Investments and Asset Pricing in a World of Satisficing Agents *

Tony Berrada† Peter Bossaerts‡ Giuseppe Ugazio§

August 25, 2024

Abstract

In 1955, Herbert Simon proposed that economic agents do not optimize, but instead satisfice: they optimize up to some point of satisfaction. But Simon did not provide a formal model. Here, we develop a formal theory of a satisficing investor and consequent financial market equilibrium borrowing a technique from robust control in engineering, namely, Model Reference Based Adaptive Control (MRAC). Instead of optimizing a portfolio in terms of, say, a mean-variance trade-off, the MRAC agent chooses portfolios that generate return distributions that minimize surprise with respect to a desired reference distribution. Surprisingly, the satisficing agent mostly acts "as if" optimizing, but we discover important – and realistic – deviations, such as willingness to accept risk even in the absence of a risk premium. This also implies that asset pricing may at times differ substantially from traditional theory. We motivate our modeling approach not only by pointing to benefits of robustness (robust control), but also with reference to developments in computational neuroscience since the 1950s that have recently inspired the idea of "efficient irrationality" in behavioral economics, thus going beyond what was known in psychology at the time Herbert Simon proposed his model of human behavior.

*The authors thank participants at a seminar at GFRI, University of Geneva, for valuable comments and suggestions. Bossaerts acknowledges generous support from the Leverhulme Trust.

†University of Geneva - Geneva School of Economics (GSEM), Geneva finance research institute (GFRI) and Swiss Finance Institute, Geneva, Switzerland. tony.berrada@unige.ch

‡Faculty of Economics, University of Cambridge, Cambridge, United Kingdom. plb32@cam.ac.uk

§University of Geneva - Geneva School of Economics (GSEM), Geneva finance research institute (GFRI), Geneva, Switzerland. giuseppe.ugazio@unige.ch

1 Introduction

We develop a theory of investments and derive corresponding equilibrium predictions when investors cast their preferences in terms of a target mean return and return volatility. Investors may not be able to attain their target, in which case they choose portfolios that minimize expected surprise relative to a world in which they could implement their preferences. This contrasts to the traditional approach, whereby the agent starts from a utility function (e.g., prospect theory) and then chooses a portfolio that maximizes this utility subject to a budget constraint.

The idea of choosing to match a target as close as possible could be viewed as an operationalization of "satisficing," first proposed by Herbert Simon as a unifying principal of human behavior: "[E]conomic man is a satisficing animal whose problem solving is based on search activity to meet certain aspiration levels rather than a maximizing animal whose problem solving involves finding the best alternatives in terms of specified criteria" [Simon, 1959] [p. 276].

Surprisingly, there have been few attempts to formalize the concept of satisficing. Caplin et al. [2011] proposes to identify it with the theory of sequential decision-making, which provides a formal framework to decide when to stop searching for evidence [Schervish, 2012] [Chapter 9]. This proposal is fair since it matches to some degree the informal description in Simon [1955] (though this article does not explicitly mention "satisficing"). Caplin et al. [2011] presents results from an experiment where participants were asked to choose among options the values of which were not obvious unless sufficient mathematical effort was applied. Consistent with sequential decision-making theory, participants stopped when reaching a threshold value.

Their evidence is far from conclusive, however. First, the choice that counted for earnings was randomly chosen from among the choices participants made during a trial. As such, choice improvements later in the trial mattered less; this induces sequential decision-making as the optimal strategy.[1] Second, in a setting where choice was not just complicated, but computationally "hard" (NP hard), Murawski and Bossaerts [2016] showed how participants continued searching way beyond the point at which value increased. If there was some kind of threshold, it certainly was not value.[2]

Here, we provide motivation for our approach to modeling satisficing, not by pointing to sequential analysis, but by appealing to long-established principles in engineering and to recent advances in computational neuroscience. For finance experts, we hasten to add that our approach is not related to "tracking." Our agent attempts to reach a target mean and volatility instead of tracking a particular

[1] "There is less incentive to continue search in the choice process task, given that the probability of additional effort raising the payoff falls over time" (p. 2904 of Caplin et al. [2011]).

[2] In the experiment, only the eventually submitted choice mattered for earnings. With few exceptions, time to submission was almost always within the (generous) time limit.

portfolio. To track a portfolio may not make much sense anyway, when this portfolio may produce dismal performance.[3]

In robust control theory, decision-making based on minimizing surprise relative to a reference world has become known as *Model Reference Based Adaptive Control* (MRAC) [Landau et al., 1998, Nguyen, 2018]. We use MRAC as the formal environment in which to derive our predictions. Unlike in many applications of MRAC, however, we will ignore policy (choice) adaptation *per se*, and instead focus on the policy obtained at any point in time giving the agent's beliefs about the environment at that moment. As will become clear, we derive properties of choices that are generic for large sets of possible beliefs. In this sense, the fact that beliefs, and hence, choices adapt hardly matters for our predictions to be have empirical content. For instance, as long as the agent targets a Sharpe ratio below one, she will invest in a risky asset even if she expects to earn a zero risk premium, regardless what else she believes about the true return generating process.

Surprisingly, in many circumstances the MRAC agent acts "as if" optimizing. Specifically, when the agent's desired Sharpe ratio is outside the feasible range (the target mean return and return volatilities are above the capital market line), choices may be indistinguishable from those of a mean-variance optimizer. This implies that CAPM may obtain even if all agents "satisfice" (in our operationalization of the term). Consequently, CAPM can thus be explained even in terms of "satisficing" behavior.

But when the market's maximal Sharpe ratio increases towards and beyond the agent's aspirations, choices will generically be distinct from those of a mean-variance optimizer. Key ways in which the MRAC agent's choices deviate from an optimizer are the following: (i) multiplicity: indifference between two choices that would imply vastly different values for any optimizing agent whose preferences are increasing in mean return and decreasing in return volatility; (ii) backward-bending demand: as the market's Sharpe ratio increases because more risk may need to be absorbed, the agent could pull back from risk – thus rendering equilibration more difficult; (iii) risk as a "good:" even when there is no risk premium (the best risky portfolio earns an expected return equal to the riskfree rate), the MRAC agent may choose to be exposed to risk. The latter emerges because, fundamentally, the agent is not averse to risk, but to surprise (relative to the stochastics in her reference model).

These three features of portfolio demand of an MRAC agent invalidate the idea that she somehow chooses to reach a value threshold and then stops. In other words, the MRAC agent cannot be modeled *as if* implementing sequential analysis. As we shall see, the agent's choices cannot fully be explained in terms of traditional utility optimization theory either, whether based on expected utility, prospect

[3]Portfolio tracking makes sense in certain equilibria, such as CAPM (Capital Asset Pricing Model), where a particular portfolio can be identified that is optimal. So, it relies on belief in the equilibrium that justifies the portfolio one is tracking.

theory, or ambiguity aversion. But, as emphasized before, under the right circumstances, choice is *as if* the agent is a mean-variance optimizer. In those circumstances, traditional asset pricing theory such as CAPM does obtain.

The reader may find some aspects of the behavior of the MRAC agent to be counterintuitive. Perhaps the strangest aspect is that the MRAC agent will not try to improve performance even if she believes that it is possible to invest at higher expected rates of return given the volatility that the MRAC agent chooses. Her attitude is the result of her desire for robustness. Indeed, while the agent may be able to do better in the present situation, the improvement may make her more vulnerable if the environment suddenly changes. The agent considers potential improvements to be purely circumstantial; she reasons that, once circumstances change, she may have adapted too much to the original situation that she is no longer in the position to change policy fast enough.

The remainder of this paper is organized as follows. In the next Section, we provide a broader justification for MRAC as a model of human choice under uncertainty, with references to the relevant literature. Section 3 defines the key concept in MRAC, namely, expected surprise. Analysis of generic properties of investments under MRAC, along with equilibrium implications, follow in Section 4. In Section 5, we study demand functions when there is only a single risky asset, as a function of, first, risk (keeping expected return constant), and second, price (which induces changes in both expected returns and return volatility). Section 6 compares choices under MRAC with those under reference-point based utility maximization, α-maxmin ambiguity sensitive utility maximization, and traditional expected utility maximization. Section 7 concludes with suggestions for further research.

2 Why MRAC?

In engineering, the use of reference model emerged as early as the 1950s. It was proposed as a more reasonable approach to robust control than the traditional maximin control model, which has inspired economic modeling [Hansen and Sargent, 2001, Gilboa and Schmeidler, 2004]. There is a moderator (principal) who monitors an agent who acts upon the environment. The moderator has control only over a limited set of parameters – called "hyperparameters" – that modulate the action policies of the agent. The agent is otherwise fixed. The moderator adapts the hyperparameters to ensure that outcomes produced by the actions of the agent do not deviate too much from those in a reference model. That is, the principal minimizes surprise. Robustness results when both the choice of the reference model and the adaptation of the hyperparameters are such that stability is obtained for a wide variety of environments. Stability means: given an environment, the principal's adaptation of the agent's hyperparameters converges. Fig 1 displays a schematic overview of adaptation and control in MRAC.

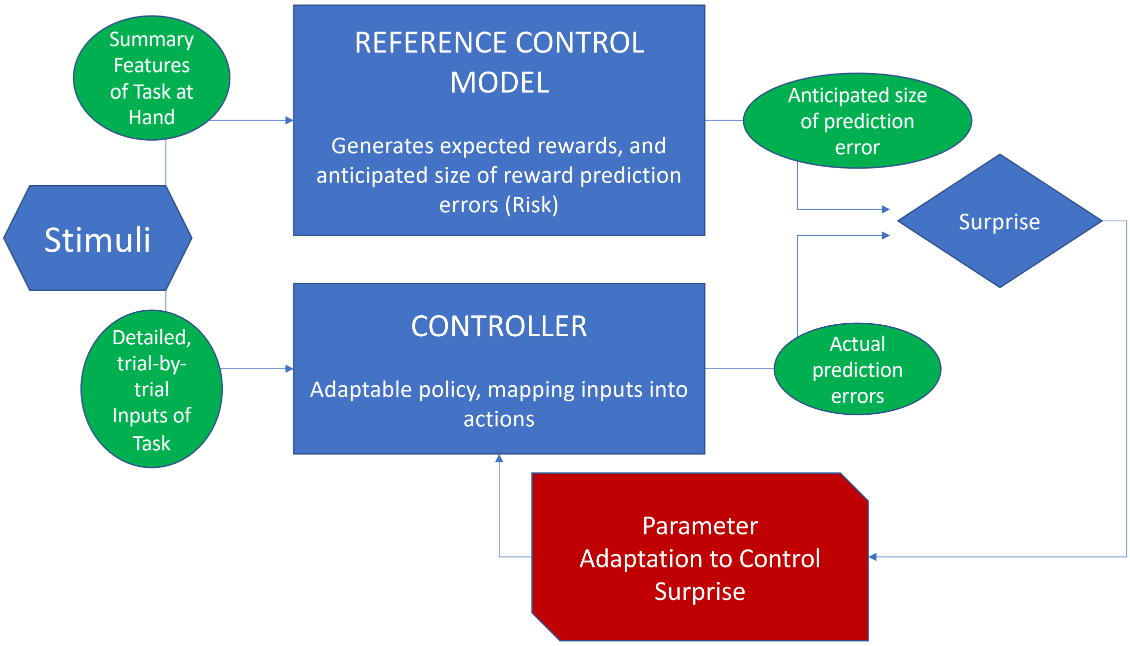

Figure 1: Figure adapted from Bossaerts [2018].

MRAC can be viewed as a special case of the principal-agent model in economics, but without the conflict of interest: there is a moderator (principal) who monitors the outcomes an agent produces, occasionally adjusting the hyperparameters with which this agent controls the environment; the monitor's goal is to become minimally surprised given the outcomes from a reference model. Details, including ways to ensure stability when monitoring the agent, along with applications can be found in standard textbooks, such as Nguyen [2018]. Engineers prefer MRAC for another reason: it obviates the need to formulate a utility function; it only requires one to specify desired, measurable outcomes [Landau et al., 1998]. In the language of classical decision theory: the engineer is to prescribe what observable outcomes she chooses, rather than determining the utility function the optimization of which represents her desired choices.

The idea of a reference model has made its way into computational neuroscience as well. There, it emerged as a result of the discovery of limitations to neural processing of sensory stimuli (signals) such as visual or auditory cues. It soon became apparent that the brain was using simplified reference models – simplified to make its computations feasible given the neurobiological constraints – adopted to minimize loss of information. The brain was thus found to implement *efficient decoding*. This was first discovered in simple sensorimotor tasks [Attneave, 1954, Barlow et al., 1961]. But more recently, evidence has been accumulating that it applies to more complex tasks, such as value-based decision-making. There, limited – but efficient – decoding of the information in a choice task has been found to explain a large fraction of well-known cognitive biases such as probability weighting. As such, cognitive biases may emerge not only because humans choose to adapt simple heuristics [Tversky and Kahneman, 1974], but also because of neural limitations in decoding, from the available information, what the problem is that one has to decide on [Yamada et al., 2018, Azeredo da Silveira and Woodford, 2019, Polanía et al., 2019, Vieider, 2024].

It has even been suggested that some of the principles to come out of efficient decoding of sensory stimuli, e.g., divisive normalization, also apply to *value encoding* [Louie et al., 2013, Glimcher and Tymula, 2023]. That is, the principles would also apply when, e.g., deciding which of a plethora of assets to invest in. The principles are thought to simplify the otherwise impossibly complex task of choosing the best option. The choices one observes are thought to inherit as much information as possible when contrasted with optimal choices; they maximize mutual information between actual choices and optimal choices.

A further development in the literature has been to start from a putative model of neural representation of values and derive the principles that allow actual choices to inherit as much information as possible from optimal choices. The agent chooses in a way that maximizes mutual information between actual choices and optimal choices [Payzan-LeNestour and Woodford, 2022, Frydman and Jin, 2022]. The resulting behavior has been referred to as "efficient irrationality" [Glimcher, 2022].

We too propose that a reference model is used in the encoding stage of value-based decision-making. We conjecture, however, that the reference model is chosen, not so much as a way to reduce complexity when choice bandwidth is limited, but to make choices more robust in the face of an ever-changing world. That is, we appeal to *robustness* instead of (informational) efficiency as a source of reference-model based decision-making.

Only targeted controlled experiments will determine the origin of the use of reference models, whether it is robustness or efficiency, or perhaps both. Here, we focus on the behavioral implications of the use of a reference model. We do so in the standard framework of MRAC, where actions are chosen to minimize surprise. One expects our findings to be of relevance for analyses based on efficient irrationality as well, since there exists a close relationship between surprise minimization and maximization of mutual information.[4]

[4]In *active inference*, a form of Bayesian inference where the agent chooses the data rather than merely conditioning on the data, the link between surprise and mutual information is usually made explicit; see, e.g., Friston [2010], Sajid et al. [2021], Bossaerts and Rayo [2023]. Active inference is an extension of Variational Bayes Analysis (VBA). There, the reference model is called the "recognition model;" for applications and extensions of VBA to explain cognitive biases in behavioral economics, see Samuelson and Steiner [2024].

3 Defining Expected Surprise

3.1 Expected Surprise

In MRAC (Model-Reference Based Adaptive Control), the goal is to control the "hyperparameters" of a controller (agent) to minimize expected surprise with respect to the principal's "reference model." Here, we assume that the reference model is characterized by a desired level of mean return and return volatility. MRAC does not aim at maximizing (hard to identify) utility, instead insisting on matching observable features of a reference model; here: mean and volatility of portfolio returns.

There are many ways to define expected surprise.[5] Here, we follow tradition in engineering and define surprise as the average "risk prediction error," i.e., the difference between the average realization of the risk and its anticipation, where realized risk equals the distance between outcome and its subjective expectation, and where anticipation equals the subjective expectation of this distance, also known as "variance". Mathematically, define Expected Surprise S as follows:

$$S = (E(\tilde{R} - \mu)^2 - \sigma^2)^2. \tag{1}$$

Here, $\tilde{R}$ is the return on the chosen portfolio, E is taken with respect to the agent's beliefs about the true return distribution, and μ and σ are the desired return mean and volatility, respectively.

As mentioned before, we will not be interested in adaptation dynamics *per se*, but in the nature of the control, and hence, choices, that would result given particular beliefs, unlike in standard MRAC analysis in engineering. There is a second dimension in which we deviate from traditional MRAC analysis. Traditionally, the control itself is delegated to an agent how follows a rigid control strategy which the principal (the investor) can only manipulate to a certain extent.[6] That is, the principal can modulate certain parameters of the control strategy; control is indirect. Here, we will allow the principal to have full control over the investment. Specifically, we allow the principal to set the optimal investment policy given the desire to minimize expected surprise. In future work, we could restrict the principal's choice to a few standard investment strategies (e.g., when there are multiple risky securities, the agent invests only in a value-weighted or an equally-weight portfolio). We could also allow the agent to re-balance portfolios at a higher frequency than the principal evaluates surprise. The principal could then use rebalancing frequency as a control parameter.

[5] See, e.g., Loued-Khenissi and Preuschoff [2020].
[6] This strategy could be the result of value maximization; see previous footnote.

4 Choice properties and equilibrium implications

In this section we derive the choice properties of a surprise minimizing agent and its implication for equilibrium prices in a competitive financial market.

Surprise is defined as

$$S = (E(\tilde{R} - \mu)^2 - \sigma^2)^2. \tag{2}$$

Here, $\tilde{R}$ is the return on the chosen portfolio, E is taken with respect to the true distribution, and μ and σ are the desired return mean and volatility, respectively. The agent chooses an investment strategy with volatility x and expected return y. Surprise can be equivalently written as

$$\begin{aligned} S =& (E(\tilde{R} - y + y - \mu)^2 - x^2 + (x^2 - \sigma^2))^2 \\ =& (E(\tilde{R} - y)^2 + (y - \mu)^2 + 2E(\tilde{R} - y)(y - \mu) - x^2 + (x^2 - \sigma^2))^2 \\ =& (x^2 + (y - \mu)^2 - x^2 + (x^2 - \sigma^2))^2 \\ =& ((y - \mu)^2 + (x^2 - \sigma^2))^2 \\ =& ((y - \mu)^2 + (x - 0)^2 - \sigma^2))^2. \end{aligned} \tag{3}$$

We first consider a market with many risky assets (number of assets $N > 1$) and a risk free asset with return R_F. In the standard deviation—expected return space, the capital market line (efficient frontier) is a line originating at R_F with slope equal to the maximum attainable Sharpe ratio θ. We label it EF. We further label nEF the line originating at R_F with slope equal to $-\theta$. Feasible investments are in the cone within EF and nEF.

Lemma 4.1. *The set of portfolios which generate surprise $S = 0$ is a circle with center $(0, \mu)$ and radius σ. We label it the **0-surprise circle**.*

Proof. When $S = 0$,

$$(y - \mu)^2 + (x - 0)^2 = \sigma^2$$

This is the formula for a circle with center $(0, \mu)$ and radius σ. □

The 0-surprise circle is depicted in Figure 2. Notice that the origin is moved to the right compared to the usual plots of the mean-variance space.[7] Of course, negative standard deviations are non-existent. Nevertheless, for clarity we plot the entire circle, including for negative values on the horizontal axis.

The following is a trivial extension.

[7] We continue tradition when referring to mean-standard-deviation plots as mean-variance plots.

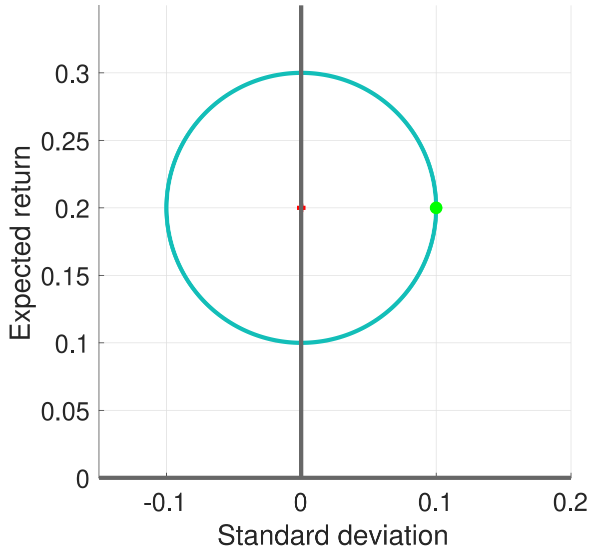

Figure 2: 0−surprise circle. Desired (μ, σ) is the green dot.

Lemma 4.2. *The set of portfolios which generate surprise $S = \varepsilon$ is a circle with center $(0, \mu)$ and radius $R = \sqrt{\sqrt{\varepsilon} + \sigma^2}$. We label it the ε-**surprise circle**.*

Proof. When $S = \varepsilon$,

$$((y-\mu)^2 + (x-0)^2 - \sigma^2)^2 = \varepsilon$$

and therefore

$$(y-\mu)^2 + (x-0)^2 = \sqrt{\varepsilon} + \sigma^2$$

This is again the formula for a circle with the same centre, but higher radius. □

Equipped with the definitions of lemma 4.1 and 4.2, we can now provide the conditions under which the choice of the surprise minimizing agent is unique and if so, what this choice is.

Proposition 4.3. *When the efficient frontier is such that*

$$0 < \theta \leq \frac{y_{tan} - R_F}{x_{tan}}$$

where

$$y_{tan} = \frac{\mu^2 - \mu R_F - \sigma^2}{\mu - R_F}$$
$$x_{tan} = \sqrt{(y_{tan} - \mu)(y_{tan} - R_F)}$$

*the surprise minimizing portfolio is at the tangency point of an ε-**surprise circle** and EF, and is unique. The surprise minimizing portfolio has risk*

$$x^* = \frac{(\mu - R_F)\theta}{1+\theta^2}$$

and expected return

$$y^* = \mu - \frac{(\mu - R_F)}{1+\theta^2}$$

.

Proof. A surprise minimizing portfolio necessarily belongs to a ε-**surprise circle**. EF is tangent to the 0-**surprise circle** if $\theta = \frac{y_{tan} - R_F}{x_{tan}}$ and the surprise minimizing portfolio is at the tangency point (x_{tan}, y_{tan}). Otherwise, the feasible portfolio closest to $(0, \mu)$, is the orthogonal projection of $(0, \mu)$ on EF, which is the tangency point of an ε-**surprise circle** and EF. The line originating at $(0, \mu)$ and orthogonal to EF has equation

$$y = \mu - \frac{1}{\theta}x.$$

Its intersection with EF is at the point (x^*, y^*). □

When the agent is reaching for a target mean and standard deviation which are overly optimistic, i.e. located strictly above the efficient frontier, proposition 4.3 shows that the surprise minimizing choice is unique. Figure 3 illustrates different market configurations which generate unique choices. When the choice is unique, it is located on the efficient frontier. We use this observation to obtain corollary 4.3.1, which provides insight into the equilibrium implications of surprise minimizing behavior.

Corollary 4.3.1. *In an economy populated by surprising minimizing agents, when all desired (μ, σ) satisfy the conditions of proposition 4.3, CAPM holds.*

Intuitively, the corollary argues that CAPM will obtain when agents have sufficiently high desired mean return and sufficiently low desired return volatility so that the 0-surprise circle fails to intersect with the set of feasible portfolios (the cone defined by EF and nEF).

Proposition 4.3 further implies that surprise minimizing agents make choices that depend on their target mean but not on their target volatility.

Corollary 4.3.2. *When the desired desired (μ, σ) satisfy the conditions of proposition 4.3, the surprise minimizing portfolio is independent of the desired level of risk σ.*

When the efficient frontier intersects the $0-$surprise circle, we expect the solution to no longer be unique. Proposition 4.4 provide a formal characterization of that situation and figure 4 illustrates it.

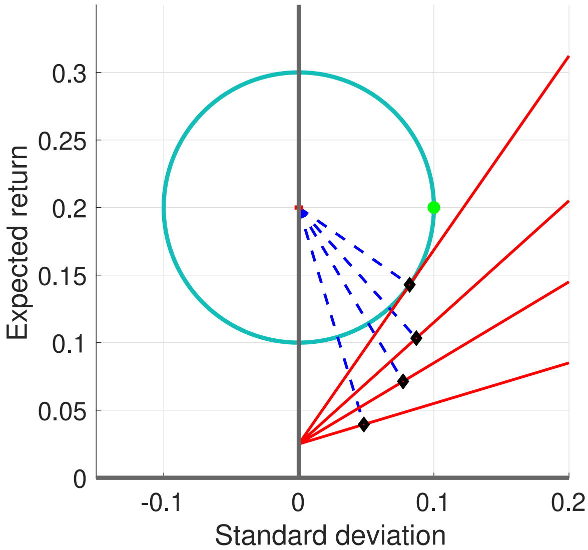

Figure 3: EF lines in red. Black diamonds indicate corresponding (unique) surprise minimizing choices. Desired (μ, σ) is the green dot.

Proposition 4.4. *When the efficient frontier is such that*

$$\theta > \frac{y_{tan} - R_F}{x_{tan}}$$

where

$$y_{tan} = \frac{\mu^2 - \mu R_F - \sigma^2}{\mu - R_F}$$
$$x_{tan} = \sqrt{(y_{tan} - \mu)(y_{tan} - R_F)}$$

there exist multiple surprise minimizing portfolios. They are on the arc of the ***0-surprise circle*** *inferior to EF and superior to nEF.*

Hence, when desired mean and volatility are sufficiently low so that the 0-surprise circle intersects the cone defined by EF and nEF, the solution is no longer unique. In those cases, the agent generically holds an inefficient portfolio. Corollary 4.4.1 highlights the implication of non-uniqueness and portfolio inefficiency for equilibrium.

Corollary 4.4.1. *In an economy populated by surprise minimizing agents, if at least one of the desired pairs (μ, σ) satisfies the conditions of proposition 4.4, there are multiple equilibria and the CAPM generically does not obtain.*

When the solution is not unique, the choice portfolio depends both on the desired mean and volatility,

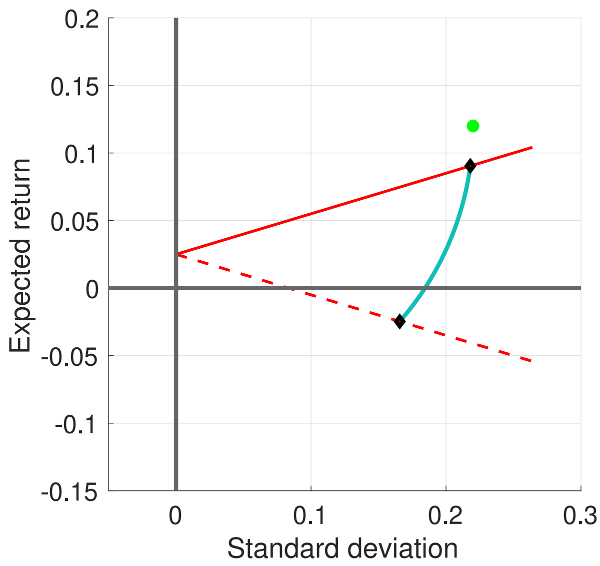

Figure 4: EF line in red, nEF line in red dashed. All surprise minimizing choices are on the green arc between EF abd nEF. Desired (μ, σ) is the green dot.

and the agent reduces volatility relative to her desired level.

Corollary 4.4.2. *When the desired pair (μ, σ) satisfies the conditions of proposition 4.4, the surprise minimizing portfolios have risk inferior to the desired level of risk σ.*

When the risk premium in the market is equal to 0, i.e. when the efficient frontier is a horizontal line, a particular behavior emerges for the surprise minimizing agent. Corollary 4.4.3 formalizes the argument.

Corollary 4.4.3. *When the desired Sharpe ratio is strictly inferior to one, i.e. $\frac{\mu - R_F}{\sigma} < 1$, the demand for risky assets is generically positive when the risk premium is 0.*

Proof. When the risk premium $\theta = 0$, if $\mu - R_F \geq \sigma$, then the **0-surprise circle** doesn't intersect EF and the demand for risky asset is 0. If $\mu - R_F < \sigma$ then the **0-surprise circle** intersects EF and the demand for risky asset is the on the arc of the **0-surprise circle** inferior to EF and with volatility greater or equal to 0. □

5 Demand Analysis

We now turn to studying demand for a single risky security ($N = 1$). This case may seem uninteresting at first since the agent will always act *as if* optimizing (with only one riskfree and one risky asset, any portfolio that returns at least the riskfree rate is mean-variance optimal). Still, the shape of the demand curve is distinctly different from that under mean-variance optimization. As a result, the demand function cannot be modeled *as if* optimizing some utility function without allowing preference parameters to change, through, e.g., reference points. (In the next section, we shall discuss to what extent reference point variations can credibly capture the peculiarities of the demand function of a MRAC agent.)

We first study demand as a function of risk, keeping the expected return on the risky asset constant. We study this case since it provides intuitive demand functions in view of our previous propositions, even if demand may be non-unique. When we turn to studying demand as a function of price, keeping constant the dollar payoff distribution of the single risky security, demand functions will no longer be intuitive, in the sense that they will violate a desired property of economic theory, namely, that they are downward sloping (the lower the price, the higher the demand). This will come on top of non-uniqueness.

5.1 Demand as a Function of Risk

With only one risky asset, the line EF and nEF coincide. It follows from proposition 4.4 that the arc of minimum surprise collapses to the two points of the intersection of EF and the **0-surprise circle**. The following proposition provides the asset demand as a function of the risk of the single risky asset.

Proposition 5.1. *When there is one risky asset with expected return m and risk s, assuming R_F and m are fixed, the demand function $d(s)$ of the surprise minimizing agent with desired (μ, σ) is given by*

$$\overline{d}(s) = \frac{(\mu - R_F)\theta(s)}{1 + \theta(s)^2} \frac{1}{s}$$

when

$$s \geq (m - R_F) \frac{x_{tan}}{y_{tan} - R_F}$$

and by

$$\underline{d}(s) = \frac{(\mu - R_F)\theta(s) \pm \sqrt{\Delta(s)}}{1 + \theta(s)^2} \frac{1}{s}$$

when $s < (m - R_F)\frac{x_{tan}}{y_{tan} - R_F}$. where $\theta(s) = \frac{m - R_F}{s}$ and $\Delta(s) = (\mu - R_F)^2\theta(s)^2 - (1 + \theta(s)^2)((\mu - R_F)^2 - \sigma^2)$.

Proof. The results follows directly from propositions 4.3 and 4.4 and the fact that EF and nEF coincide. □

5.2 Demand as a Function of Price

So far, we have kept the expected return on the risky asset fixed, allowing volatility (s) to change. We now study demand as a function of price, keeping expected (dollar) payoff and payoff volatility constant, allowing the price to change. As price decreases, both the return volatility (s) and the mean return (m)

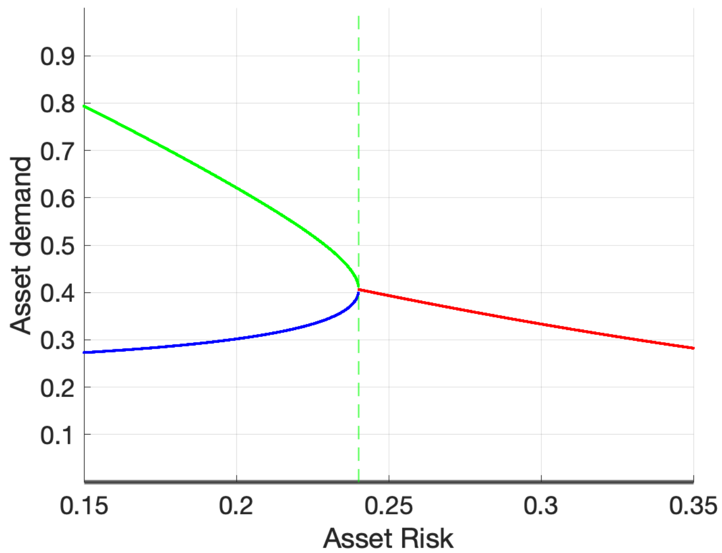

Figure 5: Risky asset demand as a function of risk ($N = 1$). The red line displays unique solutions when surprise is positive, the blue and red lines indicate the two solutions for all level of risk where surprise is 0. The green vertical dash line indicates the frontier between the two regions. We fix the asset expected return and risk free rate, and vary the asset risk. The desired mean and volatility are 0.25 and 0.125 respectively. The risk free rate is 0.05 and the asset expected return is 0.35.

on the risky security change; its location in mean-variance space is no longer constrained to be on a horizontal line. As Figure 6 shows, its location moves linearly upwards, towards the upper-right corner. This has profound implications for the demand curve of an MRAC agent, since such an agent either picks one of two points on a circle of zero expected surprise, or, when this is not feasible, picks the location on the mean-variance frontier of the orthogonal projection of the midpoint of this circle of zero expected surprise.

The asset payoff is a random variable with expected value M and standard deviation Σ. It follows that the expected return of the risky asset is given by $m = M/P - 1$ and its standard deviation is given by $s = \Sigma/P$. The following proposition provides a full characterization of the risky asset demand function.

Proposition 5.2. *Let P_b denote the price boundary level defined as*

$$P_b = \frac{M - \Sigma \frac{y_{tan} - R_F}{x_{tan}}}{1 + R_F}.$$

The demand function of the risky asset for the surprise minimizing agent with desired (μ, σ) is given by

$$\overline{d}(P) = \frac{(\mu - R_F)\theta(P)}{1 + \theta(P)^2} \frac{P}{\Sigma}$$

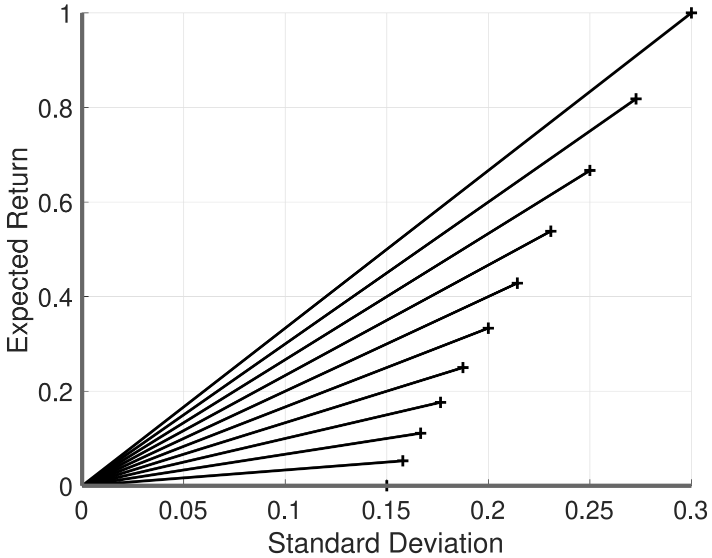

Figure 6: Location of riskfree security and risky security in mean-variance space (crosses) as price decreases, along with corresponding mean-variance frontiers (solid lines). Here, $R_F = 0$, payoff volatility equals \$0.15, and expected payoff equals \$1. As price decreases, the slope of the mean-variance frontiers (i.e., the Sharpe ratio) increases. The Sharpe ratio (ratio of expected return in excess of the risk free rate and the return volatility) decreases linearly as the price increases, with slope determined solely by R_F and payoff volatility.

when $P \geq P_b$, and by

$$\underline{d}(P) = \frac{(\mu - R_F)\theta(P) \pm \sqrt{\Delta(P)}}{1 + \theta(P)^2} \frac{P}{\Sigma}$$

when $P < P_b$, where $\theta(P) = \frac{\frac{M}{P} - 1 - R_F}{\frac{\Sigma}{P}}$ and $\Delta(P) = \theta(P)^2(\mu - R_F)^2 - \left((\mu - R_F)^2 - \sigma^2\right)(1 + \theta(P)^2)$.

Proof. The results follows directly from propositions 4.3 and 4.4 and the facts that EF and nEF coincide, $m = M/P - 1$ and $s = \Sigma/P$. □

The proposition states that the demand curve consists of two sections. This is clear from the example in Figure 7. Expected payoff on the risky asset continues to be \$1, while its payoff volatility remains \$0.15. One section lies above the dashed green line, positioned at a price equal to P_b; see red section. There, the agent chooses a unique position on the mean-variance frontier, at the point of orthogonal projection of the center of the circle of zero expected surprise. This is because prices are high, and hence, the maximum feasible Sharpe ratio is below the 0-surprise circle.

Below the critical price P_b, the maximum feasible Sharpe ratio is high enough for the capital market line (EF) to intersect the 0-surprise circle. There are two intersections, and hence, there is a bifurcation: the green line depicts the locus of higher demand; the blue line depicts the locus of lower demand. The former implies a higher demand for risk; the latter implies a lower demand for risk. Emergence of two sections, with bifurcation at the critical price, mirrors the pattern of the demand as a function of risk (s); see Figure 5. (The same color coding is used in both figures.)

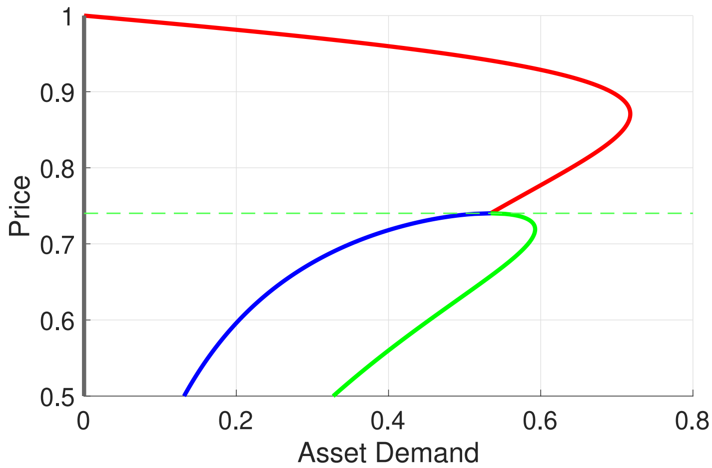

Figure 7: Demand as a function of price. For high prices, the demand curve is downward-sloping. However, below a certain level, demand bends backward: a decrease in price may lead to a decrease in demand for the risky asset. At some point, there is a bifurcation. Below this point, indicated with the horizontal dashed line, the agent can find investments with zero expected surprise. Parameters are as in Figure 6. Desired mean return (μ) equals 0.25; desired volatility (σ) equals 0.125.

At the point of bifurcation, the demand function is continuous but non-differentiable. This is stated formally in the following corollary.

Corollary 5.2.1. *$\overline{d}(P)$ and $\underline{d}(P)$ are continuous and $\overline{d}(P_b) = \underline{d}(P_b)$. The demand function is continuous on its domain, but not differentiable at $P = P_b$.*

Proof. The result follows from the observation that $\Delta(P_b) = 0$ and $\Delta'(P_b) \neq 0$ □

Notice also that the demand curve can be backward-bending: as price decreases, the agent decreases demand for the risky asset, and hence, decreases exposure to risk, despite the fact that price decreases cause the maximal Sharpe ratio to increase. This does not always occur; we will provide examples later on. The following corollary provides a sufficient condition for it to occur in the section above P_b.

Corollary 5.2.2. *If M, Σ, μ, σ and R_F are such that*

$$\arg\max_P \overline{d}(P) > P_b$$

then the demand function is non-monotonic on the interval $[P_b, 1]$

Why is the demand curve backward-bending? This is because the surprise minimizing agent does not perceive risk as a bad *per se*. Instead, she is averse to surprise relative to the reference model that sets its target mean return and return volatility. This means that she will tolerate risk, as long as it is in line with the reference model.

In fact, when it is feasible to acquire a portfolio with zero expected surprise, the agent will never take on a portfolio with more risk than desired (Corollary 4.4.2). This is the case below the bifurcation point P_b. There, as the price of the risky asset drops further below P_b, risk will increase without bound. The agent reduces exposure to the risky asset, in order to avoid taking on more risk than desired. Consequently, the demand curve is backward-bending.

When it is not possible to reduce surprise to zero, i.e., above the bifurcation point, the agent may take on more risk than desired. However, she will stay within limits imposed by the projection of the center of the 0-surprise circle onto the capital market line EF. In Figure 5, she indeed reduces demand for the risky asset below a price of 0.84. At 0.84, the agent invests 70% in the risky. As a result, her portfolio is exposed to a risk of 0.125 (= (0.70)(0.15/0.84)), the maximum she tolerates. As the price drops below 0.84, she necessarily reduces investment in the risky asset, to avoid going over her risk tolerance limit.

The backward-bending feature implies that, as price decreases, it is *as if* the agent were to become so much more risk averse that she withdraws from the market. It is as if she is reading from the price that things are going to be really terrible: she is over-pessimistic. In a dynamic world, backward-bending demand curves can exacerbate counter-cyclical risk aversion. Price drops that result from an increase in payoff volatility because of worsening economic conditions may lead some agents to reduce their exposure to risk since this is their best way to control expected surprise. Aggregate behavior is *as if* average risk aversion had increased disproportionately.

Backward-bending demand functions are not only inconsistent with mean-variance optimization; they are inconsistent with most expected-utility preferences. There exist exceptions when the coefficient of relative risk aversion varies excessively as a function of wealth [Lanier, 2020].

However, backward-bending demand curves could emerge under <u>loss aversion</u> in prospect theory [Kahneman and Tversky, 1979]. There, a price decrease may lead to decreased demand for a risky security *if* it pushes a higher proportion of return outcomes to the gain domain, where the agent is relatively more risk averse (than in the loss domain). This requires the reference point to decrease as the price decreases. If the reference point reflects aspiration, this means that aspiration level decreases as price decreases. In this story, reduction in the reference point as the result of price drops explains the bearish behavior. For us, control of surprise explains the disproportionate increase in apparent risk aversion. We come back to reference-dependent utility theory in Section 6.

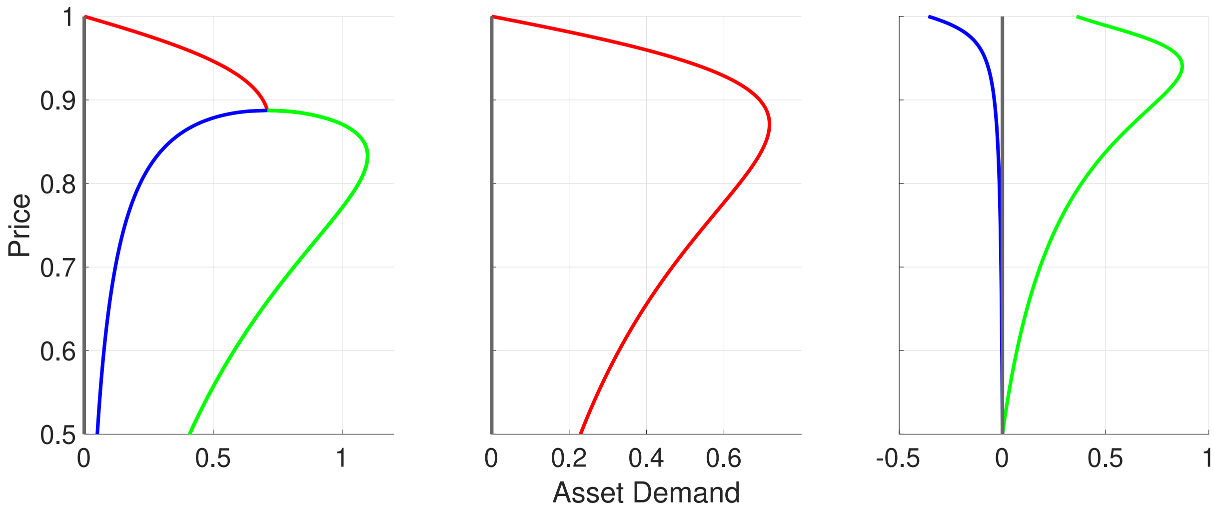

Figure 8: Demand as a function of price. Parameters are as in Figure 6. Left: desired mean return (μ) equals 0.25, desired volatility (σ) equals 0.20; middle: $\mu = 0.25$, $\sigma = 0.05$; right: $\mu = 0.14$, $\sigma = 0.15$.

When desired volatility is increased, to 0.20, demand no longer bends backward above the bifurcation point. See the left panel of Figure 8. When lowering target volatility instead, to 0.05, the bifurcation point decreases, to the point that it drops below the plotted range (0.5–1.0): the price has to drop below 0.5 for the capital market line EF to intersect with the 0-surprise circle. The entire (plotted) curve is in the red section, and it is backward-bending. See middle panel of Figure 8.

The right panel of Figure 8 displays the demand curve when the surprise minimizing agent targets a Sharpe ratio below 1 ($\mu = 0.14$, $\sigma = 0.15$, so Sharpe Ratio $= 0.93 < 1$). The demand is below the bifurcation point. Notice that neither demand curve features zero demand at zero risk premium (i.e., when price $= 1$). The blue demand function shorts at zero risk premium; the green demand function goes long at zero risk premium. Corollary 4.4.3 predicts this. Notice how the amount shorted on the blue demand function equals the amount long on the green demand curve, to attain the same risk.

Why would the agent take on risk even if it is not compensated? Again, for a surprise minimizing agent, risk is not necessarily a bad that should be taken on only if it is compensated for. The agent does tolerate risk; she expects risk; it is part of the reference model. To choose an investment that does not entail risk causes surprise: the outcomes are not as expected. Risk will then be accepted because the agent targets a risk premium that is small anyway (the Sharpe ratio is below 1).

6 Comparison With Other Utility Functions

6.1 Mean-variance Preferences

We have analyzed the choices of an MRAC agent. It is obvious her demand can be substantially different from that of a mean-variance optimizer. In a context with multiple risky securities, she will invest in an efficient portfolio, like a mean-variance optimizer, as long as her target mean return and return volatility are sufficiently high relative to the maximum feasible Sharpe ratio (slope of the capital market line). From the moment she tempers her target Sharpe ratio, moving closer to or even below the maximum feasible Sharpe ratio, she will generically not buy a mean-variance portfolio.

In a 2-asset environment (one risky and one riskfree asset), the agent necessarily buys on the frontier. But she may be on the inefficient part of the frontier, indifferent to positioning herself on the efficient or inefficient part, as long as it entails zero expected surprise. Her demand curve may be backward-bending (she reduces risk exposure even if compensation for risk – the available Sharpe ratio – increases). Her demand function exhibits a bifurcation, at which point it is not differentiable. Finally, even at zero risk premium, she may want to be exposed to risk, because she is not averse to risk; she is only averse to surprise.

6.2 Reference-Point Based Utility Function Maximization

One of the most frequently encountered deviation of actual preferences revealed in financial markets prices is the presence of a reference point. While at times this has been associated with framing, and hence, violations of rational choice axioms (see Prospect Theory), others consider reference points to be generated either by internal habit or external benchmarks (such as other investors' consumption relative to one's own). Such reference-point based expected utility models have been used mostly to accommodate the counter-cyclical risk aversion observed in historical data from the field; see, e.g., Barberis et al. [2001]. Such counter-cyclical risk aversion as revealed in prices is also consistent with expected surprise minimization, as pointed out in Section 5.2.

The reference point of Prospect Theory or of expected utility theory with first-order risk aversion lead to zero risk taking for a range of expected returns around R_F when the reference point is set equal to the wealth attained with riskfree investments. Yet, an MRAC investor with target Sharpe ratio strictly less than one will do exactly the opposite, namely, to invest in risky securities even if expected returns are equal to the riskfree rate R_F; see Corollary 4.4.3. This is because the MRAC investor is not averse to risk; she is only averse to expected surprise.

6.3 α-Maxmin Maximization

The core motivation for MRAC is robustness in the face of non-ergodic (non-stationary) changes in the environment. Another approach to deal with lack of robustness is maxmin [Gilboa and Schmeidler, 2004], or its generalization, $\alpha-$maxmin [Ghirardato and Marinacci, 2002]. There, the investor plans for the worst environment she could encounter (maxmin), or for some environment a distance "α" between the worst and the best ($\alpha-$maxmin). The ($\alpha-$) maxmin investor may learn to narrow down the range of environments she encounters, but in a truly non-stationary, i.e., non-ergodic, environment, this should not happen.

If we assume mean-variance preferences, then the ($\alpha-$)maxmin investor will always buy a mean-variance efficient portfolio, determined though by beliefs that are some convex linear combination (with weight α on the worst environment) of the worst and best environment. The MRAC agent, instead, may not invest in a mean-variance efficient portfolio. In addition, the MRAC agent may want to be exposed to risk even if there is no risk premium. The ($\alpha-$)maxmin investor never invests in risky securities when they do not provide compensation for risk. This does not depend on which beliefs see uses to perform mean-variance optimization.

6.4 Expected Utility Maximization

Under traditional expected utility preferences (such as quadratic utility used to justify mean-variance preferences), the investor's demand curve will generically be downward sloping, but never backward-bending (we mentioned some exceptions before). At zero risk premium, the risk-averse expected-utility maximizer will never want to be exposed to risk. MRAC behavior may be different in both respects, and not only in trivial cases.

7 Conclusion

We have presented a mathematical model of a satisficing agent based on a core concept in robust control theory, namely, Model Reference Based Adaptive Control (MRAC). The agent aims at minimizing expected surprise when targeting a particular expected return and return volatility. The agent often ends up choosing *as if* mean-variance optimizing, but there are well-defined situations in which she doesn't, which provides a unique opportunity to re-visit the many empirical anomalies recorded since the emergence of the first formal asset pricing model, the CAPM. Indeed, while MRAC agents may generate CAPM equilibrium without explicitly wanting to optimize, there are specific situations in which they produce very different choice. These include choice multiplicity, backward-bending demand, and demand for risk even in the absence of a risk premium.

We have been assuming that our MRAC investor directly minimizes expected surprise. As explained in Section 3, the MRAC investor could let an agent make the actual investment decisions, limiting interference by the delegate to choosing the right hyper-parameters of the agent, to assure surprise is minimized. The delegate itself could then be of any of the kinds we have been studying before. As an example, he could be of the $\alpha-$maxmin type while the investor (principal) controls the value of α. Or he could be a traditional expected utility maximizer and the principal controls his risk aversion. One more possibility is an AI (Artificially Intelligent) delegate who executes an online reinforcement learning algorithm with parameters (discount rate; learing rate) that the investor adapts to minimize expected surprise.

We have only looked at choices of the MRAC agent given her beliefs at a particular point in time, ignoring adaptation. We motivated this approach since many properties we obtained apply to a wide set of possible beliefs, and hence, apply even if beliefs – and hence choices – adapt. In future work, we plan to focus on choice adaptation itself: how do choices change when beliefs are updated? How is the reference model chosen to reach the overarching goal of robustness?

The MRAC literature has provided many examples of how to organize adaptation and how to ensure stability (convergence of control policy as more evidence of the true distribution of outcomes emerges). But the MRAC literature tends to focus on cases where uncertainty is limited to lack of knowledge of parameter values rather than the typical case in finance, where a random shock occurs the value of which will never be fully learnable, even if the latter does not cause harm because the impact of the shock is only temporary. Our future work therefore will aim at expanding MRAC algorithms in order to facilitate adaptive coding for finance applications.

Evidence is needed that human adaptation can be modeled more accurately using MRAC than traditional optimization-based reinforcement learning (RL) or Bayesian learning. First evidence in favor of MRAC is discussed in Bossaerts [2018]. There, a task is discussed where participants had to predict movements of a target that moved like a stock price: its movements were drawn from a leptokurtic distribution. Outliers either reverted or did not, and participants could easily figure out which treatment they were in. MRAC explained choices better than either RL or Bayesian updating. In particular, participants behaved as if minimizing surprise relative to a world without leptokurtosis and without reverting outliers. More recently, Pei [2024] presents evidence consistent with MRAC in a meta-analysis of learning in a collection of financial markets price prediction experiments.

We defined expected surprise as is standard in the MRAC literature, in terms of first and second moments of the distribution of outcomes. In finance, alternative expected surprise measures exist, such as VaR. As long as one stays in a gaussian world, we would argue that our metric of expected surprise encompasses these alternative measures, rendering our analysis more general than may appear at first. We leave analysis of the scope of our expected surprise metric for further research.

Finally, there is a need to better understand the commonalities and differences in the various strands of the literature on human behavior that appeal to the use of reference models in decision-making: efficient irrationality [Glimcher, 2022], active Bayesian inference [Bossaerts and Rayo, 2023], constrained data-fitting [Samuelson and Steiner, 2024], MRAC (this paper). In our setting, the agent chooses a reference model in order to ensure robustness. Elsewhere, the reference model is chosen because of alleged cognitive limitations that force the agent to choose a model that inherits as much information as possible compared to the objectively available data (active Bayesian inference; constrained data-fitting) or objectively optimal choices (efficient irrationality). Further investigation will require not only targeted controlled behavioral experiments, but also an exploration of the neural foundations of reference models in value-based decision-making.

References

F. Attneave. Some informational aspects of visual perception. *Psychological review*, 61(3):183, 1954.

R. Azeredo da Silveira and M. Woodford. Noisy memory and over-reaction to news. In *AEA Papers and Proceedings*, volume 109, pages 557–561. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203, 2019.

N. Barberis, M. Huang, and T. Santos. Prospect theory and asset prices. *The quarterly journal of economics*, 116(1):1–53, 2001.

H. B. Barlow et al. Possible principles underlying the transformation of sensory messages. *Sensory communication*, 1(01):217–233, 1961.

P. Bossaerts. Formalizing the function of anterior insula in rapid adaptation. *Frontiers in integrative neuroscience*, 12:61, 2018.

P. Bossaerts and L. Rayo. Modeling behavior when inference is complex. *Cambridge Working Paper*, 2023.

A. Caplin, M. Dean, and D. Martin. Search and satisficing. *American Economic Review*, 101(7):2899–2922, 2011.

K. Friston. The free-energy principle: a unified brain theory? *Nature reviews neuroscience*, 11(2): 127–138, 2010.

C. Frydman and L. J. Jin. Efficient coding and risky choice. *The Quarterly Journal of Economics*, 137 (1):161–213, 2022.

P. Ghirardato and M. Marinacci. Ambiguity made precise: A comparative foundation. *Journal of Economic Theory*, 102(2):251–289, 2002.

I. Gilboa and D. Schmeidler. Maxmin expected utility with non-unique prior. In *Uncertainty in economic theory*, pages 141–151. Routledge, 2004.

P. W. Glimcher. Efficiently irrational: deciphering the riddle of human choice. *Trends in cognitive sciences*, 26(8):669–687, 2022.

P. W. Glimcher and A. A. Tymula. Expected subjective value theory (esvt): a representation of decision under risk and certainty. *Journal of Economic Behavior & Organization*, 207:110–128, 2023.

L. P. Hansen and T. J. Sargent. Robust control and model uncertainty. *American Economic Review*, 91 (2):60–66, 2001.

D. Kahneman and A. Tversky. Prospect theory: An analysis of decision under risk. *Econometrica*, 47 (2):263–291, 1979.

I. D. Landau, R. Lozano, M. M'Saad, et al. *Adaptive control*, volume 51. New York: Springer, 1998.

J. Lanier. Risk, ambiguity, and giffen assets. *Journal of Economic Theory*, 186:104976, 2020.

L. Loued-Khenissi and K. Preuschoff. Information theoretic characterization of uncertainty distinguishes surprise from accuracy signals in the brain. *Frontiers in artificial intelligence*, 3:5, 2020.

K. Louie, M. W. Khaw, and P. W. Glimcher. Normalization is a general neural mechanism for context-dependent decision making. *Proceedings of the National Academy of Sciences*, 110(15):6139–6144, 2013.

C. Murawski and P. Bossaerts. How humans solve complex problems: The case of the knapsack problem. *Scientific reports*, 6:34851, 2016.

N. T. Nguyen. *Model-reference adaptive control*. New York: Springer, 2018.

E. Payzan-LeNestour and M. Woodford. Outlier blindness: A neurobiological foundation for neglect of financial risk. *Journal of Financial Economics*, 143(3):1316–1343, 2022.

J. Pei. Reference model based learning in expectation formation: Experimental evidence. *arXiv preprint arXiv:2404.08908*, 2024.

R. Polanía, M. Woodford, and C. C. Ruff. Efficient coding of subjective value. *Nature neuroscience*, 22 (1):134–142, 2019.

N. Sajid, P. J. Ball, T. Parr, and K. J. Friston. Active inference: demystified and compared. *Neural computation*, 33(3):674–712, 2021.

L. Samuelson and J. Steiner. Constrained data-fitters. *Yale university working paper*, 2024.

M. J. Schervish. *Theory of statistics*. Springer Science & Business Media, 2012.

H. A. Simon. A behavioral model of rational choice. *The quarterly journal of economics*, pages 99–118, 1955.

H. A. Simon. Theories of decision-making in economics and behavioral science. *The American Economic Review*, 49(3):253–283, 1959.

A. Tversky and D. Kahneman. Judgment under uncertainty: Heuristics and biases: Biases in judgments reveal some heuristics of thinking under uncertainty. *science*, 185(4157):1124–1131, 1974.

F. M. Vieider. Decisions under uncertainty as bayesian inference on choice options. *Management Science*, 2024.

H. Yamada, K. Louie, A. Tymula, and P. W. Glimcher. Free choice shapes normalized value signals in medial orbitofrontal cortex. *Nature communications*, 9(1):162, 2018.