

Lopes, Hedibert Freitas and Mike West (2004). "Bayesian model assessment in factor analysis". In: *Statistica Sinica*, pp. 41–67.

Lopez-Lira, Alejandro and Nikolai L Roussanov (2020). "Do common factors really explain the cross-section of stock returns?" In: *Jacobs Levy Equity Management Center for Quantitative Financial Research Paper*.

Loseto, Marco and Federico Mainardi (2023). "Oligopolistic competition, fund proliferation and asset prices". In: *Fund Proliferation and Asset Prices (February 19, 2023)*.

Markowitz, H. (1952). "Portfolio selection". In: *The Journal of Finance* 7.1, pp. 77–91.

Nakajima, Jouchi and Michael A. West (2013). "Dynamic factor volatility modeling: A Bayesian latent threshold approach". In: *Journal of Financial Econometrics* 11, pp. 116–153.

Neal, Radford M et al. (2011). "MCMC using hamiltonian dynamics". In: *Handbook of Markov Chain Monte Carlo* 2.11, p. 2.

Neuhann, Daniel and Michael Sockin (2024). "Financial market concentration and misallocation". In: *Journal of Financial Economics* 159, p. 103875.

Nirei, Makoto, Theodoros G Stamatiou, and Vladyslav Sushko (2012). "Stochastic herding in financial markets evidence from institutional investor equity portfolios". In: *BIS Working Paper*.

Pavlova, Anna and Taisiya Sikorskaya (2023). "Benchmarking intensity". In: *The Review of Financial Studies* 36.3, pp. 859–903.

Pedersen, Lasse Heje (2022). "Game on: Social networks and markets". In: *Journal of Financial Economics* 146.3, pp. 1097–1119.

Pedraza, Alvaro and Fredy Pulga (2019). "Asset price effects of peer benchmarking: Evidence from a natural experiment". In: *International Review of Economics & Finance* 62, pp. 53–65.

Peralta, Gustavo and Abalfazl Zareei (2016). "A network approach to portfolio selection". In: *Journal of Empirical Finance* 38, pp. 157–180.

Petersen, Gesa-Kristina and Theresa Spickers (2023). "Hiding in the herd: Acute stress, conformity, and bubbles". In: *Hiding in the Herd: Acute Stress, Conformity, and Bubbles: Petersen, Gesa-Kristina—Spickers, Theresa*. [Sl]: SSRN.

Polson, Nicholas G. and James G. Scott (2011). "On the half-cauchy prior for a global scale parameter". In: *Bayesian Analysis* 7, pp. 887–902.

Puckett, Andy and Xuemin Sterling Yan (2008). "Short-term institutional herding and its impact on stock prices". In: *Available at SSRN 1108092*.

Ren, Fei et al. (2017). "Dynamic portfolio strategy using clustering approach". In: *PloS one* 12.1, e0169299.

Rey, Hélene et al. (2024). *Elephants in equity markets*. Tech. rep. National Bureau of Economic Research.

Ross, Stephen A (1976). "The arbitrage theory of capital asset pricing". In: *Journal of Economic Theory* 13.3, pp. 341–360.

Rostek, Marzena and Ji Hee Yoon (2023). "Imperfect competition in financial markets: Recent developments". In: *Journal of Economic Literature*.

Ruth, Ronald D (1983). "A canonical integration technique". In: *IEEE Trans. Nucl. Sci.* 30.CERN-LEP-TH-83-14, pp. 2669–2671.

Sammon, Marco and John J Shim (2024). "Who clears the market when passive investors trade". In: *Available at SSRN*.

Sharpe, William F (1964). "Capital asset prices: A theory of market equilibrium under conditions of risk". In: *The journal of finance* 19.3, pp. 425–442.

Shyu, Jonchi and Hsin-Ming Sun (2010). "Do institutional investors herd in emerging markets? Evidence from the Taiwan stock market". In: *Asian Journal of Finance & Accounting* 2.2, p. 1.

Sigauke, Caston (2016). "Volatility modeling of the JSE all share index and risk estimation using the Bayesian and frequentist approaches". In: *Economics, Management, and Financial Markets* 11.4, pp. 33–48.

Tan, Lin, Xiaoyan Zhang, and Xinran Zhang (2023). "Retail and institutional investor trading behaviors: Evidence from China". In: *Annual Review of Financial Economics* 16.

Teh, Lilyn L, Werner FM De Bondt, et al. (1997). "Herding behavior and stock returns: An exploratory investigation". In: *Revue Suisse D Economie Politique Et De Statistique* 133, pp. 293–324.

Thomas, Samuel and Wanzhu Tu (2021). "Learning Hamiltonian Monte Carlo in R". In: *The American Statistician* 75.4, pp. 403–413.

Turing, Alan Mathison (1950). "Computing machinery and intelligence". In: *Mind* 59.236, pp. 433–460.

Uwilingiye, Josine et al. (2019). "A note on the technology herd: Evidence from large institutional investors". In: *Review of Behavioral Finance* 11.3, pp. 294–308.

Van der Beck, Philippe (2022). "On the estimation of demand-based asset pricing models". In: *Swiss Finance Institute Research Paper* 22-67.

Van der Beck, Philippe and Coralie Jaunin (2021). "The equity market implications of the retail investment boom". In: *Swiss Finance Institute Research Paper* 21-12.

Ward, Eric J. et al. (2019). "Modeling regimes with extremes: The bayesdfa package for identifying and forecasting common trends and anomalies in multivariate time-series data". In: *The R Journal*.

Yang, Wang and Ooi Kok Loang (2024). "Systematic literature review: Behavioural biases as the determinants of herding". In: *Technology-Driven Business Innovation: Unleashing the Digital Advantage, Volume 1*, pp. 79–92.

Young, Laura Jackson et al. (2015). "Specification and estimation of Bayesian dynamic factor models: A Monte Carlo analysis with an application to global house price comovement". In: *Econometrics: Econometric Model Construction*.

Zhao, Longfeng et al. (2018). "Stock market as temporal network". In: *Physica A: Statistical Mechanics and its Applications* 506, pp. 1104–1112.

A Factor

The system of equations (32) and (33) considers only one state variable to represent what would be a vector of factors. And so for $K = 1$ I assumed that $\boldsymbol{\lambda}_m' \boldsymbol{\kappa}_{m,t} = \lambda_m \kappa_{m,t}$. The expressions were also written only for a particular stock belonging to an exposure portfolio $\boldsymbol{e}_{m,t} \in \mathbb{R}^{N \times 1}$. Thus, the parameters in matrix form Equations (23), (24), and (25) can be written in a simpler form by

$$\xi_{m,t} = \frac{\mu_m}{\gamma_m \sigma_\epsilon^2}, \tag{48}$$

$$\lambda_m = -\frac{\beta}{\gamma_m \sigma_\epsilon^2}, \tag{49}$$

$$\kappa_{m,t} = \left(\frac{\beta \sigma_{\zeta,t}^2}{\sigma_\epsilon^2 + \beta^2 \sigma_{\zeta,t}^2} \right) \mu_m. \tag{50}$$

The fixed effect of the platform ξ_m depends on the beliefs of the platforms about the expected value of the returns, weighted by the product of risk aversion multiplied by the variance of the idiosyncratic error of the returns. The factor load λ_m is just a different way to write the risk factor load β. This different form depends on the risk aversion of each platform and can provide an opportunity to quantify different perceptions of a parameter that should be common to all investors. Finally, the dynamic factor $\kappa_{m,t}$ is also a way of rewriting beliefs about the expected value of a stock's return. The dynamics of the factor is due to the volatility of the risk factor over time $\sigma_{\zeta,t}^2$.

I study the dynamics of the factor function $\kappa_{m,t}$ in the Appendix B. I show the existence of fixed points and analyze the stability of the dynamics. The implications for the estimates discussed in the empirical analysis chapter are that the parameter governing the transition of states must be $|\phi| < 1$. With the possibility that it is either $-1 < \phi < 0$ or $0 < \phi < 1$.

In a simpler case, for a single asset and a single factor, it is easy to verify that when $\beta = 1$, the factor loadings and the dynamic factor can be written as follows

$$\lambda_m = -\frac{1}{\gamma_m \sigma_\epsilon^2},$$

$$\kappa_{m,t} = \left(\frac{\sigma_{\zeta,t}^2}{\sigma_\epsilon^2 + \sigma_{\zeta,t}^2} \right) \mu_m$$

and thus the dynamic factor $\kappa_{m,t}$ has a transformation term of the expected value of the returns that is bounded between $[0, 1]$, because the only bounded cases to evaluate are when $\sigma_\epsilon^2 \gg \sigma_{\zeta,t}^2$ and $\sigma_\epsilon^2 \ll \sigma_{\zeta,t}^2$. In the case of multiple factors, the dynamic linear transformation matrix, $\boldsymbol{B}_t$, will depend on the structure of the covariance matrix of the risk factors in the expression (27).

When considering an exact error decomposition with $\kappa_{m,t} \perp\!\!\!\perp \tau_{m,t}$, there is an important restriction that can be tested. If there are no heterogeneous beliefs about the expected values of the returns, then the factor model by platform is a common demand shock model, because in this case[33]

$$\kappa_{m,t} = \frac{\beta \mu \, \sigma_{\zeta,t}^2}{\sigma_\epsilon^2 + \beta^2 \sigma_{\zeta,t}^2},$$

$$\kappa_{m,t} = \eta_t. \tag{51}$$

[33]This possibility of a common factor across platforms is discussed in Ferraresi (2025b).

B Factor Dynamics

The fixed points of the factor function $\kappa_{m,t}$ and their stability for the iteration given by

$$\sigma^2_{\zeta,t+1} = f(\sigma^2_{\zeta,t}) = \frac{\beta\mu_m\,\sigma^2_{\zeta,t}}{\sigma^2_\epsilon + \beta^2\,\sigma^2_{\zeta,t}},$$

where the parameters μ_m, β, and σ^2_ϵ are constants are examined. A fixed point $\overline{\sigma}^2_\zeta$ satisfies

$$\overline{\sigma}^2_\zeta = \frac{\beta\mu_m\,\overline{\sigma}^2_\zeta}{\sigma^2_\epsilon + \beta^2\,\overline{\sigma}^2_\zeta}.$$

Solving this equation explicitly, I obtain:

$$\overline{\sigma}^2_\zeta\left(\sigma^2_\epsilon + \beta^2\,\overline{\sigma}^2_\zeta\right) = \beta\mu_m\,\overline{\sigma}^2_\zeta.$$

Simplifying, I have:

$$\overline{\sigma}^2_\zeta\left(\sigma^2_\epsilon + \beta^2\,\overline{\sigma}^2_\zeta - \beta\mu_m\right) = 0,$$

resulting in two candidate fixed points:

$$\overline{\sigma}^2_{\zeta,1} = 0, \quad \text{and} \quad \overline{\sigma}^2_{\zeta,2} = \frac{\beta\mu_m - \sigma^2_\epsilon}{\beta^2}.$$

Considering the domain of interest ($\sigma^2_{\zeta,t} \geq 0$), the second point $\overline{\sigma}^2_{\zeta,2}$ exists and is economically meaningful only if $\beta\mu_m \geq \sigma^2_\epsilon$. Otherwise, this fixed point is negative and outside the relevant domain.

To determine the stability of these fixed points, I analyze the absolute value of the derivative:

$$f'(\sigma^2_{\zeta,t}) = \frac{\beta\mu_m\sigma^2_\epsilon}{\left(\sigma^2_\epsilon + \beta^2\sigma^2_{\zeta,t}\right)^2}.$$

Evaluating at each fixed point yields the following:

At $\overline{\sigma}^2_{\zeta,1} = 0$:

$$f'(0) = \frac{\beta\mu_m}{\sigma^2_\epsilon}.$$

Thus, the stability at zero is characterized by:

$$|f'(0)| = \left|\frac{\beta\mu_m}{\sigma^2_\epsilon}\right| \begin{cases} < 1, & \text{(stable) if } \beta\mu_m < \sigma^2_\epsilon, \\ > 1, & \text{(unstable) if } \beta\mu_m > \sigma^2_\epsilon, \\ = 1, & \text{(bifurcation) if } \beta\mu_m = \sigma^2_\epsilon. \end{cases}$$

At the second fixed point $\overline{\sigma}^2_{\zeta,2} = \frac{\beta\mu_m - \sigma^2_\epsilon}{\beta^2}$, assuming $\beta\mu_m > \sigma^2_\epsilon$:

$$f'\left(\frac{\beta\mu_m - \sigma^2_\epsilon}{\beta^2}\right) = \frac{\sigma^2_\epsilon}{\beta\mu_m} < 1,$$

indicating that whenever it exists (positive), this second fixed point is stable.

In summary, the dynamics of the factors is crucially dependent on the ratio $\frac{\beta\mu_m}{\sigma^2_\epsilon}$:

1. For $\beta\mu_m < \sigma_\epsilon^2$, only the zero fixed point is stable, while the second fixed point is negative and irrelevant.

2. For $\beta\mu_m = \sigma_\epsilon^2$, a bifurcation scenario with neutral stability at zero.

3. For $\beta\mu_m > \sigma_\epsilon^2$, the zero fixed point becomes unstable, and a new positive stable fixed point emerges.

Thus, the transition at $\beta\mu_m = \sigma_\epsilon^2$ constitutes a bifurcation point, marking the transition from stable equilibrium at zero to a new positive equilibrium. Finally, the stable positive fixed point given by $\overline{\sigma}^2_{\zeta,2} = \frac{\beta\mu_m - \sigma_\epsilon^2}{\beta^2}$ is relevant because real data naturally produce $\beta\mu_m > \sigma_\epsilon^2$.

C Hamiltonian Monte Carlo

The Hamiltonian Monte Carlo (HMC) algorithm is a technique for sampling in high-dimensional parametric spaces. Unlike traditional Monte Carlo methods based on Metropolis-Hastings, HMC uses an approach inspired by classical mechanics to explore parametric space more efficiently. This algorithm introduces an auxiliary moment variable for each parameter in the model, which allows the construction of Hamiltonian dynamics to guide the sampling (Duane et al., 1987; Neal et al., 2011; Girolami and Calderhead, 2011; Betancourt, 2017; Thomas and Tu, 2021). The basic idea of HMC is to combine the parameter of interest, usually represented by the parametric vector $\boldsymbol{\theta}$, with the moment variables $\boldsymbol{r}$ sampled from a standard normal distribution. The joint system has a non-normalized joint density proportional to $\exp\left(\log \boldsymbol{\theta} - \frac{1}{2}\boldsymbol{r}^T\boldsymbol{r}\right)$. This model allows us to interpret $\boldsymbol{\theta}$ and $\boldsymbol{r}$ as the position and momentum of a fictitious particle, respectively, and $\log \boldsymbol{\theta}$ as the potential energy associated with position $\boldsymbol{\theta}$.

The evolution of this system is simulated using the leapfrog integrator, which preserves the volume and time reversibility properties of Hamiltonian dynamics. The integrator iteratively updates to $\boldsymbol{\theta}$ and $\boldsymbol{r}$ in small ϵ steps. The solution of Hamiltonian dynamics is usually addressed by Euler's discretization method. The process is described by three main equations that alternate between updating the momentum based on the gradient of the potential energy and updating the position based on the momentum. This makes it possible to simulate a trajectory in a parametric space that approximately conserves the total energy of the system. After executing L steps of the leapfrog integrator, a new proposal $(\tilde{\boldsymbol{\theta}}, \tilde{\boldsymbol{r}})$ is generated. This proposal is accepted or rejected according to the Metropolis probability, which depends on the energy difference between the proposed state and the current state. Theoretically, time reversibility and volume preservation ensure that the resulting Markov chain is ergodic and converges to the target distribution.

One of the key advantages of HMC is its ability to efficiently explore the high-dimensional parametric space. The generated samples tend to be farther apart, reducing the autocorrelation between samples and thus improving statistical efficiency. This feature makes HMC particularly useful for problems where the target density has complex correlations or difficult geometries, such as highly curved surfaces. However, the performance of the HMC is strongly dependent on the choice of the hyperparameters ϵ (step size) and L (number of leapfrog steps). Very large values for ϵ can lead to inaccurate simulations and high rejection rates, while very small values lead to high computational cost, since many iterations are needed to cover the parametric space. Similarly, an inappropriate choice of L can result in trajectories that do not sufficiently explore the parametric space, wasting computational resources.

However, even with a good configuration of ϵ and L, traditional HMC can struggle in situations where parametric space geometries generate long and inefficient trajectories. This limitation is particularly relevant for target densities with multiple modes or regions of high curvature, which require more targeted explorations. To mitigate these problems, the No-U-Turn Sampler (NUTS) algorithm was introduced as an extension of HMC with the aim of eliminating the need to manually select the number of L steps (Hoffman, Gelman, et al., 2014). The NUTS automatically builds trajectories in parameter space until it detects a "reversal" in the motion. This reversal is detected when the momentum vector $\boldsymbol{r}$ starts to point back in the initial direction, signaling that the trajectory is no longer moving efficiently. Mathematically, this can be expressed by checking whether $(\boldsymbol{\theta}^+ - \boldsymbol{\theta}^-)^T\boldsymbol{r}^+ < 0$, where $\boldsymbol{\theta}^+$ and $\boldsymbol{\theta}^-$ represent the ends of the trajectory generated so far, and $\boldsymbol{r}^+$ is the moment associated with the final end. This adaptive approach avoids the redundancy of tracing trajectories that have

already been explored, thus optimizing the use of computational resources.

The NUTS implements a recursive construction of binary trees to explore trajectories. At each step, the tree is expanded simultaneously in two opposite directions using the leapfrog integrator. Suggestions are accepted based on the probability of Metropolis, but the size of the trajectory is dynamically adjusted to avoid redundancy. In addition, NUTS uses stopping criteria to ensure that exploration remains within a high-probability region, reducing the need for manual parameter adjustments. Among the advantages of NUTS is its ability to dynamically adjust the size of the trajectories, eliminating the need to choose L beforehand. This is particularly useful for high-dimensional problems, where manually choosing L can be difficult and inefficient. However, NUTS still relies on an appropriate choice of step size ϵ, which can be adjusted using heuristics based on preliminary simulations. The probabilistic programming language STAN[34] uses NUTS to implement HMC as a default in any statistical model (Carpenter et al., 2017).

Although Gibbs Sampler and Metropolis-Hastings are robust and widely applicable algorithms, their limitations make HMC a superior alternative in certain contexts. The Gibbs sampler requires the existence of analytically tractable complete conditional distributions for each variable. This can be a challenge in complex or high-dimensional models, where such distributions are unknown or computationally prohibitive. In addition, Gibbs tends to suffer from high autocorrelation in parameter spaces with strong dependencies between variables, resulting in slower exploration. On the other hand, Metropolis-Hastings allows greater flexibility in the choice of proposed distributions, but its efficiency can be severely limited when the target density has high correlations or low probability regions that must be traversed to adequately explore the parameter space. In these situations, the selection of an appropriate proposal becomes a critical problem. HMC overcomes these difficulties by using gradients to guide transitions, resulting in more informed and exploratory jumps, even in highly correlated spaces.

A practical example of where HMC stands out relative to Metropolis-Hastings is illustrated in Figure 20. Consider a posterior distribution that has a three-dimensional parametric space with multiple modes:[35]

$$\mathrm{p}(\boldsymbol{\theta}) = \mathrm{A} \exp\left(-\frac{1}{2} \left((\theta_1^2 - 4)^2 + (\theta_2^2 - 4)^2 + (\theta_3 - 2)^2 - 4\cos(2\pi\theta_1) - 4\cos(2\pi\theta_2) \right) \right)$$

In this scenario, random walk Metropolis-Hastings may struggle to converge, while HMC takes advantage of gradients to efficiently navigate the complex relationships between the modes. The example was built with burning steps of 10.000 and steps after burning of 10.000 in both algorithms. I manually adjusted the size of the steps $\epsilon = 0.010$ and the number of jump steps $L = 100$ in the HMC algorithm to produce Figure 20.

[34] https://mc-stan.org/about/

[35] The normalization constant is estimated numerically as $\mathrm{A} = \frac{1}{40.20996}$ using 1.000.000 Monte Carlo samples

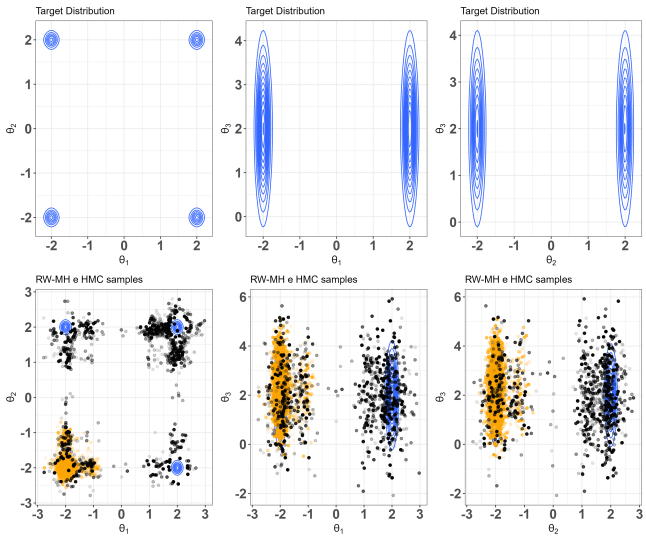
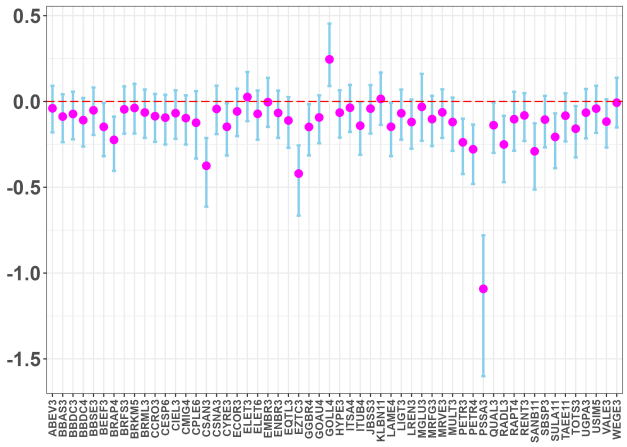
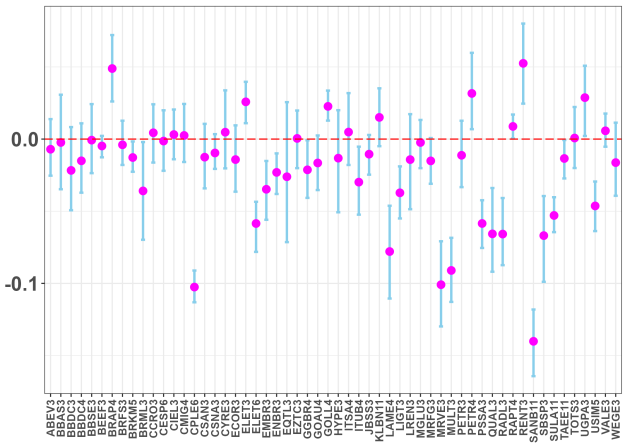


Figure 20: Yellow samples are RW-MH and black samples are HMC. HMC is not trapped in high probability regions in the posterior density showing the relevance of this new type of algorithm.

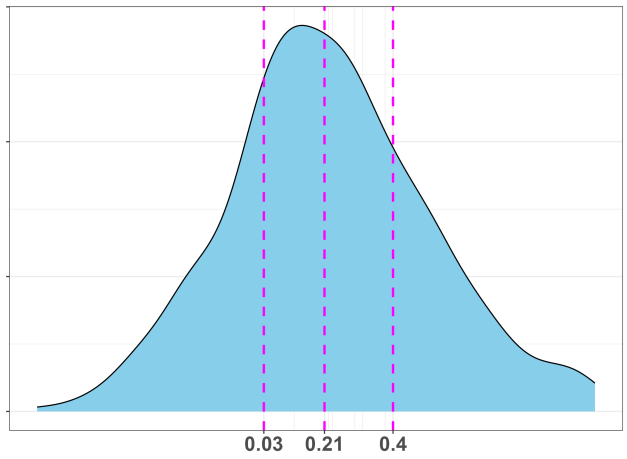
D Platforms' Parameters



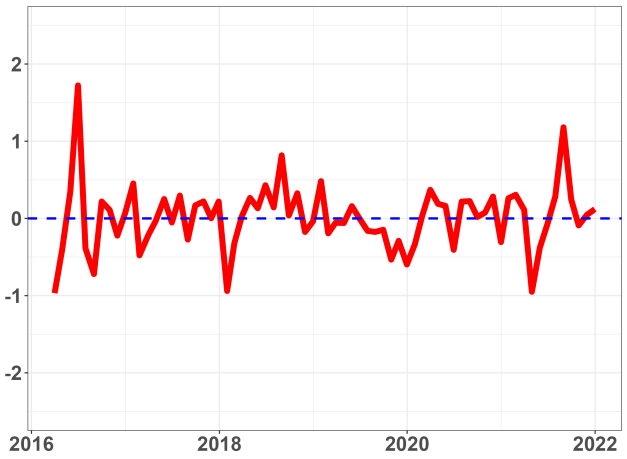
(a) loading



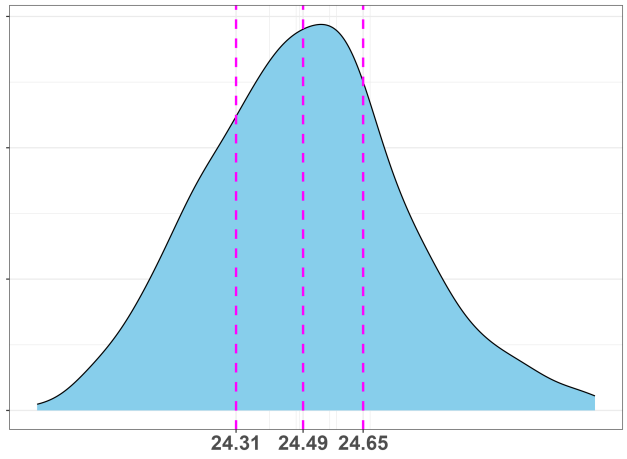
(b) exposure elasticity



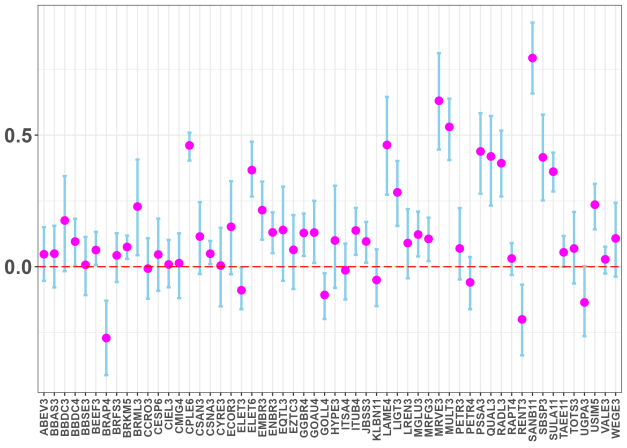
(c) ϕ 0.21 CI[0.03,0.4]



(d) factor

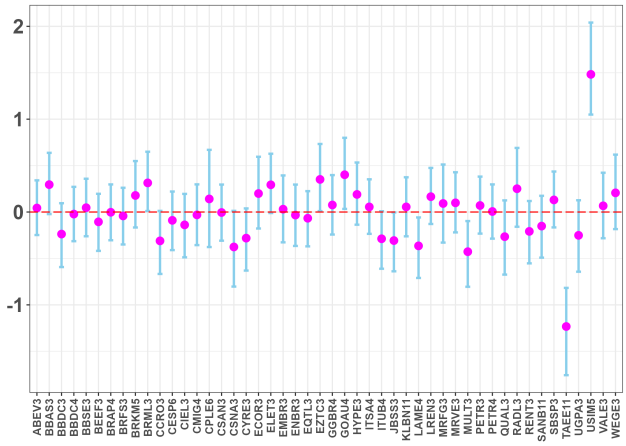


(e) σ_τ 24.49% CI[24.31%,24.65%]

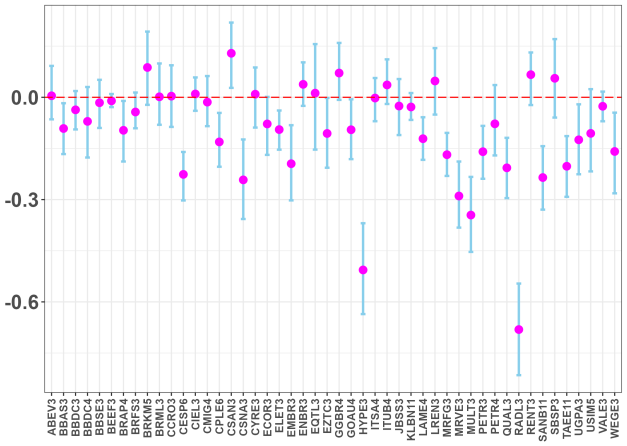


(f) intercept

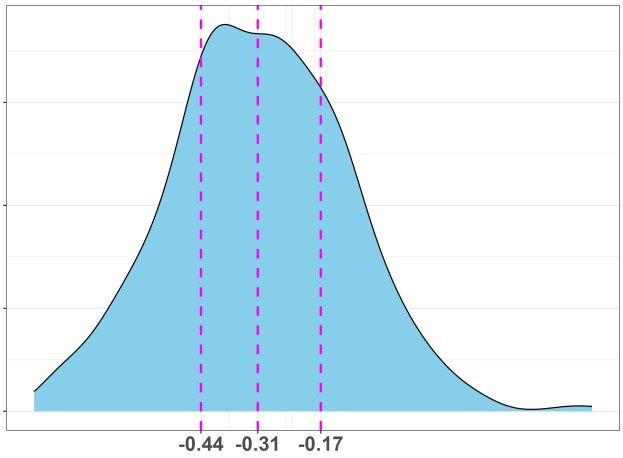
Figure 21: Banco do Brasil



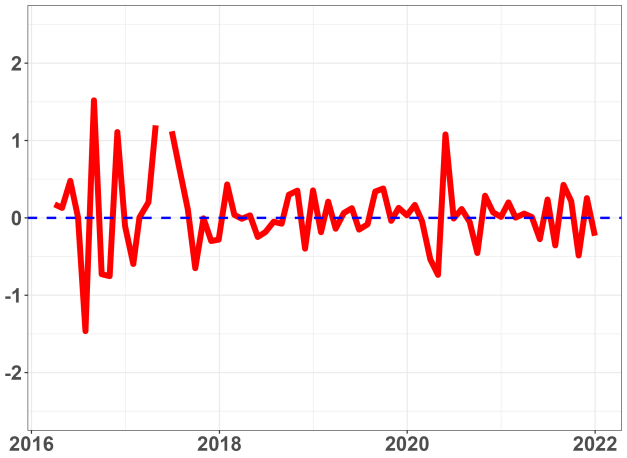
(a) loading



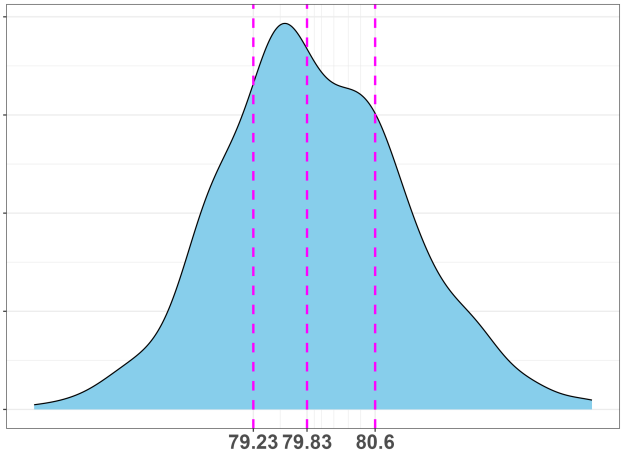
(b) exposure elasticity



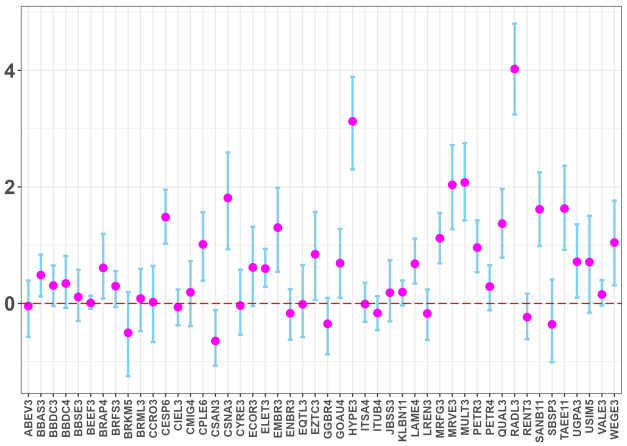
(c) ϕ -0.31 CI[-0.44,-0.17]



(d) factor

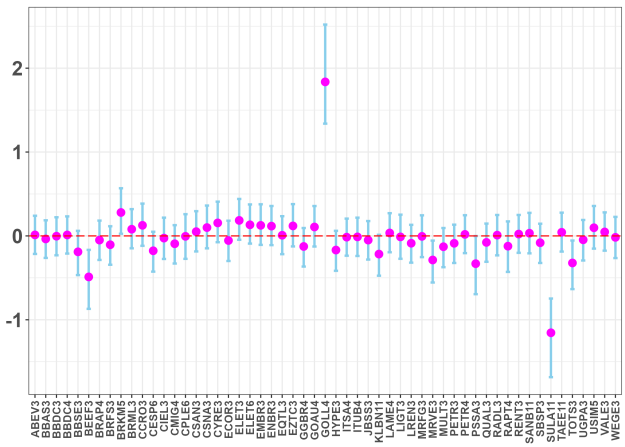


(e) σ_τ 70.83% CI[79.23%,80.6%]

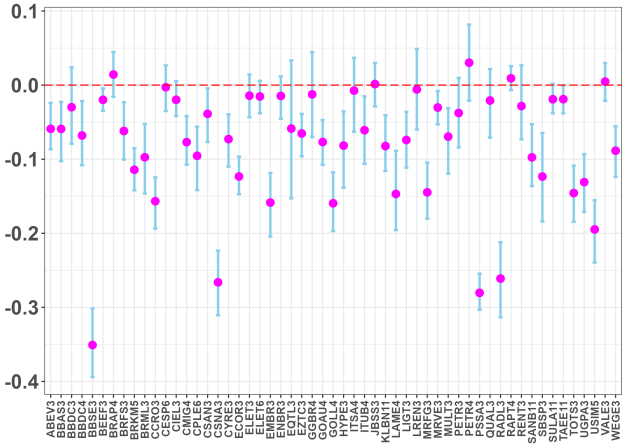


(f) intercept

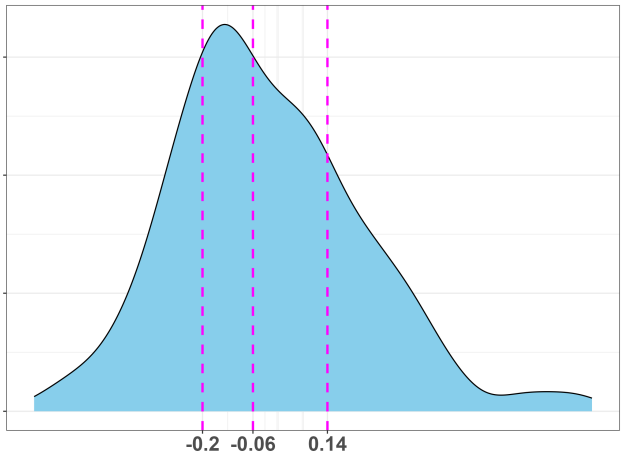
Figure 22: BNP Paribas



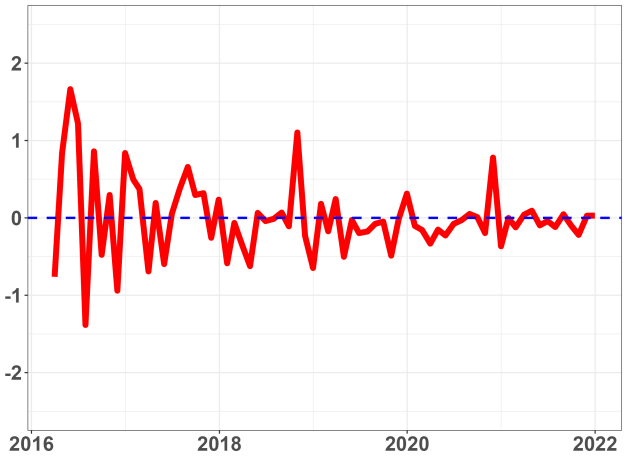
(a) loading



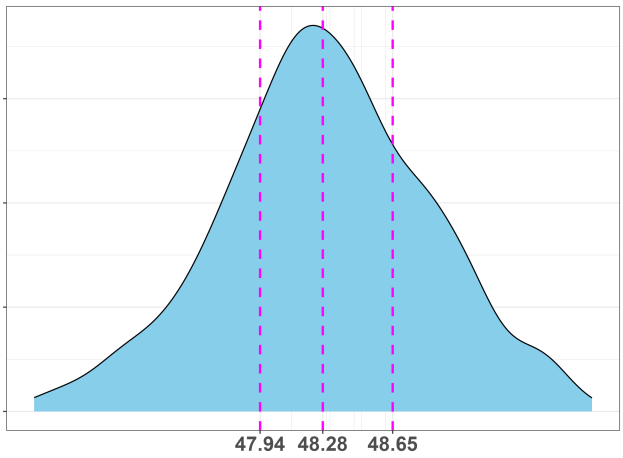
(b) exposure elasticity



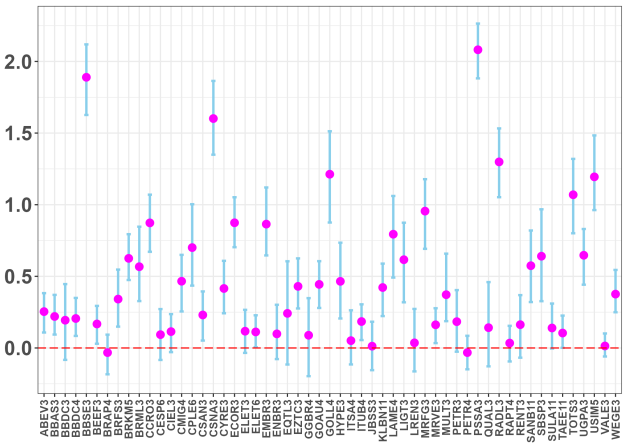
(c) ϕ -0.06 CI[-0.2,0.14]



(d) factor

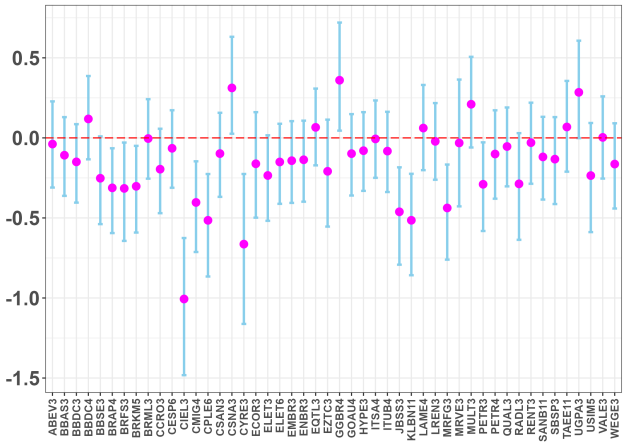


(e) σ_τ 48.28% CI[47.94%,48.65%]

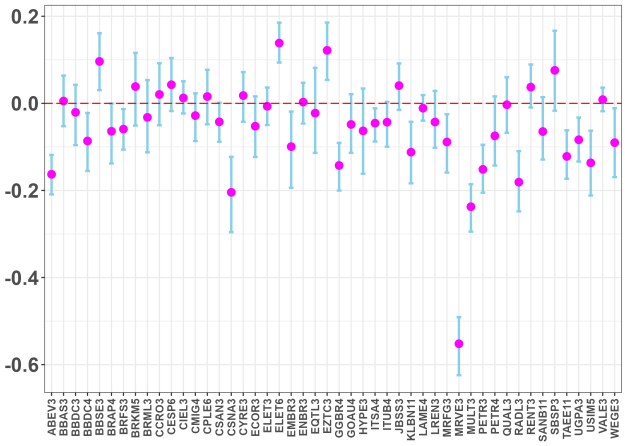


(f) intercept

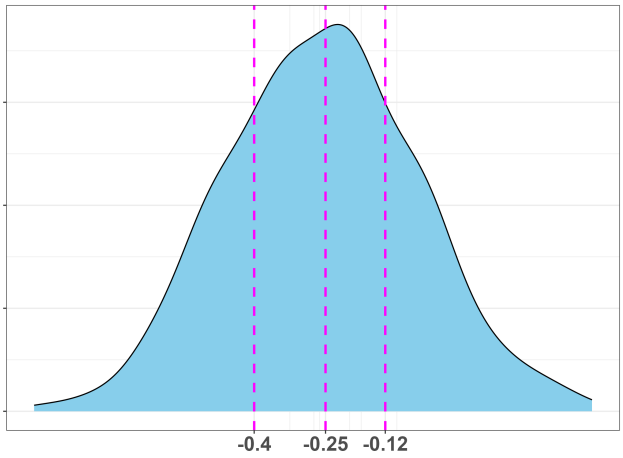
Figure 23: Bradesco



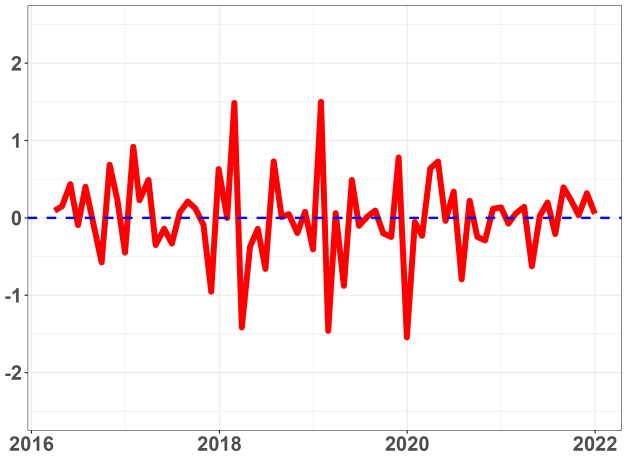
(a) loading



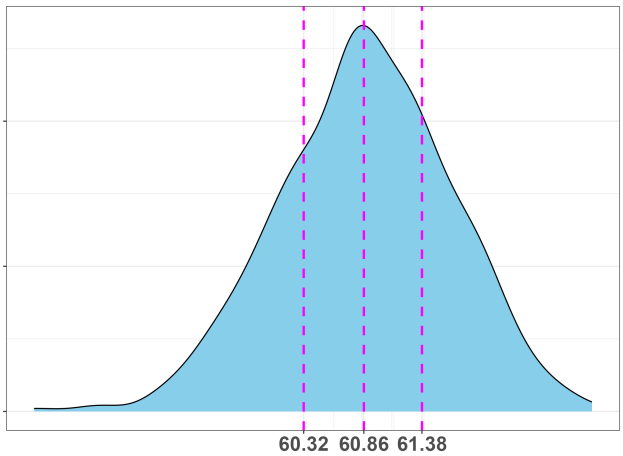
(b) exposure elasticity



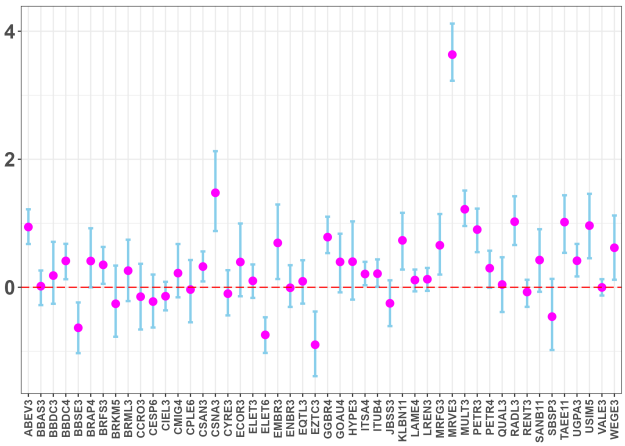
(c) ϕ -0.25 CI[-0.4,-0.12]



(d) factor

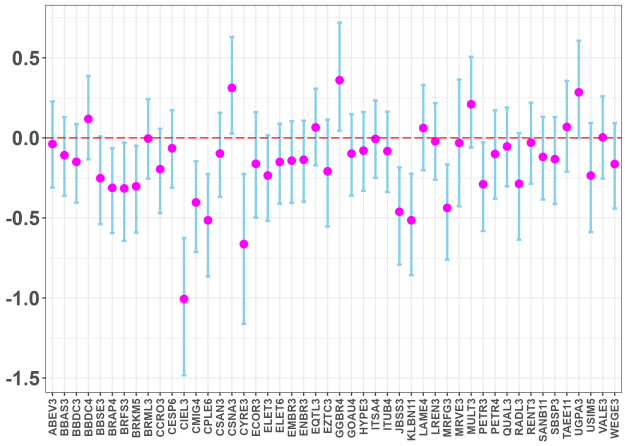


(e) σ_τ 60.86% CI[60.32%,61.38%]

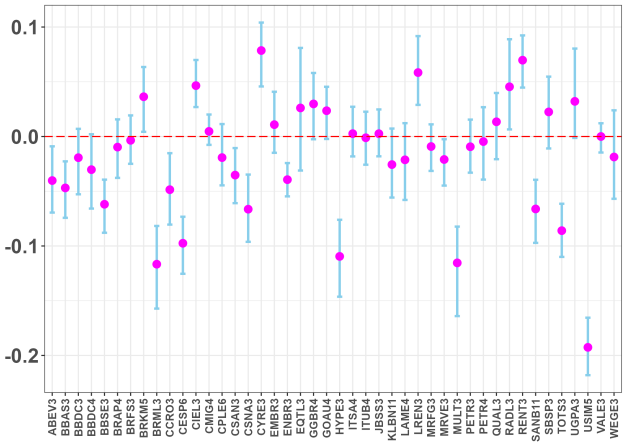


(f) intercept

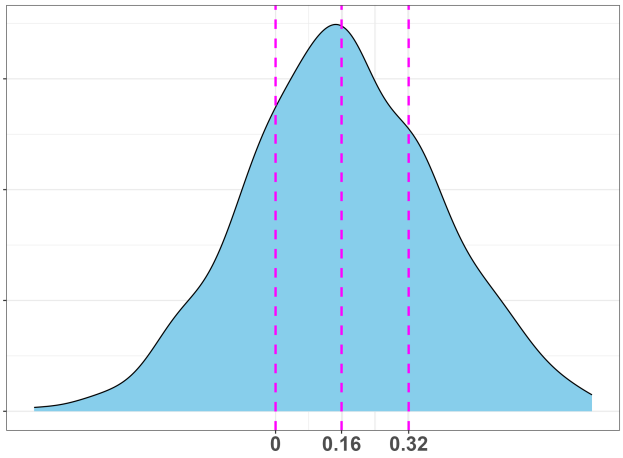
Figure 24: BTG Pactual



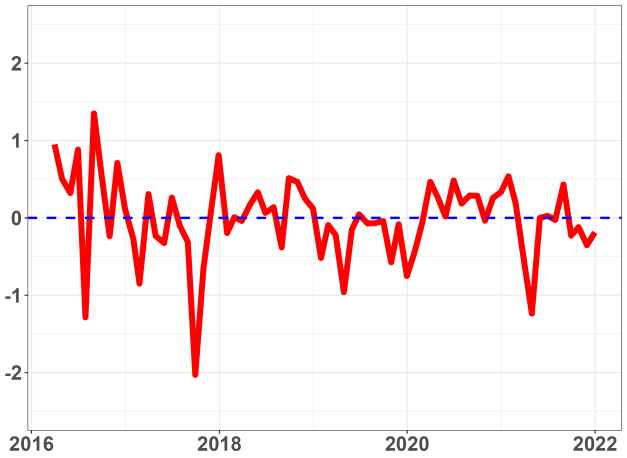
(a) loading



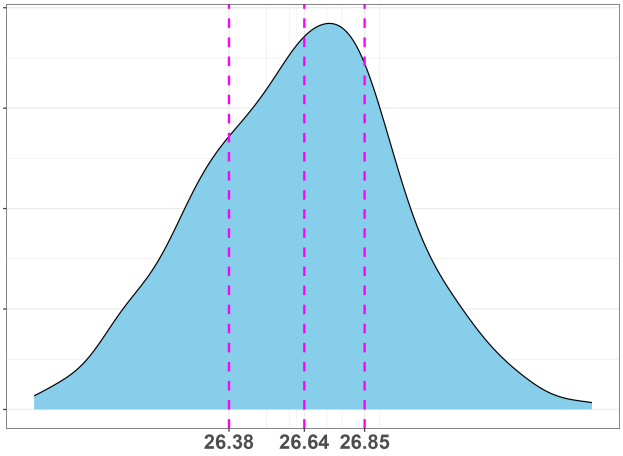
(b) exposure elasticity



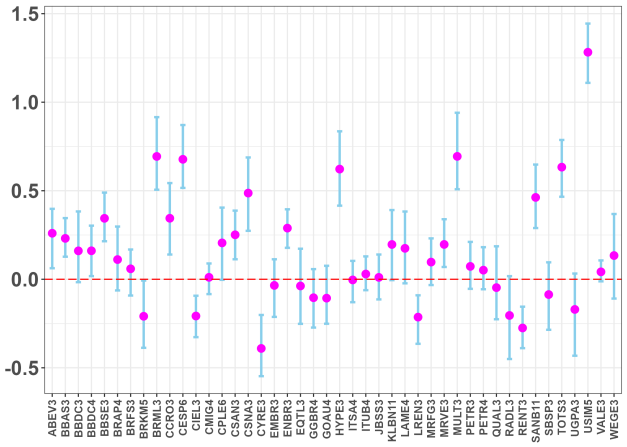
(c) ϕ 0.16 CI[0,0.32]



(d) factor

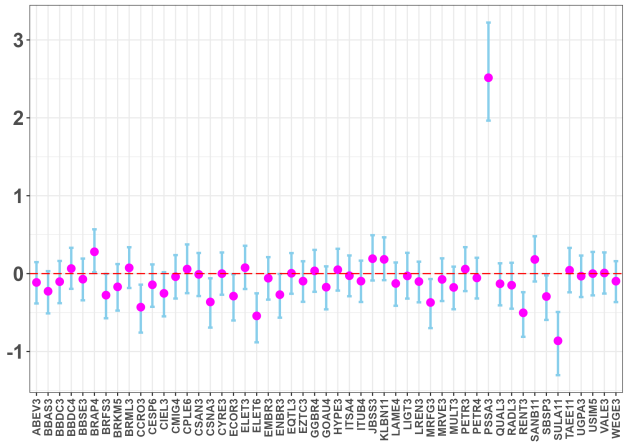


(e) σ_τ 26.64% CI[26.38%,26.85%]

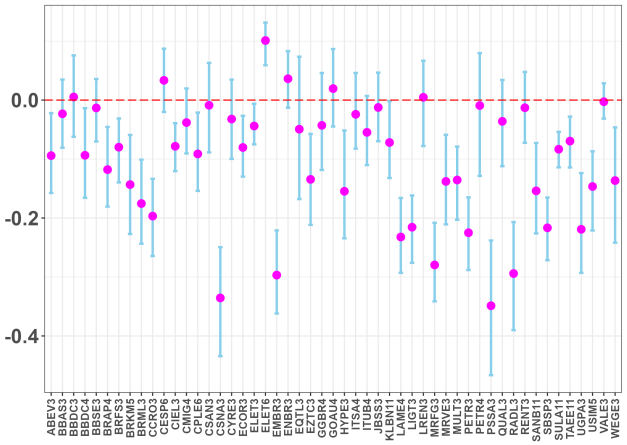


(f) intercept

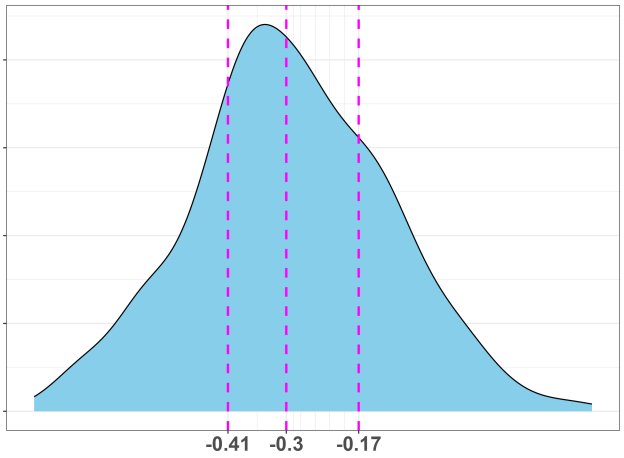
Figure 25: Caixa



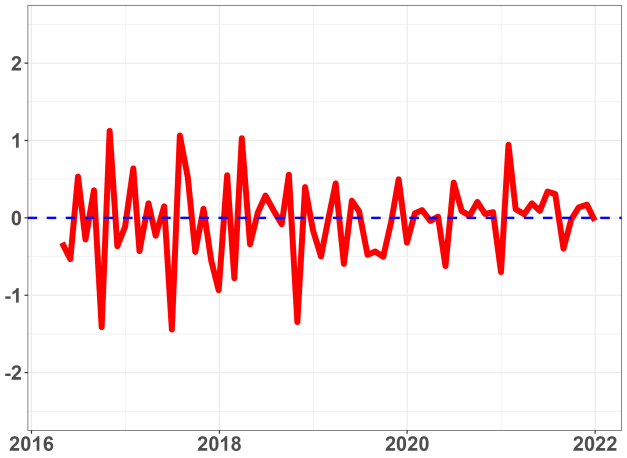
(a) loading



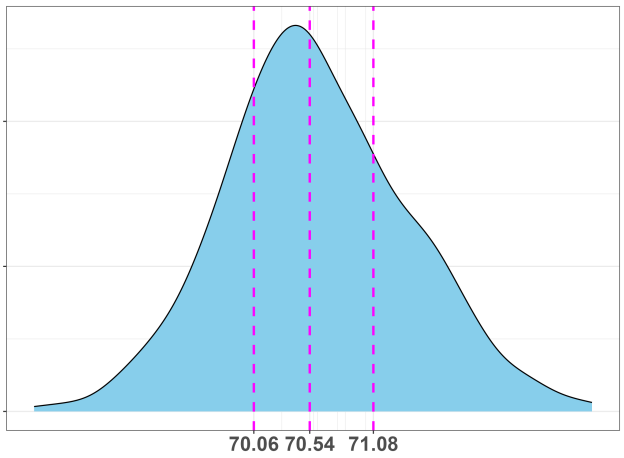
(b) exposure elasticity



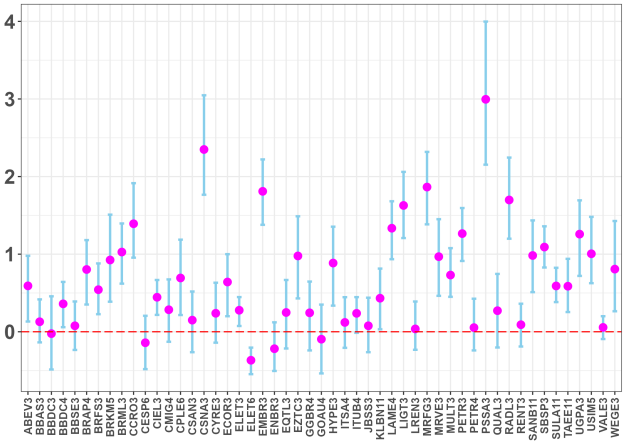
(c) ϕ -0.3 CI[-0.41,-0.17]



(d) factor

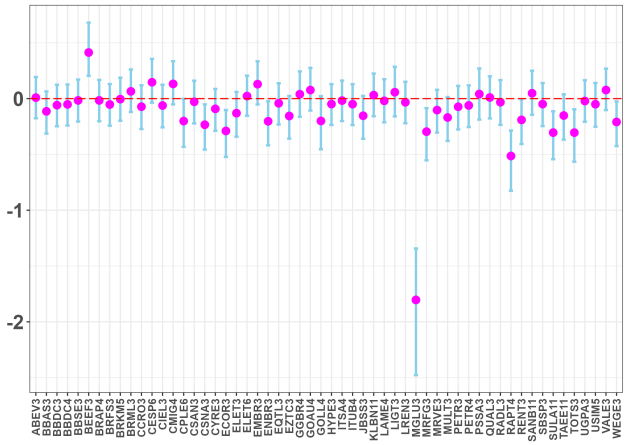


(e) σ_τ 70.54% CI[70.06%,71.08%]

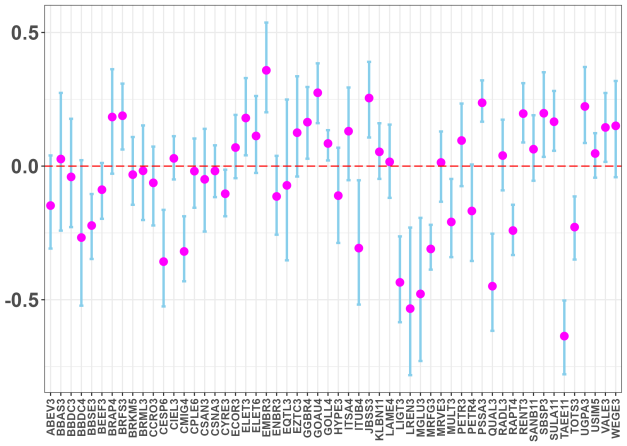


(f) intercept

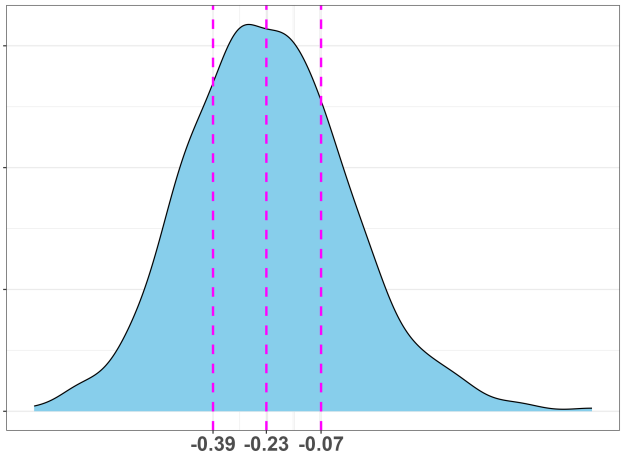
Figure 26: Credit Suisse



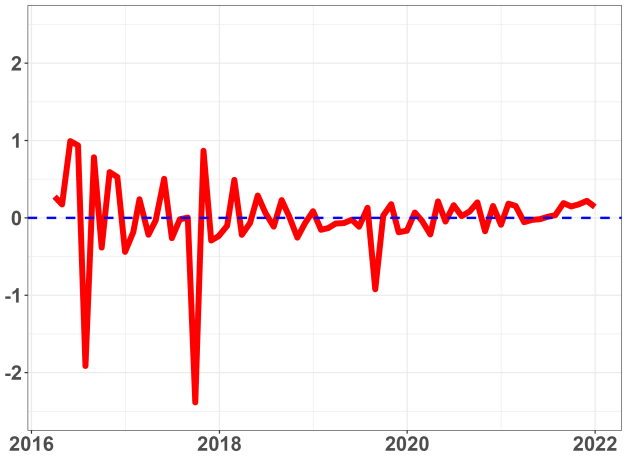
(a) loading



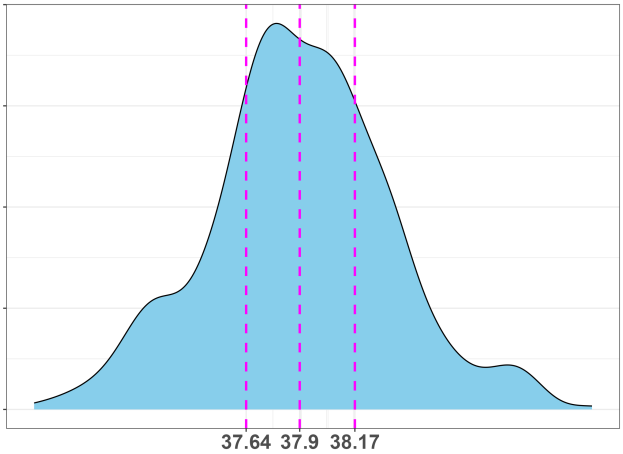
(b) exposure elasticity



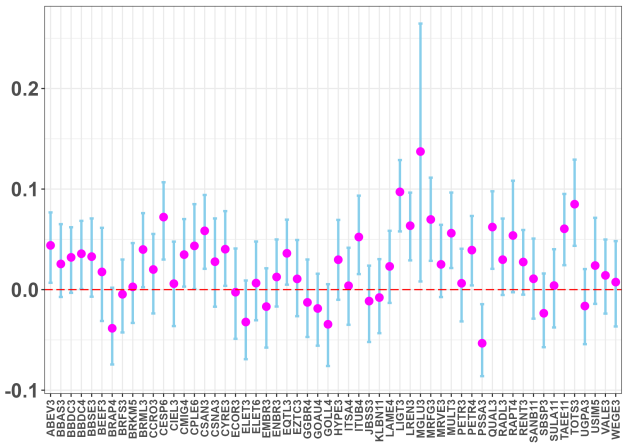
(c) ϕ -0.23 CI[-0.39,-0.07]



(d) factor

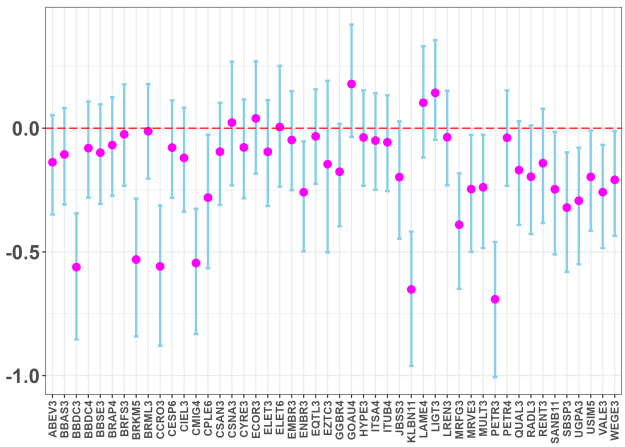


(e) σ_τ 37.9% CI[37.64%,38.17%]

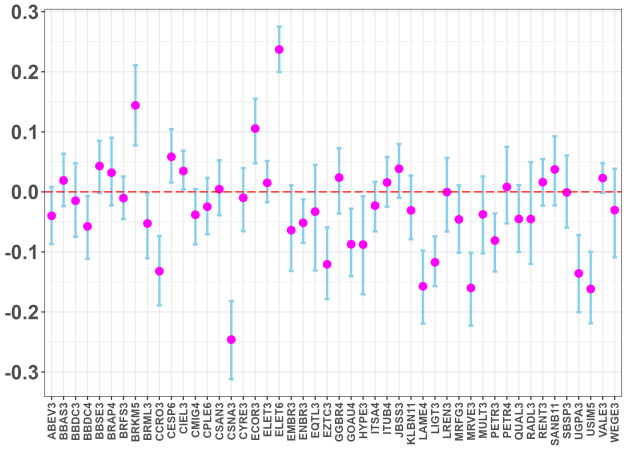


(f) intercept

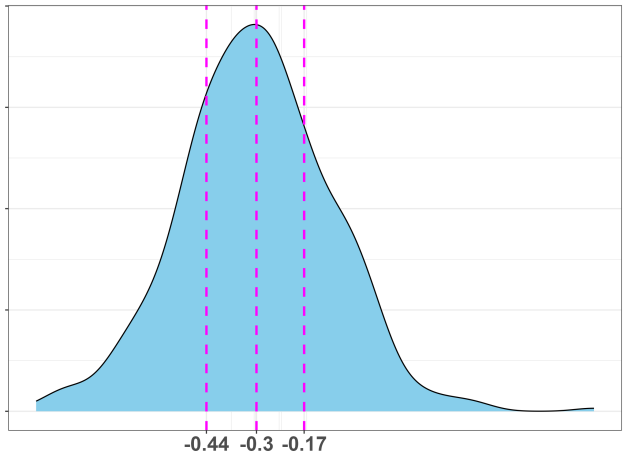
Figure 27: Itau



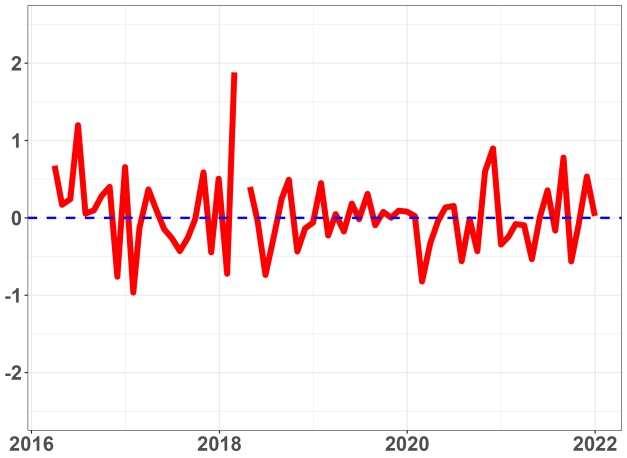
(a) loading



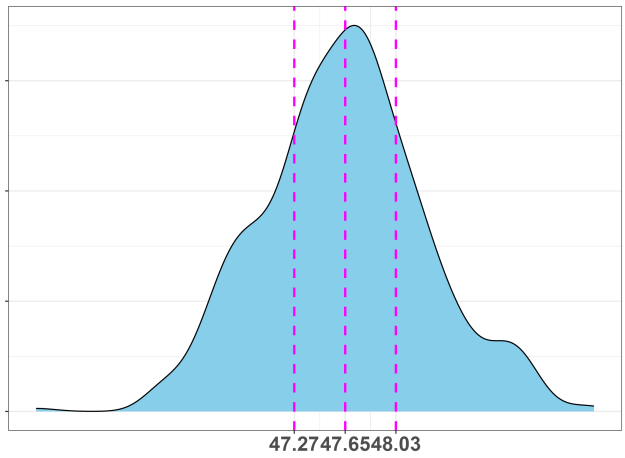
(b) exposure elasticity



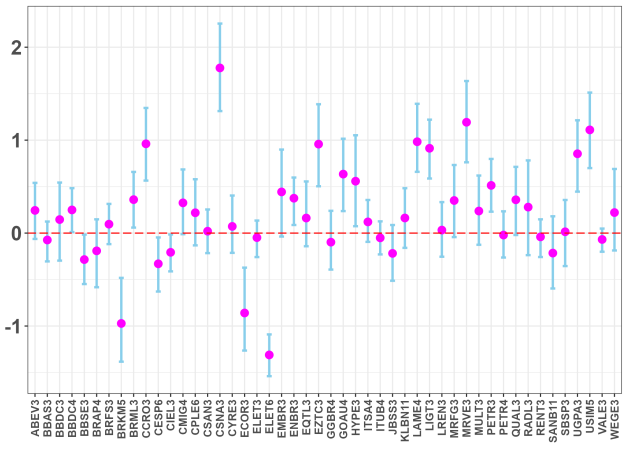
(c) ϕ -0.3 CI[-0.44,-0.17]



(d) factor

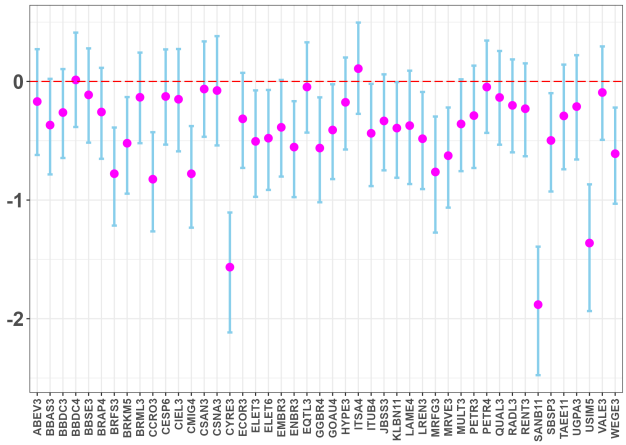


(e) σ_τ 47.65% CI[47.27%,48.03%]

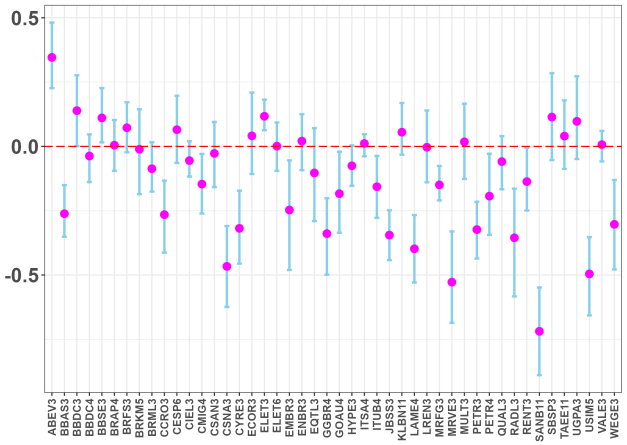


(f) intercept

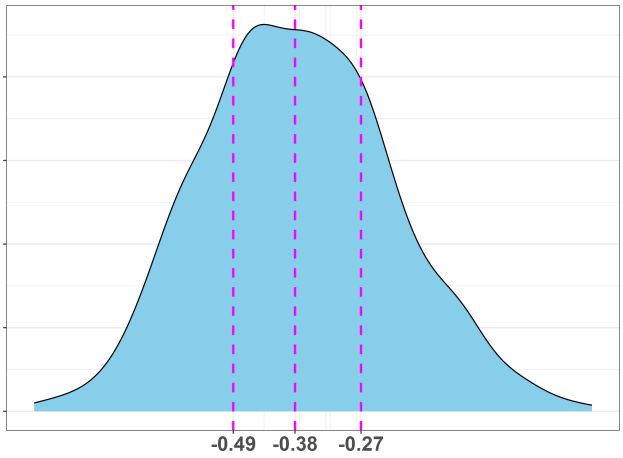
Figure 28: Julius Baer



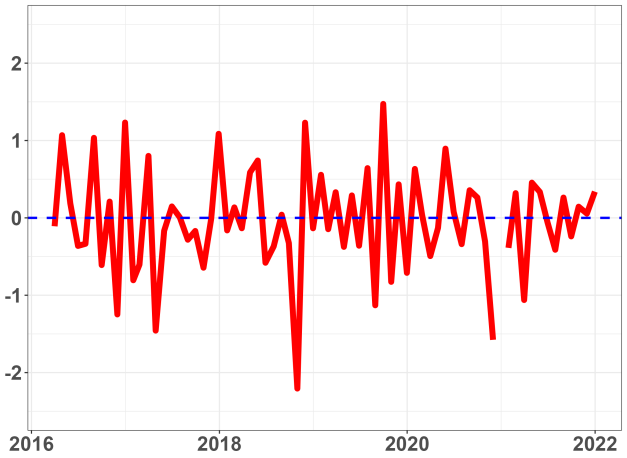
(a) loading



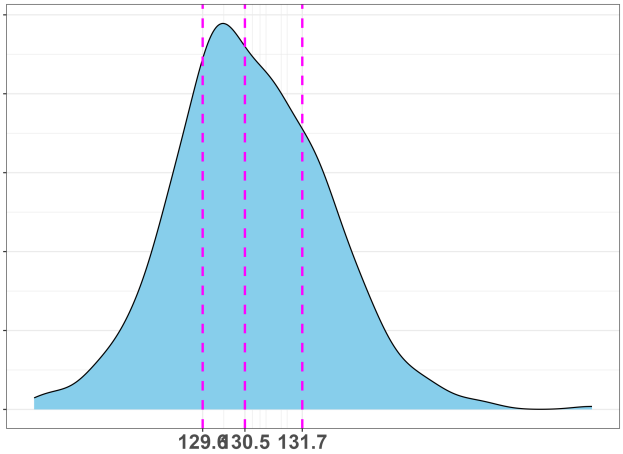
(b) exposure elasticity



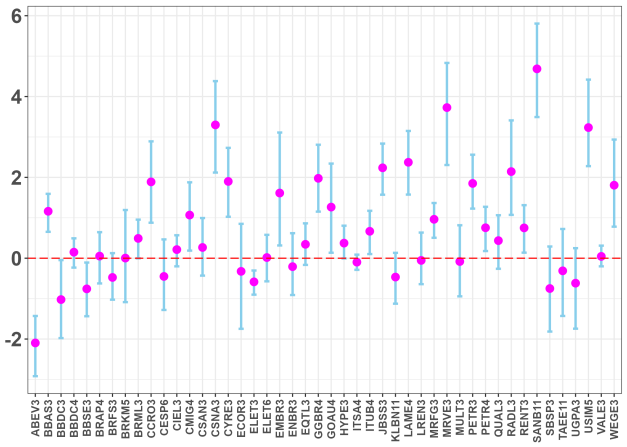
(c) ϕ -0.38 CI[-0.49,-0.27]



(d) factor

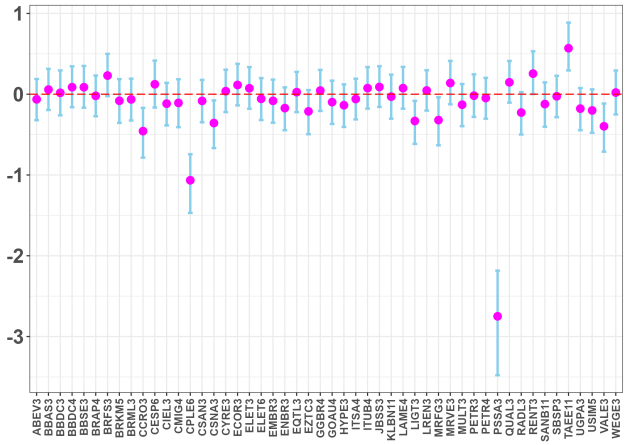


(e) σ_τ 130.5% CI[129.6%,131.7%]

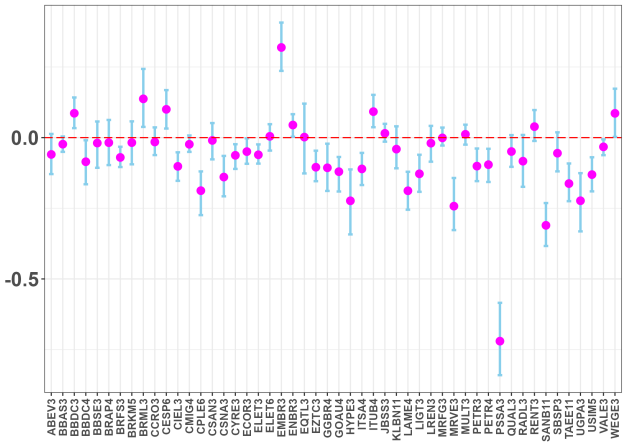


(f) intercept

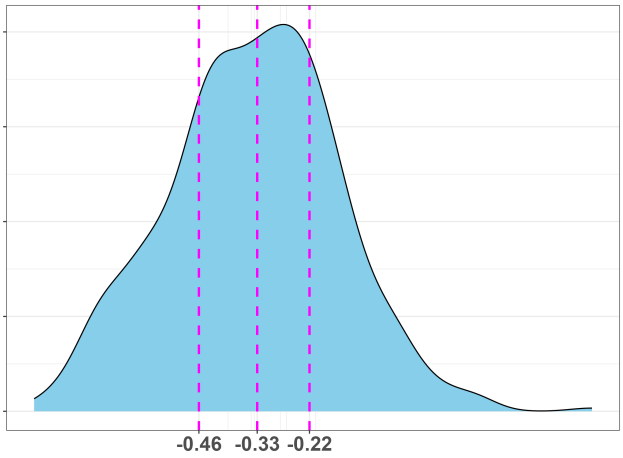
Figure 29: Opportunity



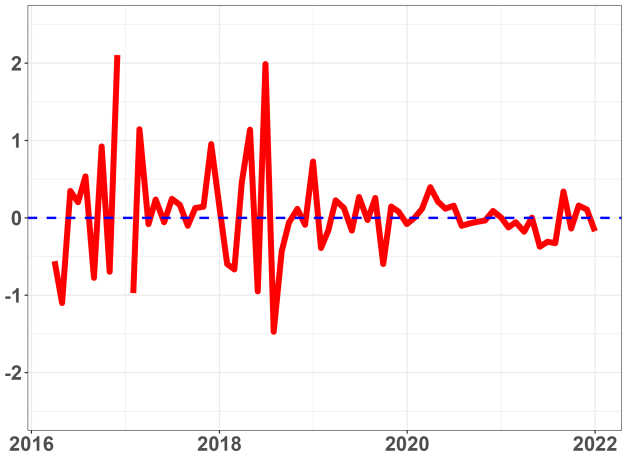
(a) loading



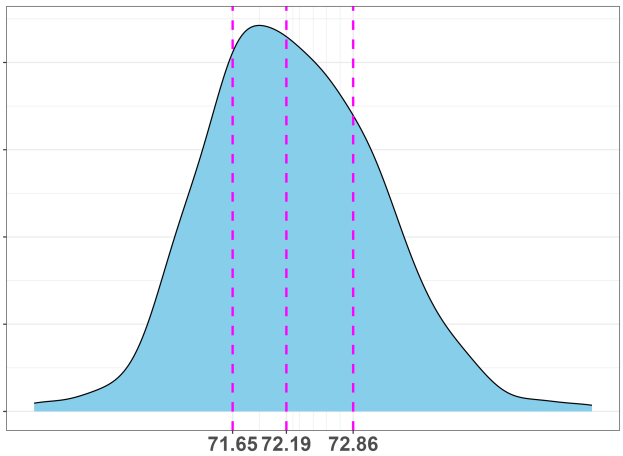
(b) exposure elasticity



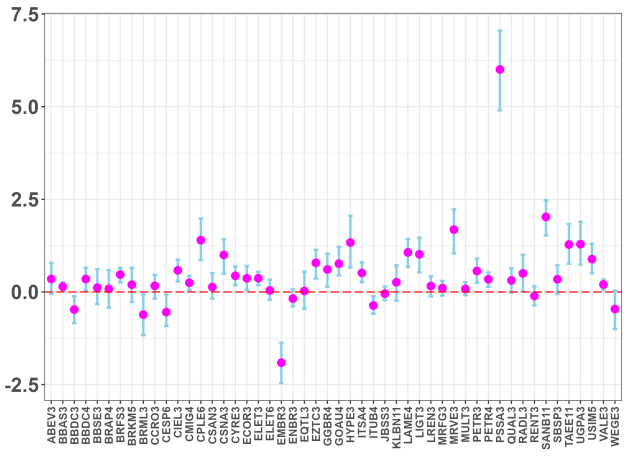
(c) ϕ -0.33 CI[-0.46,-0.22]



(d) factor

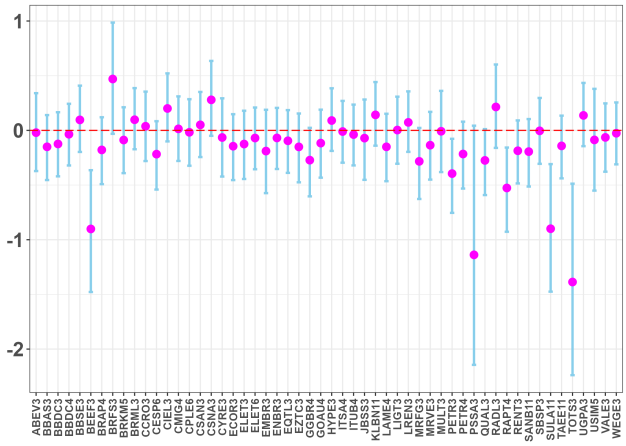


(e) σ_τ 72.19% CI[71.65%,72.86%]

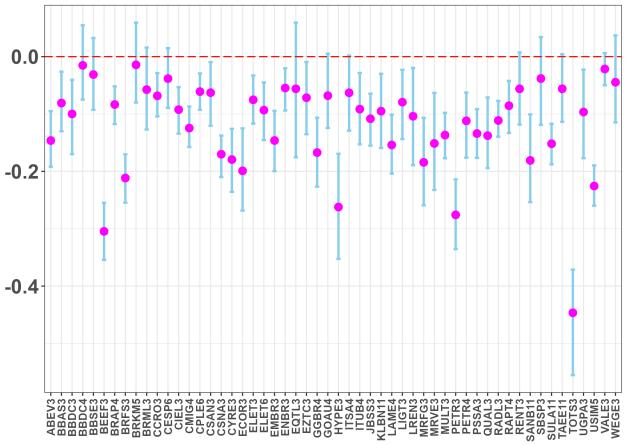


(f) intercept

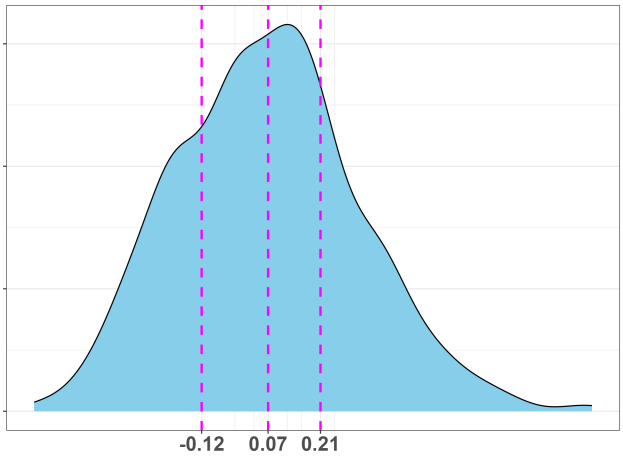
Figure 30: Safra



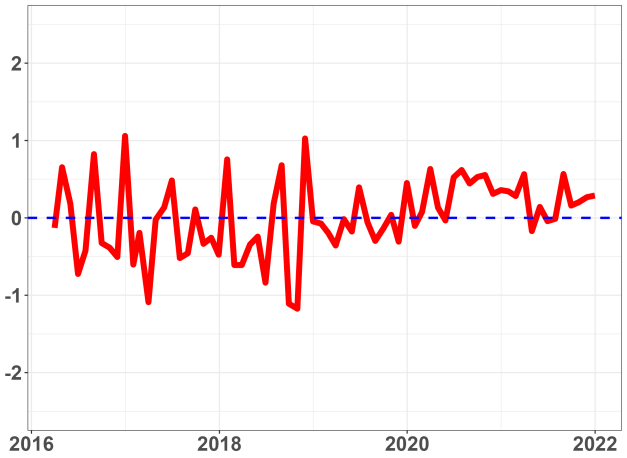
(a) loading



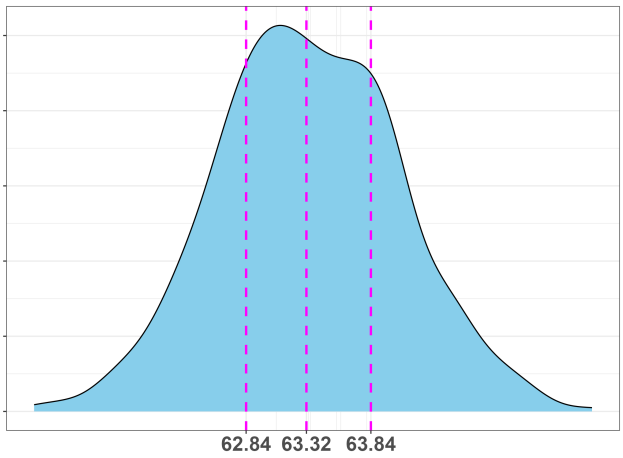
(b) exposure elasticity



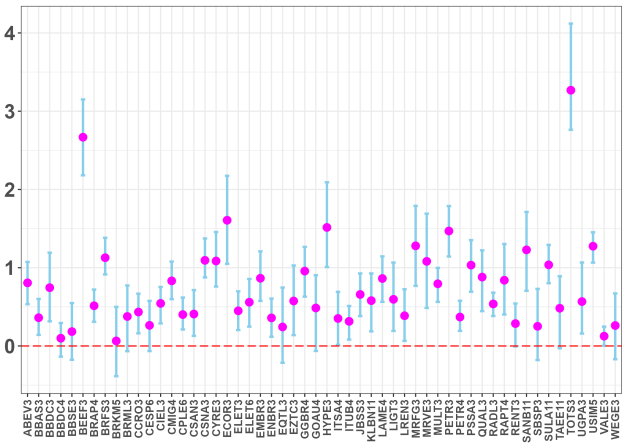
(c) ϕ 0.07 CI[-0.12,0.21]



(d) factor

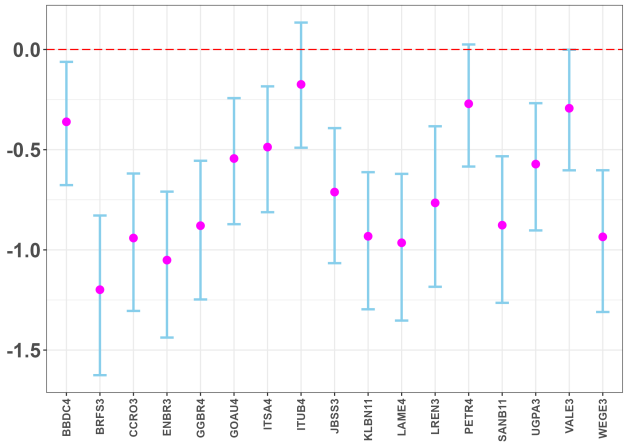


(e) σ_τ 63.32% CI[62.84%,63.84%]

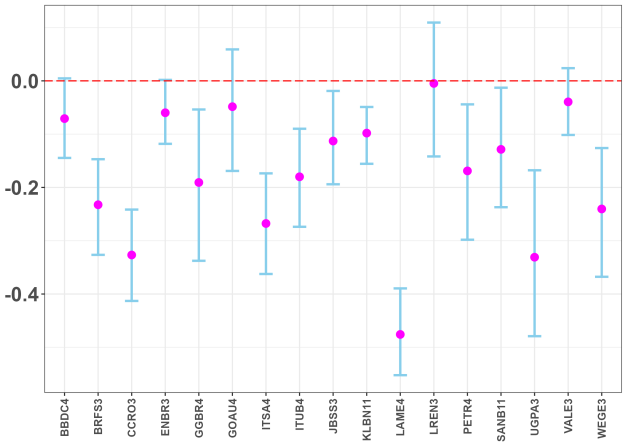


(f) intercept

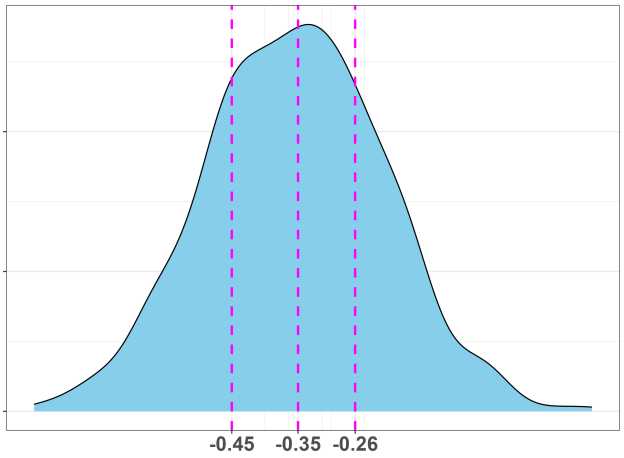
Figure 31: Santander



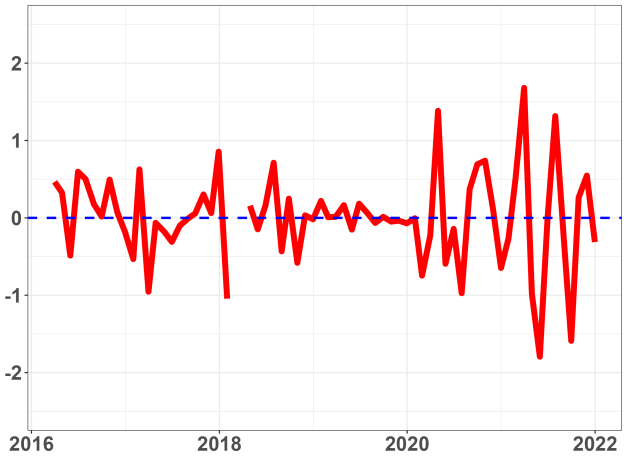
(a) loading



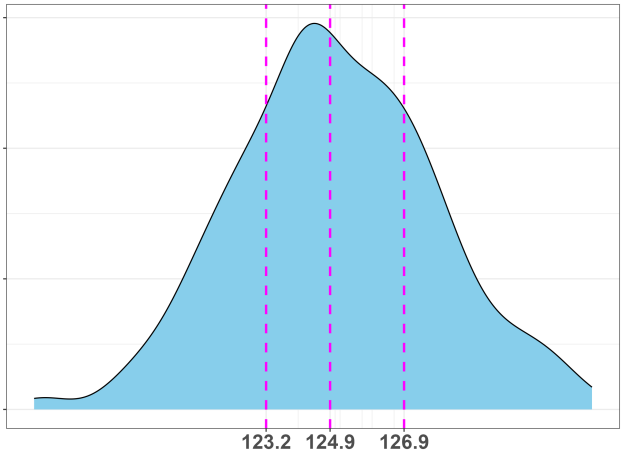
(b) exposure elasticity



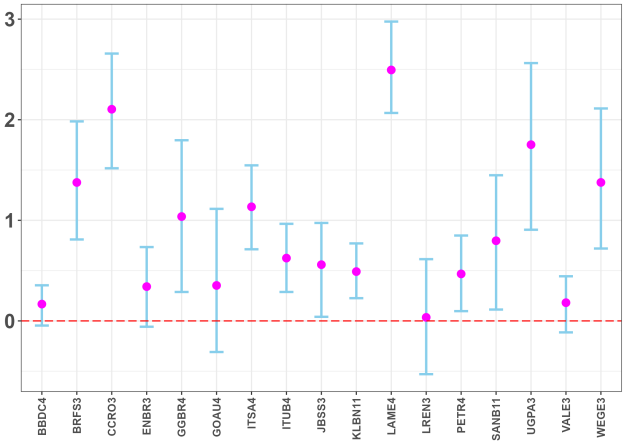
(c) ϕ -0.35 CI[-0.45,-0.26]



(d) factor

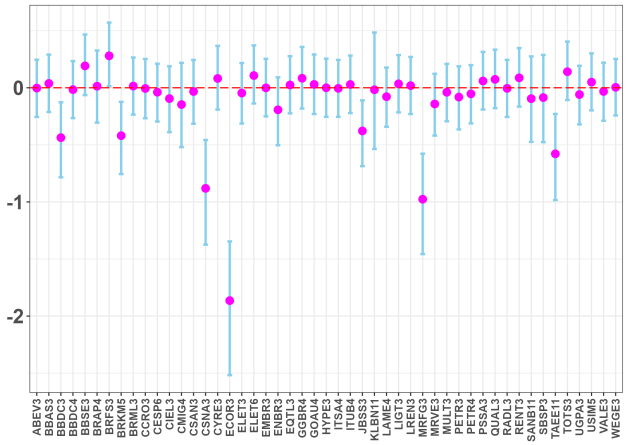


(e) σ_τ 124.9% CI[123.2%,126.9%]

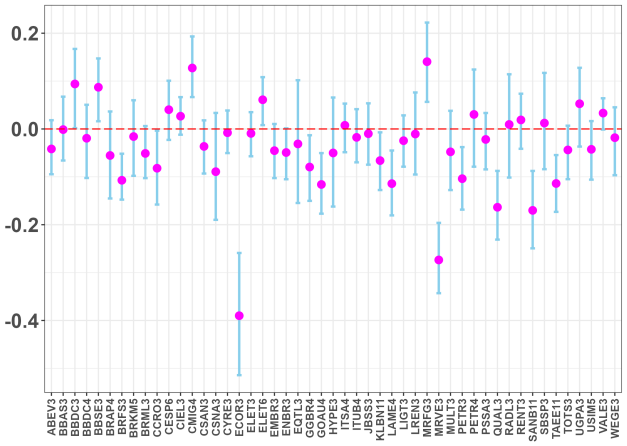


(f) intercept

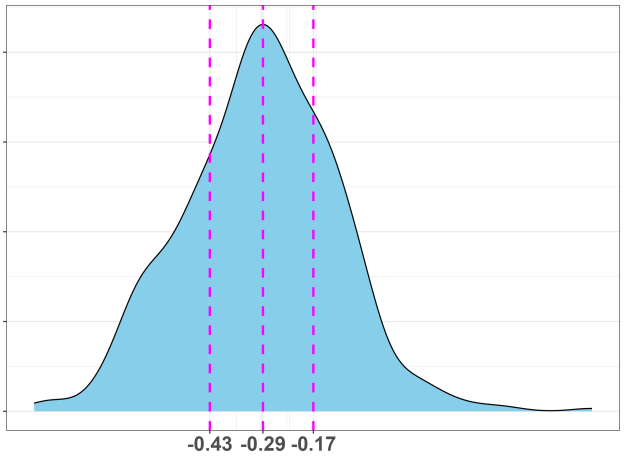
Figure 32: SPX Capital



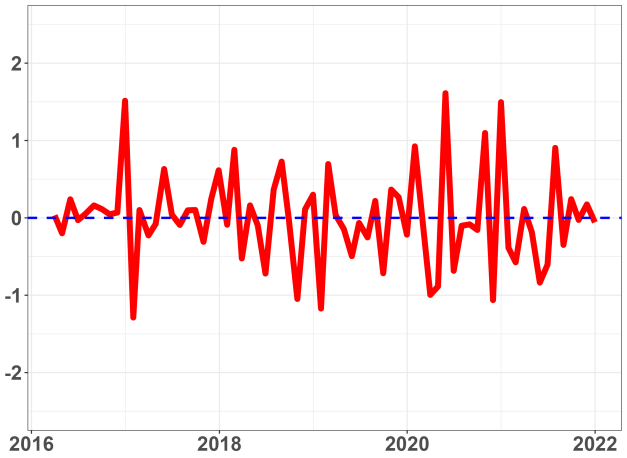
(a) loading



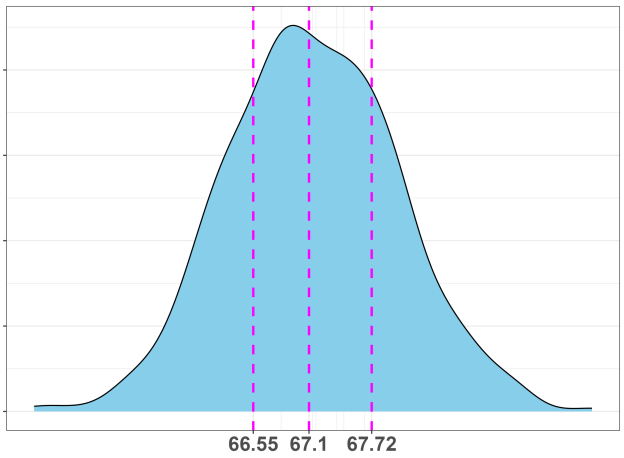
(b) exposure elasticity



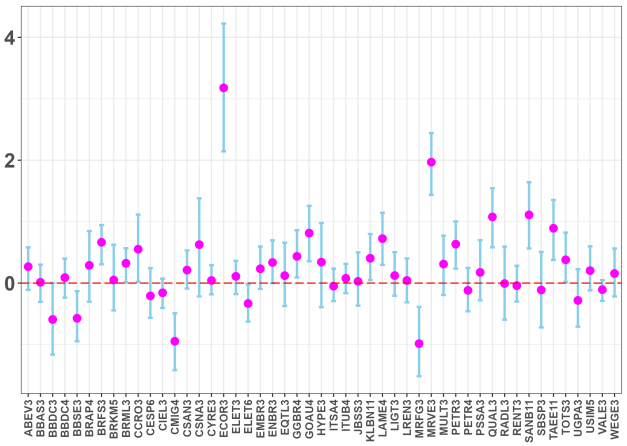
(c) ϕ -0.29 CI[-0.43,-0.17]



(d) factor

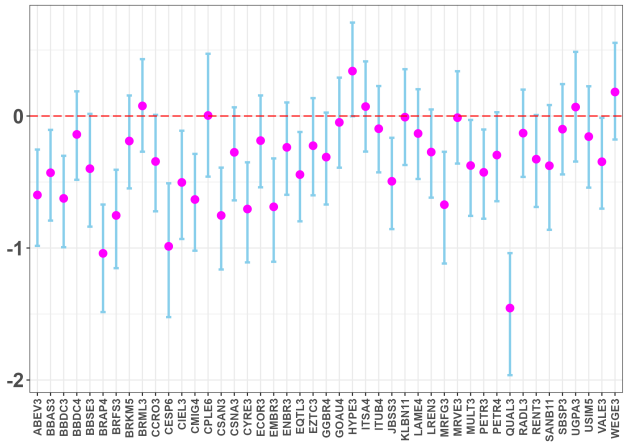


(e) σ_τ 67.10% CI[66.55%,67.72%]

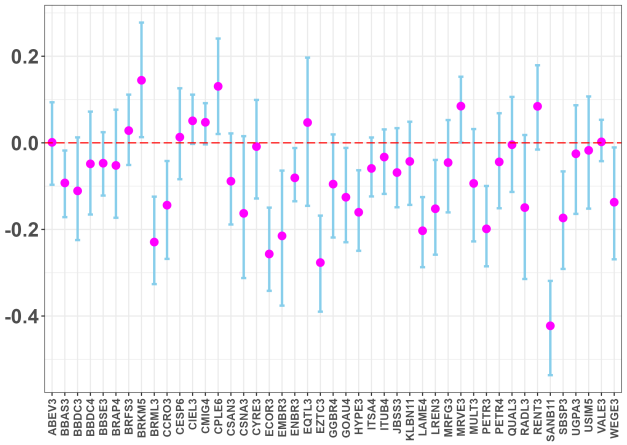


(f) intercept

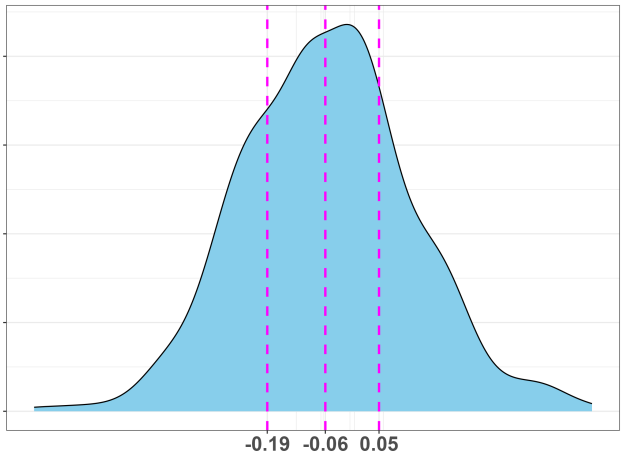
Figure 33: TNA Wealth Management



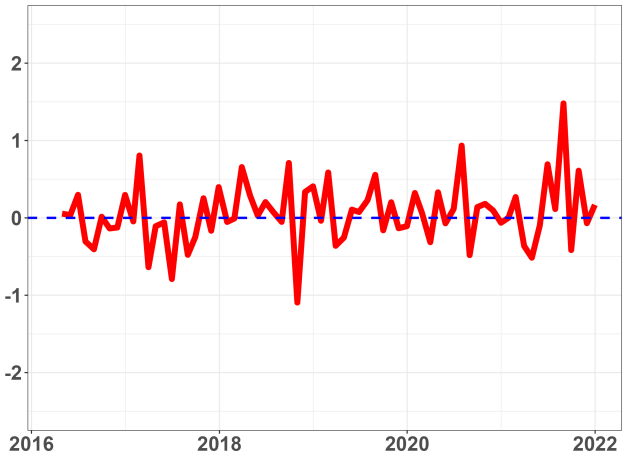
(a) loading



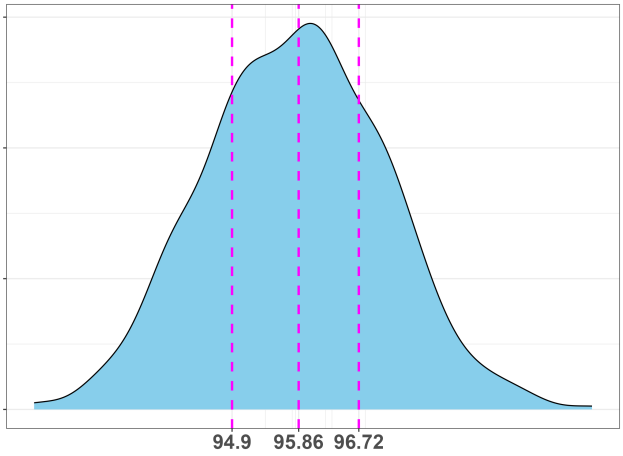
(b) exposure elasticity



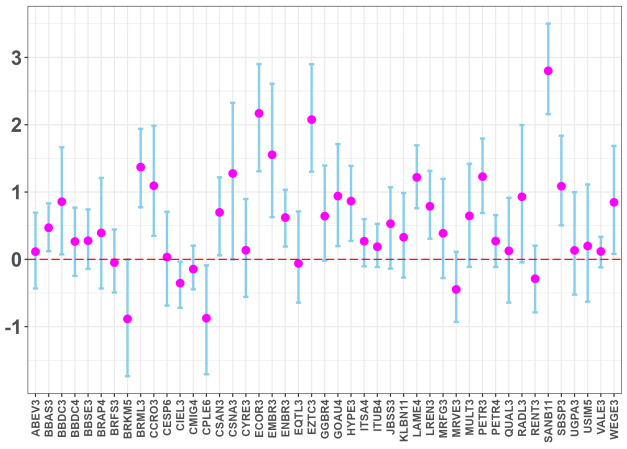
(c) ϕ -0.06 CI[-0.19,-0.05]



(d) factor

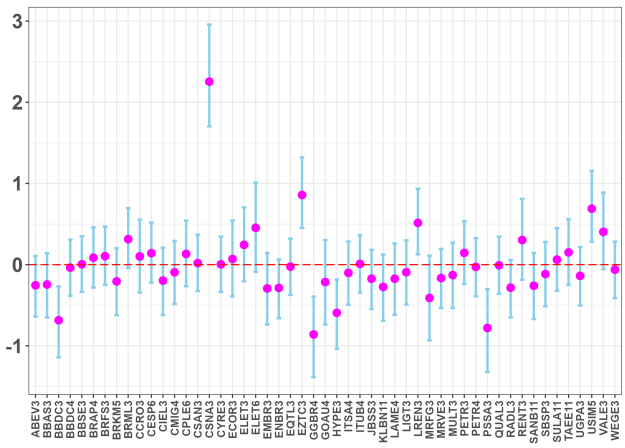


(e) σ_τ 95.86% CI[94.9%,96.72%]

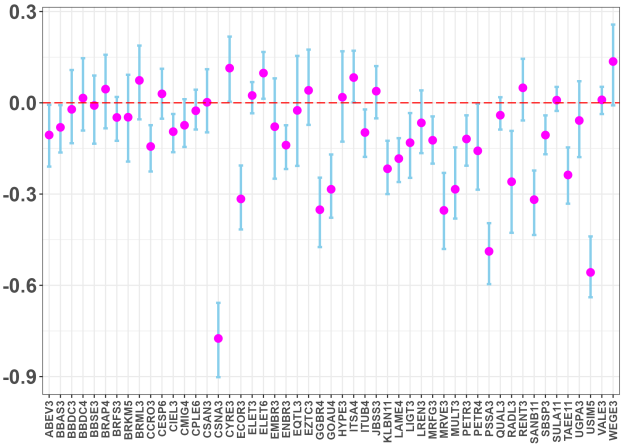


(f) intercept

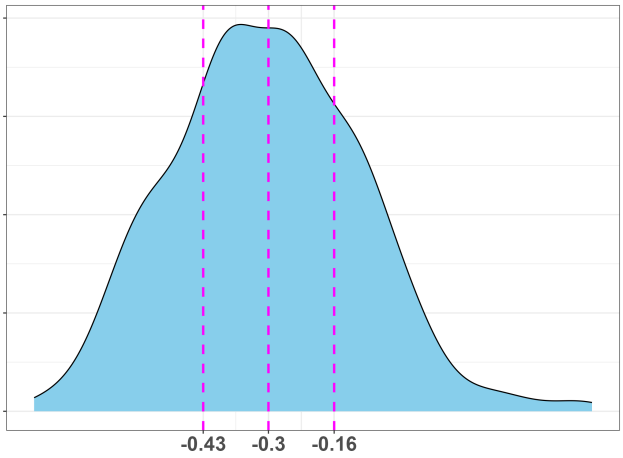
Figure 34: Vinci Partners



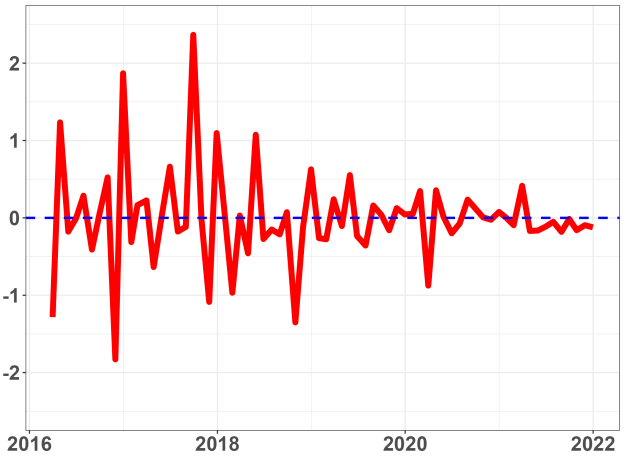
(a) loading



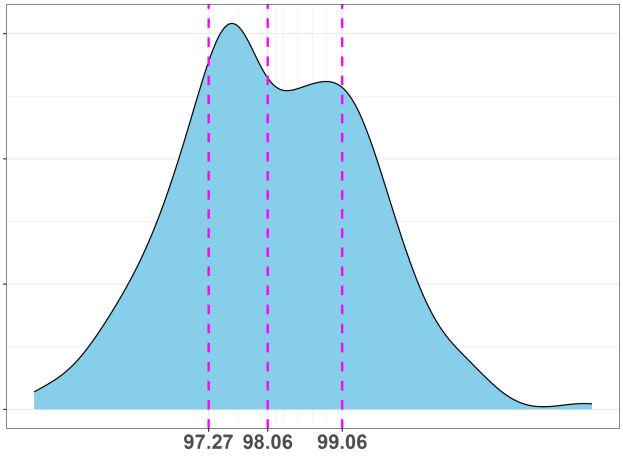
(b) exposure elasticity



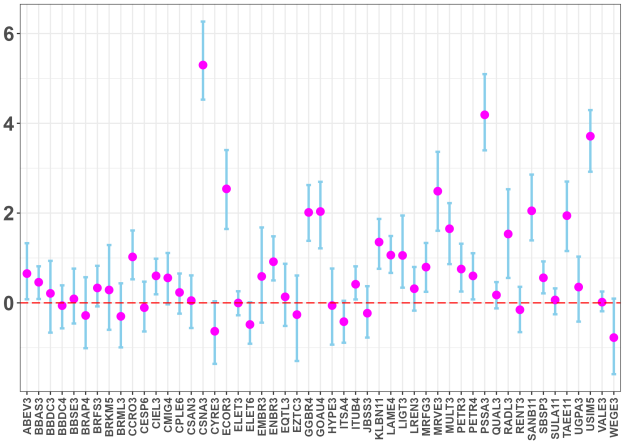
(c) ϕ -0.3 CI[-0.43,-0.16]



(d) factor



(e) σ_τ 98.06% CI[97.27%,99.06%]



(f) intercept

Figure 35: XP