Mental Capabilities and Heterogenous Trading Styles

Andreas Hefti\(^*\)  Steve Heinke\(^\dagger\)  Frédéric Schneider\(^\ddagger\)

June 9, 2017

Abstract

We propose that two distinct, non-convertible mental capabilities—“analytical” and “mentalizing”—jointly determine asset trading behavior. These two dimensions combined predict four mental types, with distinct trading patterns and revenues. Individuals possessing both capabilities trade most successfully; being strong in just one dimension does not assure trading success and can actually be self-defeating. In contrast to our framework, one-dimensional modeling approaches conflate trading styles and, crucially, ignore how both capabilities interact. Laboratory tests, where we elicit subjects’ capabilities and observe their trading behavior, confirm our theoretical predictions. We argue that our findings have important implications for research and practice.

Key words: Asset Markets; Heterogeneity; Mental Capabilities

JEL-Codes: G02, C92

\(^*\)Department of Economics, University of Zürich and School of Management and Law, Zurich University of Applied Sciences, andreas.hefti@econ.uzh.ch
\(^\dagger\)Department of Economics, University of Zürich, steve.heinke@econ.uzh.ch
\(^\ddagger\)School of Management, Yale University, frederic.schneider@yale.edu. FS and AH gratefully acknowledge financial support from the Swiss National Science Foundation, project grant no. 100018_156787. SH gratefully acknowledges financial support from the University of Zürich CanDoc grant.
1 Introduction

Recent research provides ample evidence that a person’s set of cognitive and non-cognitive abilities influences her economic performance (Heckman, Stixrud, Urzua 06; Borghans, Duckworth, Heckman, ter Weel 2008; Borghans, Weel, Weinberg 2008; Burks, Carpenter, Goette, Rustichini 2009; Lindqvist, E., Vestman, R. 2011; S Lundberg - Journal of Labor Economics, 2013) and private life outcomes (Chiteji - American economic review, 2010; Lundberg, Shelly. 2012; Dupuy, Galichon 2014). The specific skill set necessary to succeed depends on the job that is performed. For example, trading on financial markets seems to be affected by mental abilities, such as intelligence (Korniotis and Kumar 2010; Grinblatt, Keloharju, and Linnainmaa 2011; Luik and Steinhardt 2016). In this paper, we address how two distinct sets of skills, analytical ability and mentalizing, interact to affect asset trading behavior. Asset trading is an occupation where performance is arguably a function of analytical abilities of the trader, an intuition that recent research corroborates (Baghestanian, Lugovsky, and Puzzello 2012; Corgnet et al. 2013; Noussair, Tucker, and Xu 2014). At the same time, asset trading is an inherently interactive task, where traders observe other traders’ behavior and need to infer beliefs and intentions of others in order to react appropriately.\footnote{The ability to mentalize affects subjects’ strategic performance in experimental games (Coricelli, Nagel 2009; Devaine, Hollard, Daunizeau 2014). Some recent studies find a correlation between trading behavior and “Theory of Mind” (e.g., Baghestanian, Lugovsky, and Puzzello 2012; Corgnet et al. 2013; Noussair, Tucker, and Xu 2014).} It therefore seems plausible that the the endowment and combination of analytical and mentalizing affects an individual’s trading strategy and that successful traders need a combination of analytical and mentalizing abilities.

We propose that only an integrated model of both analytical and mentalizing abilities can adequately explain why certain traders are successful while others fail, despite the availability of the same information to all traders. We develop a unifying framework that explains the interaction between both mental capabilities and organizes the substantial heterogeneity of observed trading patterns in asset markets. In our model, analytical
ability is valuable to determine the fundamental value of an asset, and mentalizing helps with finding opportunities for price speculation. A crucial prediction of our model is that price speculation can only be successful if a trader has both abilities, while a one-sided deficiency produces a characteristic (costly) bias in trading behavior. Because abilities jointly determine behavior, our model also suggests that any theory of individual trading behavior that relies on a one-dimensional measure of mental capabilities may lead to an incomplete analysis of trading behavior, and to potentially biased conclusions about what causes successful trading.

We test the main predictions of our model in a standard experimental asset market, and find that the data support our theoretical conjectures.

1.1 The Framework

To make sense of asset market data, researchers commonly assume different behavioral trader types (Boswijk, Hommes, and Manzan 2007). Modeling the source of this heterogeneity is a central issue of behavioral research in economics and finance. Some research suggests that information asymmetry drives heterogeneous trading behavior (XXX Cite key papers). We hypothesize that, even if information is perfectly symmetric among market participants, observed trading behavior is heterogeneous due to differences in mental capabilities because different mental dispositions lead to systematic differences in how humans process and evaluate the same information. Specifically, we posit that the way traders think about investment decisions depends on two fundamentally different mental capabilities. Analytical capability (A-Dimension) refers to a person’s grasp of the quantitative aspects of a decision problem. This includes logical reasoning and mathematical or probabilistic calculations. Mentalizing capability (M-Dimension) is the ability to impute mental states of others and to understand their beliefs and intentions.

2. As Hommes (2011) notes: “An important challenge to a research program in behavioral economics and finance based on bounded rationality is to come up with a plausible and general theory of heterogeneous expectations.”
This ability can be used to make predictions about other people’s behavior (Premack and Woodruff 1978).³

Our framework conceptualizes how these two dimensions may determine behavior. We posit that each capability influences, in a distinctive way, how a situation is represented in the mind of the decision-maker. Our core assumption is that neither capability can be used as a substitute for the other. That is, having a strong analytical capability cannot compensate for insufficient mentalizing, and vice-versa. Therefore, a deficiency in one capability will then lead to a characteristic distortion of behavior.

![Figure 1: The four basic cognitive types](image)

This figure shows the four mental types. Due to the independence of the analytical (A)- and mentalizing (M)-abilities, one can plot both mental abilities as an ordered pair of perpendicular lines. Each point in this plane represents a specific mental ability mix. In our analysis, we concentrate on four stylized types with four distinct mental ability mixes.

Our setting yields a classification of four different mental types, according to their ability in each of the two dimensions (Figure 1). “Technocratic” types (TE) have high analytical but low mentalizing ability. This type comprehends the logical and quantitative aspects of a situation, but cannot sufficiently account for the “psychological factor” in a strategic decision. For example, a technocratic type can correctly deduce the Nash

---

3. Terms such as Theory of Mind, mentalizing, and cognitive empathy are often used interchangeably in the literature. The unifying aspect is the ability to put oneself in “the shoes of the others” (Frith and Singer 2008; Van Overwalle and Baetens 2009; Reniers et al. 2011). Analytical reasoning and mentalizing ability are stable and mostly independent traits among people (Reniers et al. 2011; Van Overwalle and Baetens 2009).
equilibrium of a game, but has difficulties in making sense of deviations from equilibrium behavior. In contrast, “semiotic” types (SE) are keenly aware of others’ behavioral patterns and can deduce intentionality from these patterns, but they lack a conceptual understanding of the decision situation.4 “Featureless” types (FL) have low levels of both analytical and mentalizing capabilities, while “Sophisticated” (SO) types are strong in both dimensions. They understand both the logical aspects of a decision situation and the possibility of intentional distortions by other agents.

We apply this framework to the asset market by suggesting that each person forms a mental model of the asset value based on her individual mental capabilities. We posit that each of the two mental dimensions corresponds to one aspect of valuing an asset: a trader needs analytical ability to correctly understand the fundamental value of an asset; she needs mentalizing ability to correctly judge market sentiments. This dichotomy is the basis for a testable model of mental capabilities and ensuing trading behavior.

The main predictions of our theory can be summarized as follows. The two-dimensional, non-convertible nature of mental capabilities generate distinguishable, characteristic trading patterns of the four mental types, which could be described as noise trading, fundamental trading, trend chasing, and bubble riding, respectively. Consequently, the four types vary in their success in the market. For example, the most successful trader needs both mental abilities, because only the combination of the two allows her to not only trade the asset according to the fundamental but also to exploit systematic deviations of the price from the fundamental. Therefore, we predict that SO will be the most profitable type.

Crucially, our theory predicts a non-monotonicity between capabilities and outcomes; being more skilled in one dimension does not necessarily imply higher profit. Analytical capability alone does not generate substantial trading gains because a lack in mentalizing capability leads to a misinterpretation of price deviations from the fundamental.

4. We call this type semiotic because this type tries to read the behavioral “signs” of intentions in observables (such as past asset prices).
Therefore, TE cannot speculate successfully and should earn less than SO. Conversely, a strong mentalizing capability alone is even more detrimental: semiotic traders will detect, and follow, an upward price trend but miss the optimal exit point since they do not sufficiently account for the fundamental. Accordingly, being more skilled in only one dimension is not sufficient to develop a successful trading strategy, and may even decrease profits. Indeed, we show that when analytical capability is weak, not having mentalizing capabilities (i.e., being featureless) mitigates the expected trading losses. Therefore, we predict SE to incur the highest trading losses.

In terms of trading patterns, the “off-diagonal” types (TE and SE) should display the starkest contrast because they lack the other type’s capability. This distinctive prediction hinges crucially on our non-convertibility assumption of mental capabilities. If capabilities were substitutes, such that only the aggregate capability level matters rather than the type of capability, then behavior on the off-diagonal should be indistinguishable.

1.2 The Experiment

We conduct a laboratory experiment to validate our main hypotheses. Beyond the opportunity to measure participants’ mental types, the laboratory offers the necessary control over the decision environment and the parameters. Most importantly, we can confine trading to a single asset, whose expected value we control. Moreover, we can make sure that all subjects have access to exactly the same information. This means that observable differences in behavior cannot be attributed to asymmetric information of the market participants, but rather to asymmetric information processing.

Our experimental design consisted of two independent phases. In a first phase, we elicited participants’ ability in both the analytical and the mentalizing dimension using separate, incentivized tasks (see Section 3.2. Participants’ behavior in these tasks allows us to independently classify them into our four mental types. In a second phase, we observed participants’ behavior in a standard experimental asset market game, which is
widely known to yield a price bubble (Smith, Suchanek, and Williams 1988). We then used the collected data to test the various predictions of our model.

The empirical findings are consistent with our theoretical predictions. Analyzing subjects’ trading patterns, we find that technocratic types largely trade on the fundamental, buying when the asset price is below or at the fundamental value, and selling when the price rises above it. These types make money from the dividend but miss out on the profits from speculating on the bubble. Semiotic types follow the rising asset price, with peak asset holdings after the peak of the bubble. These types make the largest losses as they are unable to unload their shares profitably after the peak. The sophisticates anticipate both the rising and the bursting of the bubble. These types make the most money by having the best market timing. Finally, featureless types show no pronounced trading pattern.

1.3 Related literature

Within the large literature on the determinants of traders’ behavior in (experimental) asset markets, we review those that examine the role of analytical and mentalizing abilities. Analytical capability seems to influence individual trading behavior, profits, and bubble size. For example, participants scoring well in the cognitive reflection test (Frederick 2005) achieve higher profits in laboratory spot markets (Corgnet et al. 2013), and in spot markets with an added futures market (Noussair, Tucker, and Xu 2014); their trading style is less focused on momentum and more on fundamental value (Baghestanian, Lugovskyy, and Puzzello 2012). On the other hand, Bruguier, Quartz, and Bossaerts (2010) do not find that cognitive ability is related to the ability to correctly predict asset prices, and Janssen, Weitzel, and Füllbrunn (2015) find no correlation between CRT and behavior in their speculation task. In terms of aggregate outcomes, markets with traders exhibiting higher average analytical abilities produce lower price volatility (Breaban and
These observations are in line with studies on actual stock market behavior. In general people with higher IQ are more likely to trade in stock markets, hold a more diversified portfolio, and achieve higher Sharpe ratios than people with lower IQ, even after controlling for socio-demographic covariates (Korniotis and Kumar 2010; Grinblatt, Keloharju, and Linnainmaa 2011; Luik and Steinhardt 2016).

Bruguier, Quartz, and Bossaerts (2010) are among the first to provide evidence that “theory of mind” (ToM) and the ability to forecast prices are correlated. In their experiment, a strong ToM allows an observer to successfully discern malicious from benevolent intent in price patterns. De Martino et al. (2013) find that a higher mentalizing ability correlates with the tendency to ride the bubble too far and to lose money.

According to our framework, these studies tell only part of the story. We contend that only the interaction of the two independent dimensions can explain the complex heterogeneous behaviors we observe in asset markets, even under symmetric information. To our knowledge, our paper provides the first theoretical and experimental documentation of how two separate and unrelated mental abilities interact to generate diverse mental types displaying distinctive behavioral patterns.

2 Mental models and trading behavior

In this section we present a simple model that formalizes our conception of mental capabilities and individually biased behavior. The main idea is that a systematic observed

---

5. A recent study suggests that bubbles (and high volatility) arise through the interaction between different analytical types (Hanaki et al. 2015).

6. In their experiment, they compare the brain activation of subjects in markets with and without insiders. Subjects are rewarded for accuracy in predictions (forecasting market prices), but do not participate themselves in the market.

7. Neuro-economic evidence suggests that activity in certain brain regions is correlated with the asset trading behavior shown by some of our mental types. In particular, De Martino et al. (2013) link activity in the dorsomedial and ventromedial prefrontal cortex with behavior we see in our “Semiotic” mental type; Smith et al. (2014) find that activity in the nucleus accumbens is related to “Semiotic”-type behavior and the anterior insular cortex is related to “Sophisticated” behavior.
heterogeneity in the behavior of individuals, who face the same information set and the same decision, is a consequence of systematic differences in individual thinking, where individual thinking is correlated with mental capabilities. Specifically, we posit that in a complex, strategic decision two diverse mental capabilities, analytical reasoning and mentalizing, influence individual decisions, reflecting that both, a logical understanding of the objective aspects of the decision and an appropriate assessment of how others may evaluate the decision problem are central aspects of such decisions. Differences in these mental capabilities then induce systematic distortions in the individual valuations of certain observables, resulting in heterogeneous trading behavior.

2.1 Mental models, capabilities and distortions

Suppose that in any period \( t \) each trader \( i \) forms a subjective mental model that determines her valuation \( V^i(t) \) for an asset, conditional on a set of observables. Specifically, let these observables be the past asset prices \( \{P(0),...,P(t)\} \) and the fundamental value \( F(t) \).\(^8\) Our core idea is that someone’s analytical and mentalizing capabilities can influence how these observables affect the mental model and the resulting trading decision. That is, while the same set of information is available to all decision makers, it may be interpreted differently by individuals with a different mental profile. There is a common understanding in the literature that both the fundamental value \( F(t) \) and the last observed price \( P(t) \) are important determinants of an asset’s value (e.g., Hellwig 1980; Kyle 1985, 1989), and we adopt this view by supposing that the ideal mental model of \( V(t) \) is given by a mapping\(^9\)

\[
V(t) = \tilde{\phi}(F(t), P(t)).
\]

\(^8\) In the experiment, the fundamental value amounts to the expected future dividend which, by design, is in the information set of the subjects. In reality, the fundamental value would be derived from (possibly noisy) observables, such as financial statements or balance sheets.

\(^9\) We include only the last observed price for reasons of simplicity and tractability. A longer history would not change the nature of our predictions, and we can empirically test for several price lags (which turn out to be irrelevant).
Our theoretical framework builds on the postulate that each individual mental model is calibrated around (1), where systematic deviations from (1) can occur in accordance with someone’s profile of mental capabilities. Intuitively, the ideal model represents the optimal way to evaluate the given information at any point in time. That is, if someone’s decisions were based on (1) this would result in the highest possible trading gains. The assumption that such an ideal model hypothetically exists simplifies the derivation of the relative valuation biases between agents with different mental capabilities. In fact, we do not need to know the details of the ideal model to obtain such differential predictions, nor do we require that some traders actually behave according to (1).\footnote{It may be possible to fit an ideal mental model \textit{ex-post}, e.g., by estimating a (linearized) version of (1) on past data.}

Suppose that the mental model of trader $i$ is

$$V^i(t) = \tilde{\varphi}(d_F^i, F(t), d_P^i, P(t)),$$

where $d_F^i, d_P^i \in [0, 1]$ are trader-specific distortion coefficients. These distortion coefficients capture how a trader’s mental model may deviate from (1), where we impose additive separability as a simplifying assumption.

**Assumption 1** (Additive separability). The function $\tilde{\varphi}()$ in (2) is twice continuously differentiable and verifies

$$\frac{\partial^2 \tilde{\varphi}}{\partial d_F \partial P} = \frac{\partial^2 \tilde{\varphi}}{\partial d_P \partial F} = \frac{\partial^2 \tilde{\varphi}}{\partial d_F \partial d_P} = 0.$$  \hspace{2cm} (3)

Condition (3) is equivalent to $\tilde{\varphi}$ being additively separable in $(d_1, F)$ and $(d_2, P)$, where additive separability between fundamental and price is common in the literature.\footnote{For example, if $\tilde{\varphi}_1(\cdot), \tilde{\varphi}_2(\cdot)$ are \textit{linear} functions, our model corresponds to the asset demand in the rational noise trader environment (e.g., Kyle 1985, 1989).}
sequently, we have

\[ V^i(t) = \tilde{\varphi} \left( (d^i_F, F(t)), (d^i_P, P(t)) \right) = \tilde{\varphi}_F \left( d^i_F, F(t) \right) + \tilde{\varphi}_P \left( d^i_P, P(t) \right) + \varepsilon_i. \]  

(4)

The residual term \( \varepsilon_i \) captures idiosyncratic noise that may either result from (4) being an incomplete specification of the actual mental models,\(^\text{12}\) or because of coding noise in the mental representation of information. The contribution of our model is to identify systematic differences in the average trading behavior of certain mental types, and the precise probabilistic structure of the \( \varepsilon_i \) turns out not to be important. For example, we do not require that \( \varepsilon_i \) is iid across traders. As a normalization, we let \( d_j \in [0, 1], j \in \{F, P\}, \) where \( d_j = 1 \) means that observable \( j \) enters \( i \)'s mental model in an unbiased way.

Accordingly, the ideal mental model (1) is the unbiased version of (4):

\[ \tilde{\varphi} (F(t), P(t)) = \tilde{\varphi}_F (1, F(t)) + \tilde{\varphi}_P (1, P(t)) + \varepsilon. \]  

(5)

The functions \( \tilde{\varphi}_j, j \in \{F, P\} \) embody the direction and strength of the possible bias. For example, if \( d^i_F < 1 \) and \( \frac{\partial \tilde{\varphi}_F}{\partial d^i_F} > 0 \), there is undervaluation of \( F \) in \( i \)'s mental model compared to the ideal model, or if \( d^i_P < 1 \) and \( \frac{\partial \tilde{\varphi}_P}{\partial d^i_P} < 0 \) there is overvaluation of \( P \). Moreover, our model can capture that sometimes irrelevant observables may be included in valuations by decision-makers, which can be seen as an extreme form of over- or undervaluation.\(^\text{13}\) Likewise, our model includes the case where a distorted valuation does not respond to observables when it should.

**Mental capabilities** Our core conjecture is that mental capabilities influence individual valuations by affecting the distortions \( d_F, d_P \). Specifically, we posit that \( d_F^i \) depends

---

\(^{12}\) Despite additive separability, the residual term \( \varepsilon_i \) could depend on \( d_F, d_P \) or \( P(t), F(t) \) in specific ways. For example, \( \varepsilon_i = \varepsilon_1(d^i_F) + \varepsilon_2(d^i_P) + \varepsilon_3(F, P) \).

\(^{13}\) For example, if \( \phi_P(d^i_P, P) = (1-d^i_P)\alpha P + k, \alpha \neq 0 \), then variations in \( P \) should not affect valuations, but for individuals with \( d^i_P < 1 \) they still do. Particularly, if \( \alpha > 0 \) and \( d^i_P < 1 \), then an increase in \( P \) would induce an overvaluation of the asset, making traders respond to prices where they should not.
only on the capability “analytical reasoning” \((A)\), while \(d^i_P\) depends only on the “mentalizing” \((M)\) ability of trader \(i\). The main intuition is as follows. In a sufficiently complex, strategic environment, the observables \(F(t)\) and \(P(t)\) convey information of a vastly different type, and therefore their evaluation requires a substantially different mental processing. First, the fundamental value \(F(t)\) of the asset represents its objective, intrinsic worth. In the SSW market, \(F(t)\) corresponds to the expected stream of future dividends. Correctly accounting for this type of information in an asset’s valuation typically requires logical thinking, stochastic reasoning or related quantitative processing. In the context of our experiment, this means to correctly calculate the expected dividend, which needs a solid understanding of the probabilistic structure of the dividends and backward-inductive thinking, which includes a close monitoring of the remaining periods. The ability to adequately perform such calculations is broadly related to someone’s (abstract) analytical capabilities. Second, each trader can try to infer from the observed price how others have thought about the asset because the realized market price, \(P(t)\), reflects the average valuation of all market participants. Given that the market price is the only observable through which a trader can obtain an impression about the others’ valuations, a trader must interpret the observed prices in a “semiotic” way. We therefore posit that a stronger ability to mentalize about others leads to a less biased evaluation of the information expressed by the past prices.

The above discussion is formalized as follows. Let the mental capabilities \(A\) (analytical) and \(M\) (mentalizing) be quantified by \(c^i_h \in \mathbb{R}_+, h \in \{A, M\}\), where the pair \((c^i_A, c^i_M)\) is trader \(i\)’s mental profile. We suppose that the distortion coefficients in each dimension depends separately on the relative difference between the available mental capacity \(c^j_i\) and the minimal level of capacity \(\bar{c}_j\) required by the ideal mental model.

**Assumption 2 (Non-convertibility).** For \(h \in \{A, M\}\) each distortion coefficient \(d^i_F, d^i_P\) is determined by a function \(d_F, d_P : \mathbb{R}_+^2 \to [0, 1]\), \(d^i_F = d_F(c^i_A, \bar{c}_A)\) and \(d^i_P = d_F(c^i_M, \bar{c}_M)\).
\[ d_F(c^i_A, \bar{c}_A) = \begin{cases} 
\frac{\bar{c}_A}{c^i_A}, & c^i_A \leq \bar{c}_A \\
1, & \text{else} \end{cases} \]

\[ d_P(c^i_M, \bar{c}_M) = \begin{cases} 
\frac{\bar{c}_M}{c^i_M}, & c^i_M \leq \bar{c}_M \\
1, & \text{else} \end{cases} \]  \hspace{1cm} (6)

Any trader with a sufficiently strong mental profile \((c^i_A, c^i_M) \geq (\bar{c}_A, \bar{c}_M)\) would recover the ideal model (1), but because only relative distortions matter (see below), our theory works fine even if \(\bar{c}_A, \bar{c}_M\) were arbitrarily large, such that nobody fully matches the ideal model.\(^{14}\) The importance of a possible bias \((d^i_j < 1)\) in (4), due to a lack of the underlying mental capabilities, depends on the nature of the decision or, more precisely, on the shape of the ideal mental model (5) that matches a certain decision-situation. Formally, this is accounted for by the \(\hat{\varphi}\)-functions. For the type of price bubble that occurs in the SSW experiment both mental capabilities will play a decisive role, but in other circumstances the relevance of the capabilities may change (section 2.3 contains a simple example).

Further, the functional form (6) is without loss of generality, except for its increasing nature.\(^{15}\)

The important part of Assumption 2 is that each distortion coefficient depends on one and only one mental capability. This non-convertibility of mental capabilities implies that mental capabilities are orthogonal to each other: any ceteris paribus change of \(P\) has an effect on \(V^i(t)\) that is independent of \(c^i_A\), while any change of \(F\) has an effect on \(V^i(t)\) that is independent of \(c^i_M\). In particular, any possible spare capacity \(c^i_h - \bar{c}_h > 0\) cannot compensate for a deficient capacity in the other dimension \(-h\). While this non-convertibility assumption may seem strong at this point, we view it as intuitive in light of the above discussion, and at its very least it is a fairly natural benchmark assumption.

\(^{14}\) Conversely, if all traders had \(c^i_h \geq \bar{c}_h\), then everybody would have the ideal understanding of dimension \(h\). Hence our theory offers a differential prediction for different mental profiles only if the decision situation is cognitively complex enough, i.e., each \(\bar{c}_h\) is so large that \(c^i_h < \bar{c}_h\) for some individuals.

\(^{15}\) If, say, \(d_F(c_A, \bar{c}_A)\) is any arbitrary function \(d_F : \mathbb{R}_+^2 \to [0, 1]\) that is strictly increasing for \(c_A \leq \bar{c}_A\) and constant for \(c_A > \bar{c}_A\), then there is a bijection \(B\) between \(d_F\) defined in (6) and \(d_F\) such that \(\hat{\varphi}_F(d_F(c_A, \bar{c}_A), F) = \hat{\varphi}_F(d_F(c_A, \bar{c}_A), F)\), where \(\hat{\varphi}_F = \hat{\varphi}_F \circ B\).
From the expositional viewpoint, non-convertibility allows us to derive all theoretical predictions in a concise, clean-cut and tractable way. Section 2.3 clarifies that our main results survive a substantial weakening of non-convertibility. Nevertheless, this section also highlights that full convertibility would result vastly different predictions about the observable trader heterogeneity. This obviously makes the empirical analysis covered by this paper a central mean of arbitration between these two hypothesis.

Relative biases The predictions offered by our model are of a relative nature, where the traders with the comparably strongest mental profiles in a sample serve as a benchmark. Formally spoken, the fact that no trader may have a mental profile that can fully manage the cognitive load imposed by a complex asset market does not constrain the comparative-predictive power of our theory. Particularly, we do not need to know the ideal mental model \( (e_A, e_M) \) for analyzing the data.

Suppose that in a sample of \( I \) people some individual has \( 0 < (c^S_A, c^S_M) \leq (\bar{e}_A, \bar{e}_M) \), where \( c^S_h = \max\{c^1_h, ..., c^I_h\}, j \in \{A, M\} \). Call this individual a Sophisticate, and define \( \varphi_j(r_j, \cdot) \equiv \tilde{\varphi}_j(r_j d_j(c^S_h), \cdot) \), where \( r_j = d_j(c^1_h)/d_j(c^S_h) = c^1_h/c^S_h \in [0, 1] \) is the relative distortion of trader \( i \). Note that \( r_j = 1 \) iff \( c^1_h = c^S_h \). With \( r_j = r_j(c^1_h, c^S_h) \), where \( r_j(\cdot, c^S_h) \) is strictly increasing if \( c^1_h < c^S_h \) (Assumption 2), (4) can be expressed as

\[
V^i(t) = \varphi_F(r_F(c^1_A), F(t)) + \varphi_P(r_P(c^1_M), P(t)) + \varepsilon_i,
\]

(4')

Note that if the nature of the bias is undervaluation (overvaluation), then \( \frac{\partial \varphi_F}{\partial r_j} > 0 \) (\( \frac{\partial \varphi_P}{\partial r_j} < 0 \)). As an illustration, suppose that each distortion coefficient and the corresponding observable are multiplicatively separable. With undervaluation (4') then becomes

\[
V^i(t) = r_F(c^1_A)\varphi_F(F(t)) + r_P(c^1_M)\varphi_P(P(t)) + \varepsilon_i.
\]

(7)

16. Note that if \( (c^S_A, c^S_M) \geq (\bar{e}_A, \bar{e}_M) \), the mental model of a Sophisticate would simply coincide with the ideal mental model.

17. With overvaluation, an example is given by \( V^i(t) = 1/r_1(c^1_A) \varphi_1(F(t)) + 1/r_2(c^1_M) \varphi_2(P(t)) + \varepsilon_i \).
**Binary capability model** We derive our predictions within the tractable case of binary capabilities, $c_i^j \in \{0, 1\}$. Hence any mental profile verifies $(c_A^i, c_M^i) \in \{0, 1\}^2$. This yields the four distinguishable mental types depicted in Figure 1: (0, 0) FL (featureless), (0, 1) SE (semiotic), (1, 0) TE (technocratic), (1, 1) SO (sophisticate). Assumption 2 implies the following relations among the type-specific relative distortions $r_i^j$:

\[
\begin{align*}
    r_{TE}^F &= r_{SO}^F & r_{SE}^F &= r_{FL}^F & r_{FL}^F &\neq r_{TE}^F \\
    r_{TE}^P &\neq r_{SO}^P & r_{SE}^P &= r_{FL}^P & r_{FL}^P &\neq r_{TE}^P
\end{align*}
\]

(8)

It follows that the mental types FL, SE, TE have a valuation equation (4) that is distorted, relative to SO, in an essentially unique and characteristic way. Put differently, these four mental profiles map into four distinct distortion profiles, and we now derive the heterogeneous trading patterns on the basis of this four-type model.

2.2 Mental types and trading patterns: Main Hypothesis

We first establish how mental capabilities map into differential responses to changes in the observables, and then derive the main consequences of these deviations for the trading patterns and the aggregate distribution of trading gains across the four mental types identified by the binary capability model.

2.2.1 From mental capacities to stimulus responses

We concentrate on the case, where decisions are complex and all four mental types are present in the market, such that their relative distortions (8) matter. Let $\alpha_F^i = \alpha_F(r_F^i, F) \equiv \frac{\partial \phi_F(r_F^i, F)}{\partial F}$ and, analogously, $\alpha_P^i = \alpha_P(r_P^i, P) \equiv \frac{\partial \phi_M(r_M^i, P)}{\partial P}$ denote trader $i$’s stimulus response to changes in $F$ or $P$, respectively. Non-convertibility then implies the following hypothesis

---

18. This is wlog, because we could replace $\{0, 1\}$ by an arbitrary $\{c_L, c_H\}$, where $c_L < c_H$.

19. If decisions were simple, distortions would not arise according to our model, and the four mental types should behave in an essentially indistinguishable way.
(H1) (Partial symmetry) $\alpha_{TE}^{T} = \alpha_{SO}^{T}$ and $\alpha_{SE}^{S} = \alpha_{SO}^{S}$.

The hypothesis states that the off-diagonal mental types in Figure 1 each shares a stimulus-response coefficient with the Sophisticate. That is, the off-diagonal types “get it right” with respect to a specific observable.

Further, the off-diagonal types TE and SE have the most disparate distortion profiles, because according to (8) $r_{TE}^{T} \neq r_{SE}^{T}$ and $r_{TE}^{P} \neq r_{SE}^{P}$. These differences in distortions translate to an antipodal behavior of FE and SE if relative distortions matter for the stimulus-response behavior in a monotonic way, which we assume in the following.\(^{20}\)

Monotonicity means that $\alpha_{j}(\cdot, j)$ is injective for $r_{j}^{*} < 1$, implying that our model produces a systematic “upwards” or “downwards” deviation in the stimulus responses between different mental types. Technically, the monotonicity assumption requires that each $\varphi_{j}(\cdot, \cdot)$ is modular, i.e., either strictly super- or submodular. If $\varphi_{j}$ is modular, then the “$\neq$” in (8) translate into ranked differences in the stimulus response coefficients between different capability levels, where the most extreme difference thus occurs between the off-diagonal types.

(H2) (Complete asymmetry) If for $j \in \{F, P\}$ $\varphi_{j}(\cdot, \cdot)$ is strictly supermodular (submodular) then $\alpha_{TE}^{T} > (<) \alpha_{SE}^{S}$ and $\alpha_{SE}^{S} > (<) \alpha_{TE}^{P}$.

Modularity is a rather weak assumption, meaning that a wide array of possible response patterns can occur. For example, if valuations are determined by (7) and $\varphi_{1}'(\cdot), \varphi_{2}'(\cdot) > 0$, then SE responds less sensitive to a change in $F$, and more sensitive to a change in $P$ compared to TE. Intuitively, this captures that SE underestimates the importance of the fundamental for the price determination, while TE underestimates the importance of the last observed price. The most serious case of miscalibrated mental models involves an inversion of the response coefficients between types, i.e. where $\text{sign}(\alpha_{j}(r_{j}^{*}, \cdot)) \neq \text{sign}(\alpha_{j}(r_{j}'^{*}, \cdot))$. Such an inversion is consistent with modularity. In particular, for $r_{P}^{P} < r_{P}'$, the inequalities $\alpha_{P}(r_{P}^{P}, P) < 0 < \alpha_{P}(r_{P}'^{P}, P)$ are possible if $\varphi_{P}(\cdot)$

\(^{20}\) Without the discipline of monotonicity, relative distortions and the underlying mental capabilities could affect the stimulus-response behavior in any arbitrary way.
is supermodular. In the binary capability model this would mean that SE responds positively, but TE negatively, to a price increase, which means that even with perfectly symmetric information the same observable is interpreted in an essentially contrariwise way by the different mental types.

2.2.2 Heterogeneous trading behavior

We now derive our core hypothesis about individual trading behavior from the above notion of mental modeling for a market featuring a price bubble in presence of all four mental types. Specifically, we show that in an asset market with a falling fundamental and a price bubble, the model predicts all four mental types to display distinguishable trading patterns over the various market phases.\(^{21}\) This provides us with a dynamic conjecture about the asset holding patterns and an aggregate conjecture about the distribution of trading gains across the four mental types.

**Asset-trading hypothesis** In the following derivation we consider valuations \(V^\theta\) given by (4'), where \(\theta \in \{FL, SE, SO, TE\}\), and the type-specific random variables \(\varepsilon_\theta\) are distributed according to an arbitrary joint density \(f(\varepsilon_{FL}, ..., \varepsilon_{SO}) > 0,^{22}\). We impose the monotonicity assumptions that both functions \(\varphi_j(r_j, x)\) are strictly supermodular and strictly increasing in \(x\) for \(r_j = 1.^{23}\) We restrict attention to supermodularity because this turns out to be the empirically relevant case. A mental type \(\theta\) is most likely to accumulate (more) assets if this type has the highest valuation \(V^\theta\) among all four types at a given point in time. A market with a falling fundamental and a price bubble has two natural phases with a characteristic pattern of the observables. The

---

\(^{21}\) While this is the most relevant case in the context of our SSW experiment, the explanatory power of our model is not limited to this scenario (see section 2.3).

\(^{22}\) For example, we could have \(\varepsilon_\theta = \overline{\varepsilon}_\theta + \varepsilon\), where \(\overline{\varepsilon}_\theta\) is a (possibly type-specific) constant and \(\varepsilon\) is an iid random variable. We do not require the \(\varepsilon_\theta\) to be independent.

\(^{23}\) The last assumption means that both the past price and the fundamental have predictive power, in term of the ideal mental model, for the asset’s future price. This is well known to hold for the SSW bubble market we study.
pre-peak phase features a decreasing fundamental jointly with an increasing price, while the post-peak phase is characterized by a falling fundamental and a decreasing price. Our model predicts how these two market phases with the characteristic patterns of the observables affect the likelihood that a certain mental type holds the highest valuation.

To see this intuitively, consider the pre-peak phase. If the fundamental falls while the price increases, the difference between SE and TE is that the former fails to sufficiently account for the falling fundamental (or even gets the effect the “wrong” way around) but responds adequately to the increasing price, while the latter gets the falling fundamental right but fails to adequately adjust for the increasing price. SO counterbalances an increasing price against a falling fundamental, while FL fails to sufficiently account for either of the two components. That is, the valuations of FL and SO are determined by offsetting forces, while $V_{SE}$ increases relative to $V_{TE}$.

Let $Pr(\theta|F, P) \equiv Pr(V^\theta = \max\{V^{FL}, V^{SE}, V^{TE}, V^{SO}\}|F, P)$ be the probability that type $\theta$ holds the highest valuation among all four mental types, given observables $F, P$. The following proposition states how these type-specific probabilities shift as the observables $F, P$ change.

**Proposition 1.** $Pr(\theta|F, P)$ depends on the observables $F, P$ as follows

| $Pr(\theta|F, P)$ | $\theta = FL$ | $\theta = SE$ | $\theta = TE$ | $\theta = SO$ |
|-------------------|---------------|---------------|---------------|---------------|
| $dF < 0, dP > 0$  | ?             | +             | $-$           | ?             |
| $dF < 0, dP < 0$  | +             | ?             | ?             | $-$           |

Table 1: Type-specific valuation patterns

A formal proof of Proposition 1 is in the appendix. The “?" in Table 1 means that the effects of $(F, P)$ are countervailing for the respective type, and therefore cannot be signed unambiguously. The general prediction in Proposition 1 is that the *off-diagonal types* TE and SE should display a mutually reversed valuation pattern if $F$ and $P$ move
in opposite directions, while the on-diagonal types FL and SO should display a mutually reversed valuation pattern if \( F \) and \( P \) move in the same direction. This qualitative result is independent of the joint distribution of the \( \varepsilon_i \), and driven by the characteristic relative distortion profile of the four mental types (8) jointly with the supermodularity of \( \varphi_j(r_j, x) \). More generally, the fact that the four mental types produce a distinguishable pattern (Table 1) is independent of the particular modularity assumption, while the specific pattern itself is not.\(^{24}\) Further, Table 1 easily extends to the cases where either \( dF > 0, dP < 0 \) or \( dF, dP > 0 \) by inverting the respective signs. Hence Proposition 1 would predict that if both fundamental and price are increasing (such as in a Bull market), valuations of SO tend to increase most as only this type adequately accounts for both observables.

Proposition 1 has direct implications for which mental types seek to acquire or sell assets in the two phases of a price bubble with a falling fundamental. Because traders try to buy or sell shares according to their subjective valuation (4'), the number of shares \( A^\theta \) held by type \( \theta \) correlates positively with \( Pr(\theta) \). Given that \( F(t) \) decreases in \( t \) and \( P(t) \) follows a hump-shape pattern, i.e. \( P(t) \) increases strictly for periods \( \{1, \ldots, \hat{t}\} \) and decreases strictly for periods \( \{\hat{t} + 1, \ldots, T\} \), our model predicts the following type-specific patterns:\(^{25}\)

\[
\begin{array}{cccc}
\text{H3) (Asset-trading hypothesis)}
\end{array}
\]

| \( t \leq \hat{t} \) | \( A^{FL} \) | \( A^{SE} \) | \( A^{TE} \) | \( A^{SO} \) \\
|---|---|---|---|
| \( t > \hat{t} \) | + | ? | ? | - \\

Table 2: Asset holding hypothesis

Hypothesis H3 predicts that the various types should accumulate or decumulate assets

---

\(^{24}\) For example, if both \( \varphi_j \) were submodular instead, the + and − in Table 1 would be inverted.

\(^{25}\) These predictions remain valid under the slightly weaker assumption that \( P(t) \) is non-decreasing in periods \( \leq \hat{t} \).
during the two market phases according to a differential pattern. The “?” in table 2 mean that the effects of \((F, P)\) are countervailing, and therefore we expect a flat or non-monotone pattern of \(A^\theta\) in the corresponding market phase. We predict SE to increase her shares throughout the build-up of the bubble. These shares most likely will be bought from TE, because this type seeks to sell assets in this market phase. Likewise, FL should accumulate assets after the crash while SO unloads. We provide the intuition for Hypothesis H3 for the pre-peak phase \((t < \hat{t})\). A similar reasoning applies for the crash phase. The argument is that the evolution of \((F, P)\) has a differential effect on the valuations of the four mental types, making it more likely that a certain type has the lowest (highest) valuation during a market phase. These types then are most likely to become net sellers (buyers) during that phase. On the one side, the falling fundamental means that assets loose value over time, making selling gradually more attractive. On the other side, an increasing price suggests that some people have a different valuation for the asset than captured by the fundamental, which makes it profitable to acquire or keep assets to sell them later. Appropriately resolving the tension between the falling fundamental and an increasing price needs both mental capabilities according to our model. With a (weakly) increasing price, the falling fundamental is the only channel that pushes traders towards selling during this market phase. TE and SO both account for the fundamental in the same way (Hypothesis H1). Compared to TE, however, SO has a refined intuition about the other valuations as expressed in the observed price, which counterbalances the incentive to divest in this market phase. FL and SE both fail to sufficiently account for the falling fundamental, leaving these types with higher valuations compared to TE. Thus, the peculiarity of TE’s mental profile makes this type most likely to have the lowest valuation of all types, and therefore TE should be the only type that systematically seeks to divest during this market phase. Likewise, SE is the only type that centers her valuations squarely around the observed price, making it most likely that SE acquires assets in this phase.
Trading-gains hypothesis  The above difference in the asset holding patterns translate into differential aggregate trading gains for the four types. SO should achieve the highest trading gains, because this mental type has the best ability to calibrate her mind to the asset market. Given that $P(t)$ is hump-shaped in a bubble market, $V^{SO} > V^{TE}$ becomes increasingly likely towards the max of $P(t)$, meaning that TE tends to (prematurely) divest compared to SO. While the earlier exit prevents TE from incurring trading losses, it also means that TE tends to forgo some of the gains which SO realizes. Therefore, TE should realize non-negative trading gains, which are below those of SO. Likewise, the chances of SE to overvalue the asset relative to SO increases towards the bubble peak. Thus, SE buys or holds on to shares at high prices, which she later cannot sell, or only at substantial losses. Therefore, SE is bound to incur trading losses. One could prematurely conjecture that FL should incur the highest trading losses, because this type lacks both mental capabilities. This intuition is flawed, however, according to our model. Table 2 shows that if $dF < 0, dP > 0$ it becomes more likely that SE, rather than FL, acquires expensive shares. The reason is that the two biases of FL tend to offset, rather than reinforce, each other in the bubble market considered here. On the one hand, FL might accumulate shares which she later cannot sell profitably because she fails to account for the fundamental, hence FL tends towards lower trading gains than TE. On the other hand, FL incurs smaller losses than SE because FL tends to value the asset lower than SE at high prices given her lack of mentalizing capability. It follows that the trading gains of FL should be between TE and SE. We summarize these predictions in the following hypothesis:

\textbf{(H4) (Trading-gains hypothesis)} The trading gains $G_{\theta}, \theta \in \{FL, SE, SO, TE\}$, are distributed over the four mental types as follows.

1. $G_{SO} > G_{TE}, G_{FL}, G_{SE}$
2. $G_{SE} < G_{TE}, G_{FL}, G_{SE}$
3. \( G_{TE} > G_{FL} \)

An important conclusion from hypothesis H4 is that mental capabilities can have a non-monotonic effect on type-specific trading gains. The highest gains and highest losses both are assigned to types with high mentalizing capability. Given a strong mentalizing, accordingly, analytical skill alone is decisive for whether someone wins or loses most. This follows because the ability to correctly anticipate a systematic (upward) deviation of the price from the fundamental value requires a strong ability to mentalize that others may follow such a tendency, which neither FL nor TE possess. This lack of skill implies that FL and TE have more difficulty to profitably follow a price trend, which likewise protects them from trading losses due to asset overaccumulation. Such a misjudgment is the reason why SE makes the highest losses. Thus, not having strong mentalizing capabilities is better once one also is analytically weak.

2.3 Main assumptions: A critical review

Do mental capabilities matter? Our most basic theoretical presumption is that i) mental capabilities can differ across traders, and ii) mental capabilities matter differentially for trading behavior. Whether these assertions are true is an empirical question. The first assumption is principally verifiable with an appropriate elicitation method, and our experimental design comprises of such a procedure by using different established measures. The second assumption could be invalidated by the data if no evidence for any systematic behavioral differences of the four mental types were uncovered.\(^{26}\)

On non-convertibility The bijectivity between mental profiles, distortion profiles and trader types (see Figure 1) follows from the non-convertibility assumption. We now illustrate that a weak form of convertibility is consistent with most of our predictions, while our results break down if mental capabilities are one-to-one convertibility.

\(^{26}\) Note, however, that our model would also predict that all four types display the same behavior if the asset market were cognitively so simple that differences in mental capabilities play no role.
Non-convertibility is violated if mental capabilities can be at least partially substituted, meaning that, e.g., a strong analytical capability can make up for a weak mentalizing capability in determining the “right weight” of the observed price in the mental model. Formally, non-convertibility fails if
\[ r_j = r_j(c_F, c_P) \text{ with } \frac{\partial r_j}{\partial c_F}, \frac{\partial r_j}{\partial c_P} > 0. \]
In such a case, Hypothesis H1 should fail, making H1 a crucial part of the evidence that could speak in favor of non-convertibility. The remaining hypothesis are robust to moderate violations of non-convertibility. Technically, our predictions need that the four mental types map into four distinct distortion profiles in a way that preserves the partial order induced by the mental profiles. To be specific, let \( c_A', c_A, c_F', c_F \) with \( c_A' > c_A \) and \( c_F' > c_F \) be the four mental capabilities that constitute the four types in the binary capability model. A sufficient condition for the four mental profiles to induce four distinct distortion profiles, consistent with the case of non-convertibility, is that
\[ r_F(c_A', x') > r_F(c_A, x), \quad x, x' \in \{c_M, c_M'\}, \]
and
\[ r_P(y', c_M') > r_P(y, c_M), \quad y, y' \in \{c_A, c_A'\}. \]
Clearly, these inequalities include non-convertibility as a special case, but they only require that convertibility effects are dominated. Hypothesis H2 and Proposition 1 remain valid if convertibility is bounded in terms of the above inequalities.\(^{27}\) Therefore, we would expect to see the same type-specific trading dynamics (Hypothesis H3) and a similar distribution of trading gains (Hypothesis H4) as predicted by strict non-convertibility. However, Hypothesis H1 - H2 would fail under strong forms of convertibility. In particular, if only average capability \((c_A + c_M)/2\) or maximal capacity \(\max\{c_A, c_M\}\) matters, then TE and SE would have the same distortion coefficients. This means that the off-diagonal types should display an indistinguishable trading pattern, contrary to the predictions under non-convertibility. Likewise, if one capability is irrelevant for both distortion coefficients, then at least two types should be indistinguishable.\(^{28}\) In sum, these arguments show that

\(^{27}\) To see the latter, note that if \(d_F < 0\) and \(d_P > 0\) the mental type TE still underestimates the importance (possibly at a smaller magnitude). Therefore, TE still is most likely to hold the lowest valuation, and SE the highest valuation, with this change of observables.

\(^{28}\) For example, if only analytical reasoning mattered and \(\phi_2(\cdot) = k\), a constant, then \(V^{TE}\) and \(V^{SO}\) should behave identically.

---

23
if the trading behavior of at least two mental types cannot be distinguished, the data would falsify the (relaxed) non-convertibility presumption.

The above arguments clarify the decisive character of non-convertibility for our predictions, but likewise reveal a critical shortcoming of simplified approaches, should non-convertibility (possibly in its weaker form) indeed be appropriate. If non-convertibility applies, then the corresponding trading behavior and asset-holding patterns, cannot be explained by a one-dimensional measure of mental capabilities, because such a measure could not possibly span the entire “box” of mental types. Such an attempt would therefore mix the types and produce biased estimates of how mental capabilities shape the trading behavior.

There is another, more subtle, reason why one could fail to empirically distinguish between the trading behavior of certain mental types, even despite the validity of non-convertibility. This may happen if mental capabilities are strongly correlated at the population level. To see this, consider the extreme where $c_A$ and $c_M$ are perfectly correlated. In this case only two instead of four mental profiles would exist. More generally, correlation implies that if someone gets it right, say, in the analytical dimension, this is also predictive of whether she gets it right in the mentalizing dimension. In this sense, a strong analytical capability may affect the valuation of the observed price in a purely statistical way, despite the possible validity of non-convertibility.\footnote{Our experimental procedure allows us to check the correlation between mental capabilities, and we uncover at most a weak level of correlation, making correlated capabilities a less relevant concern.}

**Risk attitudes** Our model predicts that different mental profiles yield different trading patterns as a consequence of a systematic bias in the evaluation of observable information of the respective type. It is well conceivable that other individual characteristics influence valuations besides mental capabilities. A natural candidate are heterogeneous risk preferences. In terms of (4’), idiosyncratic risk attitudes could affect the two $\tilde{\varphi}$-functions. Our main predictions remain valid if risk attitudes are independent from mental capa-
bilities. The experimental approach allows us to control for risk preferences, where we shall find that risk attitudes are at most weakly related to mental capabilities.

A comment on the “ideal model” Our notion of an “ideal model” is akin to a Laplacian demon, an omniscient observer who accurately manages to predict the future market price on the basis of (5). On intuitive grounds, the ideal model encodes the best forecast of an asset’s value on the basis of what all traders think in a market (including their possible mistakes). That is, the ideal model should reflect, inter alia, the set of traders in a given market. Specifically, if all traders were perfect clones, then their mental model would coincide with the ideal model, and there would be no systematic heterogeneity in the asset’s valuations, and neither in the observable trading behavior. For this reason, heterogeneity in the individual thinking shapes the ideal model, from which then individual deviations have the interpretation of our distortions.\textsuperscript{30}

Non-bubbly markets Our model of mental types can be applied to settings other than the SSW bubble market. In this respect, one important observation is that different mental types need not always behave differently, other than predicted in the SSW market. As an example, consider a situation where the asset price is determined entirely by its fundamental $F(t)$, i.e. $P(t) = P(F(t))$, consistent with the efficient market hypothesis.

Since the past asset price then does not deviate from $F(t)$ it follows that $\tilde{\varphi}_P(\cdot, P(t)) = 0$ in (5). If the price has no separate effect on the valuation, then (ignoring $\varepsilon_i$) $V^{SO} = V^{TE} = \varphi_F(1, F(t))$ and $V^{SE} = V^{FL} = \varphi_F(r_F^{SE}, F(t))$. That is, the four mental types effectively yield only two behavioral types, where analytical capability alone induces a distinction. Moreover, if additionally $\theta_F \geq \tilde{c}_F$ already for $\theta = SE, FL$, that is, everybody’s analytical capacity is sufficient to understand the fundamental, then all four mental types should form essentially the same mental model, congruent to the ideal model, resulting in a

\textsuperscript{30} We seek to address the formal relation between the ideal model and how it may vary with the composition of mental types in future work.
single behavioral type. In other words, our theory does not predict that the four mental
types will *always* show a distinct behavior.

3 Experimental Design

The laboratory has several advantages for testing our hypotheses. First, it allows us
to measure independently each individual’s mental capacities through different, incen-
tivized tasks. Then, we can observe the same individual’s behavior in an experimental
asset market. The experimental asset market gives us control over the market environ-
ment. In particular, we can restrict trading to a single asset, and we can control the
fundamental value of this asset, which is determined by a simple stochastic dividend
process. Importantly, this prevents any form of privileged trading through asymmetric
information about the fundamental value, as all participants are equally informed
about the dividend-generating process. We can assure with monetary incentives that
every participant has the same objective, namely to maximize her cash at the end of
the final period of the asset market. Finally, since experimental asset markets like ours
reliably produce bubbles, we can relate observed trading behavior to someone’s mental
capabilities.

3.1 Procedure

We conducted 8 experimental sessions with 32 participants each at the laboratory of
the economics department at the University of Zurich.\textsuperscript{31} All experimental sessions, con-
sisting of two phases, began with participants being seated in front of their computer
terminals, and then receiving general information about the procedures.\textsuperscript{32} Then, partici-
pants started with Phase 1 of the experiment, where each participant completed a series

\textsuperscript{31} We used hroot (Bock, Baetge, and Nicklisch 2014) for recruitment and ztree (Fischbacher 2007) for
conducting the experiment.

\textsuperscript{32} See online material (Hefti, Heinke, and Schneider 2016).
of tasks designed to obtain performance-based measures of the analytical and mentaliz-
ing capacities. All tasks were incentivized, yielding CHF0.30 per correct answer/round
won. Except for the Game of Nim, participants did not receive feedback about their per-
formance before the end of the experiment. Phase 1 (capability measurement) of each
session took about 45 minutes. Phase 2 consisted of an experimental financial market
where participants could trade shares of an asset against cash. Each session had two
markets with 16 participants each. We provided participants with detailed paper in-
structions, which were read aloud in front of all participants.33 Participants then had to
answer comprehension questions.34 The asset market did not start before all participants
had answered all questions correctly. Finally, we implemented two payoff-irrelevant prac-
tice periods to make participants more familiar with the computer interface. After the
practice rounds, the 15 actual trading periods ensued. The income from phase 2 con-
sisted of the amount in their cash account at the end of the asset market. Shares of the
asset became worthless at that point. Phase 2 took on average a little over 90 minutes.
After Phase 2, we administered the exit questionnaire and paid out participants in cash
as they left the laboratory.35 One entire session lasted about 2.5 hours. The earnings for
the entire session were the sum of the earnings for each task in phase 1, the cash holdings
at the end of period 15 in the asset market plus a show-up fee of CHF 10. On average
subjects earned around CHF 70 (minimum CHF 23, maximum CHF 121).36

3.2 Phase 1: Measuring Mental Capabilities

In phase 1, instructions for each of the tasks were presented on participants’ computer
screens, before the task started. Each task was designed to capture a specific aspect of
either the analytical or the mentalizing ability dimension.37 For each of the two capacity

33. See online material (Hefti, Heinke, and Schneider 2016).
34. For details see online material (Hefti, Heinke, and Schneider 2016).
35. The questionnaire is available from the authors on request.
36. For details see online material.
37. During this phase, we additionally elicited risk attitude, using a standard Holt-Laury-type price
list. The lottery choice was fixed to a 50 : 50 chance of winning either CHF 20 or nothing, and the
dimensions, the performance measure ranges from 0 (all items incorrect) to 100 (all items correct). That is, each subject has two scores, one for analytical and one for mentalizing ability, each of which can range from 0 (lowest ability in that dimension) to 100 (highest ability in that dimension).

For the analytical dimension, we chose three frequently used experimental tasks that reflect general intelligence (Raven’s Progressive Matrices), mathematical and logical skill (SAT-style word problems, Bruguier, Quartz, and Bossaerts 2010), and strategic reasoning (Game of Nim, McKinney Jr and Van Huyck 2006).38

Mentalizing is the ability to make accurate inferences about the mental states of others. This requires at least two abilities, (1) to recognize and identify others’ intentions (“perspective taking”) and (2) develop a correct working model about the resulting behavior (“online simulation”) (Reniers et al. 2011). Incentivized tasks assessing these abilities are less ubiquitous than for the analytical dimension.39 We followed Bruguier, Quartz, and Bossaerts (2010) and operationalized the mentalizing ability with two separate tests.40 In the first one, participants had to infer other people’s mental states (reading the mind in the eye test), in the other they had to infer intentions from others’ actions (Heider-Simmel Test).41

certain payment moved upward from CHF 0 in increments of CHF 1.
38. For a more detailed description of each task see online material (Hefti, Heinke, and Schneider 2016).
39. Psychologists traditionally use self-reported measures, such as questionnaires (e.g., Reniers et al. 2011).
40. For a more detailed description of each task see online material (Hefti, Heinke, and Schneider 2016). Both performance in the Reading the mind in the eye test and the Heider-Simmel Task correlate positively with the ability to forecast price changes when insiders are present in the experimental asset market (Bruguier, Quartz, and Bossaerts 2010). This indicates that participants scoring high in these tasks are able to detect when price movements are due to intentional actions by insiders.
41. Mentalizing capabilities may be important in asset trading for several reasons. First, a market is a collective of individual traders, and the mental states and intentions determine these traders’ actions. Second, as Bossaerts (2016) suggest, traders may treat a financial market as if it was itself an intentional entity.
Table 3: Summary Statistics by Mental Type

<table>
<thead>
<tr>
<th>Mental Type</th>
<th>Participants</th>
<th>Women (%)</th>
<th>Age (years)</th>
<th>Av. risky choices</th>
</tr>
</thead>
<tbody>
<tr>
<td>Featureless (FL)</td>
<td>166</td>
<td>59.6</td>
<td>23.5</td>
<td>11.6</td>
</tr>
<tr>
<td>Semiotic (SE)</td>
<td>155</td>
<td>66.5</td>
<td>23.4</td>
<td>11.1</td>
</tr>
<tr>
<td>Technocratic (TE)</td>
<td>155</td>
<td>29.7</td>
<td>22.8</td>
<td>12.1</td>
</tr>
<tr>
<td>Sophisticated (SO)</td>
<td>164</td>
<td>43.3</td>
<td>23.0</td>
<td>12.3</td>
</tr>
<tr>
<td>Total</td>
<td>640</td>
<td>49.8</td>
<td>23.1</td>
<td>11.8</td>
</tr>
</tbody>
</table>

Age and risk attitude are similar across mental types. Men tend to score higher in the A dimension, resulting in a gender imbalance across skill types.
Table 4: Correlation Matrix for individual mental measures

<table>
<thead>
<tr>
<th></th>
<th>Raven’s Test</th>
<th>Game of Nim</th>
<th>Word Problems</th>
<th>Eye Gaze</th>
<th>Heider Simmel Task</th>
<th>Av. risky choices</th>
<th>A-Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Game of Nim</td>
<td>0.3079***</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Word Problems</td>
<td>0.3477***</td>
<td>0.354***</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Eye Gaze</td>
<td>-0.013</td>
<td>-0.007</td>
<td>0.003</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Heider Simmel Task</td>
<td>0.039***</td>
<td>0.027***</td>
<td>0.012</td>
<td>0.051***</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Av. risky choices</td>
<td>0.024**</td>
<td>0.113***</td>
<td>0.141***</td>
<td>0.013</td>
<td>-0.05***</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>A-Measure</td>
<td>0.398***</td>
<td>0.831***</td>
<td>0.814***</td>
<td>-0.03</td>
<td>0.1537***</td>
<td>0.154***</td>
<td>-</td>
</tr>
<tr>
<td>M-Measure</td>
<td>0.058***</td>
<td>0.070***</td>
<td>0.068***</td>
<td>0.486**</td>
<td>0.755***</td>
<td>0.124</td>
<td>0.089***</td>
</tr>
</tbody>
</table>

Significance levels: * p<0.1, ** p<0.05, *** p<0.01.
N=640 for all pairwise correlations.
To obtain a good approximation of the underlying distributions of the two measures, we use data from 640 participants. Figure 2 displays the joint distribution of the two dimensions.

Figure 2: Distribution of A and M performance measures

Each dot represents a participant. The horizontal axis shows the participant’s performance in the M dimension, the vertical axis the performance in the A dimension. The dot cloud suggests little relation between the two measures, as highlighted by the red line of best linear fit. The black lines indicate the median for each measure.

The red line shows the linear fit. The corresponding correlation coefficient $\rho = 0.099$ suggests a weak positive correlation between the two measures ($p = 0.012$, $N = 640$). The black horizontal and vertical lines indicate the median for each dimension. We use the resulting four quadrants to classify our participants into our four mental types, Featureless (FL), Semiotic (SE), Technocratic (TE), and Sophisticated (SO). Table 3 lists the basic summary statistics of the total group and by mental type.

The participants in each skill category are similar with respect to age, with a mean

---

42. The data come from 40 experimental asset markets including two other treatment conditions that are part of a separate paper. In these conditions, we exogenously create markets that consist of more extreme types.

43. A factor analysis confirms that the five measures can be grouped into two factors, and that one factor loads exclusively with the analytical measures while the other exclusively loads with the mentalizing factors (see appendix).
age around 23 years in all four categories, and we cannot reject the null hypothesis of no contingency (Pearson $\chi^2$ test, $p = 0.207, N = 640$). The four groups are less balanced with respect to gender ($\chi^2$ test, $p < 0.01, N = 640$) because women and men tend to perform differently on the A dimension.\footnote{On average, women score 54 points out of 100, and men score about 63 points (t-test, $p < 0.01$, $N = 640$). On the other hand, women tend to perform slightly better on the M scale, where women score on average 57 points, and men 56 points (t-test, $p = 0.05$, $N = 640$).} We also elicited participants’ risk preferences using a Holt-Laury-type choice task (Holt and Laury 2002) between a lottery that yields either CHF20 or CHF0 with equal probability and a certain amount. The last column of Table 3 shows the average number of times that a participant chose the lottery over the certain amount. The differences across mental types are small. We cannot reject the null hypothesis of no contingency ($\chi^2$ test, $p = 0.214, N = 640$). This suggests that risk attitudes and mental capabilities are independent human traits.

3.3 Phase 2: Experimental Asset Market

We used a call market version of the Smith-Suchanek-Williams asset market (Smith, Suchanek, and Williams 1988).\footnote{We implemented a slightly modified version of the call market from the GIMS program for asset market experiments in ztree (Palan 2015).} In each market, 16 participants were endowed with cash and shares of an asset.\footnote{A market had 40 assets in circulation and each participant was randomly endowed with one of four possible portfolios. The portfolio consisted of either 1, 2, 3, or 4 shares and a correspondingly decreasing amount of cash (average portfolio: 2.5 shares and 1820 Rappen in cash). 100 Rappen = 1 Swiss Franc, with 1 Swiss Franc \approx USD 0.98 at the time of the experiment.} The asset market was divided into 15 periods. Each period had a trading phase, where participants could trade shares against cash, followed by a dividend phase, where the asset payed a randomly drawn dividend.\footnote{Shares of the asset have no intrinsic value beyond the dividend stream. The dividend is drawn from the set $d \in \{0, 8, 28, 60\}$, each equal probability. The two randomly generated dividend streams used in the study were 60, 0, 0, 60, 8, 28, 60, 0, 8, 28, 60, 0, 8, 0, 28, 60, 0, 28, 8, 60, 28, 8, 60, 0, 28, 8, 60, 0, 28, 8, 60, 0, 28 for the first four sessions and 8, 8, 8, 0, 28, 60, 0, 0, 28, 60, 0, 8, 28, 60, 0, 28, 60, 0, 0, 28, 8, 60, 0, 28, 8, 60, 0, 28 for the remaining sessions.} In each period participants could trade shares by submitting one sell order and one buy order.\footnote{The online material (Hefti, Heinke, and Schneider 2016) shows the trading screen.} A buy order consisted of the maximum price that a participant is willing to pay for a share, and the number of shares that the participant is willing to buy if the
market price is equal or lower than this maximum. Conversely, a sell order consisted of the minimum price at which a participant is willing to sell a share, and the number of shares that the participant is willing to sell if the price is equal or higher than this minimum. The computer automatically collected all buy and sell orders and calculated the market-clearing equilibrium price.

For our experiment, the call market offers multiple advantages over a double auction. First, our model assumes that market participants have essentially no influence on the market price. We chose the call market procedure and the market size (16 participants is a large number compared to most other laboratory asset markets) to make it harder to make manipulative orders (Baghestanian et al. 2014). Second, the call market is the more conservative choice to test our hypotheses. Finally, compared to the double auction the call-market is faster, which helped to keep the total session duration manageable.

3.3.1 Market Price and Order Volumes

The average pattern of the market price and order volumes over the 15 periods are shown in figure 3. Due to the expected dividend payment of 24 Rappen in each period, in the first period the expected value of the asset is 360 Rappen and it decreases by 24 Rappen in each subsequent period leading to a declining fundamental value (solid grey line). The lowest expected dividend earning from holding an asset is zero over all periods. The highest expected dividend earning from holding an asset is 60 Rappen in each period with

49. Participants could leave their buy/sell order blank, in which case they would not buy/sell shares in this period. The computer did not accept orders that violate a participant’s budget constraints, that is, buying on credit and short selling are not allowed. Participants could only buy whole shares but no fractions of shares.

50. Because call markets give fewer opportunities for making offers and trading shares, they exhibit fewer price mirages (over-reactions to uninformative offers and trades), are in general closer to the rational expectation equilibrium, and have less trading price volatility than double auctions. In addition, the removal of within-period trading dynamics in call markets reduces the opportunities both for speculative trading (Baghestanian et al. 2014) and for learning about others’ trading strategies. Thus, a call market reduces the amount of offers and the size of the asset bubble compared to a double auction, which should limit the downside risk of trend followers and the upside scope for speculation by bubble riders. All effects on the M-dimension we find in a call-market environment should be at least as pronounced in a double auction.
a dividend payment. Thus, it starts at 900 Rappen in the first period and it declines by 60 Rappen in each period (dashed grey line). As discussed in section ?? given the behavior of our types, we expect that the market price stays above the downward trending fundamental value.

![Figure 3: Average Market Price (Baseline)](image)
The solid black line shows the evolution of the average market price across all baseline markets. The solid grey line shows the expected value and the dashed line shows the highest possible earnings from the dividends over period 1 to 15.

As is usual in this type of experimental asset markets, the average market price starts below the expected value of 360 (Palan 2013), here around 337 Rappen.\(^51\) Subsequently, the price increases until period 6 or 7; one market already peaked in period 1 (minimum) another in period 12 (maximum). Then, between period 7 and 12 the price stagnates or slightly declines, and few shares are traded. Finally, it crashes towards the end of the asset market. Thus, the deviation of the price from the expected value (i.e. the bubble component) is negative at the beginning and continuously increases afterwards until it bursts, see the grey dashed line in figure ??.. Due to the stagnating prices during the middle part, the bubble component is largest between period 11 and 12 - where in one market the bubble peaked in period 9 (min) and in another in period 14 (max). In the following, we refer to the period with the largest deviation of the market price from the

\(^{51}\) We will refer to the expected value of the dividend stream as the fundamental later in our analysis of the results.
expected value as the peak period. We provide a detailed graph of the 16 individual markets in the appendix.

The market price tracks closely the sell order prices. The number of assets demanded on the buy side (100–200 per period in an average market) substantially exceeds the number of assets offered to sell (4-25 per period in an average market). Moreover, while the number of assets demanded by buy offers stays almost the same over all periods, the number offered assets to sell declines from 25 to 4 towards the middle periods 6-9, till it increases again to 30 towards the end, which leads to a u-shaped pattern in the number of transacted assets itself.

4 Results

4.1 Trading Gains and Total Income

We start by examining trading gains because this outcome focuses on traders’ ability to speculate successfully on price changes. It also abstracts from profits made through random dividend draws. After the analysis of trading gains, we will investigate total income from both trading and dividends.

A central prediction of our theory is that the four mental types earn different trading gains, and we begin the empirical analysis by testing the trading gains hypothesis H4.52

Result 1 (Mental Types and Trading gains). SO realizes the highest trading gains (and overall income), SE the highest trading losses (and lowest overall income). By pairwise comparison, all mental types earn statistically different total income and trading gains, as predicted by hypothesis H4, except for the difference between TE and FL, which is not statistically significant. Therefore, including only one capability masks important heterogeneity among traders.

52. In section 3.2 we already found that risk attitudes are largely uncorrelated with mental capabilities. Consistent with this result, all our results remain valid when we control for risk attitudes.
Figure 4a depicts the average trading gains for the four mental types by the end of the experiment.\textsuperscript{53} The figure shows that the trading gains are substantially different for the four mental types. In particular, the order of the trading gains ($G_{SO} > G_{TE} > G_{FL} > G_{SE}$) in the figure is the one we predicted in hypothesis H4. While SE loses 390 Rappen on average, SO gains 403 Rappen over the course of the asset market. FL (45 Rappen loss) TE (93 Rappen gain) lie in between.

Figure 4: Trading gains across mental capability types.

(a) Difference in Trading Gains

(b) Difference in Income

The vertical axis shows the outcome variables trading income and cash at the end of the asset market. The right panel shows the deviation from average cash holdings in period 15, by type. SE earns the least amount, 177 Rappen below average, and SO the highest amount, 181 Rappen above average. The two other types fall in between, with FL slightly below average and TE a bit above average.

Subjects’ goal in the experimental asset market is to maximize their cash at the end of the final period. Subjects can earn cash either through dividends on shares they hold or through trading assets (buying shares at a low price and selling them at a high price). Figure 4b shows the incomes of the four types from the asset markets. We see the same order as in Figure 4a.\textsuperscript{54} In particular, Sophisticates earn the most of all four types, 358

\textsuperscript{53} Cash at the end of a period, $t$, comes from three different sources: initial cash, cumulative dividends from periods 1 to $t$, and cumulative trading gains from sales and purchases in periods 1 to $t$. Therefore, the cumulative trading gain in a period is simply the residual if we subtract initial cash and cumulative dividend income from current cash.

\textsuperscript{54} A subject’s income in the asset market is equal to the amount of cash at the end of the final period.
Rappen more than SEs, who have the worst outcome.\textsuperscript{55} Moreover, comparing figures
Figure 4a and 4b reveals that the difference in total income is driven by the difference in trading gains, and not by the dividend income.\textsuperscript{56}

A regression analysis confirms the impression from figure 4. The results of OLS regressions of the two outcome measures, cash and trading gains, on type dummies (FL is the omitted category) are presented in Table 5. The regression framework allows us to additionally control for risk aversion (number of lottery choices in the multiple price list), and to correct standard errors for clustering at the session level. The Sophisticated type outperforms all three other types in terms of final cash (first column). Conversely, the Semiotic type performs worse than all other types. Risk attitudes do not change this result. Table 5 provides statistical evidence for hypothesis H4.\textsuperscript{57} The only problematic comparison is between FL and TE, where we cannot reject the possibility these two types realize the same trading gains and total income. But this does not mean that FL and TE are empirically indistinguishable types. Below we show that the similar profits of FL and TE are the results of two different behavioral patterns.

The results for trading gains (second column) are qualitatively the same, only with greater magnitudes and smaller $p$-values.\textsuperscript{58} Risk aversion does not seem to play an important role.\textsuperscript{59}

Figure 4 and Table 5 imply that focusing on only \textit{one} capability measure would

\textsuperscript{55} We also find suggestive evidence that performance is most consistent among the Sophisticates and least consistent among Semiotic types, who have a 30 percent higher standard deviation in final cash (SO: 658 Rappen, TE: 682 Rappen, FL: 721, SE: 853 Rappen).

\textsuperscript{56} Dividend income tends to mitigate the difference between SOs and SEs because SEs hold more shares on average than SOs.

\textsuperscript{57} Because hypothesis H4 is of a one-sided nature, we could divide the standard errors by 2, showing that all pairwise hypotheses are as stated in H4 at least at the 5\% level, except the comparison between TE and FL.

\textsuperscript{58} In this paper, we focus our analysis on average effects. Type composition and market dynamics differ across sessions and may influence the success of the different types. However, our results are robust to using relative instead of absolute outcome measures, namely the within-session cash/trading gains rank of a subject or within-session standardized cash/trading gains.

\textsuperscript{59} Recall from section 3.2 that risk aversion is independent from mental capabilities. Furthermore, we obtain the same qualitative results when excluding those ambiguous subjects who score close to the medians of the two mental dimensions.
Table 5: Regression analysis of asset market outcomes across mental types

<table>
<thead>
<tr>
<th></th>
<th>Cash</th>
<th>Trading Gain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Semiotic</td>
<td>-169.426*</td>
<td>-349.796***</td>
</tr>
<tr>
<td></td>
<td>(101.520)</td>
<td>(118.288)</td>
</tr>
<tr>
<td>Technocratic</td>
<td>30.592</td>
<td>136.603</td>
</tr>
<tr>
<td></td>
<td>(68.473)</td>
<td>(166.107)</td>
</tr>
<tr>
<td>Sophisticated</td>
<td>188.896**</td>
<td>447.873***</td>
</tr>
<tr>
<td></td>
<td>(83.541)</td>
<td>(167.741)</td>
</tr>
<tr>
<td># Lottery choices</td>
<td>-0.249</td>
<td>-5.656</td>
</tr>
<tr>
<td></td>
<td>(9.609)</td>
<td>(15.838)</td>
</tr>
<tr>
<td>Constant</td>
<td>2840.468***</td>
<td>22.535</td>
</tr>
<tr>
<td></td>
<td>(107.861)</td>
<td>(206.381)</td>
</tr>
</tbody>
</table>

adj. $R^2$ 0.011 0.037
N 256 256
Clusters 16 16

Type comparisons:

SE=TE $\chi^2 = 3.40$ $\chi^2 = 6.40$
$p = 0.065$ $p = 0.011$

SO=TE $\chi^2 = 2.73$ $\chi^2 = 2.96$
$p = 0.099$ $p = 0.085$

SO=SE $\chi^2 = 14.12$ $\chi^2 = 23.45$
$p < 0.001$ $p < 0.001$

OLS regressions, bootstrapped standard errors in parentheses, 1000 repetitions, adjusted for clustering at the session level. Unit of observation: participant. Significance levels: * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Dependent variables: Cash and trading gains for entire asset market, in Rappen.
Independent variables: Constant: Featureless type. “Semiotic,” “Technocratic,” “Sophisticated”: dummies for mental type; # Lottery choices: number of times a participant chose the lottery over the certain amount in the Holt-Laury task.
produce a substantially biased picture by masking the interaction of the capabilities, consistent with our theory.\footnote{See appendix 6.2 for the split along one dimension.} In particular, we would wrongly conclude that high analytical capability assures higher trading gains (and total income) if we ignored mentalizing capability.\footnote{A simple regression of the A-capability, where we either use our effective measure of the A-capability or a capability dummy as a right-hand variable, shows that trading gains and total income are significantly increasing in the A-dimension.} In reality, this result is entirely driven by the high earnings of SO; the earnings of TE are not significantly different from zero. Mentalizing ability is thus crucial for the success of SO. Similarly, if we ignored analytical reasoning, we would wrongly conclude that mentalizing is entirely irrelevant in the asset market: As Figure 4 shows, this is because the gains and losses of SO and SE (TE and FL) tend to offset each other.\footnote{A simple regression of the M-capability reveals no significant effect of mentalizing on trading gains and total income.} Hence we would miss that mentalizing ability is responsible for both the highest gains and the largest losses.

### 4.2 Trading Patterns

Our theory predicts the four mental types to display distinct, characteristic trading patterns over the various market phases. Specifically, individual asset holdings should show a type-specific stimulus-response to a change in the observables “fundamental” and “last price” (hypothesis H2-H1). As a consequence of the respective stimulus-response equation (4’), the asset holdings of the four types should obey specific portfolio dynamics over the course of the experiment (hypothesis H3). We now test the hypothesis about the type-specific stimulus-response coefficients and the portfolio dynamics hypothesis, starting with the latter.

#### 4.2.1 Portfolio dynamics

**Result 2** (Portfolio dynamics). *We find empirical trading patterns that are strongly consistent with our model (hypothesis H3). While FL trades little, SE strongly increases...*
its position while the bubble builds up, and holds the highest amount of shares even beyond the peak of the bubble. TE divests before the price peak, while SO is the only type who profitably divests after the bubble bursts.

Figure 5: Asset holdings and cumulative trading gains over time, by type.

The left graph shows the different portfolio strategies of the four types, by plotting the asset holdings for each type over time. The right graph shows the resulting trading gains. All four types show distinct portfolio dynamics.

Figure 5 depicts average asset holdings over time for the four mental types, while the right figure shows how the trading gains develop over time. The figure shows a substantial difference in how the asset holdings of the four mental types evolve over time, once both observables become available.63

Of all types, FL has the flattest asset holding pattern, with a slight upwards tendency in pre-price-peak phase, which is broadly consistent with our observation from Section 2.2.2 that the two relative biases of FL tend to offset each other. SE is the only

63. In the first period there is no past asset price. At the end of the first period, we cannot reject the hypothesis that all four types hold the same number of shares (Kruskal-Wallis, $p = 0.727$, $N = 256$).
type who accumulates assets during the entire market. Both TE and SO tend to reduce their holdings, albeit in different market phases.

To test hypothesis H3, we analyze subjects’ behavior separately for the periods before the market price reaches its peak (“pre-price-peak”) and after the peak (post-price-peak). On average, prices reach their peak between period 6 and 7. Therefore, we fit two linear regressions, one for the pre-price-peak market phase and one for the post-price-peak market phase (Table 6).

The results support that the portfolio dynamics of the four mental types are consistent with hypothesis H3. First, in the pre-price-peak phase, SE accumulates while TE divests over time. Both FL and SO do not seem to systematically vary their asset holdings during this market phase. Second, in the post-price-peak phase, FL acquires shares while SO sells shares over time. In contrast to the pre-phase, TE and SE now do not seem to vary their asset holdings as time progresses, and we cannot reject the hypothesis that SE and TE behave similarly during this market phase.

The regressions in Table 6 together with Figure 5 provide additional insight into the source of the different final trading gains. We see that the accumulative tendency of FL during the post-phase tends to burden this type with some trading losses because FL acquires assets while the price is still falling. Further, SEs incur most of the trading losses due to the shares bought in the pre-phase, which they then cannot sell in the post-phase. Finally, SO achieves the highest trading gains especially because this type is the most successful seller in the post-bubble phase, where he outperforms TE.

4.2.2 Stimulus response coefficients

In the previous section, we tested our predictions about the portfolio dynamics of the different types. We derived these predictions from the model’s type-specific stimulus-response equations (4'). In this section, we test whether the stimulus response coefficients are consistent with our theory (hypothesis H2-H1).
Table 6: Regression analysis of portfolio dynamics over time, across types

<table>
<thead>
<tr>
<th></th>
<th>pre-price-peak</th>
<th>post-price-peak</th>
</tr>
</thead>
<tbody>
<tr>
<td>Period</td>
<td>0.017</td>
<td>0.054***</td>
</tr>
<tr>
<td></td>
<td>(0.060)</td>
<td>(0.017)</td>
</tr>
<tr>
<td>TE</td>
<td>1.259**</td>
<td>0.300</td>
</tr>
<tr>
<td></td>
<td>(0.494)</td>
<td>(0.636)</td>
</tr>
<tr>
<td>SE</td>
<td>-0.505</td>
<td>1.379**</td>
</tr>
<tr>
<td></td>
<td>(0.323)</td>
<td>(0.574)</td>
</tr>
<tr>
<td>SO</td>
<td>-0.171</td>
<td>1.165</td>
</tr>
<tr>
<td></td>
<td>(0.603)</td>
<td>(0.896)</td>
</tr>
<tr>
<td>TE×Period</td>
<td>-0.225*</td>
<td>-0.050</td>
</tr>
<tr>
<td></td>
<td>(0.116)</td>
<td>(0.053)</td>
</tr>
<tr>
<td>SE×Period</td>
<td>0.158***</td>
<td>-0.042</td>
</tr>
<tr>
<td></td>
<td>(0.075)</td>
<td>(0.042)</td>
</tr>
<tr>
<td>SO×Period</td>
<td>-0.042</td>
<td>-1.611***</td>
</tr>
<tr>
<td></td>
<td>(0.115)</td>
<td>(0.055)</td>
</tr>
<tr>
<td>Constant</td>
<td>2.382***</td>
<td>1.883***</td>
</tr>
<tr>
<td></td>
<td>(0.238)</td>
<td>(0.300)</td>
</tr>
<tr>
<td>overall $R^2$</td>
<td>0.020</td>
<td>0.036</td>
</tr>
<tr>
<td>N</td>
<td>1312</td>
<td>2272</td>
</tr>
<tr>
<td>Clusters</td>
<td>16</td>
<td>16</td>
</tr>
</tbody>
</table>

Assets change over time?

|                          | $\chi^2$ = 25.33 | $\chi^2$ = 0.15 |
| Period + SE×Period = 0   | $p < 0.001$       | $p = 0.703$      |
|                         | $\chi^2 = 6.29$  | $\chi^2 = 0.00$ |
| Period + TE×Period = 0   | $p = 0.012$       | $p = 0.945$      |
| Period + SO×Period = 0   | $\chi^2 = 0.01$  | $\chi^2 = 5.04$ |
|                         | $p = 0.913$       | $p = 0.025$      |

Type comparisons:

|                          | $\chi^2 = 15.39$ | $\chi^2 = 0.01$ |
| SE×Period=TE×Period      | $p < 0.001$       | $p = 0.915$      |
|                         | $\chi^2 = 2.83$  | $\chi^2 = 1.45$ |
| SO×Period=TE×Period      | $p = 0.093$       | $p = 0.228$      |
|                         | $\chi^2 = 6.41$  | $\chi^2 = 4.02$ |
| SO×Period=SE×Period      | $p = 0.011$       | $p = 0.045$      |

Random effects panel regressions, standard errors adjusted for clustering at the session level. Unit of observation: participant-period.

Significance levels: * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Dependent variable: Shares held at end of period.

Independent variables: Constant: FL, SE, SO, TE; dummies for mental type.
**Result 3** (Stimulus-response coefficients). TE and SE display different stimulus-response coefficients to both observables, as predicted by hypothesis H2. Moreover, the coefficients have opposite signs, consistent with inversion as the strongest form of miscalibration as identified by our theory. SE and SO have statistically similar response coefficients to last price, while TE and SO have statistically similar response coefficients to the fundamental, as predicted by hypothesis H1. Note that the fundamental is not the realization of the dividend; the fundamental is the expected dividend stream from the current period until the end of the asset market, and thus, in our experiment, a variable equal to 360 in the beginning, and linearly declining to 24 at the beginning of the final period.  

Figure 6: Marginal effects of fundamental and last price on asset holdings

The graphs show marginal effects of fundamental value (left) and previous market price (right) on shares held from regression M1 in table 7 by mental type. SO reacts positively to both the fundamental and the last price; SE only responds positively to the last price, TE only positively to the fundamental. Both SE and TE react inversely to their “deficient” dimension.

Table 7 shows random effects regressions of asset holdings as a function of the fundamental and the last price, allowing for differences across mental types. Figure 6 illustrates the estimated stimulus-response coefficients, showing the heterogeneity across

---

64. An alternative definition of fundamental could be the previously observed dividend. Trading in our experiment does indicate a hot hand fallacy, that is, a (mistaken) reaction to the dividend draw in the last period. Regressions of market price, buy price and sell price on period and previous dividend show that neither of the three variables react significantly to the value of the previous dividend draw (output omitted).

65. We present the results with and without controlling for risk aversion. As before, risk aversion has no influence on the estimation results.
Table 7: Stimulus response regression, across types

<table>
<thead>
<tr>
<th></th>
<th>M1</th>
<th>M2</th>
</tr>
</thead>
<tbody>
<tr>
<td>SE</td>
<td>0.793*</td>
<td>0.794*</td>
</tr>
<tr>
<td></td>
<td>(0.421)</td>
<td>(0.423)</td>
</tr>
<tr>
<td>TE</td>
<td>0.949</td>
<td>0.949</td>
</tr>
<tr>
<td></td>
<td>(1.092)</td>
<td>(1.093)</td>
</tr>
<tr>
<td>SO</td>
<td>-1.797**</td>
<td>-1.796**</td>
</tr>
<tr>
<td></td>
<td>(0.727)</td>
<td>(0.728)</td>
</tr>
<tr>
<td>Fundamental</td>
<td>-1.797**</td>
<td>-1.796**</td>
</tr>
<tr>
<td></td>
<td>(0.727)</td>
<td>(0.728)</td>
</tr>
<tr>
<td>SE×Fundamental</td>
<td>-0.003**</td>
<td>-0.003**</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>TE×Fundamental</td>
<td>0.005**</td>
<td>0.005**</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>SO×Fundamental</td>
<td>0.003*</td>
<td>0.003*</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>Market Price_{t-1}</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>SE×Price_{t-1}</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>TE×Price_{t-1}</td>
<td>-0.005*</td>
<td>-0.005*</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>SO×Price_{t-1}</td>
<td>0.002</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.003)</td>
</tr>
<tr>
<td># Lottery choices</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>2.465***</td>
<td>2.457***</td>
</tr>
<tr>
<td></td>
<td>(0.290)</td>
<td>(0.466)</td>
</tr>
<tr>
<td>overall R²</td>
<td>0.029</td>
<td>0.029</td>
</tr>
<tr>
<td>N</td>
<td>3440</td>
<td>3440</td>
</tr>
<tr>
<td>Clusters</td>
<td>16</td>
<td>16</td>
</tr>
</tbody>
</table>

Type comparisons Fundamental:

<table>
<thead>
<tr>
<th></th>
<th>M1</th>
<th>M2</th>
</tr>
</thead>
<tbody>
<tr>
<td>SE×Fun. = TE×Fun.</td>
<td>χ² = 17.31</td>
<td>χ² = 17.29</td>
</tr>
<tr>
<td></td>
<td>p &lt; 0.001</td>
<td>p &lt; 0.001</td>
</tr>
<tr>
<td>SO×Fun. = TE×Fun.</td>
<td>χ² = 0.31</td>
<td>χ² = 0.31</td>
</tr>
<tr>
<td></td>
<td>p = 0.578</td>
<td>p = 0.578</td>
</tr>
<tr>
<td>SO×Fund. = SE×Fund.</td>
<td>χ² = 10.32</td>
<td>χ² = 10.32</td>
</tr>
<tr>
<td></td>
<td>p = 0.001</td>
<td>p = 0.001</td>
</tr>
</tbody>
</table>

Type comparisons Price:

<table>
<thead>
<tr>
<th></th>
<th>M1</th>
<th>M2</th>
</tr>
</thead>
<tbody>
<tr>
<td>SE×Price_{t-1} = TE×Price_{t-1}</td>
<td>χ² = 5.10</td>
<td>χ² = 5.10</td>
</tr>
<tr>
<td></td>
<td>p = 0.024</td>
<td>p = 0.024</td>
</tr>
<tr>
<td>SO×Price_{t-1} = TE×Price_{t-1}</td>
<td>χ² = 3.24</td>
<td>χ² = 3.24</td>
</tr>
<tr>
<td></td>
<td>p = 0.072</td>
<td>p = 0.072</td>
</tr>
<tr>
<td>SO×Price_{t-1} = SE×Price_{t-1}</td>
<td>χ² = 0.13</td>
<td>χ² = 0.13</td>
</tr>
<tr>
<td></td>
<td>p = 0.715</td>
<td>p = 0.716</td>
</tr>
</tbody>
</table>

Random effects panel regressions, standard errors adjusted for clustering at the session level. Unit of observation: participant-period.

Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

Dependent variable: Shares held at end of period.

Independent variables: Constant; FL, SE, TE, SO; dummies for mental type; # Lottery choices: number of times a participant chose the lottery over the certain amount in the Holt-Laury task.

44
the four mental types. The relations among the estimated coefficients are as predicted by hypothesis H2 and H1.

The off-diagonal types TE and SE have opposite signs in their response coefficients to both observables. As predicted in hypothesis H2, TE’s response coefficient of the fundamental (last price) is significantly positive (negative), while SE responds significantly negative (positive) to the fundamental (last price). The difference in the coefficients between these two types is highly significant (see Table 7). The negative coefficients are consistent with inversion, the strongest form of miscalibration. This result highlights that the same set of observables can lead to a diametrically opposed perception and decisions if mental capabilities are different.

Figure 6 further suggests that SO and TE have similar response coefficients of the fundamental, and that SO and SE have similar response coefficients of the last price. The regression analysis in table 7 indeed shows that we cannot reject the hypothesis that these coefficients are similar, consistent with hypothesis H1.

4.3 Market Timing

Exit timing is crucial in a bubble market, both for trading gains and losses. Since Sophisticates reap the highest gains and Semiotic types incur the highest losses, we should expect that Sophisticates exhibit the best exit timing and Semiotic types the worst.

Result 4 (Market Timing). SO has the best market timing, showing the largest net sale of shares among all types right at the bubble peak. TE sells too early, and SE sells too late.

Figure 7 shows average net sales and purchases over time across mental types. We adjust the periods so that the period of the bubble peak in each market is denoted as period 0.\footnote{66. The bubble peak is the period where the over-pricing of the asset relative to the fundamental is} Before the bubble peak, FL’s do not show a systematic pattern of sales and
Figure 7: Net aggregate sales and purchases of shares by mental type. Periods are aligned so that period 0 is the bubble peak of each market.

The graph shows the order of the exit timing: TE exits first, several periods before the bubble peak; SO exits right at the bubble peak; SE misses the exit and divests only after prices have collapsed.

purchases but are net buyers at and after the bubble peak. SE’s tends to buy in almost every period before the bubble peak. TE buys early (when the price is still close to the fundamental) but also sells early, about 6 to 4 periods before the bubble peak. The largest turnover of all periods occurs right at the bubble peak, where most of the shares are sold by SO types to SE types. That is, SO’s have the best market timing of all mental types. TE has a worse timing, exiting the market too early, and SE has the worst timing, being a net buyer in almost all periods before and at the bubble peak, and a net seller only three periods after the peak, when the market price has fallen close to or below the fundamental value.

To test if SO has the best market timing, we investigate portfolio changes before an increase in the bubble component (market price minus fundamental value) and before a decrease in the bubble component. We regress the portfolio changes on an indicator maximal. The average bubble peak occurs between period 11 and 12 in the experiment.
of whether the bubble component will increase in the next period, and on interactions between this indicator and the type dummies (output omitted). We find that only Sophisticates correctly adjust their portfolio in anticipation of a change in the bubble component \( (p = 0.006) \). This adjustment is also significantly different from the adjustment of the other three types (SE: \( p = 0.002 \), TE: \( p = 0.019 \), FL: \( p = 0.064 \)). The other three types do not significantly anticipate the change. SE’s even seem to get it the wrong way, divesting before an increase in the bubble component and investing before a decrease \( (p = 0.110) \).

5 Conclusion

Anecdotal evidence suggests that a purely “fundamentalist” approach to asset trading, without an understanding of the psychology of the market, may not yield maximum profits. For example, both Julian Robertson of Tiger funds and John Templeton correctly identified that dot-com stocks were overpriced in the late 1990s. Robertson shorted them too early because he did not anticipate the size and duration of the bubble; but Templeton understood the market “sentiment” better and shorted in 2000, right before the bubble burst (Zuckerman 2000).67

We argue that traders like Robertson and Templeton do not fit in a one-dimensional spectrum of trader types. Instead, we develop a simple framework that conceptualizes how two independent mental dimensions, analytical reasoning and mentalizing, determine an agent’s mental model of the decision she faces. This framework yields a tractable model of how mental capabilities translate to asset trading behavior. Our core conjecture is that each capability determines one aspect of asset valuation. The ability to evaluate

67. The investor George Soros pointed out the importance of sentiment: sometimes market prices “do not merely reflect the so-called fundamentals; they themselves become one of the fundamentals which shape the evolution of prices” (Soros 2003). These moments of “irrational exuberance,” when prices detach themselves from the fundamental, are “the psychological basis of a speculative bubble,” according to Robert Shiller (Shiller 2015).
the fundamental value of an asset depends on analytical capability while the ability evaluate market sentiment depends on the mentalizing capability. As a consequence of the two-dimensional, non-convertible nature of mental capabilities, we predict four mental types. These types represent the four possible combinations of analytical and mentalizing capability (low-low, high-low, low-high, high-high). Each type has unique, distinguishable trading patterns, which lead to different profits. Our approach is fundamentally different from the literature about how heterogeneous information affects the trading in asset markets. We posit that traders with the same information can evaluate this information very differently if they have different mental types.

We examine our main theoretical predictions in a laboratory asset market. This approach allows us to control the environment tightly and to measure the traders’ mental types. We find that both mental capabilities determine individual trading behavior and profits. Consistent with our theory, analytical capability alone is not enough to maximize trading gains without a strong accompanying mentalizing capability. Technocratic types can avoid major losses because they correctly identify the fundamental value of the asset and exit early. On the other hand, their inability to correctly analyze the intentions of others prevents them from exploiting the price bubble. To anticipate such a bubble, a trader also needs an understanding of the “animal spirits” in a market. She needs to realize when the price deviates from the fundamental and when the bubble peaks. Consistent with our prediction, only Sophisticated types, who possess both capabilities, can ride a price bubble and time their exit correctly. As a consequence, they also make the highest profits of all four types. The Semiotic type has only mentalizing capability. This is outright detrimental as this type can only identify the upward price trend but fails to realize the departure from the fundamental. As predicted by our theory, this type therefore incurs the largest trading losses.

68. Keynes, who coined the term animal spirits in this context, was himself a successful speculator. He acknowledged that “there is the instability due to the characteristic of human nature that a large proportion of our positive activities depend on spontaneous optimism rather than mathematical expectations” (Keynes 1936).
Because our framework explains observed heterogeneity in trading behavior and success, institutional traders may be able to select traders according to their mental capabilities to increase trading gains. Concerning public policy, interventions targeting Semiotic types, who are most prone to follow a price bubble, may help to reduce the size of such bubbles.

On the theoretical side, we think that our approach can offer a new view on recent accounts of off-equilibrium behavior (such as asset market bubbles). This behavior is outside the scope of traditional game theory. In the last two decades, research has incorporated behavioral aspects into the standard framework. Some approaches introduce the possibility for random mistakes in people’s choices (McKelvey and Palfrey 1995); others assume that people’s responses are optimal, but based on flawed beliefs (Stahl and Wilson 1995; Camerer, Ho, and Chong 2004). Since we formulate (and test) a specific cognitive foundation (differential mental capabilities) for diverging behavior, our contribution provides a new take on both types of mistakes. In particular, biased beliefs could be the result of deficient mentalizing capability, while failure to best-respond could be based on lacking analytical capability.

As a framework for off-equilibrium behavior, we believe that our two-dimensional capabilities approach has potential beyond financial markets and may be adapted to explain behavioral puzzles in other domains. For example, a couple of recent studies examine the role of “strategic sophistication”, which we see as an combination of our two mental dimensions. It would therefore be interesting to investigate strategic games, such as the Beauty Contest, using our classification of mental types.

References


6 Appendix

6.1 Proof of Proposition 1

Define $\varepsilon_{\theta,\theta'} \equiv \varepsilon_{\theta} - \varepsilon_{\theta'}$ and $Z_{j}^{\theta,\theta'} \equiv \varphi_{j}(r_{j}^{\theta}, X_{j}) - \varphi_{j}(r_{j}^{\theta'}, X_{j})$, $X_{1} = F$, $X_{2} = P$. By the valuation equation (4’) we have

$$
\Pr(FL) = \Pr \left( \varepsilon_{FL,SE} < Z_{2}^{FL,SE}, \varepsilon_{FL,TE} < Z_{1}^{FL,TE}, \varepsilon_{FL,SO} < Z_{1}^{FL,SO} + Z_{2}^{FL,SO} \right)
$$

$$
\Pr(SE) = \Pr \left( \varepsilon_{SE,FL} < Z_{2}^{SE,FL}, \varepsilon_{SE,TE} < Z_{1}^{SE,TE} + Z_{2}^{SE,TE}, \varepsilon_{SE,SO} < Z_{1}^{SE,SO} \right)
$$

$$
\Pr(TE) = \Pr \left( \varepsilon_{TE,FL} < Z_{1}^{TE,FL}, \varepsilon_{TE,SE} < Z_{1}^{TE,SE} + Z_{2}^{TE,SE}, \varepsilon_{TE,SO} < Z_{2}^{TE,SO} \right)
$$

$$
\Pr(SO) = \Pr \left( \varepsilon_{SO,FL} < Z_{1}^{SO,FL} + Z_{2}^{SO,FL}, \varepsilon_{SO,SE} < Z_{1}^{SO,SE}, \varepsilon_{SO,TE} < Z_{2}^{SO,TE} \right)
$$

The effects of $dF, dP$ on $Pr(\theta|F, P)$ as stated in table 1 then follow from (9), noting that, by supermodularity, $\frac{\partial Z_{j}^{\theta,\theta'}}{\partial X_{j}} > 0$ iff $r_{j}^{\theta} > r_{j}^{\theta'}$. ■
6.2 Split along each Dimension and Trading Gains

Figure 8: Trading gains across cognitive types.

(a) Trading gains across one dimension (b) Trading gains across all four types

On the vertical axis, we plot trading income over the entire duration of the asset market for the baseline treatments (N=256). Panel (a) shows both the median split along the A dimension (left side) and along the M dimension (right side). Viewed in isolation, the A dimension shows a substantial difference in trading gains/losses, while the M dimension does not seem to predict trading success. Panel (b) displays the four different types and uncovers great heterogeneity. The semiotic type incurs most of the losses while the sophisticated type earns most of the profits and the featureless types as well as the technocrats earn zero profits.