

Generative agents and expectations: do LLMs align with heterogeneous agent models?

Questa è la versione preprint della seguente opera:

Original

Generative agents and expectations: do LLMs align with heterogeneous agent models? / Gusella, F., Vicario, E.. - (2025). [10.48550/arXiv.2511.08604]

Availability:

This version is available at: 20.500.11771/41540

Publisher:

Published

DOI:10.48550/arXiv.2511.08604

Terms of use:

This publication is made accessible in accordance with the terms for deposit in the institutional repository, as defined by the IMT School for Advanced Studies Lucca's Open Access Policy. (https://library.imtlucca.it/sites/default/files/regolamento-policy-open-access-imtlib_0.pdf).

Si prega di consultare le pagine informative dell'editore relative alle politiche di autoarchiviazione.

(Article begins on next page)

Generative Agents and Expectations: Do LLMs Align with Heterogeneous Agent Models?

Filippo Gusella

Università Cattolica del Sacro Cuore, Milano

Università degli Studi di Firenze

New York University Florence

Eugenio Vicario

IMT School for Advanced Studies Lucca

Abstract

Results in the Heterogeneous Agent Model (HAM) literature determine the proportion of fundamentalists and trend followers in the financial market. This proportion varies according to the periods analyzed. In this paper, we use a large language model (LLM) to construct a generative agent (GA) that determines the probability of adopting one of the two strategies based on current information. The probabilities of strategy adoption are compared with those in the HAM literature for the S&P 500 index between 1990 and 2020. Our findings suggest that the resulting artificial intelligence (AI) expectations align with those reported in the HAM literature. At the same time, extending the analysis to artificial market data helps us to filter the decision-making process of the AI agent. In the artificial market, results confirm the heterogeneity in expectations but reveal systematic asymmetry toward the fundamentalist behavior.

Keywords: Heterogeneous Expectations, Large Language Model, Generative Agent, Fundamentalists, Trend Followers

Jel codes: E30, E70, D84

“As economics acquires aspirations to explain behavior under these typical conditions of modern organizational and public life, it will have to devote major energy to building a theory of procedural rationality to complement existing theories of substantive rationality. Some elements of such a theory can be borrowed from the neighboring disciplines of operations research, artificial intelligence, and cognitive psychology”

Herbert A. Simon, *Rationality as Process and as Product of Thought*, 1978

1 Introduction

The role of expectations among economic agents is fundamental in economics. The ways in which agents form their beliefs not only raise fundamental questions about the observed dynamics of both real and financial variables, but also affect the role that monetary and fiscal policy is expected to play.

For several decades, the prevailing approach to analyzing the formation of expectations has been based on the foundational concept of a representative rational agent. On the one hand, the use of the representative agent excluded any consideration of heterogeneity. On the other hand, rational expectations ruled out any endogenous influence on the dynamics of macroeconomic and financial variables (Muth, 1961; Lucas Jr, 1972; Fama, 1991). Parallel to this view, an alternative perspective on the role of expectations emerged, forming the foundation of what is known as the behavioral approach (Simon, 1978). Within this field, models of heterogeneous agents (HAM) deviate from the principle of methodological rational individualism. These models incorporate the concept of heterogeneity in expectations, with different groups of agents forming their beliefs based on heuristics (Day and Huang, 1990; Lux, 1995; Brock and Hommes, 1998).¹

In the early 21st century, various estimation methods were employed to test the hypothesis of heterogeneity in expectations (Lux and Zwickels, 2018; Ellen and Verschoor, 2018). These studies support theoretical models by identifying distinct beliefs among groups of speculators (Lux, 2018; Majewski et al., 2020; Schmitt, 2021; Ter Ellen et al., 2021; Gusella and Ricchiuti, 2024; Di Francesco and Hommes, 2025). In addition, fluctuations in asset prices can derive from a switching mechanism that moves agents among different expectation-based rules; the presence and time-varying adoption of these rules lead to the endogenous emergence of boom and bust periods (Schmitt and Westerhoff, 2021; Gatti et al., 2025). In particular, while periods before the dot-com and global financial crises are associated with a high percentage of agents adopting a trend-following strategy, the subsequent post-crisis phases are instead characterized by a rising share of fundamentalists, which drives the dynamics toward a convergence of prices to their fundamental values (Hommes et al., 2017).

In recent years, a rapidly expanding body of literature has examined the interactions between humans and artificial intelligence (AI), exploring how AI can be leveraged across diverse domains — from education and healthcare to politics and economics (Tsvetkova et al., 2024; Motoki et al., 2025; Di Cagno and Lin, 2025). This line of research highlights the transformative potential of AI systems as tools capable of augmenting human decision-making, generating knowledge, and shaping social and economic outcomes. On the economic side, the rapid diffusion of large language models has opened new avenues for research on expectation formation, decision-making, and forecasting, as these models can act as artificial economic agents that generate predictions, narratives, and trading strategies. Consequently, a growing number of studies

¹For a comprehensive overview of this literature, see the seminal survey by Hommes (2006), as well as the more recent review by Dieci and He (2018) and Hommes (2021), which incorporate recent advances in heterogeneous agent modeling.

have investigated whether LLMs can serve as substitutes or complements to human forecasters, evaluating their performance against professional forecasts, survey data, or controlled laboratory experiments (Bybee, 2023; Lopez-Lira and Tang, 2023; Faria-e Castro and Leibovici, 2024; Hansen et al., 2024; Zarifhonarvar, 2024; del Rio-Chanona et al., 2025; Lopez-Lira, 2025). The present study aligns with this strand of research and advances the analysis of belief formation by introducing a generative agent driven by a Large Language Model (LLM), which serves as an active decision maker in a simulated financial environment. However, while previous research has provided valuable benchmarks on the accuracy and consistency of LLM-based forecasts, much less attention has been devoted to the behavioral mechanisms underlying the strategies adopted by LLM agents. Given the ability of LLMs to mimic human behavior and replicate cognitive biases (Binz and Schulz, 2023; Strachan et al., 2024), it becomes crucial to analyze their decision processes, particularly as they play an increasingly central role in economic and financial decision-making (Gao et al., 2024; Korinek, 2023; Horton, 2023). Their ability to replicate or diverge from the behavior of past economic agents may either reinforce existing market dynamics or counteract them. Indeed, LLMs may default to the most statistically dominant interpretation, often aligning with the prevailing mainstream or dominant view (Shen et al., 2024; Qu and Wang, 2024). This paper contributes to this growing literature by examining whether generative agent price strategies align with those documented in the heterogeneous agent modeling literature, thereby providing new insights into LLM-driven market behavior.

We proceed as follows. In a simulated financial market environment, the generative LLM agent can select between two market strategies: one assumes convergence between future prices and perceived fundamental values. This is the expectation of *fundamentalists*. The other, associated with the behavior of *trend followers*, follows a trend-based approach, where speculation is based on past price movements.² Choosing between these strategies, we examine how the LLM agent forms beliefs when forecasting S&P 500 Index over the period January 1990 to December 2020 and compare its decision with traditional econometric results in the literature on the heterogeneous agent model. In doing this, we offer opportunities to test theories and hypotheses on expectation formation, helping to determine whether LLM-driven agents align with or diverge from established HAM results, which may therefore have been inherited in ChatGPT training. At the same time, based on artificial market conditions, we aim to deepen the analysis by examining whether it is possible to identify the behavioral rules that govern the adoption of specific expectations based on available information. This helps us to understand the specific mechanisms that underlie the AI agent’s decision-making process. In other words, the analysis allowed us to filter the behavioral rules that have been internalized by the generative agent from heterogeneous agent models.

Results based on time series analysis indicate that AI-generated expectations align with findings from the HAM literature. In particular, we show that LLM-based expectations mirror key deviations from the fundamentalist expectation during major events such as the dot-com bubble and the global financial crisis. Furthermore, when the expectations of the generative agents are explicitly modeled, the values of the reaction parameters are found to be in line with those predicted in the HAMs setting. At the same time, results based on artificial data confirm the heterogeneity in expectations, but with an asymmetry toward fundamentalist behavior.

The remainder of the paper is organized as follows. Section 2 introduces the beliefs formalized in the HAM literature, reviews related empirical findings, and outlines the AI setting. Section 3 presents the simulation results. Section 4 concludes.

²The focus on these two types of expectations is motivated by the fact that most empirical studies on heterogeneous agent models adopt the classification of trend followers and fundamentalists, as established in the theoretical literature.

2 Methodology

2.1 Theoretical setting and empirical results

In the literature on heterogeneous agent models, developed through early contributions by Day and Huang (1990); Chiarella (1992); Lux (1995); Brock and Hommes (1998), the asset price dynamics is shaped by the coexistence of groups of agents with different expectations. The most common specification distinguishes between two of them: fundamentalist (F) and trend follower (TF) expectations.³ Fundamentalists rely on the efficient market hypothesis, assuming that the price (P) will eventually converge to its fundamental value (P_t^F). This expectation can be formalized as follows:

$$E^F [P_{t+1}] = P_t + \alpha_t (P_t^F - P_t), \quad 0 \leq \alpha \leq 1, \quad (1)$$

where α , the reaction parameter, captures the intensity with which the market price is driven toward the fundamental value. This implies that during periods of pronounced asset price expansions or contractions, fundamentalist agents anticipate a correction, expecting the price to realign with the underlying fundamental.

In contrast, trend followers base their forecasts on previously observed price patterns. They extrapolate recent trends and assume that increases (or decreases) in past prices are likely to persist in the near future:

$$E^{TF} [P_{t+1}] = P_t + \beta_t (P_t - P_{t-1}), \quad \beta > 0. \quad (2)$$

The strength of this behavior is governed by the reaction parameter β , which captures how aggressively trend followers respond to past price movements.

Aggregating these two behavioral rules, the market price at time $t + 1$ can be expressed as a weighted average of the expectations of the two groups:

$$P_{t+1} = \delta_t E^F (P_{t+1}) + (1 + \delta_t) E^{TF} (P_{t+1}), \quad (3)$$

where δ_t and $1 - \delta_t$ are the percentage of fundamentalists and trend followers in the market, respectively. Combining Eqs. 1, 2, and 3 yield the following reduced-form expression for price dynamics:

$$P_{t+1} = P_t + \delta_t \alpha_t (P_t^F - P_t) + (1 + \delta_t) \beta_t (P_t - P_{t-1}). \quad (4)$$

The coexistence of heterogeneous behavior, captured by differing reaction parameters and a changing distribution of agents over time, gives rise to nonlinear market dynamics that produce market booms and crashes.

Building on this theoretical framework, over the last three decades, a growing body of empirical work has investigated whether the HA models are consistent with observed asset price dynamics.⁴ These contributions generally adopt the two forecasting strategies previously defined to keep the model tractable and to prevent over-parameterization. Essentially, applying equation 4 to the data reveals two main findings: first, investors hold heterogeneous beliefs rather than a single representative expectation. Empirical findings generally indicate that fundamentalist

³Although related to the broader Agent-Based Modeling (ABM) literature (Dawid and Gatti, 2018; Dosi and Roventini, 2019; Delli Gatti et al., 2018; Monti et al., 2023), HAMs do not consider numerous interacting agents, but rely on a smaller set of behavioral rules adopted by groups of agents. This formalization allows analytical insights into how different expectations shape macro-finance outcomes.

⁴The empirical performance of HAMs in equity markets has stimulated a wider range of applications across different markets, including commodities (Reitz and Westerhoff, 2007), foreign exchange (Goldbaum and Zwinkels, 2014), housing market Bolt et al. (2014), and macro model (Cornea-Madeira et al., 2019; Kukacka and Sacht, 2023; Lux, 2024).

behavior stabilizes asset prices, whereas chartist behavior destabilizes them, causing prices to deviate from the fundamental value. This is reflected in significant, theoretically consistent estimates for both types of reaction parameters. In particular, the reaction coefficient for fundamentalists falls within the expected range of 0 to 1 (Franke and Westerhoff, 2011; Ter Ellen et al., 2021; Kukacka and Barunik, 2017). The exact value varies depending on the specific asset considered. Conversely, the reaction parameter for trend followers is typically positive (Chiarella et al., 2014; Recchioni et al., 2015; Gusella and Ricchiuti, 2024). However, when this parameter is not subject to certain restrictions, it can also have negative values, allowing contrarian behavior (Chiarella et al., 2012). Second, agents frequently switch between fundamentalist and trend-follower rules. The prevailing behavior depends on the market environment and the period under analysis. Studies focusing on the late 1990s and early 2000s show that episodes such as the dot-com bubble and the global financial crisis were preceded by a surge in trend-following behavior, which contributed to the formation of asset price bubbles. However, after such crises, the price strategy tended to revert toward fundamentalist expectations, leading to more stable price adjustments (Boswijk et al., 2007; Chiarella et al., 2014; Hommes et al., 2017; Gatti et al., 2025).

2.2 *AI setting*

As previously outlined, this paper aims to evaluate whether, and to what extent, the results and assumptions found in the heterogeneous agent modeling literature have been internalized within the emergent knowledge of major commercial large language models. The primary focus of our analysis is the composition of the market: specifically, whether the LLM chooses a relative weighting between fundamentalists and trend followers that aligns with established findings in the literature. In addition, we extract the expectations conditional on the strategy employed. From these expectations, we derive the reaction parameters of the equations that, in the HAM literature, characterize the behavior of fundamentalists and trend followers. Combining the market composition with these reaction parameters enables us to compute the expected price for the following period, under the hypothetical scenario of a market populated exclusively by generative agents relying on an LLM to make their trading decisions.

The simulation analysis employs the ChatGPT-4o-mini model, accessed via an API key through a dedicated Python script. Interaction with the model is achieved through the use of prompting techniques. Two distinct prompts are used to maintain analytical clarity and isolate the mechanisms under investigation. The first prompt is designed to elicit the market composition, that is, the relative shares of fundamentalists and trend followers within the market. Specifically, we input the following prompt:

Prompt 1 — Market Composition: *You are a financial trader and you need to choose between two alternative market strategies according to your analysis based on expectation formation: the first is a fundamentalist approach, and the second is a trend-following approach. You have access to the current value of the fundamental, the current value of the asset price, and the value of the asset price at the previous period. In conclusion, for each strategy, you only have to provide a weight that represents the probability with which you would choose the strategy. The sum of the weights must be 1. Do not provide any other comment, just the two weights divided by a comma.*

Following the prompt, three values are provided: the current price, the price of the previous period, and the current fundamental value. The price is the value of S&P 500 index at monthly frequency from 1990 to 2020, while the fundamental price is calculated through the Gordon growth model as in Chiarella et al. (2012) using S&P 500 data on dividends. Each interaction includes only these values, with no memory of the previous input. This prevents identifying the

overall market context, extrapolating the market phase, or linking responses to prior literature on market structure.

The second prompt is used to obtain the model’s expectations of the future price, conditional on the two distinct market strategies. As before, the prompt is followed by the two price values and the value of the fundamental:

Prompt 2 — Price Expectations: *You are a financial trader and you need to provide the value of price expectations for the next period using two alternative market strategies according to your analysis: the first is a fundamentalist approach, and the second is a trend-following approach. You have access to the current value of the fundamental, the current value of the asset price, and the value of the asset price at the previous period. In conclusion, you only have to provide the value of price expectations for each strategy. Do not provide any other comment or word, just the two numbers divided by a comma.*

In the existing literature, the composition of the market and the reaction parameters are typically estimated simultaneously (Lux and Zwinkels, 2018; Ellen and Verschoor, 2018). In contrast, this study deliberately employs separate prompts, ensuring that the two dimensions of the analysis remain independent. This methodological choice is motivated by our primary interest in market composition, which represents a more straightforward and interpretable concept from which implicit heuristics and behavioral rules can be more easily inferred. Conversely, deriving reaction parameters from expectations involves additional analytical complexity, as it requires the specification of functional forms for the reaction equations, introducing further uncertainty into the interpretation of results. By adopting separate prompts, we ensure that the outcomes concerning market composition are not confounded by ancillary information provided to or requested from the LLM. For completeness, in the Appendix A.1 we also report the results obtained using a single integrated prompt, which simultaneously generates both market composition and expectations. Although minor quantitative differences arise, the main qualitative findings remain robust.

To obtain the results, we employed the ChatGPT *gpt-4o-mini* model. Robustness checks using an alternative model are reported in the Appendix B.1. The temperature parameter was set to 1, corresponding to the default configuration. In models such as ChatGPT, the temperature parameter regulates the degree of randomness in text generation by influencing the likelihood of selecting less probable tokens. At low values, close to zero, the model behaves deterministically, producing stable, predictable, and coherent outputs, though at the expense of creativity and diversity. At higher values, typically between 0.8 and 1.2, the model explores a broader range of possible continuations, increasing variability and originality in the responses, but also the likelihood of inconsistencies or errors. In the Appendix B.2, we test the robustness of our results with respect to different temperature settings.

The results were obtained through 50 independent replications of the prompt for each value triplet and the average of the generated responses was used as the final result. In the Appendix A.3, we show that 50 replications are sufficient to stabilize the results. Finally, consistent with the discussion of the integrated prompt presented earlier, in the appendix A.1 we also report robustness analyzes that assess the sensitivity of the results to variations in the prompt formulation.

3 Results

In this section, we present the results of the simulations, which are divided into two parts. In subsection 3.1, we apply the proposed procedure to the time series of the S&P 500 Index and the corresponding Gordon fundamental value over the period 1990 to 2020. In this way, we

identify the market composition across different time periods (Prompt 1). We then extract the reaction parameters of the agents using the expected price for the following period (Prompt 2), which allows us to analyze the forecasting error of the generative agent. Through the use of Prompt 1, we obtain a qualitative insight into the relative presence of fundamentalists and trend followers, whereas Prompt 2 provides a quantitative result. In subsection 3.2, we investigate the implicit heuristics and behavioral rules employed by the LLM. To this end, we generate artificial market conditions and examine how the model responses vary as these conditions change. As before, we begin by simulating the market composition (Prompt 1) and then, using the agents’ expectations (Prompt 2), we derive the reaction parameters and compute the expected price variation for the subsequent period.

3.1 Standard & Poor 500 index

Figure 1 plots the dynamics of the S&P 500 Index (blue line) together with the corresponding Gordon Fundamental Value (orange line). Together with these two dynamics, the green dots represent the probability with which the generative agent selects the fundamentalist strategy when forecasting the next period’s price, based on the information provided through the data prompt. These probabilities are computed as the average result of 50 Monte Carlo simulations. The red dashed line reports the 13-month centered moving average of this probability measure, calculated over a rolling window.

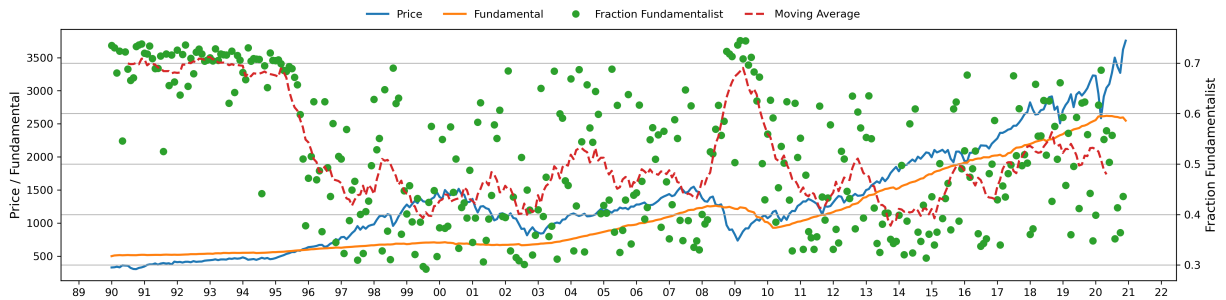


Figure 1: Time series of the S&P 500 index and the Gordon fundamental value, together with the percentage of adoption of the fundamentalist expectation.

Several interesting dynamics emerge over the sample period. Focusing on the evolution of expectation formation, the figure shows that during the first half of the 1990s, the probability of observing a generative fundamentalist behavior is relatively high, stabilizing around a value of 0.7, which is consistent with prices moving closely toward the fundamental value. However, in the second half of the same decade, there is a marked increase in the likelihood of adopting a generative trend-following expectation. This trend-following dominance persists through the late 1990s, with peaks just before the dot-com crisis, suggesting that boom dynamics were fueled by agents extrapolating from rising prices. Furthermore, this same behavior appears to dominate in the immediate aftermath, indicating that the bust phase was largely driven by bearish trend-following expectations. A similar pattern emerges in the run-up to the global financial crisis, when generative trend-following expectations again prevail. In particular, although the probability of fundamentalist behavior occasionally rises thereafter, generative trend-following strategies remain relatively strong in the years preceding the global financial crisis. In contrast, the 2007-2008 collapse is followed by a more persistent reversion to fundamentalist behavior, restoring convergence between prices and fundamentals.

Overall, the strategies adopted by the LLM agent mirror patterns often reported in the empirical literature on heterogeneous agent models. In particular, the evidence suggests that boom phases are fueled by trend-following behavior that drives prices away from fundamentals,

as seen before the dot-com bubble and the global financial crisis. Conversely, fundamentalist behavior tends to dominate after the collapse of speculative booms, restoring the link between prices and underlying values (Boswijk et al., 2007; Chiarella et al., 2014; Hommes et al., 2017; Gatti et al., 2025). It is worth noting that this last mechanism is more clearly reflected in the aftermath of the global financial crisis than in the post-dot-com period, highlighting asymmetries across different market events.

As a further analysis, we apply the Prompt 2, related to price expectations, to the historical series of the S&P 500 Index and its corresponding fundamental value. From the expectations obtained through 50 Monte Carlo simulations, we computed the reaction parameters, as previously defined in Equations 1 and 2, such that:

$$\alpha_t = \frac{E^F[P_{t+1}] - P_t}{P_t^f - P_t} \quad (5)$$

and

$$\beta_t = \frac{E^{TF}[P_{t+1}] - P_t}{P_t - P_{t-1}} \quad (6)$$

The estimated reaction parameters are presented in Figure 2. The values obtained fall within the ranges estimated in the heterogeneous agent modeling literature (Franke and Westerhoff, 2011; Ter Ellen et al., 2021; Kukacka and Barunik, 2017). In particular, the reaction parameter for fundamentalists lies between zero and one, with values very close to one, thereby formalizing the behavior typically associated with so-called dogmatic fundamentalists. In contrast, the estimated parameters for chartists take both positive and negative values, thus allowing for the possibility of contrarian behavior (Chiarella et al., 2012). Subsequently, using the relative weight of fundamentalists together with the reaction parameters, we calculate the expected price for the following period, under the assumption that the market is populated by generative agents behaving according to the patterns observed in the simulations. By subtracting the realized price from the expected price, we derived the forecast error of the generative agents, as shown in Figure 2.

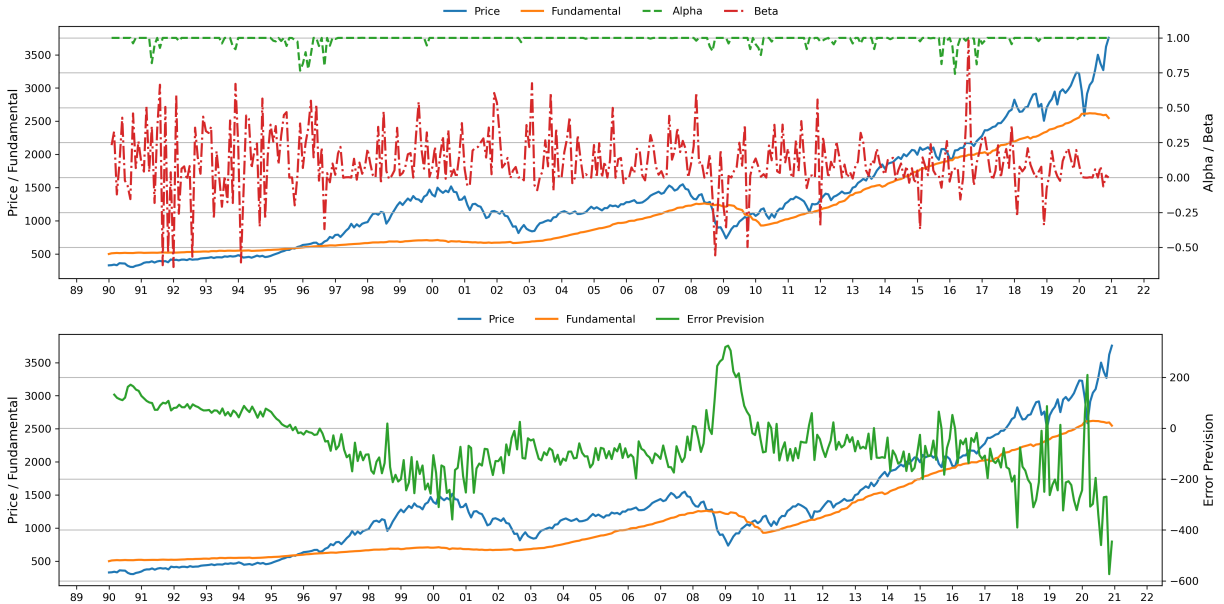


Figure 2: Top: alpha and beta; Bottom: Error prevision.

We observe that the sign of the forecasting errors systematically depends on the sign of the deviation between the fundamental value and the current price: when the market price lies

above the fundamental, the forecasting error is negative, whereas it becomes positive when the fundamental value exceeds the market price. This pattern is driven by the consistent presence of fundamentalist agents characterized by reaction parameters close to one, that is, strict fundamentalists. These agents systematically push the expected price toward the fundamental, thereby generating a systematic bias in the forecast.

3.2 Artificial market conditions

Building on the previous results, we deepen the analysis by investigating whether it is possible to identify the rules governing the adoption of specific expectations based on the information available to the generative agent. This approach allows for a closer examination of the mechanisms underlying the generative expectation formation, with particular attention to the conditions that trigger the switching between fundamentalist and trend-following behavior. To uncover the implicit switching rules and heuristics driving these outcomes, we test the AI’s behavior and corresponding price strategies by varying the fundamental value and the previous price while keeping the current price constant. Specifically, we evaluate all combinations in which the distance between the fundamental value and the current price ranges from -600 to 600 , with increments of 120 , while the difference between the current and previous prices ranges from -100 to 100 , with increments of 20 . These ranges were selected to correspond to the most frequently observed variations relative to the average price level of the S&P 500 Index, thereby ensuring that the analysis captures empirically relevant dynamics rather than extreme or atypical cases (Figure 3). For these analyses, we set the current price P_t equal to 1333 , corresponding to the average value of the S&P 500 Index over the sample period considered in this study. In the Appendix A.2, we test the robustness of the results by varying the actual price.

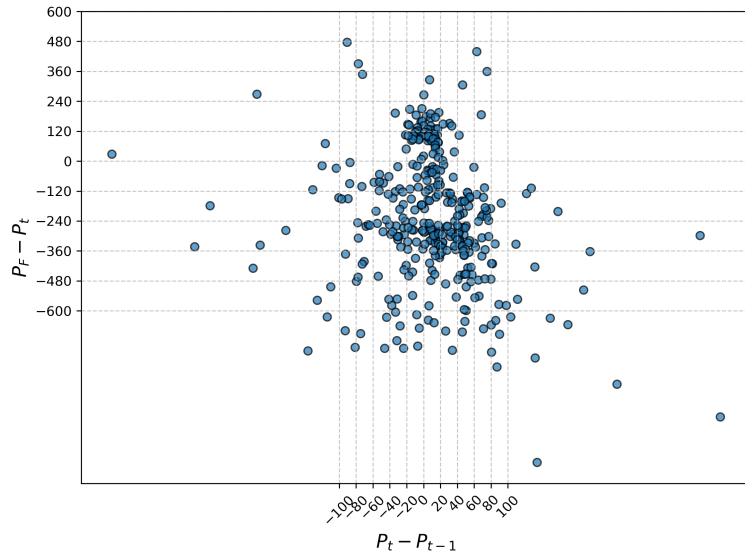


Figure 3: Scatter plot of S&P 500 price variations relative to Gordon fundamentals and lagged prices over the period 1990-2020.

For each combination of the current price, the previous price, and the fundamental value, we conduct 50 independent replications of Prompt 1 to simulate the market composition. The final results are obtained by averaging across the 50 replications. All simulations employ the *gpt-4o-mini* model with a temperature parameter set to 1. Figure 4 shows the probability of assuming a fundamentalist strategy as a function of the deviation from the fundamental value ($P_t^f - P_t$) on the vertical axis and the deviation from the lagged price ($P_t - P_{t-1}$) on the horizontal axis. Red squares indicate the prevalence of fundamentalists, conversely blue

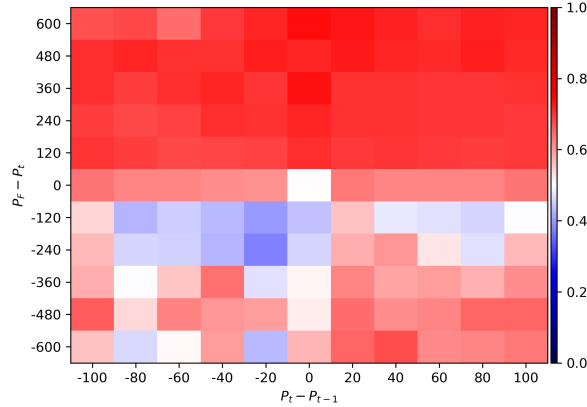


Figure 4: Probabilities of expectation adopted by the generative agent.

squares indicate the prevalence of trend followers. The results indicate that the generative agent exhibits distinct behavioral patterns. When the market price is lower than the fundamental value ($P_t \leq P_t^f$), the fundamentalist strategy dominates the market, regardless of whether the market is in an upward trend (first quadrant) or a downward trend (second quadrant). This result is consistently replicated across all robustness check. In other words, regardless of whether the two strategies offer conflicting signals or move in the same direction, the generative agent adopts a fundamentalist expectation. In contrast, when the market price exceeds the fundamental price, blue colors appear and the probability of adopting a trend-following strategy increases, particularly when $P_t > P_t^f$ and $P_t < P_{t-1}$. Under this condition, agents extrapolate recent trends, expecting the price decline to continue. Interesting, when the market price is above the fundamental and the trend is increasing ($P_t > P_t^f$ and $P_t > P_{t-1}$), the greater the difference between the fundamental value and asset price, the higher the probability of choosing the fundamentalist strategy. This strategy mirrors the one introduced by (Westerhoff, 2009), which assumes an increasing adoption of the fundamentalist strategy when prices deviate excessively from the fundamental value, indicating an unstable trajectory. In fact, as noted by (Westerhoff, 2009), the greater the discrepancy between the fundamental price and the market price, the stronger the belief in mean reversion.

To summarize the results obtained, simulations based on artificial market data confirm the presence of heterogeneity in expectations. At the same time, the prevalence of red areas indicates that fundamentalist expectations prevail across the different cases, particularly when the fundamental price exceeds the actual price. Conversely, blue regions emerge when the actual price lies below the fundamental or when recent price changes are negative, suggesting that trend follower behavior becomes more likely under conditions of positive mispricing and downward momentum.

At this point of the analysis, we can extract the reaction parameters implied by the expectations of the generative agent. To this end, we simulate Prompt 2 and use the combination of price and fundamental value, as in the previous exercise. Based on 50 Monte Carlo simulations, we use Eq. 1 and Eq. 2 to extract the average values of the reaction parameters α (for the fundamentalist expectation) and β (for the trend follower expectation). In Figure 5, we present the results obtained when a fundamentalist expectation is used (on the left) and when trend follower behavior is chosen (on the right).

As shown in the figure on the left, for the fundamentalists, the value of the reaction parameter is consistently positive and lies between zero and one, as suggested by theory. This implies that the generative agent always pushes in the direction of the fundamental value, never exhibiting contrarian behavior with respect to it. In contrast, for generative trend follower, both positive and negative values of the reaction parameter are observed. When the price lies

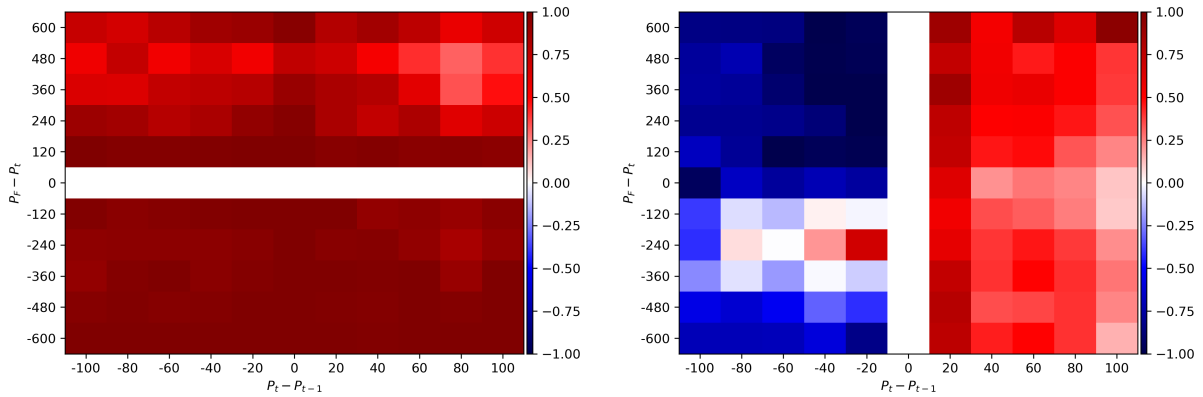


Figure 5: Average values of the reaction parameters for fundamentalists (on the left) and for trend followers (on the right).

below the fundamental, two qualitatively different states can be observed. In periods of upward trends, the reaction parameter is strongly positive, implying that the trend follower expects the appreciation to continue toward the fundamental. Conversely, during downward trends, the reaction parameter becomes negative, indicating that the trend follower behaves as a contrarian, anticipating a trend reversal, and once again expecting an appreciation toward the fundamental. Thus, when the price is below its fundamental value, both fundamental and trend followers contribute to pushing the expected price upward. This pattern is robust in all the variations tested.

Finally, we combine the results obtained from Prompt 1, which identifies the composition of the market, with those of Prompt 2, which generates the agents' expectations, to calculate the expected price for the following period, as previously illustrated in Equation 3. We derive the expected price variation ($E[\Delta(P_{t+1})]$) by subtracting the current price from the expected price:

$$E[\Delta(P_{t+1})] = E[P_{t+1}] - P_t = \delta_t \alpha_t (P_t^f - P_t) + (1 - \delta_t) \beta_t (P_t - P_{t-1}) \quad (7)$$

The resulting values are presented in Figure 6.

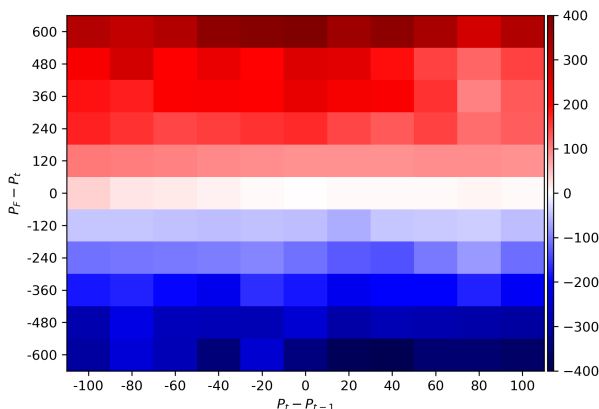


Figure 6: Expected price variation.

As shown, the expected price variation is positive when the fundamental value exceeds the current price and negative when the fundamental is below the current price. Moreover, the absolute magnitude of the expected variations increases with the distance between the current price and the fundamental value, while no significant change is observed with respect

to variations in the difference between the current and previous prices. This outcome is fully consistent with efficient market theory, according to which prices are expected to converge toward their fundamental values. This result can be explained by two main factors. First, the deviations between the fundamental value and the current price are substantially larger than those between the current and previous prices, thus assigning greater weight to the fundamental component in Equation 7. Second, the reaction parameter (α) of the fundamentalist agents, which takes values close to one, further amplifies the influence of the fundamentalist strategy in determining the expected future price.

4 Conclusions

Large language models enable social scientists to move beyond conventional methodologies and pursue innovative research approaches. This paper explored the intersection between large language models and economics, with the aim of assessing whether LLMs can replicate the heuristic expectation-formation mechanisms described in the heterogeneous agent modeling literature, in which market dynamics arise from the coexistence of agents holding different beliefs.

In our analysis, we employed the ChatGPT *gpt-4o-mini* model, accessed via a Python script, and interacted with the LLM using prompting techniques. The study focused on two main aspects: first, the composition of the market, and second, the expectations of the generative agents according to their respective belief systems. Both prompts are tested on real time series data and on artificial market data.

In the time-series analysis, both behavioral strategies formalized in HAMS emerge as relevant, with a higher probability of trend-following behavior preceding the dot-com and global financial crises. This is fully consistent with the predictions of heterogeneous agent models. At the same time, mixed results are obtained during the bust phases. In fact, following the 2007–2008 financial crisis, the adjustment phase is driven by fundamentalist behavior, whereas a negative trend following pattern dominates in the aftermath of the dot-com crisis. Furthermore, when explicitly modeling agents’ expectations, the estimated reaction parameters lie between 0 and 1 for fundamentalists and take both positive and negative values for trend followers.

In the artificial market simulations, both strategies are again identified, although generative fundamentalist behavior prevails. Moreover, the analysis of market composition reveals a marked asymmetry between cases in which the current price is above or below the fundamental value. For prices above the fundamental, the generative agent is likely to switch between the two alternatives. At the same time, for the generative agent, there appears to be no justification for the price to remain below its fundamental value. In fact, when the price lies below its fundamental value, the generative agent systematically selects the fundamentalist strategy. This interpretation is further supported by the analysis of the reaction parameters derived from the expectations of the generative trend-following agents.

In summary, the LLM successfully reproduces a market composition broadly consistent with the predictions of the heterogeneous agent modeling literature with filtered reaction parameters in the range predicted by the HAM literature. At the same time, artificial market simulations provide additional insight into the behavioral rules of the generative agent filtered through the information contained in the prompts. First, some of the heuristics underlying these results lack theoretical support, particularly with respect to the qualitatively asymmetric behavior that depends on the relative position of the fundamental value and the current price. Second, the λ parameter being close to unity and the predominant red-shaded area in the probability of selection of the strategy both point to a greater relevance of fundamentalist behavior. In support of this, by combining the composition of the market with the expectations of the agents, we derived the expected prices implied by a market composed entirely of generative agents. The resulting expected price converges toward the fundamental value, regardless of whether the

market exhibits an upward or downward trend.

In the late 1970s, Simon argued that economics should move beyond the narrow confines of substantive rationality and embrace a theory of procedural rationality, capable of capturing actual decision making under bounded information and computational resources. Crucially, he anticipated the relevance of cross-disciplinary contributions, including artificial intelligence, to advance this research agenda (Simon, 1978). In line with this vision, recent advances in artificial intelligence, particularly the emergence of large language models, have further expanded the methodological frontier of the social sciences. LLMs are increasingly employed as flexible research tools capable of supporting diverse analytical applications, ranging from qualitative text analysis and survey design to discourse modeling and theory testing. Their adaptability allows researchers to tailor them to specific empirical contexts, while their scalability facilitates large-scale analyses that would be prohibitively costly or time-consuming using traditional approaches (Bail, 2024; Bilancini et al., 2024; Zhang and Song, 2024; Argyle et al., 2025). As such, LLMs not only exemplify the integration of computational methods into social inquiry but also embody Simon’s call for a more procedural and interdisciplinary understanding of human decision-making.

References

- Argyle, L. P., E. C. Busby, J. R. Gubler, B. Hepner, A. Lyman, and D. Wingate (2025). Arti-
‘fickle’ intelligence: using llms as a tool for inference in the political and social sciences. *Nature
Computational Science*, 1–8.
- Bail, C. A. (2024). Can generative ai improve social science? *Proceedings of the National
Academy of Sciences* 121(21), e2314021121.
- Bilancini, E., L. Boncinelli, and E. Vicario (2024). Ai-powered chatbots: Effective communica-
tion styles for sustainable development goals. *arXiv preprint arXiv:2407.01057*.
- Binz, M. and E. Schulz (2023). Using cognitive psychology to understand gpt-3. *Proceedings of
the National Academy of Sciences* 120(6), e2218523120.
- Bolt, W., M. Demertzis, C. G. Diks, C. H. Hommes, and M. Van der Leij (2014). Identifying
booms and busts in house prices under heterogeneous expectations.
- Boswijk, H. P., C. H. Hommes, and S. Manzan (2007). Behavioral heterogeneity in stock prices.
Journal of Economic dynamics and control 31(6), 1938–1970.
- Brock, W. A. and C. H. Hommes (1998). Heterogeneous beliefs and routes to chaos in a simple
asset pricing model. *Journal of Economic dynamics and Control* 22(8-9), 1235–1274.
- Bybee, J. L. (2023). The ghost in the machine: Generating beliefs with large language models.
arXiv preprint arXiv:2305.02823.
- Chiarella, C. (1992). The dynamics of speculative behaviour. *Annals of operations re-
search* 37(1), 101–123.
- Chiarella, C., X.-Z. He, W. Huang, and H. Zheng (2012). Estimating behavioural heterogeneity
under regime switching. *Journal of Economic Behavior & Organization* 83(3), 446–460.
- Chiarella, C., X.-Z. He, and R. C. Zwickels (2014). Heterogeneous expectations in asset pricing:
Empirical evidence from the s&p500. *Journal of Economic Behavior & Organization* 105,
1–16.

- Cornea-Madeira, A., C. Hommes, and D. Massaro (2019). Behavioral heterogeneity in us inflation dynamics. *Journal of Business & Economic Statistics* 37(2), 288–300.
- Dawid, H. and D. D. Gatti (2018). Agent-based macroeconomics. *Handbook of computational economics* 4, 63–156.
- Day, R. H. and W. Huang (1990). Bulls, bears and market sheep. *Journal of Economic Behavior & Organization* 14(3), 299–329.
- del Rio-Chanona, R. M., M. Pangallo, and C. Hommes (2025). Can generative ai agents behave like humans? evidence from laboratory market experiments. *arXiv preprint arXiv:2505.07457*.
- Delli Gatti, D., G. Fagiolo, M. Gallegati, A. Russo, M. G. Richiardi, et al. (2018). *Agent-based models in economics: A toolkit*. Cambridge University Press.
- Di Cagno, D. and L. Lin (2025). How do individuals interact with an ai advisor in strategic reasoning? an experimental study in beauty contest. *Journal of Economic Behavior & Organization* 237, 107159.
- Di Francesco, T. and C. Hommes (2025). Sentiment-driven speculation in financial markets with heterogeneous beliefs: A machine learning approach. *Journal of Economic Dynamics and Control* 175, 105092.
- Dieci, R. and X.-Z. He (2018). Heterogeneous agent models in finance. *Handbook of computational economics* 4, 257–328.
- Dosi, G. and A. Roventini (2019). More is different... and complex! the case for agent-based macroeconomics. *Journal of Evolutionary Economics* 29(1), 1–37.
- Ellen, S. t. and W. F. Verschoor (2018). Heterogeneous beliefs and asset price dynamics: a survey of recent evidence. *Uncertainty, Expectations and Asset Price Dynamics: Essays in Honor of Georges Prat*, 53–79.
- Fama, E. F. (1991). Efficient capital markets: Ii. *The journal of finance* 46(5), 1575–1617.
- Faria-e Castro, M. and F. Leibovici (2024). Artificial intelligence and inflation forecasts. *Federal Reserve Bank of St. Louis Review* 106(12), 1–14.
- Franke, R. and F. Westerhoff (2011). Estimation of a structural stochastic volatility model of asset pricing. *Computational Economics* 38(1), 53–83.
- Gao, C., X. Lan, N. Li, Y. Yuan, J. Ding, Z. Zhou, F. Xu, and Y. Li (2024). Large language models empowered agent-based modeling and simulation: A survey and perspectives. *Humanities and Social Sciences Communications* 11(1), 1–24.
- Gatti, D. D., F. Gusella, and G. Ricchiuti (2025). Endogenous vs exogenous fluctuations: unveiling the impact of heterogeneous expectations. *Macroeconomic Dynamics* 29, e125.
- Goldbaum, D. and R. C. Zwinkels (2014). An empirical examination of heterogeneity and switching in foreign exchange markets. *Journal of Economic Behavior & Organization* 107, 667–684.
- Gusella, F. and G. Ricchiuti (2024). Endogenous cycles in heterogeneous agent models: a state-space approach. *Journal of Evolutionary Economics* 34(4), 739–782.

- Hansen, A. L., J. J. Horton, S. Kazinnik, D. Puzzello, and A. Zarifhonarvar (2024). Simulating the survey of professional forecasters. *Available at SSRN*.
- Hommes, C. (2021). Behavioral and experimental macroeconomics and policy analysis: A complex systems approach. *Journal of Economic Literature* 59(1), 149–219.
- Hommes, C. et al. (2017). Booms, busts and behavioural heterogeneity in stock prices. *Journal of Economic Dynamics and Control* 80, 101–124.
- Hommes, C. H. (2006). Heterogeneous agent models in economics and finance. *Handbook of computational economics* 2, 1109–1186.
- Horton, J. J. (2023). Large language models as simulated economic agents: What can we learn from homo silicus? Technical report, National Bureau of Economic Research.
- Korinek, A. (2023). Generative ai for economic research: Use cases and implications for economists. *Journal of Economic Literature* 61(4), 1281–1317.
- Kukacka, J. and J. Barunik (2017). Estimation of financial agent-based models with simulated maximum likelihood. *Journal of Economic Dynamics and Control* 85, 21–45.
- Kukacka, J. and S. Sacht (2023). Estimation of heuristic switching in behavioral macroeconomic models. *Journal of Economic Dynamics and Control* 146, 104585.
- Lopez-Lira, A. (2025). Can large language models trade? testing financial theories with llm agents in market simulations. *arXiv preprint arXiv:2504.10789*.
- Lopez-Lira, A. and Y. Tang (2023). Can chatgpt forecast stock price movements? return predictability and large language models. *arXiv preprint arXiv:2304.07619*.
- Lucas Jr, R. E. (1972). Expectations and the neutrality of money. *Journal of economic theory* 4(2), 103–124.
- Lux, T. (1995). Herd behaviour, bubbles and crashes. *The economic journal* 105(431), 881–896.
- Lux, T. (2018). Estimation of agent-based models using sequential monte carlo methods. *Journal of Economic Dynamics and Control* 91, 391–408.
- Lux, T. (2024). Lack of identification of parameters in a simple behavioral macroeconomic model. *Journal of Economic Dynamics and Control* 168, 104972.
- Lux, T. and R. C. Zwinkels (2018). Empirical validation of agent-based models. In *Handbook of computational economics*, Volume 4, pp. 437–488. Elsevier.
- Majewski, A. A., S. Ciliberti, and J.-P. Bouchaud (2020). Co-existence of trend and value in financial markets: Estimating an extended chiarella model. *Journal of Economic Dynamics and Control* 112, 103791.
- Monti, C., M. Pangallo, G. De Francisci Morales, and F. Bonchi (2023). On learning agent-based models from data. *Scientific Reports* 13(1), 9268.
- Motoki, F. Y., V. P. Neto, and V. Rangel (2025). Assessing political bias and value misalignment in generative artificial intelligence. *Journal of Economic Behavior & Organization*, 106904.
- Muth, J. F. (1961). Rational expectations and the theory of price movements. *Econometrica: journal of the Econometric Society*, 315–335.

- Qu, Y. and J. Wang (2024). Performance and biases of large language models in public opinion simulation. *Humanities and Social Sciences Communications* 11(1), 1–13.
- Recchioni, M. C., G. Tedeschi, and M. Gallegati (2015). A calibration procedure for analyzing stock price dynamics in an agent-based framework. *Journal of Economic Dynamics and Control* 60, 1–25.
- Reitz, S. and F. Westerhoff (2007). Commodity price cycles and heterogeneous speculators: a star-garch model. *Empirical Economics* 33(2), 231–244.
- Schmitt, N. (2021). Heterogeneous expectations and asset price dynamics. *Macroeconomic Dynamics* 25(6), 1538–1568.
- Schmitt, N. and F. Westerhoff (2021). Trend followers, contrarians and fundamentalists: Explaining the dynamics of financial markets. *Journal of Economic Behavior & Organization* 192, 117–136.
- Shen, S., L. Logeswaran, M. Lee, H. Lee, S. Poria, and R. Mihalcea (2024). Understanding the capabilities and limitations of large language models for cultural commonsense. *arXiv preprint arXiv:2405.04655*.
- Simon, H. A. (1978). Rationality as process and as product of thought. *The American economic review* 68(2), 1–16.
- Strachan, J. W., D. Albergo, G. Borghini, O. Pansardi, E. Scaliti, S. Gupta, K. Saxena, A. Rufo, S. Panzeri, G. Manzi, et al. (2024). Testing theory of mind in large language models and humans. *Nature Human Behaviour* 8(7), 1285–1295.
- Ter Ellen, S., C. H. Hommes, and R. C. Zwinkels (2021). Comparing behavioural heterogeneity across asset classes. *Journal of Economic Behavior & Organization* 185, 747–769.
- Tsvetkova, M., T. Yasseri, N. Pescetelli, and T. Werner (2024). A new sociology of humans and machines. *Nature Human Behaviour* 8(10), 1864–1876.
- Westerhoff, F. (2009). Exchange rate dynamics: A nonlinear survey. *Handbook of research on complexity*, 287–325.
- Zarifhonarvar, A. (2024). Experimental evidence on large language models. *Available at SSRN 4825076*.
- Zhang, S. and J. Song (2024). A chatbot based question and answer system for the auxiliary diagnosis of chronic diseases based on large language model. *Scientific Reports* 14(1), 17118.

A Robustness input

In this appendix, we perform robustness checks by varying the inputs provided to the LLM. In subsection A.1, we vary the prompt, while in subsection A.2 we modify the current price, which also affects the fundamental value and the price of the previous period. For price modification, the analysis is repeated both for strategy identification (first script) and for expectation formation (second script). Finally, we test the robustness of the results by increasing the number of repetitions in subsection A.3.

A.1 Prompt

Through interactions with ChatGPT, we generate five variations of the original prompt. These variations preserve its overall meaning while altering the sentence structure and replacing some terms with synonyms. For each version, we run 10 repetitions with the current price set to 1333, the temperature set to 1, and using the model gpt-4o-mini.

Anticipating the results, prompt variations confirm the asymmetry between periods when the price is above versus below the fundamental. As before, fundamentalists dominate when the price is below the fundamental, while trend followers increase when the price is above it. The prompts are reported below and the results are in Figure 7.

Version 1: *You are acting as a market trader who must decide between two strategies based on expectation formation: a fundamentalist strategy and a trend-following strategy. You are given the current fundamental value, the current asset price, and the asset price from the previous period. Your task is to output only two weights, representing the probabilities assigned to each strategy. The weights must sum to 1. Respond only with the two weights separated by a comma, with no additional text.*

Version 2: *Imagine you are a financial trader selecting between two possible trading strategies: fundamentalist or trend-following. Your choice should be guided by expectation formation and the data provided (fundamental value now, current price, and past period price). For each strategy, assign a probability weight. The two weights must add up to 1. Your response must include only the two weights separated by a comma, nothing else.*

Version 3: *As a trader, you are tasked with choosing between two market approaches: the fundamentalist strategy and the trend-following strategy. You base your decision on the given information: the current fundamental value, the current asset price, and the asset price from the prior period. Provide only two numbers, each representing the probability of selecting one of the strategies. These numbers must sum to 1. Your reply should contain only the two weights separated by a comma.*

Version 4: *You are a trader deciding between two alternative strategies informed by expectation formation: (1) a fundamentalist approach and (2) a trend-following approach. The available data are the current fundamental, the present asset price, and the asset price from the last period. Output exactly two weights—probabilities for each strategy—that sum to 1. Provide only the two weights separated by a comma, with no extra commentary.*

Version 5: *Take the role of a financial trader who must allocate probabilities between two strategies: a fundamentalist strategy and a trend-following strategy. Use the given information (current fundamental, current price, and last period's price) to form expectations. Output only two values representing the weights for the two strategies. They must sum to 1. Your answer should consist only of the two weights separated by a comma.*

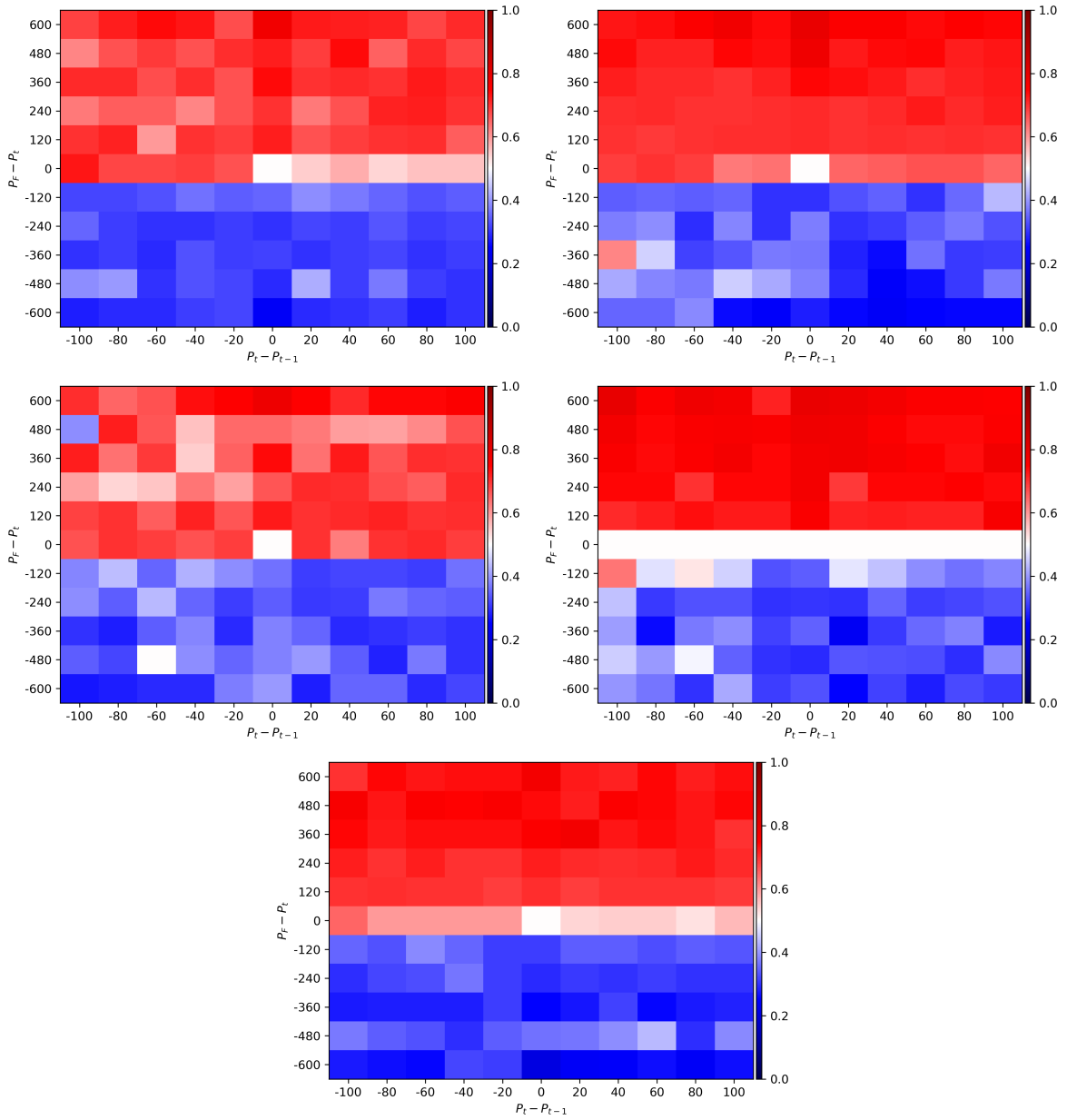


Figure 7: Weight of the fundamentalist strategy. Results are averaged over 10 repetitions with *gpt-4o-mini* at temperature 1 and $P_t = 1333$. Top left: **Version 1** of the prompt; Top right: **Version 2** of the prompt; center left: **Version 3** of the prompt; center right: **Version 4** of the prompt; bottom: **Version 5** of the prompt.

As an additional robustness check, we modify the prompt to simultaneously obtain both the weights associated with the two strategies and the expectations linked to them. Instead of using the two separate prompts presented in the main analysis, we employ the following prompt:

Prompt: *You are a financial trader and you need to choose between two alternative market strategies according to your analysis based on expectation formation: the first is a fundamentalist approach, and the second is a trend-following approach. You have access to the current value of the fundamental, the current value of the asset price, and the value of the asset price at the previous period. You have to provide a weight that represents the probability with which*

you would choose the strategy. The sum of the weights must be 1. Do not provide any other comment, just the two weights divided by a comma and included in double quotes each one. Moreover, you need to provide the value of price expectations for the next period using the two alternative market strategies. Do not provide any other comment or word, just the two numbers divided by a comma and included in double quotes each one. You must write the entire answer on a single line.

As shown in Figure 8, the graphs for the values of α and β closely resemble those obtained in the main analysis. However, the plot for the relative weight of the fundamentalists shows some differences; nevertheless, the main result is confirmed: there is an asymmetry between periods when the price is above or below the fundamental value. Specifically, the predominance of fundamentalists when the price is below the fundamental decreases once the price moves above it. Finally, we present the results for the expected price variation, confirming the robustness of the findings when using an integrated prompt that jointly generates strategies and expectations.

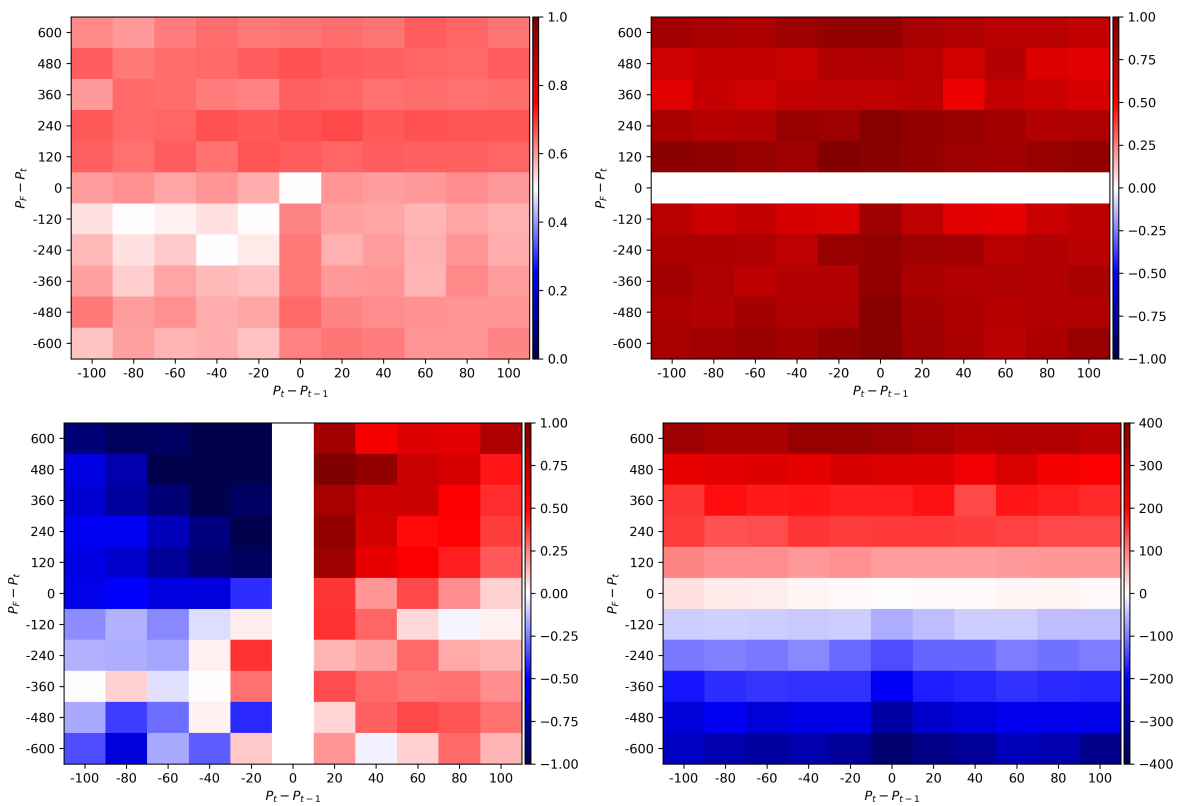


Figure 8: Top left: weight of the fundamentalist strategy; Top right: α ; Bottom left: β parameter; Bottom right: Expected price variation. Results are averaged over 50 repetitions with *gpt-4o-mini* at temperature 1 and $P_t = 1333$, using the unified prompt.

A.2 Current price

In this subsection, we vary the current price. To ensure comparability, we keep $P_F - P_t$ and $P_t - P_{t-1}$ constant; thus, changing P_t also implies changes in P_F and P_{t-1} . The new current price values are: 833, 1083, 1583, and 1833. All other parameters are kept constant, and we perform 50 repetitions for each set of values using *gpt-4o-mini* with temperature 1. The results are replicated both for identifying the strategies used by the LLM and for determining expectations about the future price under the two strategies.

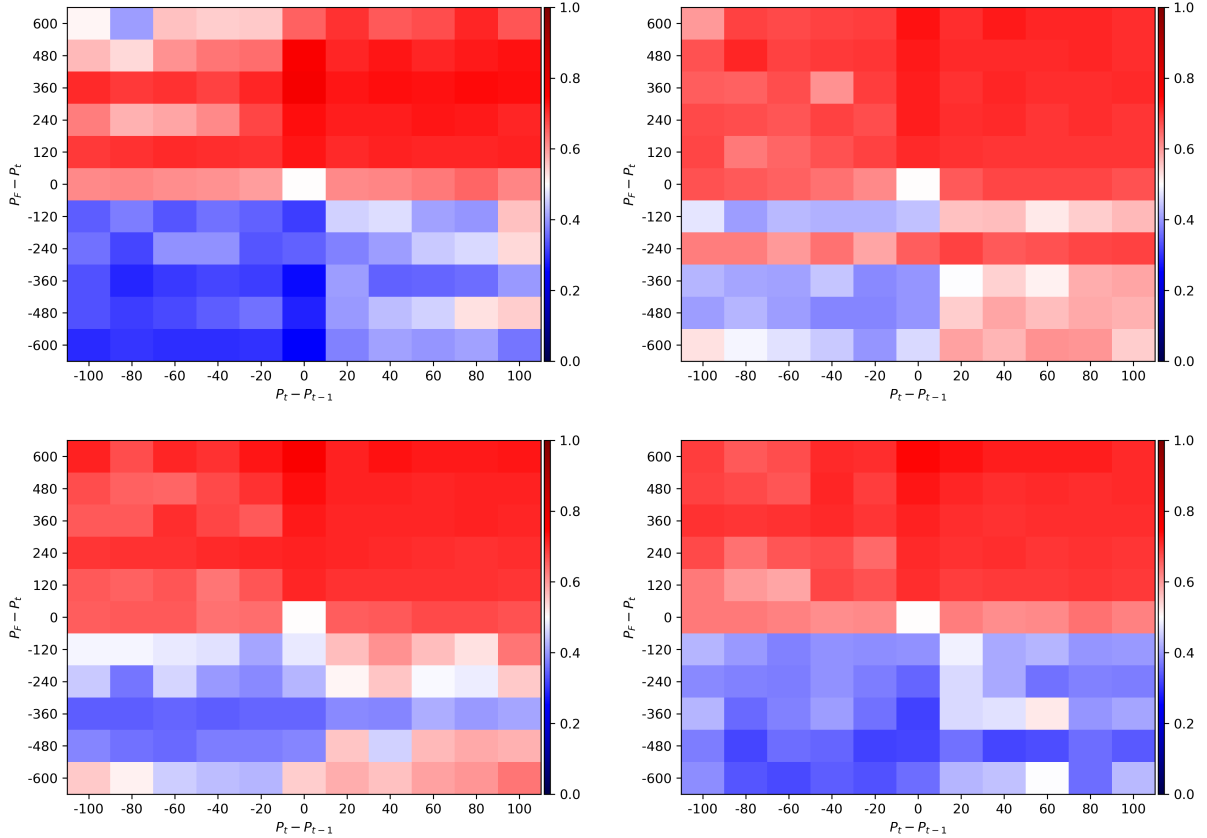


Figure 9: Weight of the fundamentalist strategy. Top-left: simulations for $P_t = 833$; top-right: $P_t = 1083$; bottom-left: $P_t = 1583$; bottom-right: $P_t = 1833$. Each subfigure reports averages over 50 repetitions with *gpt-4o-mini* at temperature 1.

The simulation results for the weights assigned to the two strategies are reported in Figure 9. Once again, we observe an asymmetry between cases where the current price is above or below the fundamental value.

In the following, we report the results of the reaction parameters as the current price varies. As before, we present the average values of the reaction parameters α and β . The results for the reaction parameter α in equation 1, related to fundamentalist expectations, and for the reaction parameter β in equation 2, related to trend follower expectations, are reported in Figure 10. Consistent with previous findings, the parameter α always takes values between 0 and 1, while β shows greater variability. In line with earlier results, when the price is below the fundamental ($P_t^F - P_t > 0$), and the trend is negative ($P_t - P_{t-1} < 0$), the trend follower acts as a contrarian, expecting a reversal. In contrast, when the trend is positive ($P_t - P_{t-1} > 0$) and approaches the fundamental, β takes values greater than zero, indicating that the trend follower expects the upward trend to continue. When the current price is above the fundamental ($P_t^F - P_t < 0$), there is a stronger tendency to follow an upward trend ($P_t - P_{t-1} > 0$), with $\beta > 0$. When the trend is downward ($P_t - P_{t-1} < 0$), both positive and negative values of β coexist, leaving room for expectations that oppose the declining trend.

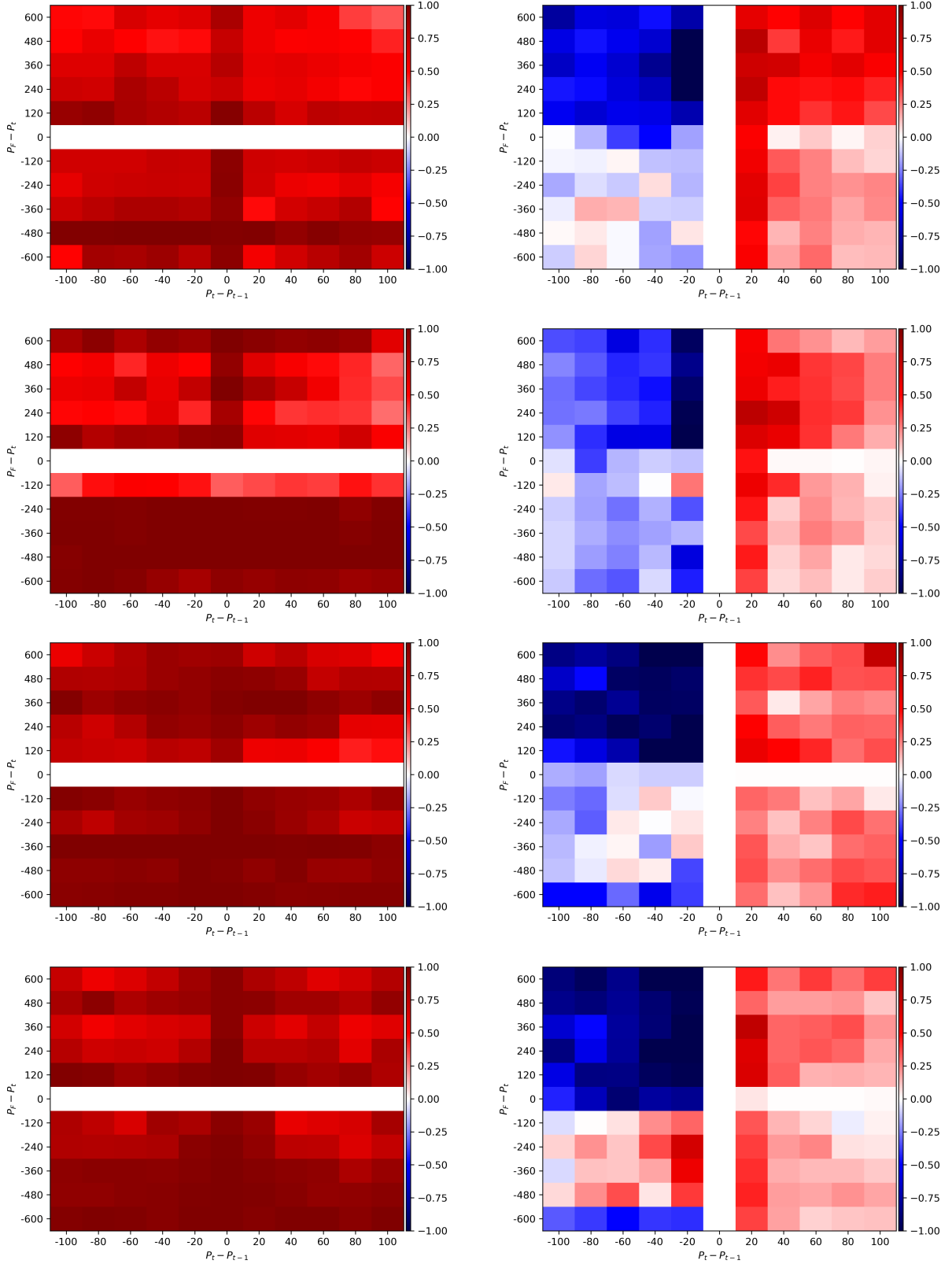


Figure 10: Left: values of α ; Right: values of β ; First line: simulations for $P_t = 833$; Second line: $P_t = 1083$; Third line: $P_t = 1583$; Fourth line: $P_t = 1833$. Each subfigure reports averages over 50 repetitions with gpt-4o-mini at temperature 1.

Finally, we present the results for the expected price change as the reference price at time t varies. Specifically, we use the weights (Figure 9) and the values of the reaction parameters

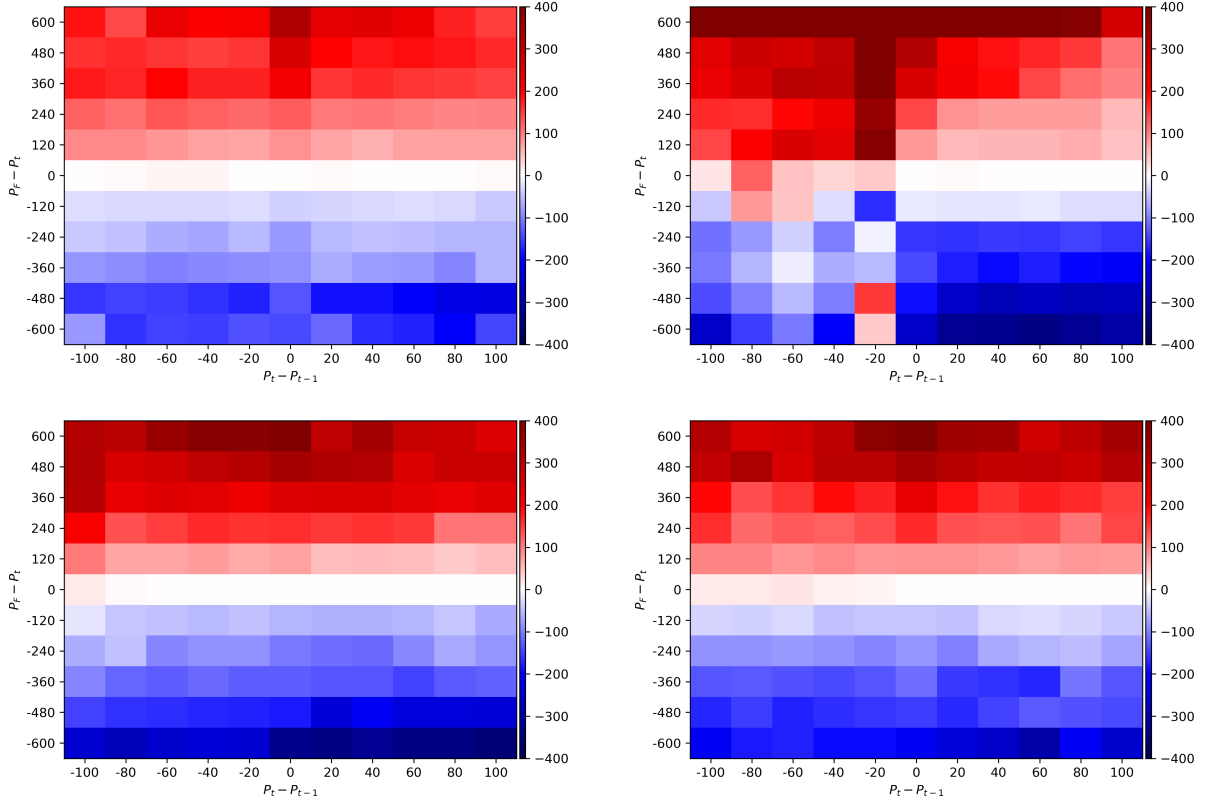


Figure 11: Expected price variation. Top-left: $P_t = 833$; top-right: $P_t = 1083$; bottom-left: $P_t = 1583$; bottom-right: $P_t = 1833$.

(Figure 10) to compute the expected price; by subtracting P_t from this value, we obtain the expected price variation. The results, shown in Figure 11, indicate that the result remains robust to changes in the value of P_t .

A.3 Dimension of simulations

In this subsection, we examine the stability of the results with respect to the number of repetitions. In the main analysis, the results were averaged over 50 repetitions of the prompt. Here, we show the results using 100 repetitions, keeping all other parameters constant: gpt-4o-mini with temperature 1, the original prompt and $P_t = 1333$. As shown in Figure 12, the results with 50 repetitions (left) are stable and fully replicated with 100 repetitions (right).

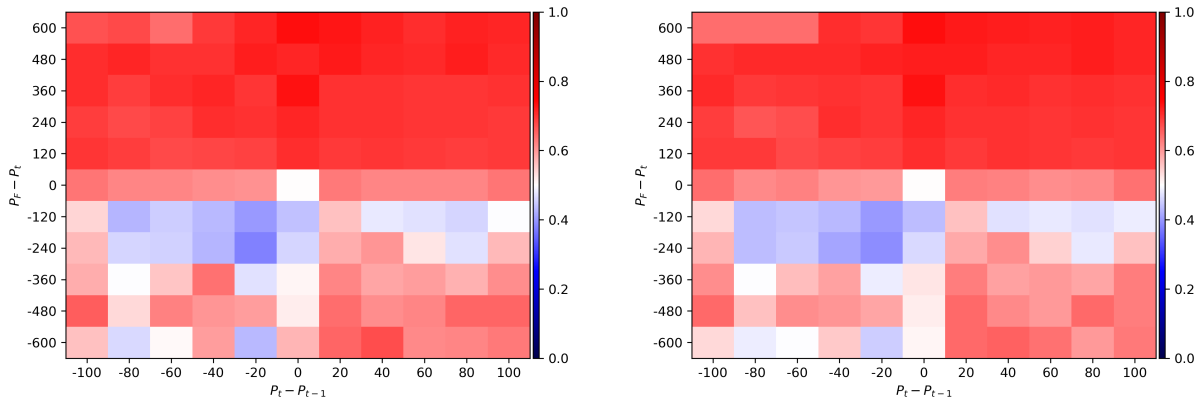


Figure 12: Weight of the fundamentalist strategy. On the left: 50 repetitions; On the right: 100 repetitions. Each subfigure reports results with *gpt-4o-mini* at temperature 1 and $P_t = 1333$.

B Robustness GPT parameters

In this section, we test the robustness of the results by varying the LLM parameters specified in the Python code. Specifically, in subsection B.1 we change the model to *gpt-3.5-turbo*, and in subsection B.2 we vary the temperature.

B.1 GPT model

In this subsection, we present the main results obtained by varying the chatGPT model. We use *gpt-3.5-turbo*, keeping the temperature at 1, the original prompt, and $P_t = 1333$. As before, results are averaged over 50 repetitions. As shown in Figure 13, the asymmetry between the cases where the price is above or below the fundamental persists even with the model change. The absolute values are lower, reflecting greater variability in the responses across the 50 runs.

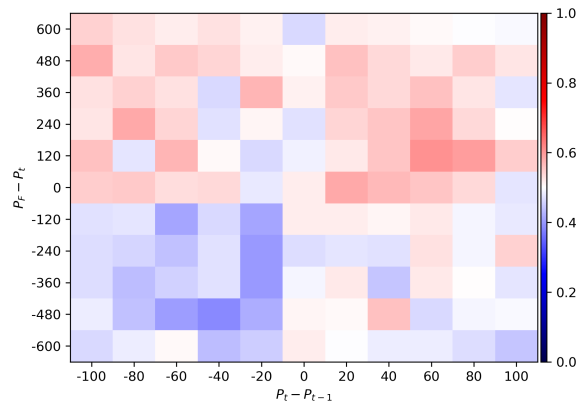


Figure 13: Weight of the fundamentalist strategy. The figure reports the averages over 50 repetitions with *gpt-3.5-turbo* at temperature 1, using the original prompt and $P_t = 1333$.

B.2 Temperature

In this subsection, we examine the robustness of the results by varying the ChatGPT temperature to 0.4 and 0.7, compared to 1 used in the main analysis. As previously mentioned, in

models such as ChatGPT, the temperature parameter controls the randomness of text generation by influencing how likely the model is to select less probable tokens. At low values, close to zero, the model becomes more deterministic, consistently choosing the most probable words and producing stable, predictable, and coherent outputs, though at the expense of creativity and diversity. At higher values, typically between 0.8 and 1.2, the model explores a wider range of possible continuations, which increases variability and originality in the responses, but also increases the likelihood of inconsistencies or errors.

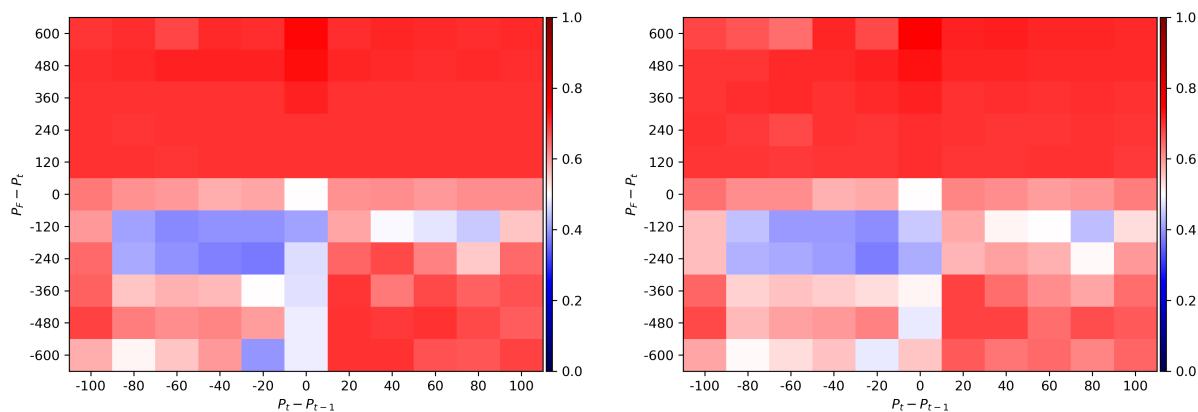


Figure 14: Weight of the fundamentalist strategy. On the left: temperature equal to 0.4; On the right: temperature equal to 0.7.

The results are shown in Figure 14. All other parameters are kept constant, results are averaged over 50 repetitions using gpt-4o-mini, using the original prompt, and with $P_t = 1333$. For both temperatures -0.4 on the left and 0.7 on the right - the main results are confirmed: the asymmetry between cases where the price is above or below the fundamental is evident, with a predominance of fundamentalists when the price is above.