A Theory of Complexity Aversion*

Xavier Gabaix

August 21, 2026

Abstract

This paper proposes a tractable model of "complexity aversion." The key ingredient is first-order complexity aversion: when people know they are making a mistake because the situation is complex, they experience dread—a utility loss proportional to the expected size of the error in their action. This adds a new complexity-aversion term to the traditional utility function, governed by a single parameter that appears to be portable across contexts. This simple enrichment has a range of implications, which I illustrate: (i) Whereas the traditional model predicts that contracts should be indexed to aggregate factors (e.g., inflation), with sufficient complexity aversion contracts become non-indexed. (ii) If complexity aversion is high enough, the price of a good will be constant over time—even when marginal cost varies—to avoid burdening consumers with a complex pricing system. (iii) Traditional economics often recommends tax systems far more complex than what we observe. With complexity aversion, the optimal tax system is simpler—e.g. has uniform taxes or a few tax brackets. (iv) Complexity leads to no, or very delayed, decisions, and choice overload. Models based only on information costs yield a muted but nonzero response to incentives; in contrast, when it is large enough, first-order complexity aversion generates no response at all, "full simplicity". Finally, I discuss how this model of complexity aversion can underpin a useful "behavioral mechanism design," leading to more realistic—and simpler—mechanisms.

*xgabaix@fas.harvard.edu. For helpful comments, I thank Abhijit Banerjee, Oren Bar-Gill, John Cochrane, Stefano DellaVigna, Ben Enke, Ed Glaeser, Thomas Graeber, Brian Hall, Oliver Hart, Louis Kaplow, David Laibson, Giacomo Lanzani, Shengwu Li, Florencia Marotta-Wurgler, Ulrike Malmendier, Eric Maskin, Emi Nakamura, Indira Puri, Alex Rees-Jones, Josh Schwartzstein, Andrei Shleifer, Bennett Smith-Worthington, Kathy Spier, Jeremy Stein, Jón Steinsson, Dmitry Taubinsky, 10amDhara Yu, Ariel Zetlin-Jones, and participants at various seminars. I gratefully acknowledge financial support from the Ferrante Economics Research Fund.

1 Introduction

I propose a tractable notion of "complexity aversion" and explore its implications for "simple mechanisms" in several settings: simple (non-indexed) contracts, simple price systems, simple taxes.[1]

Actual economic systems are typically much "simpler" than a frictionless rational model would predict. For instance, basic economic theory suggests that most contracts should be indexed to aggregate risks—such as inflation, GDP, or the stock market—because this reduces the risk faced by consumers. But that is typically not the case in practice. Intuitively, this is because doing so would be "too complicated." The model formalizes this and, sensibly, implies that contracts will generally not be indexed unless the risk-reduction effects are large enough.

Likewise, prices for Whole Foods delivery are the same across hours and days of the week, even though they "should" vary in a frictionless world, as the marginal cost of deliveries changes over the day.[2] A barber typically charges prices that do not depend on the time of day, even though the 5pm slot is much more popular than the 2pm slot, so (generically, at least) their prices should differ.

The core contribution is a broadly applicable model of "complexity aversion." I attend to the psychology while ensuring the model is sufficiently general to be portable across contexts. I also show that other devices, such as information costs, or smooth ambiguity aversion, cannot generate "full simplicity": for example, they predict some indexation rather than none.

Two key properties of "first-order complexity aversion" yield full simplicity. The model is essentially a half-equation: a complexity aversion cost which is subtracted from the traditional utility function:

$$C^{\text{CA}} = \overline{\xi} \sum_i \left| a^r_{x_i} \right| \sigma_{x_i} \left(1 - m_i\right) \tag{1}$$

Let me explain its main features. An agent has a problem which rationally would be $\max_a u(a, x)$, where x is a vector of disturbances with mean 0. I call $a^r_{x_i}$ the rational slope: a fully rational agent would react to x_i by an amount $a^r_{x_i} x_i$, to the leading order for a small x_i. As σ_{x_i} is the typical size (standard deviation) of x_i, $\left| a^r_{x_i} \right| \sigma_{x_i}$ gives the typical size of the

[1] A companion paper explores the complexity cost of inflation and the business cycle, and the resulting optimal monetary policy.

[2] In the spring 2024, this was \$9.99 for 2-hour slots, e.g. 6-8am, 8-10am, etc. There are also 1-hour slots (6-7am, 7-8am etc.), that all cost \$14.99.

impact of x_i on the action—it is the importance of x_i for a rational agent.

However, the agent does not fully understand x_i, or how to react to it. I call (and microfound) $m_i \in [0,1]$ the degree of understanding or attention to x_i. Here $m_i = 1$ means that the agent fully understands x_i, and reacts to it rationally; conversely $m_i = 0$ means the agent does not understand x_i, and x_i is entirely not paid attention to. Hence, $e_i \coloneqq \left|a^r_{x_i}\right| \sigma_{x_i} (1 - m_i)$ is the typical size of the mistake, in action space.

The complexity aversion cost (1) is first order in the size of the mistake e_i—rather than second order, as the traditional utility consequences are. This "first-order complexity aversion" has a worthy psychological motivation: it is a close cousin, or perhaps indeed an offspring, of the "first-order risk aversion" found necessary to explain the behavior under risk (Kahneman and Tversky, 1979; Rabin, 2000). It is Kahneman and Tversky (1979)'s first-order risk aversion, but for the risks of making a mistake "in the inner world" of the agent, rather than (as Kahneman and Tversky) for the risks in the outside world.

The agent experiences a complexity "dread", the sense that she is making a mistake, which directly lowers her utility (above the traditional cost from lowering one's utility because of a mistake). Unlike, say, cognition or information costs, this "dread" cost is paid even when attention is 0 (whereas with zero attention, no cognition cost is paid). Hence, even for a very behavioral consumer who pays zero thinking costs, lower complexity (and the scope for mistake) will be useful to lower her dread. So cognitive costs and complexity dread are distinct: cognitive costs generate mistakes, which in turn generate complexity aversion. Paying higher mental effort (increasing m_i) decreases complexity aversion costs, but increases cognition costs.

This first-order complexity aversion has sharp consequences, forcing simplicity in economic outcomes. Pure information economics would not lead to full simplicity, because it would lead to a second-order penalty for small mistakes (proportional to e_i^2). To clarify the discussion, I submit criteria that, I think, a complexity aversion cost should satisfy. I find that all the formulations satisfying them are (1), or a very minor variant of it, replacing $1 - m_i$ by a general function $h(m_i)$, which basically would not change the core economics. So not only does (1) have good psychological motivation, it also satisfies a few criteria that I think are reasonable.

The main virtue of (1) is not only its psychological motivation, its general applicability, and its good properties with respect to those criteria: it is also its ability to tractably generate simple economic outcomes. I turn to those.

Complexity aversion, when high enough, leads to very simple optimal mechanisms

Contracts are not indexed. I start with non-indexed contracts. Indexing reduces risk, but also (if the index pays off at some point in the future) creates extra complexity, as a consumer should keep track of the value of the index. Hence, if complexity aversion is high enough, no indexation is better. To illustrate this, I take the simplest contract — the choice of buying a hedge, or not. This is just a solo decision problem (against nature), and this way I eschew all complications of strategic interactions.

Price schemes are more rigid. When complexity is high enough, the price of a good (e.g. Whole Foods deliveries) will be constant across time-periods, even though the marginal cost might be variable. Intuitively, this is to avoid annoying the consumer with a very complex price system.

A price system is optimally simpler when complexity aversion is higher, when the complexity of thinking about the price is higher, when the marginal cost is less variable, and when the price elasticity is lower (which often means that competition is lower). So Whole Foods deliveries will have a constant price (low elasticity of demand), while the price of strawberries will be more variable (as the marginal cost is more variable, and the price elasticity is higher). This result of a fully simple price system may seem intuitive, but it does require some care in the modeling of complexity aversion.

Optimal tax systems are much "simpler" than in the traditional model, e.g. with one or a few tax brackets, rather than a different tax for each good and state of the world. I revisit the (Ramsey) theory of optimal goods taxation, with now complexity-averse agents. In the traditional Ramsey model, the optimal tax on each good is (generically) different for each good (as it is inversely proportional to the elasticity of demand for the good). This is ridiculously complex—with ten thousand different goods, rational theory wants ten thousand different tax rates. This is probably one of the reasons why those strict recommendations of optimal taxation theory are scarcely followed in practice. With complexity-averse agents, things are much more sensible. Indeed, if complexity aversion is high enough, the optimal tax system is "simple", e.g. just features a uniform tax rate, and with milder complexity aversion, we sensibly obtain say a few tax brackets.[3]

Complexity aversion creates "aversion to deciding" and "choice overload". If an "active decision" (e.g., rebalance one's portfolio) entails complexity aversion, while a "passive course"

[3]Generating simple taxes is non-trivial. One of our original hopes, when working on Farhi and Gabaix (2020), was that optimal taxes with behavioral agents would be simple—but, to our disappointment, they were not: optimal taxes in the Ramsey problem were changed by limited attention, but were not simple, e.g. uniform. The new ingredient of first-order complexity aversion is what allows the present paper to obtain simple taxes.

(e.g. do nothing) does not, then when complexity aversion is high enough, and the decision is hard enough, people will stay passive.[4] Likewise, under conditions detailed below, the model delivers "choice overload": faced with many confusing options, people will often choose the simple option.

Unified calibration of complexity aversion. At the core of the model, there is a unitless parameter ξ that calibrates the strength of complexity aversion (it is a unitless counterpart of the $\overline{\xi}$ in (1)). Drawing from the analogy with prospect theory, I calibrate it to $\xi \simeq \frac{2}{3}$. Then, I find that this calibration works, or at least, gives quite reasonable orders of magnitude, for a wide range of empirical phenomena, including the aversion to indexation, the rigidity of prices, the simplicity of tax systems, and choice overload. Hence, the coefficient of complexity aversion may be a new, portable behavioral parameter.

The advantage of a unified framework. I found it useful to have several applications, for three reasons. First, to verify that the theory is expressive enough, and simple enough, that it can be used in a number of situations. Second, for a given situation, there could be a number of explanations: for instance, the rigidity of prices might come from fairness (Eyster et al., 2020). But fairness, presumably, cannot explain why consumers do not wish to index contracts. Likewise, the lack of indexation might be due to the fact that the bank, say, knows more than the consumer (Spier (1992), Hartman-Glaser and Hébert (2020)), but that does not explain why Whole Foods would have uniform delivery prices across time windows. So it is useful to show how complexity sheds light on outcomes across a domain of situations. Third, it is intriguing that the same complexity aversion parameter seems to work across different domains.

I view these results as potentially useful since they give the hope that one can explain existing institutions (and their degree of "simplicity"). The underlying model yields a principled way to do "behavioral mechanism design", i.e. design optimal institutions (e.g. price or tax systems) that aim both at high allocative efficiency and not too high complexity. This will augment the power of mechanism design, which is often hampered by the fact that it generates too complex recommendations to be used in practice.

Related literature There is relatively little work on explaining why contracts are simple. One is Tirole (2009), which offers a cognitive interpretation of the incomplete contracts with renegotiation à la Hart (1995), but not why simple, static contracts are "simple". Another literature, on the lack of indexation, grounds it on asymmetric information: the person who proposes to index the contract may have more information than the one who has to

[4]Years ago, I described this to Robert Lucas. He told me that indeed he never rebalanced his retirement portfolio, and just let it drift. In his worldly life, the great theorist of rational behavior acted like the simple complexity averse agent modeled here.

accept or reject the indexation (Spier (1992), Hartman-Glaser and Hébert (2020)).[5] The present mechanism is very different, and centered on complexity, rather than asymmetric information: the model presented here predicts that even without asymmetric information, there will still be a lack of indexation. Carroll (2015) proposes conditions under which optimal contracts are linear, using robustness under the worst-case behavior, which is a form of infinite risk aversion. In contrast, this paper uses "smoother" preferences, and emphasizes complexity aversion.

On the theoretical side, this work relates to a number of attacks on complexity. Some are conceptually interesting, but sadly not very tractable, e.g. using discrete automata (Rubinstein, 1998), or sequences of discrete decisions (Gabaix et al., 2006). This paper is closer to the literature on limited attention, and crucially it adds a new, "complexity aversion" cost to the psychological description of the economic agent. I draw on earlier work (Gabaix, 2014), whose emphasis on "sparsity" I share — but the technology in that paper does not generally generate simple contracts. This model is meant to be a reasonable compromise between naive bounded rationality and fully Bayesian information processing, and is very tractable, which allows it to give in a unified manner a behavioral version of basic micro and Arrow-Debreu (Gabaix (2014)), classical taxation (Farhi and Gabaix (2020)), and monetary/fiscal policy (Gabaix (2020)). Other models of attention or skewed perception of the world include Bordalo et al. (2027), Gagnon-Bartsch et al. (2026), Schwartzstein and Sunderam (2021). At the opposite end, a fully Bayesian processing approach is developed in a series of papers by Sims (2003), Matějka and McKay (2015), Caplin et al. (2020), Khaw et al. (2021). The analysis I propose here could be applied to those styles of models. A cumulative body of papers shows that inattention (in particular in its behavioral flavor) is a good candidate for one unifying force in behavioral economics (see Gabaix (2019) for a survey/manifesto, and Enke et al. (forthcoming) for recent systematic evidence).

de Clippel et al. (2025) provide evidence qualitatively supportive of the cost I model here. In their large experiments, people assign a lower dollar value to an option because they find it complex. In their setup, if a person held any posterior and maximized expected utility, she could not value both of two complementary bets below a 50/50 lottery. So they show a distaste for the complexity itself. They suggest that a model subtracting a first-order complexity cost from expected utility would fit their evidence, and point to Gabaix and Graeber (2024).[6] The present paper develops such a model and puts it to work.

[5]Those asymmetric information theories predict "stickiness of institutions" rather than simplicity per se. For instance, if the social norm was, for some reason, to index contracts by 2/3 of inflation, then any deviation from that would be "suspicious": that arrangement would be hard to get away from, and society would be stuck with indexing at 2/3 of inflation.

[6]Gabaix and Graeber (2024) actually explores a different point, the impact of compositional complexity,

This paper belongs to the "behavioral theory" tradition, but it has some connection with the "decision theory" line, in particular with its emphasis on complexity. It relies on a continuous measure of complexity, different from the 0/1 measure in Li (2017) and the integer measure in Nagel and Saitto (2027). It relates to Puri (2025)'s axiomatic analysis of complexity, in the context of lotteries.[7] The setups are extremely different: I focus on continuous actions and continuous variables, while those papers focus on choice over lotteries, or discrete actions in games. One may hope that the two lines of research can be unified, but I do not pursue that here. Fudenberg and Mudekereza (2026) propose a model where "complexity" is the entropy of the beliefs of the agent within a robustness framework, again in a setup with discrete actions. They also show how to modify their model to import this paper's first-order penalty, and to replicate its reduced forms—offering some hope of unification.

Within decision theory, complexity has been studied through preferences over menus. Ortoleva (2013) takes preferences over lotteries of menus as primitive and axiomatizes "thinking aversion," representing the agent as an affine evaluation of a menu's content net of an anticipated cost of choosing from it; relatedly, Ergin and Sarver (2010) model costly contemplation, and Sarver (2008) anticipated regret. The present paper differs in its primitive and its target: I model an agent facing a decision problem $\max_a u(a, x)$ and price the complexity of the policy $a(\cdot)$, which allows me to take the model to standard economic environments (indexation, price schedules, Ramsey taxation). Those papers however do not trace this phenomenon to some "mistake aversion". It would be interesting to revisit them with that angle, especially as "complexity aversion" and mistake aversion might be (in part) a way to generate risk aversion.

There is a recent line of work on complexity, which is largely empirical. A growing body of work measures how "intuitively complex" setups lead to more "mistakes", or more precisely, deviations from a baseline simple economic model (Martínez-Marquina et al., 2019; Oprea, 2020). Huck and Weizsäcker (1999), Fudenberg and Puri (2022), Enke and Shubatt (2023), Goodman and Puri (2025) find lotteries are less tempting when they have more outcomes (so are more "complex"). Colliard and Georg (2025) measures the complexity of rules via text analysis.

In law and economics, contracts tend to be "simple", standardized and sticky (Hart (1995), Bar-Gill (2012), Gulati and Scott (2012), Dari-Mattiacci and Marotta-Wurgler (2022)).[8] This is true for the "shape" of the contracts, but also for some important economic features.

its measurement, and its consequences for the cost of cognition.

[7] That influential paper started in Puri (2018).

[8] This literature emphasizes the "legal certainty" of using a conventional "boilerplate" contract—prior litigation has stabilized its interpretation; and the transaction costs of simply using an existing template.

For instance, many contracts are "rigid in percentage space": historically, most hedge funds have charged "2 and 20"— 2% of the assets under management, plus 20% of the upside, in a rigid manner. Likewise, sharecropping contracts tend to have a rigid split (often, 50%) between landlord and tenants (Young and Burke, 2001). In finance, firms and analysts often use simple, quite rigid rules of thumb to assess cost of capital, and valuations (Graham, 2022; Gormsen and Huber, 2025; Ben-David and Chinco, 2024). The present model might apply to those setups.

Outline The paper first lays out the basic model, then explores successively: aversion to indexation, simple price schemes, simple tax systems, and choice aversion and choice overload. The first three applications are ordered by who chooses the degree of complexity: a person deciding for herself (indexation), then the market mechanism making a firm decide for its customers (pricing), then a planner deciding for everyone (Ramsey taxes). The paper then illustrates complexity aversion in a discrete-choice setting, for two well-documented phenomena: inertia and choice overload. Many proofs are in the appendix.

2 Complexity aversion: the core model

2.1 Background material: A model of imperfect thinking

I first review a model that does not feature complexity aversion yet (it is laid out in Gabaix (2014) and Gabaix and Graeber (2024), with an easy review in Gabaix (2019), but the presentation here is self-contained). The task is to maximize over a continuous action a an objective function $u(a, x)$, where u is smooth (twice continuously differentiable) and concave in a. The vector of disturbances x is drawn from a distribution with mean normalized to 0,[9] and its components are perceived by the agent to be uncorrelated. Action a is just a scalar for now, but it is easy to extend to a multidimensional a. I call $a^r(x) = \text{argmax}_a u(a, x)$ the rational action. I call default action $a^d = a^r(0)$ the optimal action "at the default", i.e. when all x are equal to 0, and I normalize it to 0; $a^r_{x_i}$ is the partial derivative at a default. The rational answer is thus (after linearization, so up to second-order terms in x_i, assuming that the deviations x_i are "small")

$$a^r = \sum_i a^r_{x_i} x_i \tag{2}$$

[9] Otherwise, the same reasoning applies to $x_i - x_i^d$, where x_i^d is a default value.

The agent's objective is to cognitively construct those $y_i \coloneqq a^r_{x_i} x_i$, which indicates by how much dimension i of the problem should change the rational action.

The agent does not perceive x_i exactly, but rather has a subjective, imperfect perception of it x^s_i:[10]

$$x^s_i = m_i x_i + \eta_i \tag{3}$$

where $m_i \in [0,1]$ is the precision of the signal, and $\sigma^2_{\eta_i} = m_i (1 - m_i) \sigma^2_{x_i}$.[11] If all the shocks are jointly Gaussian, with η_i mean-zero and uncorrelated with x_i, we have $\mathbb{E}[x_i | x^s_i] = x^s_i$. The rational case corresponds to $m_i = 1$, and the extremely inattentive case, to $m_i = 0$.[12] The s signifies "using subjective perceptions", or sometimes "as a sparsity-seeking agent" (who economizes on attention m). Accordingly, I posit that if the decision maker sees x^s, the subjectively perceived value of x, she takes the decision $a^r(x^s)$. Linearizing:

$$a^s = \sum_i a^r_{x_i} x^s_i.$$

The traditional utility losses from an imperfect decision are then, to second order and assuming uncorrelated innovations to all variables:

$$\mathbb{E}[u(a^r, x)] - \mathbb{E}[u(a^s, x)] = \frac{1}{2} \left| u''\left(a^d\right) \right| \sum_i \left(a^r_{x_i} \sigma_{x_i}\right)^2 (1 - m_i) \tag{4}$$

Attention m is easy to endogenize, with a cognition cost $C^{\text{Cognition}}$, but most of the economics in this paper obtains with exogenous attention.

2.2 Complexity aversion: The innovation in this paper

The consumer's full utility, given attention policy m, is:

$$V(m) = \mathbb{E}[u(a^s, x)] - C^{\text{Cognition}}(m) - C^{\text{CA}}(m) \tag{5}$$

where $a^s = a^r(x^s(x, m))$ is the action given the perceived x (with $a^r(x)$ the rational action given x) and where $C^{\text{Cognition}}(m)$ is a cognitive cost.

[10] Another interpretation is that the agent does not know well how to react to x_i, i.e. does not know $a^r_{x_i}$.

[11] This comes from a more conventional "signal = truth + noise" framework, but here I reason directly with a post-shrinkage signal (so "signal = $m_i \times$truth + noise") which gives the functional form for $\sigma^2_{\eta_i}$. Section A.1 gives the derivation.

[12] I do not assume that agents are Bayesian (as traditional information economics)—instead, I simply use that benchmark as an inspiration for the model (as e.g. in Gabaix, 2014, Woodford, 2020). For instance, when y_1 and y_2 are correlated, a Bayesian agent would use $\mathbb{E}[y_1 | y^s_1, y^s_2]$ to infer y_1, but instead I model the agent as being a "limited Bayesian", who simply performs $\mathbb{E}[y_1 | y^s_1]$.

The key new term in this paper is the "complexity aversion" term (1), and more fully specified here:

$$C^{\text{CA}} = \xi \left| s_a u_{aa} \right| \sum_i \left| a^r_{x_i} \right| \sigma_{x_i} \left(1 - m_i\right) \tag{6}$$

I call $\xi \geq 0$ the parameter of complexity aversion, which is unitless. This makes it potentially portable across settings, something I explore in Section 7.1.

Motivation for the formulation of complexity aversion (6) Let me motivate it slowly, starting with a simple, intuitive version. I start with one disturbance, x. The agent feels she is making an error of size e (in absolute value): this is an error in action space, i.e. a sense of the distance between her action and the optimal action. Then, the CA is

$$C^{\text{CA},1} = \bar{\xi} e$$

for some parameter $\bar{\xi}$ that I will detail shortly. This is "first-order mistake aversion" (hence complexity aversion), because the penalty is proportional to e, rather than e^2.

The rational answer is, to the leading order, $a^r = a^r_x x$; while the actual action is $a^s = a^r_x x^s$. So the average error is:

$$a^\delta \coloneqq \mathbb{E}\left[a^r - a^s | x\right] = a^r \left(1 - m\right)$$

and the typical error size is[13]

$$e \coloneqq \sigma_{a^\delta} = \left| a^r_x \right| \sigma_x \left(1 - m\right)$$

Now, let us turn to the units. In (5), C^{CA} is in utils, $[u]$ (where brackets mean "units of") while the error is in units of action, $[a]$. So, to keep the formulation as invariant as possible, I set $\bar{\xi} = \xi \kappa^u_a$, where ξ is a unitless parameter that can be the same across problems, and κ^u_a fixes the units: it has the units of utils over action. To proceed at a high level of generality, there are not many choices. I take $\kappa^u_a = s_a \left| u_{aa} \right|$, where s_a is the "natural scale" of a. Typically, it will be the typical "magnitude" of the action—e.g. the typical size of consumption of the good. In the analysis below, the specific value of s_a will play almost no role, except to fix the orders of magnitude.

[13] I use "typical error size" with some mild abuse of language. Another measure would be the root-mean-square error, $\left(\mathbb{E}\left[\left(a^r - a^s\right)^2\right]\right)^{1/2} = \left| a^r_x \right| \sigma_x \sqrt{1 - m}$. This is still linear in the stakes $\left| a^r_x \right| \sigma_x$, which is the important part, though the last term is $\sqrt{1 - m}$ rather than $1 - m$. That square root makes things less tractable, but the important economics would not change. Section A.4 discusses such small variants.

Hence, we arrive at the specification $C^{\text{CA}} = \bar{\xi} e = \xi \kappa_a^u \sigma_{a^\delta}$, hence:

$$C^{\text{CA}} = \xi \left| s_a u_{aa} \right| \left| a_x^r \right| \sigma_x \left(1 - m \right) \tag{7}$$

The "degree of complexity aversion" ξ is unitless. Later I will propose to calibrate it to $\xi \simeq 2/3$. Importantly for portability, C^{CA} is invariant if we change the units of a or x, and scales correctly if we change the units of u.[14] In that sense, we achieve a formulation that gives a reasonable degree of invariance.

To build the key intuition, consider a very boundedly rational agent that has $m = 0$. Then, the cognition cost is 0 ($C^{\text{Cognition}}(m) = 0$ in (5)), but the complexity aversion cost C^{CA} is strictly positive (see (7)), as the agent is aware enough that she is making a mistake—though not analytical enough to know how to act rationally.[15] Hence, there is an incentive for the planner to create a "simple scheme", that will not create too much "dread", i.e. complexity aversion. Indeed, adding complexity will not guide the very boundedly rational agent towards greater allocative efficiency (as $m = 0$, so she does not respond to incentives), but it diminishes her utility, via the dread term. Adding complexity just adds deadweight loss.

On the other hand, when the consumer understands fully the situation (so $m_i = 1$), complexity aversion becomes 0 in (7): we recover the rational case, at least for "easy" decisions.

To go from a one-dimensional to a multidimensional x, we take (6): the complexity aversion is simply the sum over each dimension x_i.[16]

Psychological foundation: Complexity aversion as a form of loss aversion for risk in the inner world of the mind, rather than in the outside world Complexity lowers utility via two channels in (5). First, it requires extra thinking (or in other language, more information). Second, if the agent makes a mistake (or can fear she is making a mistake), complexity leads to first-order "mistake aversion cost".

One can have the first without the second: suppose that entering the decision requires

[14] For instance, if we change x from units of meters to units of centimeters, then x is multiplied by 100, but $|a_x^r|$ is divided by 100, so that $|a_x^r| \sigma_x$ is invariant by the change in units of x.

[15] See Enke and Graeber (2023) for evidence that people have a good sense of their cognitive uncertainty, and Section A.9 for a simple formulation with partial awareness.

[16] The complexity measure hence depends on the "frame" of the problem—e.g. which x or, rotation of the x vector is presented to the agent. This is realistic: when the problem is presented with useful summary statistics, rather than as a unedited deluge of numbers; or when it is in nominal vs real terms (plus inflation), in this model, as surely in the real world, agents' decision, and feeling of complexity, will vary. More technically, Section A.2 offers a way to recover the frame that agents use—so the frame is not purely based on intuition, but also could be elicited. An interesting extension would be to allow the planner to choose the frame (as in Li, 2017, Gonczarowski et al., 2023).

each time filling a long form; this is a transaction cost akin to cognition cost. But, if the task is straightforward enough (though burdensome), there will be no mistake, and no complexity aversion cost.

This first-order mistake aversion, for the risk of a mistake in the inner world of the decision maker, is best seen as the mirror of Kahneman and Tversky (1979)'s first-order risk aversion for gambles in the outside world. Recall that in decision-making, people tend to reject a gamble yielding with equal probability +\$10 or −\$9. More generally, they reject a gamble with standard deviation σ when the risk premium is less than $\lambda\sigma$, for some λ around 0.2 to 0.6 (Kahneman and Tversky, 1979). This "first-order risk aversion" (where the risk penalty is proportional to the standard deviation) has been found many times to describe people better than the rational model's prediction of second-order risk aversion (where the risk penalty is proportional to the squared standard deviation); see e.g. Rabin (2000). This idea motivates a calibration, in Section 7.1.

Should this complexity aversion term be used in welfare? At first, this is not obvious, as many behavioral deviations are arguably just "mistakes", and would not count in welfare. I think that a fair delineation is the following.[17]

1. If there are cognition costs, or emotional costs while making a decision, they are welfare-relevant. The dread cost is a form of emotional cost, so it is welfare-relevant.

2. If the agent misperceives the future utility $\mathbb{E}[u(a)]$ (i.e. forgets that smoking is bad for her) the welfare-relevant utility is the expected utility under the correct beliefs.[18]

Hence, I will take here dread to be welfare-relevant.

I note that in some cases it does not matter, because there is no "planner", and we are just computing the agent's decision, which involves a subjective utility. For instance, in solo decision-making problems (Sections 3 and 6), the agent wants to optimize her perceived welfare (including dread), so taking dread into account is necessary for the positive analysis. In the baguette example of Section 4, the consumer chooses the best bakery "all inclusive" (including low price and low complexity cost), so to derive her decision, calculating total surplus, inclusive of dread, is necessary. The only delicate case might be the Ramsey planning case of Section 5. There, however, it is plain, I think, that planners should take into account the complexity costs associated with the tax systems—both normatively, and positively, as

[17]For more discussions of welfare in behavioral economics, see Bernheim and Taubinsky (2018) and Farhi and Gabaix (2020).

[18]With "anticipatory" utility, this is less clean—one would like to take into account the anticipatory utility at time 0, and, mostly, the rational utility at time 1.

they might be pressured by complexity-averse voters. I conclude that complexity aversion should count in welfare.

Why this formulation? (preview) The formulation has some advantages, beyond the psychological connection with prospect theory. It is invariant to the units of action, variables x_i, and homogeneous of degree 1 in the units of utility, which makes it potentially portable. It also generates simple mechanisms—as we shall see, but keeps the CA penalty finite and moderate in all situations, including when there is a myriad of small elements that rationally should matter a bit. This is formally discussed in Section 7.2-7.3, and it makes more explicit why CA, as formulated here, has good properties, and in some sense is the unique model with those properties.

Link with the Lasso The complexity cost (6) is a Lasso-like (Tibshirani, 1996) regularization penalty, that guides towards a simple optimal policy for the agent: this is the term $\left|a^r_{x_i}\right|$.[19] The term $(1-m_i)$ ensures that obvious or familiar features do not exhibit a complexity-aversion (they have $m_i = 1$). This is quite different from a Lasso penalty in information space (as opposed to action space), which would look like $\sum_i |m_i|$, without the crucial term $\left|a^r_{x_i}\right|$.[20]

Extensions and other alternative formulations The online appendix develops CA for multidimensional actions and problems with budget constraints (Section A.3) and for discrete choice (Section A.5). Section 7.3 contrasts the basic formulation with other formulations, e.g. that do not generate full simplicity.

But first I present a few applications, which will give a good feel for the concrete economics of the model.

3 Aversion to indexation

I present the application of the general model to a simple situation: the aversion to indexation. It provides some key economics of complexity aversion, valid in later applications.

When a risk-averse agent deals with a risk-neutral firm, in basic rational economic theory, contracts should be indexed, to remove some extraneous risk from the consumer: for instance, debt or mortgage contracts should be indexed on inflation, or on other macroeconomic variables. However, the stylized facts on indexation are arguably the following. (i) Most

[19] Indeed, in the applications, the agent, or a planner, will de facto choose $a^r_{x_i}$.

[20] That penalty $\sum_i |m_i|$ was however important in Gabaix (2014).

consumer contracts are not indexed, contradicting the rational prediction. (ii) On the other hand, in extreme cases, e.g. hyperinflations, contracts become indexed again.

Intuitively, the reason for the lack of indexation is that an indexed contract would be too "complicated", compared to the traditional gains (lower risks). I formalize that now.

3.1 Basic analysis

There are three periods, $t = 0, 1, 2$. Consumption happens at times 1, 2, and utility is

$$V = v(c_1) + v(c_2), \qquad v(c) = \left(c - \frac{\Gamma}{2}c^2\right)/\Gamma \tag{8}$$

The consumer receives income at time 1 only, which is $Y_1 = 2(\bar{y} + \hat{y})$, where $\hat{y}$ is a mean 0 random variable. The real interest rate is 0.

At time 0, the consumer can choose to "index", i.e. to enter into a contract adding $-2bx$, which is received at time 2, where b is the consumer's choice of indexation, and x is a random variable, with mean 0 and variance 1, positively correlated with $\hat{y}$. So, her full intertemporal income is

$$Y(b) = Y_1 - 2bx = 2(\bar{y} + \hat{y} - bx)$$

The action a is the time-1 consumption. So full utility is

$$u(a, b) = v(a) + v(Y(b) - a) \tag{9}$$

The value of $\hat{y}$ and x is announced at time 1. The rational decision at time 1 is $a^r = \frac{Y}{2}$, i.e.

$$a^r = \bar{y} + \hat{y} - bx \tag{10}$$

The expected utility is

$$u^r = \mathbb{E}_0[u(a)] = 2v(\bar{y}) - \text{var}(\hat{y} - bx) \tag{11}$$

Hence, to maximize utility $\mathbb{E}[u(a, b)]$ from the point of view of time 0, a rational consumer would minimize the variance of income, i.e. choose the rationally optimal indexation level: $b^r = \text{argmin}_b \text{var}(\hat{y} - bx)$, i.e.[21]

$$b^r = \text{cov}(\hat{y}, x) \geq 0 \tag{12}$$

But intuitively, a behavioral consumer will dislike the additional complexity of having to

[21] Indeed: $\text{var}(\hat{y} - bx) = \sigma_y^2 - 2\,\text{cov}(\hat{y}, x)\,b + b^2\sigma_x^2$. Optimizing over b, using $\sigma_x = 1$, gives the announced result, $b^r = \text{cov}(\hat{y}, x)$.

think through bx at time 1 (the values of $\hat{y}$, x are realized at time 1). To model this, let us say that at time 1, her time-1 income y_1 is visible at no cost; but anticipating the value of the hedge ($-2bx$) is effortful and mistake-prone: it is knowable at time 1, but received at time 2.[22] The perceived value of the hedge at time 1 is $x^s = mx + \varepsilon$ (with ε uncorrelated with x and $\hat{y}$, and $\sigma_\varepsilon^2 = m(1-m)\sigma_x^2$, see (3)). Hence the consumer chooses the time-1 consumption:

$$a^s = \bar{y} + \hat{y} - bx^s$$

So, her utility is

$$V(b) = \mathbb{E}[u(a^s, b)] - C^{\text{CA}} \tag{13}$$

This leads to the following explicit utility. To simplify the expressions, I normalize $s_a = 1$—the reader can otherwise replace all the ξ by ξs_a.

Proposition 1. (Indexation example: Consumer's utility with complexity aversion) *The expected utility of the consumer choosing indexation level b is:*

$$V(b) = V(0) + \underbrace{(2b^r b - b^2)}_{\substack{\textit{Allocative gains} \\ \textit{from indexation}}} - \underbrace{(1-m)b^2}_{\substack{\textit{2nd order losses} \\ \textit{from mistakes}}} - \underbrace{2\xi|b|(1-m)}_{\substack{\textit{1st order losses} \\ \textit{from CA}}} \tag{14}$$

where $V(0) = 2v(\bar{y}) - \sigma_y^2$ is the utility without indexation, $b^r = \text{cov}(\hat{y}, x)$ is the optimal hedge for a rational consumer, and m is the consumer's attention to the future income from the hedge.

The term $2b^r b - b^2$ is the rational benefit of indexation for a rational consumer, and it is maximized for $b = b^r$: more generally, this represents the traditional allocative gains from having a more complex scheme. The term $-(1-m)b^2$ is the loss coming from the fact that the consumer does not understand the hedge well at time 1, hence does not choose the optimal consumption. The term $2\xi|b|(1-m)$ is the complexity aversion penalty.

Proof. I derive it slowly for clarity. The full consumption utility is in (13), which can also be written as:

$$V(b) = u^r + \frac{1}{2}u_{aa}\mathbb{E}\left[(a_x x)^2\right](1-m) - C^{\text{CA}} \tag{15}$$

i.e. the full rational utility, minus the misoptimization costs due to limited attention m, minus the complexity aversion dread. We calculate them in turn. The rational utility (11) is:

$$u^r = 2v(\bar{y}) - \text{var}(\hat{y} - bx) = 2v(\bar{y}) - \left(\sigma_y^2 - 2b\,\text{cov}(\hat{y}, x) + b^2\right) = V^0 + 2bb^r - b^2$$

[22] One could expand to the case where $\hat{y}$ also is hard to see, without changing the conclusions.

where $V^0 = 2v(\bar{y}) - \sigma_y^2$ is the utility with no hedge.

The misallocation term is (given that $a_x = -b$, and $u_{aa} = 2v''(c) = -2$):

$$\frac{1}{2}u_{aa}\mathbb{E}\left[(a_x x)^2\right](1-m) = -b^2(1-m) \tag{16}$$

The complexity aversion term is (6) with just one variable x :[23]

$$C^{\text{CA}} = \xi\left|s_a u_{aa}\right|\left|a_x^r\right|\sigma_x(1-m) = 2\xi|b|(1-m) \tag{17}$$

Hence, the full utility obtains by adding the three terms in (15). □

We move on to the optimal indexation level b.

Proposition 2 (How complexity dampens or annuls the incentive to index)**.** *Suppose first-order complexity aversion, and exogenous attention m. The optimal indexation level is*

$$b = \frac{\max\left(b^r - \xi(1-m), 0\right)}{2-m} \tag{18}$$

Hence, the consumer chooses not to index ($b = 0$) iff

$$b^r \leq \xi(1-m) \tag{19}$$

i.e. iff the traditional allocative benefits of indexation ($b^r = \text{cov}(\hat{y}, x)$) are less than the complexity aversion costs of indexation ($\xi(1-m)$).

Put simply, the consumer does not index iff complexity aversion ξ is large enough, and $m < 1$.

Proof. Maximizing utility (14) with respect to b gives:

$$0 = \frac{1}{2}V_b = b^r - (2-m)b - \xi\,\text{sign}(b)(1-m)$$

where I use $\text{sign}(b) = \frac{d}{db}|b|$, which is the generalized derivative of the function $|b|$. So it is 1 if $b > 0$, -1 if $b < 0$, and some number between -1 and 1 if $b = 0$.[24] Hence, if $b > 0$

$$b = \frac{b^r - \xi(1-m)}{2-m}$$

[23]The factor of 2 here is specific to this example, and comes from $|u_{aa}| = 2$ (as $u_{aa} = 2v''(c) = -2$), $|a_x^r| = |b|$, and $\sigma_x = 1$. In the pricing application of Section 4, the same general formula (6) instead yields $C^{\text{CA}} = \xi s_a \sigma_x |b|(1-m)$ (see (26)), without the factor of 2, because there $|s_a u_{aa} a_x^r| = s_a|b|$. The unitless coefficient ξ is the same across the two settings; only the problem-specific scaling differs.

[24]The machinery of subgradients gives the general framework for this.

But this requires $b^r > \xi(1-m)$. If that inequality fails, then $V_b < 0$ for $b > 0$. As $V_b > 0$ for $b < 0$, the optimal b is then $b = 0$. So the solution is (18). □

Proposition 2 is qualitatively consistent with the main facts mentioned earlier, and indeed may explain why most contracts (e.g. debt contracts) are not indexed. When the traditional allocative benefits of indexation ($\text{cov}(\hat{y}, x)$) are small, and complexity aversion ξ is large enough, there is no indexation. However, when stakes increase (e.g. in hyperinflations), then there is some indexation.

Pure "information economics" does not yield full simplicity, and we need first-order complexity aversion Proposition 2 has the following consequence.

Proposition 3. (Pure information costs do not create Full Simplicity) *If there is no complexity aversion ($\xi = 0$), the optimal index is (via (18)):*

$$b = \frac{b^r}{2-m} \in \left[\frac{b^r}{2}, b^r\right], \qquad b > 0 \tag{20}$$

So, importantly, there is always some indexation, $b > 0$. As a result, a pure "information cost" model always generates some indexation, $b > 0$.

This means that a pure "information cost" model (with exogenous or endogenous m) cannot generate "full simplicity" (i.e. $b = 0$): It only creates dampening compared to the rational case, $0 < b \leq b^r$ (indeed, any value of m creates this).[25]

I now show why second-order complexity aversion does not generate full simplicity.

Proposition 4. (Second-order complexity aversion does not generate full simplicity). *Suppose that instead of $C^{CA} = K|b|$, we have a complexity aversion penalty equal to $k|b|^\beta$ for some $k > 0$. If $\beta > 1$, there is always a positive amount of indexation, $b > 0$. On the other hand, a $\beta \in [0^+, 1]$ (where $\beta = 0^+$ is a dread cost paid iff $b \neq 0$) does generate zero indexation ($b = 0$) when k is large enough.*

Proof. In the consumer's value function (14), replace the term $2\xi|b|(1-m)$ in C^{CA} with $k|b|^\beta(1-m)$. For no indexation to be optimal, we need $V(b) \leq V(0)$ for small b's. Given the allocative gains are positive for a small $b > 0$, and proportional to b, we need the CA term to be large enough to offset the allocative gains, so we need (with $m \in (0,1)$): $\beta \leq 1$. It is then easy to verify that if $k > 0$ is large enough, and $\beta \in [0^+, 1]$, then $b = 0$ is optimal globally, not just locally around $b = 0$. □

[25] The hedge still helps stabilize her income, so is useful even when $m = 0$: in period 2, when she receives the payoff of the hedge, the agent must react to it, consume the remainder of a stabilized income.

This example made the point that complexity-averse agents will not want to index, under a wide range of parameters. More generally, they will not want to contract on some observables, which might be a way to generate incomplete contracts, in the spirit of Hart (1995) and Segal (1999), but with a more explicit cognitive structure, which would lend itself to measurement.

3.2 Extension: Endogenous attention

We can extend the model to endogenous attention. Then, the problem is to maximize the same utility (13)-(14), minus cognition costs $C^{\text{Cognition}}(m)$. So, the value function is (using (14)):

$$\max_{b,m} V(b,m) = 2b^r b - (2-m)\, b^2 - 2\xi\,|b|\,(1-m) - C^{\text{Cognition}}(m) \tag{21}$$

The following proposition gives a necessary and sufficient condition for $b = 0$ to be optimum, with endogenous attention.[26]

Proposition 5. (Indexation problem, with endogenous attention). *The jointly optimal choice of indexation and attention m is to take $b(m)$, the value in (18), and m to maximize $\mathcal{V}(m) \coloneqq V(b(m), m)$, which is:*

$$\mathcal{V}(m) = b(m)^2 (2-m) - C^{Cognition}(m) \tag{22}$$

Hence, if cognition costs are high enough, in the sense that for all m,

$$C^{Cognition}(m) \geq b(m)^2 (2-m)$$

then the optimal indexation is 0.

This means that we have zero indexation, with endogenous attention, if the cognition costs are large enough. Again, complexity aversion is crucial. Without it, when $m = 0$, some indexation is better than none.[27]

This indexation example illustrated the core economics, with a minimum of math. I now turn to a richer, market setting.

[26] The online appendix (Section A.6) adds explicit closed form solutions.

[27] To see why, consider the completely naive consumer with $m = 0$. The index will not help her at time 1, as she does not take the index into account. But it will help her at time 2: the index will decrease the volatility of her final consumption.

4 Simple pricing schemes

I apply the general model to a slightly richer situation: the pricing of one good—which, to make things concrete, I will call a baguette, sold by a bakery. This will show how complexity aversion interacts with market behavior, and deliver a few simple comparative statics.

The motivation is that many pricing schemes look "simple", compared to a frictionless benchmark. For instance, prices for Whole Foods delivery are the same for all hours and days of the week, even though they "should" vary in the frictionless world, as the marginal cost to Whole Foods must be different across the slots. Why does Whole Foods adopt this very simple pricing scheme? Most likely, this is because of some form of complexity aversion by the consumer.[28] This sort of situation is extremely prevalent: for instance, restaurants typically charge the same price for meals at different hours of the evening; bars do not implement "congestion pricing", etc.

4.1 Setup and basic analysis

A rational competitive bakery (maximizing social welfare) sells baguettes to a behavioral consumer.[29] The bakery's marginal cost is an exogenous $c(x) = 1 + x$ where x is a random variable with mean 0 and variance σ_x^2. The endogenous price per baguette is $p(x) = 1 + bx$, where b is chosen by the benevolent bakery, i.e. the planner. If consumers were rational, the efficient value would be to have the price equal marginal cost, i.e. $b = 1$, but with behavioral consumers the bakery might choose another b. For instance, if $b = 0$, the price of the baguette is constant over time, which "simplifies" the consumer's choice.

The consumer chooses to buy a (continuous) quantity $C = 1 + a$ of baguettes, where a is the deviation of consumption from its average amount, which is 1. Her total utility, net of the cost of the baguettes, is:

$$u(a, x) = U(1 + a) - p(x)(1 + a) \tag{23}$$

[28] This could also logically be a form of complexity aversion by the producer—i.e. some employee at Whole Foods—who does not know all the marginal costs. The modeling would be similar to the one proposed here, and might apply to the situations documented in DellaVigna and Gentzkow (2019). It could not plausibly be a monetary "menu cost" for Whole Foods—their sales are so enormous that any cost of changing the web site should be comparatively trivial.

[29] Here, I take the benchmark of firms forced to maximize social welfare by competitive forces, hence to be "benevolent". Section A.7 gives a microfoundation. This is the opposite tack of the literature where firms exploit the naive consumers' bounded rationality (Gabaix and Laibson (2006), Heidhues and Kőszegi (2018)), because there is a "shrouded attribute" (e.g. the spare part to be bought later) in addition to a "visible base good" (the car). As those issues are now well-understood, I focus on complexity with a benevolent firm, and discuss the case of exploitation in Section A.9.

where the gross utility from consuming a quantity $1+a$ of baguettes is $U(1+a) = 1+a-\frac{1}{2\psi}a^2$, with $\psi > 0$ the elasticity of demand (when the price is centered at 1). The rational action is $a^r = \text{argmax}_a\, u(a,x)$ i.e.

$$a^r = -\psi b x \tag{24}$$

4.1.1 The behavioral consumer's utility

The consumer perceives the price as $p^s = 1 + bx^s$ with $x^s = mx + \varepsilon$ with $\sigma_\varepsilon^2 = m(1-m)\sigma_x^2$ and $m \in [0,1]$ is the precision of the perception (exo- or endogenous). She takes the action $a = -\psi b x^s$, see (24). Her utility gains from taking the right action ($a^r = -\psi b x$) are:[30] $-\frac{1}{2}u_{aa}\mathbb{E}\left[(a_x^r x)^2\right] = \frac{1}{2}\psi\sigma_x^2 b^2$. Traditional consumption utility is increasing in attention:[31]

$$U^{\text{cons}} = \mathbb{E}\left[u(a,x)\right] = \frac{1}{2}\psi\sigma_x^2 b^2 m.$$

The complexity aversion term (6) is:

$$C^{\text{CA}} = \xi s_a \left|u_{aa}\right| \left|a_x^r\right| \sigma_x (1-m) = \xi s_a \frac{1}{\psi}\psi \left|b\right| \sigma_x (1-m)$$

hence

$$C^{\text{CA}} = \xi s_a \sigma_x \left|b\right| (1-m) \tag{26}$$

The consumer's full utility is, from (5):

$$V^{\text{cons}}(m) = \frac{1}{2}\psi\sigma_x^2 b^2 m - C^{\text{Cognition}}(m) - C^{\text{CA}}(m). \tag{27}$$

With endogenous attention, I posit that the consumer optimizes cognition, i.e. chooses m to maximize $V^{\text{cons}}(m)$.

[30] It may be surprising that a more volatile price makes the consumer better off, even though the consumer is risk-averse. This is always true with quasi-linear utility $u(c_0,\ldots,c_n) = c_0 + U(c_1,\ldots,c_n)$, with U concave, which is the case here. This is not true for general utilities, as can be seen for the one-good case: the indirect utility is $u\left(\frac{w}{p}\right)$, which shows the tension between the convex $\frac{1}{p}$ and the concave u in determining whether indirect utility is locally convex or concave in p.

[31] The derivation is, with $\chi \coloneqq \frac{1}{2}\psi\sigma_x^2 b^2$

$$\mathbb{E}\left[u(a,x)\right] = \mathbb{E}\left[u(a^r,x)\right] + \frac{1}{2}u_{aa}\mathbb{E}\left[(a_x^r x)^2\right](1-m) = \chi - \chi(1-m) = \chi m \tag{25}$$

4.1.2 Social welfare

Social welfare is:

$$W = V^{\text{cons}}(m) + \mathbb{E}\left[\left(p(x) - c(x)\right)\left(1 + a\left(x^s\right)\right)\right] \tag{28}$$

i.e. the consumer's full utility (including dread and cognition cost), plus profits from consumption $1 + a\left(x^s\right)$, which reflects that the social cost of a baguette is c, not p.[32]

So, the social planner (or benevolent bakery) wants to maximize total social surplus, while taking into account that the variability of prices (indexed by b) affects the endogenous attention m. Mathematically: $\max_b W\left(m\left(b\right), b\right)$. In the exogenous attention case, $m\left(b\right)$ is just constant; with endogenous attention, m optimizes $V^{\text{cons}}(m)$ in (27). We next calculate that social welfare.

Proposition 6. *Social welfare is*

$$W(m, b) = \underbrace{(b - \frac{b^2}{2})\psi\sigma_x^2 m}_{\substack{\textit{Allocative gains from pricing} \\ \textit{closer to marginal cost}}} \underbrace{- \xi s_a \sigma_x |b|(1 - m)}_{\substack{\textit{1st order losses} \\ \textit{from CA}}} - C^{Cognition}(m) \tag{29}$$

The first term verifies that, with a rational consumer, social welfare is maximized when $b = 1$, i.e. when price equals marginal cost.[33]

4.2 When is the optimal price system "simple"?

Proposition 7. (Optimal price simplicity, first-order complexity aversion) *Suppose that consumers have first-order complexity aversion. Then, the optimal sensitivity of price to marginal cost is:*

$$b = \max\left(1 - \frac{\xi s_a \sigma_x (1 - m) - W_m \frac{dm}{db}}{\psi \sigma_x^2 m}, 0\right) \tag{30}$$

With endogenous attention this is a first-order condition rather than a global characterization, and it presumes $m > 0$; Section A.8 gives sufficient conditions for a global optimum. So, with exogenous attention ($\frac{dm}{db} = 0$), the optimum entails rigid prices if and only if:

$$\xi s_a \sigma_x (1 - m) \geq \psi \sigma_x^2 m \tag{31}$$

[32] Adding $(p(x) - c(x))(1 + a(x^s))$ to (23) makes the baguette evaluated at its social cost of production, which is the relevant welfare criterion.

[33] In particular, there is no motive for the bakery to "insure" the consumer. The consumer has quasilinear utility, so is risk-neutral to income shocks.

i.e. if marginal complexity aversion cost, $\xi s_a \sigma_x (1-m)$, exceeds the marginal increase in allocative benefits, $\psi \sigma_x^2 m$.

Proof. The total derivative of $W(m(b), b)$ in (29) is:

$$\frac{d}{db}W = (1-b)\psi\sigma_x^2 m - \xi s_a \sigma_x \operatorname{sign}(b)(1-m) + W_m \frac{dm}{db}$$

Solving for b gives the announced expression (30). □

Let us take the simplest case of exogenous attention, so that $\frac{dm}{db} = 0$, and use Proposition 7, especially (31).

Contracts are more likely to be "simpler" (weakly lower b) when

1. complexity aversion ξ is high;
2. the degree of understanding m is low—we want "Simple contracts for simple people"
3. the neoclassical gains from variable pricing ($\psi\sigma_x^2/2$) are low. This is why Uber has variable prices—the neoclassical gains of variable prices are high, as high prices direct the cabs to key destinations (which vary over the day), and lessen waiting times;
4. the elasticity of demand ψ is low—hence, if competition is low (in many models, higher competition leads to a higher elasticity of demand ψ, see e.g. Atkeson and Burstein (2008), Gabaix et al. (2016)).

The important point is that we get a simple price system if the complexity aversion parameter is large enough. We see why Whole Foods has a uniform pricing system for delivery, but variable prices between goods: the demand for delivery is reasonably inelastic, while the demand for an individual good is much more elastic (e.g. raspberries vs blueberries, or raspberries at Whole Foods vs another store), and the marginal cost of perishable goods (σ_x) is very variable (e.g. depending on harvests). Likewise, in very competitive markets (e.g. airlines), the elasticity of demand is very high, and as a result the schemes are "complex".

The necessity of having a complexity aversion cost The next proposition shows that the CA is essential to obtain simple mechanisms.

Proposition 8. (We need complexity aversion to explain full simplicity) *Suppose that there is no complexity aversion ($\xi = 0$), and there is either exogenous or endogenous attention. Then social welfare is optimized for $b = 1$, so that prices are not rigid at the optimum.*[34]

[34] Strictly so, unless attention is 0 for all b, in which case welfare does not depend on b at all, and a fringe of rational consumers restores strictness (see the proof).

Proof. First, with exogenous attention m, we can see (29): it is maximized at $b = 1$, regardless of m. But for endogenous attention, the problem is more delicate, as the planner might wish to manipulate b to distort the consumer's attention. Actually, this is not the case, as the complement to the proof in the online appendix shows. □

We conclude that, with first-order CA, a high CA parameter ξ generates simplicity, $b = 0$, with exogenous attention. Section A.8 shows how this is also true with endogenous attention, provided the cost of cognition is high enough.

5 Simple tax systems: Ramsey problem with complexity-averse taxpayers

The Ramsey (1927) problem is the prototypical optimal taxation problem, and it leads to very complex taxes: with 10,000 goods, there are 10,000 different tax rates.[35] This complexity explosion is true of all optimal tax problems, where generically the tax scheme depends on each good, history, and state of the world.

It should by now be obvious qualitatively that with high enough complexity aversion, the optimal tax system should be simple. But it is instructive to see how to set this up, and how to compute the optimal tax system.

5.1 The Ramsey problem with rational agents

I start with the traditional framework for goods taxation. There are $G+1$ goods indexed by g. The agent has utility $U(c) = \sum_{g=1}^{G} u^g(c_g) + c_0$ where $u^g(c_g) = c_g - \frac{1}{2\psi_g}(c_g - 1)^2$ and c_0 is a residual good with price 1 (which is utility from residual income), untaxed. The rational demand is $c_g = 1 - \psi_g(p_g - 1)$, so that ψ_g is the sensitivity of demand of good g to its price.

Each good g is taxed at a rate of τ_g, so the price paid by the consumer is $\tilde{p}_g \coloneqq p_g(1 + \tau_g)$. The pre-tax price p_g is independent of taxes, e.g. fixed by a linear production technology, and it is normalized to be $p_g = 1$.

The government has a social welfare function $W = U(c) + (1 + \lambda)\sum_g \tau_g c_g$ where $\lambda > 0$ is the excess marginal value of public funds, which are the collected taxes, $\sum_g \tau_g c_g$.

Following the tradition, I take the limit of a small λ, so that taxes are small. Taking a

[35] This is true even when agents are behavioral, in the sense that they misperceive the tax, but are not complexity averse (Farhi and Gabaix (2020), Rees-Jones and Taubinsky (2020)).

Taylor expansion up to second-order terms, social welfare is

$$W = W^0 + \sum_g \left(-\frac{1}{2}\psi_g \tau_g^2 + \lambda \tau_g \right)$$

where W^0 is the welfare with 0 taxes. Optimizing over τ_g gives the traditional Ramsey (1927) result:

$$\tau_g = \frac{\lambda}{\psi_g} \tag{32}$$

Each good has a different tax: the tax system is very "complex".[36]

5.2 The Ramsey problem with complexity averse agents

The government sets up an endogenous number n of tax brackets G_k, for $k = 1 \dots n$. The tax in bracket k is $\tau_{G_k} = \bar{\tau} + \hat{\tau}_{G_k}$, where $\bar{\tau}$ is the average tax and $\hat{\tau}_{G_k}$ is a deviation.

I assume that the consumer sees a tax[37]

$$\tau^s_{G_k} = \bar{\tau} + \hat{\tau}^s_{G_k} \tag{33}$$

where the subjective perception $\hat{\tau}^s_{G_k}$ is an inattentive version of the true $\hat{\tau}_{G_k}$:[38] the consumer sees average tax $\bar{\tau}$ perfectly, but sees the deviation $\hat{\tau}_g$ only partially. The consumer also sees for free the bracket $k(g)$ of a good g.

The consumer's demand is:

$$c_g = 1 - \psi_g \tau^s_g \tag{34}$$

[36]This is generically true, even if the government knows only imperfectly the elasticity of each good.

[37]This model applies directly to the US, where it is customary for firms to show the pre-tax price p_g, but not the price inclusive of the tax, $\tilde{p}_g \coloneqq p_g(1+\tau_g)$. However, it also applies to other countries, where the price inclusive of taxes is shown: simply then, the cognitive difficulty is on the side of firms. Indeed, consider a world where what is shown is the price $\tilde{p}_g$ inclusive of the tax, as in European countries; and where, as in all sensible systems, intermediate production is not distorted, so that firms do not pay the VAT on intermediate inputs. Suppose that firms see the price $\tilde{p}_g$ but the price they pay is $\frac{\tilde{p}_g}{1+\tau_g}$, as the VAT that a firm pays is rebated to it. Then, a firm h's problem is to maximize over the intermediate inputs X^h_g of good g, $\max_X p_h F^h\left(X^h_g\right) - \sum_g \frac{\tilde{p}_g}{1+\tau_g} X^h_g$. So, it is the firm that now faces the complexity cost of knowing each τ_g. As decision-makers in firms are people, if prices inclusive of taxes are shown, those people will pay the complexity cost instead of the consumer. Their attention may be higher than that of consumers, but (at least for non-giant firms) it is likely to be imperfect. In either case, the qualitative mechanism carries over, though the relevant elasticities and welfare incidence differ, and a production-side reduction would be needed to transport the formulas literally.

[38]It is intuitive that it is hard to react to the varying tax rates. Chetty et al. (2009), Taubinsky and Rees-Jones (2017) and Rees-Jones and Taubinsky (2020) provide experimental evidence that people do not fully rationally react to taxes, and indeed typically under-react to them, consistent with limited attention. Glaeser and Shleifer (2001) propose that quantities rather than prices are sometimes regulated (e.g., no alcohol should be sold on Sundays), when quantities are easier to observe, a form of simplicity. Aghion et al. (2024) find significant frictions to optimal reaction to taxes, consistent with complexity costs.

Calling m_{G_k} the attention to the special tax $\hat{\tau}_{G_k}$ of bracket k, the (first-order) complexity aversion cost associated with a good in bracket k is as in (26):[39]

$$C_{G_k}^{\text{CA}} = \xi \left|\hat{\tau}_{G_k}\right| (1 - m_{G_k}) \tag{35}$$

Social welfare is:

$$W = \mathbb{E}\left[U\left(c\right)\right] - C^{\text{CA}} - C^{\text{Cognition}} + (1+\lambda)\sum_g \tau_g c_g \tag{36}$$

i.e. the consumer's utility (inclusive of the complexity-aversion cost and the cognition cost), plus the government revenues (with weight $1+\lambda$, where $\lambda \geq 0$). As in the rational case, it is useful to study the limit of small taxes.

Lemma 1. (Welfare in the Ramsey problem with CA agents: Taylor expansion) *In the limit of small taste for government funds λ, hence small taxes, social welfare is, up to higher order terms:*

$$W = \sum_{k;g\in G_k} [\frac{-1}{2}\psi_g \mathbb{E}\left[\left(\bar{\tau} + \hat{\tau}_{G_k}^s\right)^2\right] - \xi \left|\hat{\tau}_{G_k}\right| (1 - m_{G_k}) + \lambda\left(\bar{\tau} + \hat{\tau}_{G_k}\right)] - C^{Cognition} + W^0 \tag{37}$$

where W^0 is the social welfare when all taxes are zero. Also, $\mathbb{E}\left[\left(\bar{\tau} + \hat{\tau}_G^s\right)^2 \left|\hat{\tau}_G\right.\right] = \bar{\tau}^2 + m_G\left(2\bar{\tau}\hat{\tau}_G + \hat{\tau}_G^2\right)$.

The government problem is to maximize social welfare W, subject to the average tax being $\bar{\tau}$, i.e. $\sum_{G_k} s_{G_k}\hat{\tau}_{G_k} = 0$, so that $\bar{\tau}$ is truly the average tax.

To solve the problem,I will use the thresholding function

$$T\left(y,\kappa\right) = \text{sign}\left(y\right)\max\left(\left|y\right| - \kappa, 0\right) \tag{38}$$

which shrinks the argument y towards 0, by an amount that depends on the threshold κ. In this subsection, I will take the simple case where all brackets receive the same exogenous attention, $m_{G_k} = m$.[40] For simplicity, I consider the case with a mass 1 continuum of

[39]Strictly speaking, in (35), the right-hand side should be multiplied by s_c, which is the size modulus for the variation in consumption of the goods in the group—assumed to be independent of k, but that could be relaxed. As usual, it will play no role except giving the correct units. Hence, to streamline expressions, I normalize it to $s_c = 1$—the reader can recover the un-normalized expressions by replacing below ξ by ξs_c.

[40]This is less restrictive that it may appear. In Section B.2, where attention is endogenous and chosen bracket by bracket, all the brackets with non-zero tax deviation will indeed, endogenously, receive the same attention.

goods, whose elasticities are uniformly distributed in $[\psi_{\min}, \psi_{\max}]$.[41] I call $\bar{\psi} \coloneqq \frac{\psi_{\min}+\psi_{\max}}{2}$ the average elasticity, and $v_\psi \coloneqq \frac{\psi_{\max}-\bar{\psi}}{\bar{\psi}}$ the half-width of the range of elasticities. I take a Taylor expansion in the limit of small v_ψ (the proof tracks the $O(v_\psi^j)$ terms, but I omit them here for brevity).

Proposition 9. (Optimal tax rate with complexity aversion) *The optimal tax for bracket G_k is, to leading order in v_ψ:*

$$\tau_{G_k} = \bar{\tau} + T\left(\tau^R_{G_k} - \bar{\tau}, \bar{\tau}\bar{\xi}\right), \tag{39}$$

where $\tau^R_{G_k} = \frac{\lambda}{\psi_{G_k}}$ is the traditional Ramsey tax for bracket G_k, with ψ_{G_k} the average elasticity of the goods in bracket G_k; T is the thresholding function (38), and $\bar{\xi} \coloneqq \frac{\xi}{\lambda}\frac{(1-m)}{m}$. Hence, if for all tax brackets, $\left|\frac{\tau^R_{G_k}}{\bar{\tau}} - 1\right| \leq \bar{\xi}$, then the optimal tax is uniform: $\tau_{G_k} = \bar{\tau}$.

In each tax bracket, the tax is given by (39). In particular, the taxes are always more compressed (towards $\bar{\tau}$) than under a Ramsey model with the same tax brackets, but no complexity aversion.[42] I now calculate welfare.

Proposition 10. (Welfare gain from n tax brackets, with complexity-averse consumers) *Consider the Ramsey problem with a continuum of goods, and complexity averse consumers. Suppose that the government implements n tax brackets, chosen optimally and with optimal taxes, and that attention is positive, $m > 0$. When moving from 1 to n brackets, welfare increases by, to leading order in v_ψ:*

$$V(m,n) = \frac{1}{6}\lambda\bar{\tau}v_\psi^2 m \max\left(1 - \frac{\xi}{\lambda v_\psi}\frac{1-m}{m}, 0\right)^3\left(1 - \frac{1}{n^2}\right) \tag{40}$$

To interpret this result, let us consider the case of endogenous attention, with two cases: the simple case of an attention budget constraint, and the richer case of costly attention.

Optimal number of taxes with an attention budget constraint. I first explore the case where attention is subject to a simple "budget constraint" but otherwise free. To fix ideas, the constraint could be $mn \leq \mathcal{M}$ for some constant $\mathcal{M}$, so that $m(n) = \min(\frac{\mathcal{M}}{n}, 1)$. Then, the optimal number of brackets is as follows.

[41] Proposition 27 works out a case with a finite number of goods, and with a general distribution of elasticities, but it is less tractable.

[42] The proof of the proposition (in Section D) describes the shape of the tax brackets. The central bracket, around the average elasticity, has a width $2b^*v_\psi$ with b^* given in (124); and otherwise the other tax brackets have equal size.

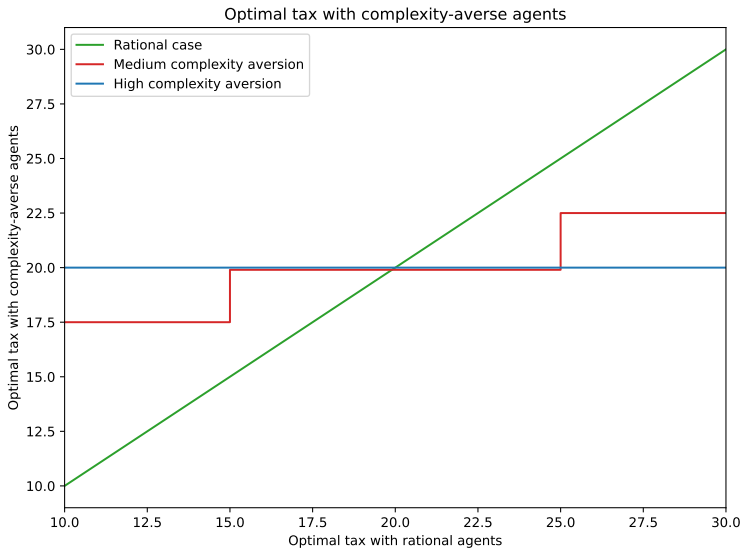


Figure 1: Optimal tax system for three levels of complexity aversion. With rational agents, the optimal tax is just the Ramsey tax τ^R: there are generically as many tax brackets as goods. With very high complexity aversion, there is a uniform tax: there is just one tax bracket. For an intermediate complexity aversion, there are three tax brackets.

Proposition 11 (Optimal number of tax brackets)**.** *If agents* are *complexity averse: (i) With exogenous attention $m < 1$, and ξ large enough, the optimum is just to have a uniform tax; (ii) With endogenous attention $m(n)$, the optimum number of brackets (weakly) decreases with the degree of complexity aversion, and becomes 1 with large complexity aversion.*

If agents are not *complexity averse, one never gets a simple tax system with just one tax bracket, provided attention is positive—if attention to the deviations is 0, extra brackets bring no allocative gain, and every number of brackets is equally good. Indeed, (iii) with exogenous attention $m > 0$, there are as many tax brackets as goods; (iv) with endogenous attention $m(n)$ such that $m(2) > 0$, there must be at least two tax brackets.*

The situation is shown in Figure 1. With rational agents, the optimal tax is just the Ramsey tax $\tau_g^R = \frac{\lambda}{\psi_g}$. With very high complexity aversion, there is a uniform tax: there is just one tax bracket. For an intermediate complexity aversion, there are three tax brackets in this example.

The biggest step here was to find the conditions (namely, first-order complexity aversion) that generate full simplicity. This is not trivial. Indeed, without complexity aversion, if there is a fringe of rational, zero-cost agents, then again the planner wants a different tax for every good.

Optimal number of taxes with a smooth cost of attention. I consider the case where attention is costly, with a smooth cost of attention. I take the cost to be linear in the attention level: $C^{\text{Cognition}}(m) = \kappa m$, but with the constraint that $m \leq \bar{m}$ for some positive constant $\bar{m} < 1$. This captures the idea that attention is costly, and that there is a limit to how much attention can be paid to the tax code, at least for some agents. Hence, the consumer chooses a separate attention m_{G_k} for each bracket deviation:

$$C^{\text{Cognition}} = \kappa \sum_k m_{G_k}, \qquad m_{G_k} \in [0, \bar{m}], \qquad 0 < \bar{m} < 1. \tag{41}$$

The planner is free to optimize over all tax bracket structures, and he takes into account the endogeneity of the attention level m_{G_k} for each bracket k. This structure is convenient, as it allows the planner to solve the problem in closed form.

Proposition 12. *Suppose the linear cost of attention, and suppose that $\xi \geq \lambda \bar{m}$. Then, the planner's problem with endogenous attention has the same solution as the previous problem with exogenous attention (Propositions 9-10), with $m = \bar{m}$: every bracket with a non-zero tax deviation receives attention $\bar{m}$. In addition, the optimal number of brackets solves*

$$\max_n \; V(\bar{m}, n) - \kappa \bar{m}(n-1) \tag{42}$$

where $V(m, n)$ is the welfare gain of (40)*.*

The cost is $\kappa \bar{m}(n-1)$ because only $n-1$ brackets exhibit a tax deviation, hence non-zero attention. There is a subtlety in the problem (42). In general, the problem would be more complex, as when the planner changes taxes, this changes attention (a sort of "fiscal externality"). However, with the cost function above, attention is stuck at $\bar{m}$, which simplifies the problem.

Proposition 13 (Optimal number of tax brackets with a linear cost of attention)**.** *Under the assumptions of Proposition 12, the optimal number of tax brackets is:*

$$n^* = \max\left(\left(1 - \frac{\xi\,(1-\bar{m})}{\lambda \bar{m} v_\psi}\right)^+ \left(\frac{\lambda \bar{\tau} v_\psi^2}{3\kappa}\right)^{1/3}, 1\right) \tag{43}$$

This is the optimum of the continuous relaxation; the integer optimum is the better of the two neighboring integers.

Hence, we obtain a simple formula for the optimal number of tax brackets. As expected, the optimal number of tax brackets is lower with: high complexity aversion, high cost of

attention κ, low maximum attention level $\bar{m}$; and is higher with a strong incentive to differentiate taxes, namely a high dispersion of elasticities v_ψ, and a high taste for public funds λ.[43]

The online appendix (Section B) contains a more general treatment, and the proofs. There are some subtleties when complexity aversion ξ is small, but this is only advanced material. As always, the most useful case is that of high complexity aversion.

Other potential ways to derive a uniform tax, or a small number of tax brackets. Let me discuss other ways to potentially derive the result.

Cost for the sheer number of taxes. I discuss the counting measure $G(n)$ in Section 7.3.2, and do not adopt it: in the real world there are always thousands of small disturbances, so the counting measure would counterfactually predict almost infinite angst.

Administrative costs. They would work, and would be a form of cost $G(n)$. I do not pursue them here, because the conceptually interesting reason here is the limited processing power of the people, not of the government. Still, complexity aversion costs for planners would be interesting to explore.

Purely limited attention, without complexity aversion. This would not work, as suggested in Proposition 11 and formalized in Proposition 32: the number of brackets is always at least two.

Let me conclude this analysis of the Ramsey problem. It is meant to be a prototype for much more, e.g. Ramsey taxes with redistributive motives, Pigouvian taxes, a Mirrlees problem. More generally, this approach suggests a way to do "behavioral mechanism design", i.e. design optimal institutions (e.g. price or tax systems) that aim both at high allocative efficiency and not too high complexity.

6 Choice aversion and choice overload

I close by returning to individual decisions, and to two phenomena that are well documented: inertia, the tendency not to decide at all, and choice overload, the tendency to fall back on the simple option when there are many options. This is also the occasion to show the model in a discrete choice setting, as all the previous applications used a continuous action. In

[43] One might worry that Proposition 12 requires the cost of attention κ to be small: if attention is very costly, why would the consumer pay any? The number of brackets adjusts, and takes care of this. A higher κ gives fewer and wider brackets, hence larger tax deviations, so the consumer's stake in a bracket rises exactly as fast as the cost of attending to it: at the optimum (43), the stake in each bracket is at least 6κ, whatever the value of κ. And when κ is so large that (43) gives $n^* = 1$, the tax is uniform, and there is nothing to pay attention to.

both cases complexity aversion delivers sharp and testable comparative statics.

6.1 Choice aversion and inertia in decision-making

Very often, people prefer not to decide at all. For instance, they do not rebalance their portfolio. Complexity aversion offers a natural hypothesis for this: thinking is aversive, especially if one dreads taking the wrong decision.

Suppose that the active decision yields the traditional utility, compared to the default action of "doing nothing"—that is, sticking to a^d:

$$U^{\text{active}} = \mathbb{E}\left[u\left(a^s\left(x^s\right), x\right)\right] - \mathbb{E}\left[u\left(a^d, x\right)\right] - C^{\text{Cognition}}(m) - C^{\text{CA}}(m) \tag{44}$$

where a^s is the noisy decision given the imperfect perception x^s. The passive decision is evaluated at $U^{\text{passive}} = 0$. Here I assume that the passive decision does not bear any complexity aversion. This is in line with the psychology of commission (the active decision) versus omission (the passive decision). The omission feels like "nothing", whereas the commission feels like "utility consequences from deviating from the default", and so carries some complexity aversion.

To complete this simple psychology, we can add "motivated cognition" with[44]

$$\tilde{U}^{\text{active}} = U^{\text{active}} + \mu \sum_i u_{x_i} x_i^s$$

where $\mu > 0$ is the weight on motivated attention. This means that reacting to "bad news" (news x_i with $u_{x_i} x_i^s < 0$, i.e. that lowers utility) is penalized, an "ostrich effect" (Karlsson et al., 2009).[45]

To see the effects, take the model $u(a, x) = -\frac{1}{2}(a - x)^2 + vx$. The agent takes the passive course of action if and only if $U^{\text{passive}} \geq \max_m \tilde{U}^{\text{active}}(m)$. To get a closed-form solution, I analyze the case $C^{\text{Cognition}}(m) = \kappa m$, and the behavior of the median agent (i.e. the agent receiving the median signal x^s).

Proposition 14. (When the agent fails to act, and passively follows the default) *The median agent takes the passive course of action if and only if:*

$$\min\left(\xi s_a \sigma_x, \kappa - \frac{\sigma_x^2}{2} - \mu v x\right) \geq \frac{1}{2}\left(a^d\right)^2 \tag{45}$$

[44] This modeling of motivated cognition is in Gabaix (2019), Section 4.3.2.

[45] Sensibly, real ostriches do not exhibit that effect—Pliny the Elder spread the myth.

Condition (45) means that people are more likely to be passive if: complexity aversion ξ is high, cognition cost κ is high, the news is "bad news" (vx is very negative), and the passive allocation is not too far from the optimum ($\left|a^d\right|$ small). If there is no complexity aversion, people are never passive.

Take the lack of rebalancing of one's portfolio. The active decision involves thinking about the correct determinants of portfolio choice, a complicated vector x, involving e.g. the expected returns of U.S. versus international stocks, or long- versus short-term bonds.[46] Hence, many people just choose not to rebalance at all (Giglio et al., 2021; Gabaix et al., 2024). They will do so only when the portfolio is very off, and when they have low complexity aversion.

6.2 Choice overload

Suppose that we have a simple option and n complex options. The complex options suffer the CA penalty, but the simple option does not. Hence, the simple option is more likely to be chosen. This accounts, qualitatively, for the findings in the "choice overload" literature (e.g. Iyengar and Lepper, 2000, Iyengar and Kamenica, 2010): as people face more options, they more often choose the simple option—with nuances I discuss soon.

I illustrate how the model of CA quite naturally delivers choice overload. This also lets me show the model in a discrete choice setting: all the previous examples used the continuous action setting of the model. I use the discrete-choice machinery of Section A.5, but this exposition here is self-contained.

There are $n+1$ options, indexed by $k \in \{0, \ldots, n\}$. Option 0 is the "simple" one, with value 0. The other options $k \in \{1, \ldots, n\}$ have value $X_k = x_{k,1} + \cdots + x_{k,p}$, i.e. are "complex" as made of p parts, which are drawn independently—across components and across options—from a standard normal distribution. To go to the core economics, I assume that attention is m for all components. When thinking about a complex option k, the agent sees each of the p components fuzzily, as $x^s_{k,i}$, as in the general model. Hence the total CA penalty for the option is:

$$C^{\mathrm{CA},k} = \xi p \left(1 - m\right) \tag{46}$$

[46] There could be other explanations. The thinking cost is paid now and the benefits are received later, so a hyperbolic discounter will fail to act (Laibson, 1997). The act of choosing itself could be aversive (Bernheim et al., forthcoming), or entail some adjustment cost (see the models reviewed in Gabaix (2019)). Choice could activate loss aversion, as looking up information can reveal bad news (Kőszegi and Rabin, 2006; Andriès and Haddad, 2020). Still, complexity aversion and the dread of mistakes seem like a plausible alternative route.

So, the option is valued as:

$$v_k = X_k^s - \xi p\,(1-m)\,, \qquad X_k^s = x_{k,1}^s + \cdots + x_{k,p}^s \tag{47}$$

where X_k^s is the perceived value of the raw payoff of option k (before the CA penalty). The agent chooses the option with the highest valuation. This leads to the following result.[47]

Proposition 15. *The probability of choosing the simple option, rather than any of the n complex options, is*

$$P_0 = \left[\Phi\Big(\frac{\xi\sqrt{p}\,(1-m)}{\sqrt{m}}\Big)\right]^n \tag{48}$$

where $\Phi(x)$ is the CDF of the standard normal distribution. In particular, it increases with complexity aversion ξ and the number of components p in each option, and decreases with attention m.

Proof. As in the basic model (3), $\text{var}(x_{k,i}^s) = m$; the perceived payoff of a complex option k, $X_k^s = x_{k,1}^s + \cdots + x_{k,p}^s$, has variance pm. So we can represent it as $X_k^s = \sqrt{pm}\epsilon_k$, where ϵ_k is a standard normal random variable. The probability that each v_k is less than 0 is $q = \Phi\big(\frac{\xi p(1-m)}{\sqrt{pm}}\big)$. Hence, the probability of choosing the simple option is $P_0 = q^n$. □

I now assume an "attention" function $m = f(n)$ that decreases with n, which may come from an attention budget. Then, there are two effects in (48), as the number of complex options n increases. First, the attention m to any option decreases, so the CA penalty increases, which makes the simple option more attractive: P_0 goes up. Second, for a given attention m, more options decrease the probability of choosing the simple option, as it has more competitors: P_0 goes down. The first and main effect requires complexity aversion, $\xi > 0$.[48]

Figure 2 conveys the effects.[49] When ξ is at the typical value used in this paper, as the number of complex options n increases, the probability of choosing the simple option P_0 increases: this is the choice overload effect. However, without complexity aversion ($\xi = 0$), the probability falls: as more options are available, the simple option is less likely to be chosen.

One can show a number of general results. First, the probability of choosing the simple option goes to 1 in the large n limit, at least if attention $m(n)$ decays to 0 at a faster than glacially slow rate, as it surely does in practice.

[47] I imagine a very large dataset, with many draws of all variables and choices.
[48] With 0 complexity aversion or 0 inattention (i.e. if $\xi = 0$ or $m = 1$), $P_0 = 2^{-n}$ falls with n.
[49] I take $\xi = 2/3$, $p = 4$ components, and $m(n) = 1/n$.

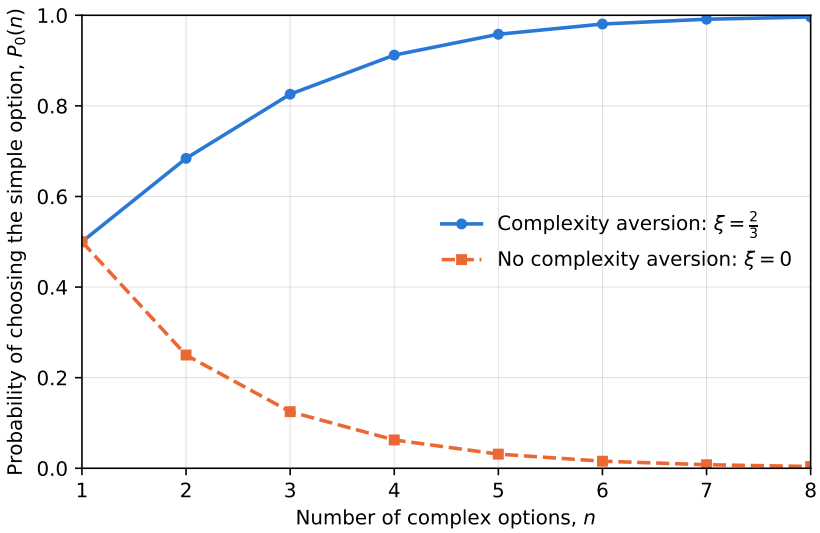


Figure 2: This figure shows the probability of choosing the simple option as a function of the number n of complex options (Proposition 15). When complexity aversion is high enough, the probability of choosing the simple option increases with the number of complex options.

Proposition 16 (Choice overload: simple asymptotics)**.** *The probability of choosing the simple option P_0 goes to 1 when ξ becomes large with $m < 1$ and n fixed; or, for a given $\xi > 0$, as n becomes large and $m(n)$ goes to 0 fast enough, in the sense that $\lim_{n\to\infty} m(n) \ln n = 0$.*

The trade-off between the two forces identified above can be pinned down precisely.

Proposition 17 (Choice overload, small increase in number of complex options)**.** *Assume $m(1) = 1$. There is "local choice overload" (the simple option is chosen more often with two complex options than with one, i.e. $P_0(2) > P_0(1)$) if and only if the complexity aversion parameter ξ and the number of components p are large enough, namely*

$$\xi\sqrt{p}\,\frac{1-m(2)}{\sqrt{m(2)}} \;>\; \Phi^{-1}\big(2^{-1/2}\big) \simeq 0.54. \tag{49}$$

With $m(n) = 1/n$, this reads $\xi\sqrt{p} > \sqrt{2}\,\Phi^{-1}(2^{-1/2}) \simeq 0.77$.

Those propositions explain transparently some controversies in the psychology literature, namely that sometimes choice overload does not happen (Scheibehenne et al., 2010), but it happens more frequently when the decision is very "complex", and the decision makers have little time (Chernev et al., 2015). Indeed, Proposition 17 shows that there is "local choice overload" (the simple option is chosen more often when the number of complex options slightly increases away from 1) if and only if the choice is complex enough ($\xi\sqrt{p}$ high enough) and "attention pressure" is strong enough ($m(2)$ sufficiently low). More generally, when

things are very easy, $m = 1$, there is no choice overload—this is like the lower curve for no complexity aversion in Figure 2, and also like in Proposition 15. However, when choice is hard, m is low, and the advantage of the simple option grows with more options (under the simple condition of Proposition 16)—this is like the upper curve for high complexity aversion in Figure 2.

In conclusion, complexity aversion gives a natural explanation for the choice overload effect: complexity is penalized; and as more options are available, attention to any option decreases, which increases the CA penalty, making the simple option more attractive.

Link with the literature We see how the "first-order complexity aversion" advocated in this paper is useful: it ensured here that the model was scale-independent, i.e. independent of the standard deviation of the components.[50] A penalty that is not homogeneous of degree one in the scale of the components would break that scale-independence, very likely counterfactually; the sharper identification of (6) comes from the criteria of Theorem 1, not from scale-independence alone. Hence, it supports the conjecture in de Clippel et al. (2025) that behavior is best described by first-order complexity aversion.

There are now a number of explanations of choice overload, based on anticipated regret (Sarver, 2008; Buturak and Evren, 2017), on inference from the seller's product line (Kamenica, 2008), and on search and evaluation costs (Caplin et al., 2011). Models of limited consideration (Manzini and Mariotti, 2014) are related, but their consideration structure is menu-independent, so it does not by itself generate overload. I do not claim that complexity aversion is a "better" mechanism than those. The point is rather that choice overload need not have a new mechanism of its own: it comes out of the same model, with the same ξ, that delivers the absence of indexation, rigid prices, and few tax brackets in the preceding sections. Hence, we are closer to a unified account of behavior across domains.

7 Discussion

7.1 A very simple calibration of complexity aversion

I propose a simple calibration of the complexity aversion parameter ξ, around

$$\xi \simeq \frac{2}{3} \tag{50}$$

A baseline based on simple gambles Suppose a very simple gamble, with equal

[50] Proposition 15 would still hold if all the noises were multiplied by some σ_x.

probabilities of payoffs $\mu+\sigma$, $\mu-\sigma$, with $\sigma>\mu>0$. A prospect-theoretic agent with loss aversion λ attributes to it a value v with $2v=\mu+\sigma+\lambda(\mu-\sigma)=(1+\lambda)\mu-(\lambda-1)\sigma$. So, to be indifferent between accepting the gamble or not ($v=0$), she will require a Sharpe ratio: $S=\frac{\mu}{\sigma}=\frac{\lambda-1}{\lambda+1}$. With the typical calibration $\lambda\simeq 2$ (Brown et al., 2024), this gives $\frac{\mu}{\sigma}=\frac{1}{3}$.

The complexity interpretation[51] is that the certainty equivalent is $\mu-\xi\sigma(1-m)$, so that people require a Sharpe ratio $S=\frac{\mu}{\sigma}=\xi(1-m)$.[52] We are led to $\xi(1-m)=\frac{\lambda-1}{\lambda+1}\simeq\frac{1}{3}$. With $m\simeq\frac{1}{2}$ (so that people understand the situation only partially), that leads to $\xi\simeq\frac{2}{3}$.

Now, I will examine the earlier examples, and we shall see that this value of $\xi\simeq\frac{2}{3}$ broadly calibrates well. In most examples, I will take $m=\frac{1}{2}$, a sort of prior over $m\in[0,1]$.[53]

Lack of indexation We have seen that there is no indexation if and only if (Proposition 2)

$$\frac{b^r\sigma_x}{s_a}=\frac{\text{s.d. of income risk removable by the index}}{\text{average consumption}}\leq\xi(1-m)\simeq\frac{1}{3}. \tag{51}$$

For instance, suppose the income risk is $\hat{y}=b^r x+\epsilon_y$, where $b^r x$ is the hedgeable part of the risk, and ϵ_y is the non-hedgeable part. The criterion says that there is no indexation if and only if the volatility of the hedgeable income risk (say, inflation) is less than $\frac{1}{3}$ of average consumption. Hence, roughly speaking, for someone whose nominal income is reset just once a year, hedging occurs iff inflation volatility is above about 30% per year.[54] This is in line with the empirical evidence.[55]

Price rigidity Prices are rigid if and only if

$$\psi\,\mathrm{CV}(\mathrm{mc})\leq\xi\frac{1-m}{m}\simeq\frac{2}{3}, \tag{52}$$

where CV(mc) is the coefficient of variation of marginal cost, i.e. its standard deviation divided by its average. For a typical good, one could imagine that the coefficient of variation

[51] This interpretation is not uncontroversial (Oprea, 2024; Banki et al., 2025; Oprea, 2025), but *in fine* seems to be indeed part of the psychology of the aversion to gambles (Puri, 2025; Enke and Shubatt, 2023; Shubatt and Yang, 2024).

[52] Note that this theory predicts that there will be more "effective loss aversion" in complex gambles, and under cognitive load. This is qualitatively true in Iyengar and Kamenica (2010) (in their experiment, when there are more lotteries to choose from, the simple, lower expected value lottery is chosen more). A variant is that the CA parameter ξ could be higher when in an adversarial relationship, i.e. when the agent knows that there is the possibility for an Akerlof lemons effect.

[53] For the "scale of the decision" parameter s_a, I take the average level of consumption in the applications where the decision is about consumption, and a natural scale, specified in each case, elsewhere—the average equity share in the inaction application.

[54] I assume that income and consumption have the same average.

[55] In the US and euro area, inflation volatility has been 1 to 3 percentage points, an order of magnitude below the threshold, and contracts are not indexed; in the Brazilian, Israeli and Chilean inflations it was far above, and indexation was pervasive.

is around 0.1, and the short-run elasticity of demand is $\psi \simeq 2$. Hence, the price is typically rigid: $0.2 < 2/3$. But say for fresh produce, which spoils, and is easy to substitute (one substitutes easily from strawberries to raspberries), one can imagine $\psi \simeq 5$, and CV $\simeq 0.3$. Hence, the price is flexible: $1.5 > 2/3$. The point is that this is the right order of magnitude.

Ramsey The tax is uniform if and only if (by Proposition 9) for all tax categories k:[56]

$$\bar{\psi}|\tau^R_{G_k} - \bar{\tau}| \leq \xi \frac{1-m}{m} \simeq \frac{2}{3} \tag{53}$$

With the same value $\bar{\psi} = 2$ as just above, the planner will keep the tax system uniform, unless he wants to tax or subsidize a particular category by more than 33%. This is a reasonable order of magnitude.[57] It explains why, with complexity aversion but without redistribution motives, the tax system should be kept uniform. In a richer model, one might add redistribution motives, which could lead to a somewhat less uniform tax system.[58]

Inaction Let us apply the inaction model to the rebalancing of the equity share, so that a^d represents how much the equity share deviates from the rational optimal equity share. The agent stays passive within the band (Proposition 14) $\left|a^d\right| \leq \sqrt{2\xi s_a \sigma_x}$. With $s_a \simeq 0.6$ (a typical equity share) and $\sigma_x \simeq 0.1$ (a typical volatility of the equity share if one lets it drift passively), this means that the agent will stay passive if the equity share drifts by less than about 28 percentage points. This is in the right range. For instance, Ameriks and Zeldes (2004) track TIAA-CREF participants over a ten-year period and find that 44% made "no changes whatsoever to either their flow or asset allocations" across the entire decade; 73% never once altered the allocation of their accumulated assets. Conditional on an asset transfer occurring, the mean absolute transfer is 23.5% of assets. Agnew et al. (2003) find the same in a large 401(k) plan: no trade occurs in 87.6% of participant-years, and 35% of pre-sample entrants remain entirely in the guaranteed-income fund, never trading over the full 52-month panel. Brunnermeier and Nagel (2008) report that 41% of PSID stock market participants over 1984–1999, and 43% over 1999–2003, undertook no purchase or sale of risky assets at all.

Choice overload I calibrated Figure 2 with $\xi = 2/3$, and 4 components, which gives reasonable-looking values. One can also tease this out more finely. There is local choice overload (Proposition 17) if and only if $\xi\sqrt{p} > 0.77$ i.e. with $m(2) = 1/2$, this is if and

[56] I express it in terms of the Ramsey desired tax, which might make the model more portable when calibrating to tax data, as it might include other motives, e.g. redistribution. Also, using $\bar{\tau} = \lambda/\bar{\psi}$ (to leading order in v_ψ), the value of public funds parameter λ drops out.

[57] Observed VAT systems have between one and four rates, against the thousands the traditional Ramsey formula prescribes.

[58] One could also calibrate the optimal number of brackets, which requires calibrating the cost of thinking κ, which I defer to another paper.

only if the number of components p is greater than 1.3. The calibrated model thus predicts overload when options have strictly more than one attribute to weigh, and none otherwise. The "complexity of the options" p is a moderator isolated by Chernev et al. (2015), and the reason why Scheibehenne et al. (2010) find a null when pooling simple and complex tasks.

Assessment of the calibration Let me conclude. The strong version of this exercise is that a single value $\xi \simeq \frac{2}{3}$ works well across lots of domains.[59] Now, those are just rough orders of magnitude, and it would be interesting to calibrate more finely. The parameters in economics (e.g. intertemporal elasticity of substitution, coefficient of loss aversion, etc.) typically take many papers to measure, and they often seem a bit unstable, as a variety of factors (e.g. distractions, here m) make the estimate difficult. I still think that it is encouraging that the theory delivers a unitless parameter, potentially "portable" across domains, and that it appears to give reasonable orders of magnitude in a broad range of domains.

7.2 Why this formulation?

I now propose some desirable criteria for a model of complexity aversion. They motivated the model proposed here. I prefer the "psychological" route to lead to the model, but for the discussions of the merit of other models, the language of criteria may be useful.

7.2.1 Some proposed criteria for complexity aversion

Suppose a generic decision $\max_a u(a, x)$, and assume that the CA cost has the form:

$$C^{\text{CA}} = G\left(\left(|a^r_{x_i}|, \sigma_{x_i}, m_i\right)_{i\in I}\right) \tag{54}$$

where G is some function with values in $[0, \infty]$. Here I assume that it depends only on the "germ" of the decision around $x = 0$, i.e. the rational slope $a^r_{x_i}$ (and I assume that it does not depend on its sign), and the magnitude of the typical innovation, σ_{x_i}. The idea is that the agent uses her "local" (around $x = 0$) knowledge to assess the situation, including complexity aversion.[60] The most generic criterion is the following.

"Unit invariance" criterion: The CA cost should be invariant in the units of x_i (i.e. are they in dollars or in cents). This means that it is invariant when changing $\left(|a^r_{x_i}|, \sigma_{x_i}\right)$ into $\left(|a^r_{x_i}|\lambda_i, \frac{\sigma_{x_i}}{\lambda_i}\right)$, for any $\lambda_i > 0$.

[59] Section C also analyzes the results of de Clippel et al. (2025) with the complexity aversion model proposed here, and finds a fine fit with their data.

[60] This is the tack taken generally in the sparse max (Gabaix, 2014); it leads to much tractability, and arguably to psychological realism.

Immediately, then, that forces the function to be (indeed, apply this to $\lambda_i = \sigma_{x_i}$):

$$C^{\text{CA}} = G\left((b_i, m_i)_{i \in I}\right) \tag{55}$$

where b_i is the total impact of the variable x_i in action space:

$$b_i \coloneqq |a^r_{x_i}| \sigma_{x_i} \tag{56}$$

I now propose two criteria specific to complexity aversion.

"Generation of full simplicity" criterion: The CA cost yields full simplicity for small incentives. Mathematically,

$$\left.\frac{\partial C^{\text{CA}}}{\partial b_i}\right|_{b_i = 0^+} > 0 \quad \text{(possibly infinite)} \quad \text{when } m_i < 1. \tag{57}$$

Indeed, as we saw, we need a first-order penalty for small incentives that are poorly understood ($m_i < 1$).[61] As we see lots of "perfect simplicity" in practice (rigid prices, no indexation, fairly uniform tax systems, etc.), and want to explain it with CA, we need this criterion.

"Finiteness" criterion. The CA cost is finite when the error is finite (i.e. $\sum_i b_i$ is finite).

$$\sum_i b_i < \infty \text{ implies } C^{\text{CA}} < \infty. \tag{58}$$

Indeed, in most situations (for instance, when watching say a room, with its objects, shades of color, luminosity, etc.), the number of variables is essentially infinite, but the CA cost is finite. We do not want a measure that would predict infinite CA costs in such situations. To capture that the true importance of the variables is finite, I impose $\sum_i b_i < \infty$. This says that, even if all variables were higher or lower by one standard deviation—by one "modulus", or "order of magnitude", of each deviation—the rational action would still change by a finite amount. Hence, the CA should also be finite.

The above three criteria are the most substantive ones. Now, I will impose one more criterion, solely for tractability.

"Separable homogeneity" criterion: The CA cost is homogeneous of degree 1 in b, i.e. for all $\lambda > 0$, $C^{\text{CA}}(\lambda b, m) = \lambda C^{\text{CA}}(b, m)$, and it is quasi-separable in the variables, i.e.

[61] To fight the first-order traditional efficiency gains of going from 0 incentive to a small incentive, one needs a penalty that is first order or larger. See e.g. Proposition 4.

has the form

$$C^{\mathrm{CA}} = g\left(\sum_i f\left(b_i, m_i\right)\right) \tag{59}$$

where $f(\cdot, m)$ and g are C^2 on $(0, \infty)$, g can take infinite values, and $f(0, m)$ is finite.

Finally, I will sometimes impose one last criterion, which is useful to pin down the "prefactor" $\overline{\xi}$ in (1).

"Unit invariance" for utility and actions criterion: the CA cost is invariant when changing the units of actions, and choices are invariant to affine transformations of the utility function. Mathematically, when we change the units of actions from a to λa, and utility from u to $\mu u + \nu$, where $\lambda > 0$ and $\mu > 0$ are positive scalars (these are changes of units) and ν is any real number, then the CA cost is simply multiplied by μ. In addition, the CA cost depends only on the "germ" at $x = 0$, i.e. depends on $u(a^d, 0)$, $u_{aa}(a^d, 0)$ and s_a, which expresses the "modulus" (typical size) of the action:[62]

$$C^{\mathrm{CA}}\big(\mu u(a) + \nu,\ \mu\lambda^{-2} u_{aa},\ \lambda s_a,\ \lambda b,\ m\big) = \mu\, C^{\mathrm{CA}}(u(a),\ u_{aa},\ s_a,\ b,\ m)\,. \tag{60}$$

7.2.2 Unique representation result

Then, we have the following proposition:

Theorem 1. *Suppose there are at least two variables. A CA cost satisfies the criteria of Unit invariance, Generation of full simplicity, Finiteness, and Separable homogeneity if and only if it has the form*

$$C^{CA}(b, m) = \sum_i b_i \bar{h}(m_i) \tag{61}$$

where $b_i \coloneqq |a^r_{x_i}| \sigma_{x_i}$, and $\bar{h}$ is a bounded function such that $\bar{h}(m) > 0$ for all $m < 1$. If in addition we impose the Unit invariance for utility and actions criterion, this becomes

$$C^{CA}(b, m) = |u_{aa} s_a| \sum_i b_i h(m_i) \tag{62}$$

for a function h with the same properties, that does not depend on u.

In light of Theorem 1, I took $h(m) = \xi(1 - m)$ for tractability, which led to the main proposal, (6). But in general, h can be any bounded function with $h(m) > 0$ for $m < 1$ and $h(1) = 0$: perfect understanding leads to no complexity aversion.

[62] This comes from basic dimensional analysis: a change of units that multiplies a by a factor λ will divide u_{aa} by a factor λ^2.

7.3 Discussion of alternative formulations

7.3.1 Some variants of complexity aversion that satisfy the main criteria

Given Theorem 1, I took $h(m) = \xi(1-m)$ for tractability. One variant is a step function, $h(m) = \xi 1_{m<1}$. Then, we are in a dichotomous world where "obvious things" ($m_i = 1$) have no penalty, and not completely obvious things ($m_i < 1$) incur a full penalty. This leads to a somewhat simpler model, where one may take

$$C^{\text{CA}} = \xi \sum_i \left|a^r_{x_i}\right| \sigma_{x_i} 1_{i \in I^{\text{Complex}}} \tag{63}$$

where $I^{\text{Complex}} = \{i : m_i < 1\}$ is a set of "complex" features, and the remaining features are "simple" ($m_i = 1$).

The separable homogeneity criterion is mostly for tractability—it is conceptually the least essential of all the criteria. Indeed, one could also admit any transformation $C^{\text{CA}\prime} := |u_{aa}| s_a^2 \Phi\left(\sum_i \frac{b_i h(m_i)}{s_a}\right)$ for some function Φ with strictly positive derivative. Then, essentially all the economics would still hold.

7.3.2 Some formulations that do not satisfy the criteria

Here I list some formulations that, though appealing in some respects, do not satisfy the two key criteria of "Generation of full simplicity" and "Finiteness".

Non-linear scalings One could have:

$$\tilde{C}^{CA} = \xi \left|s_a^{2-\beta} u_{aa}\right| \sum_i \left(\left|a^r_{x_i}\right| \sigma_{x_i}\right)^\beta h(m_i) \tag{64}$$

The criteria imposed $\beta = 1$ (Theorem 1), and it is worth seeing more directly why. Any $\beta \in [0,1]$ would deliver the "Generation of full simplicity" criterion. However, a $\beta < 1$ would violate the "Finiteness" criterion.[63] Likewise, a $\beta > 1$ would violate the "Generation of full simplicity" criterion, as it would generate a smaller than first-order penalty for small slopes b_i. However, it would satisfy the "Finiteness" criterion.

Information cost This approach does not satisfy the "Generation of full simplicity" criterion, as we saw (e.g. Propositions 3 and 8). This also includes the fixed cost of thinking,

[63] The proof gives a rigorous argument, but here is a clear example of the problem with $\beta < 1$. Take $\sigma_{x_i} = 1$ and $b_i = 1/n$ for $i = 1, \ldots, n$, and $b_i = 0$ otherwise. So $\sum_i b_i = 1$: the decision has bounded stakes. But the CA cost goes to infinity as n grows whenever $\beta < 1$ (as $\sum_{i=1}^n \left(\frac{1}{n}\right)^\beta = n^{1-\beta} \to \infty$).

which is just a type of information cost.

Counting measure of number of parameters This would be $C^{\text{CA,counting}} = G(n)$, where n is the number of variables, and G is some weakly increasing convex function. It does not satisfy the "Finiteness" criterion. Indeed, generically, a situation has a myriad of small elements, so $n \to \infty$. Of course, a counting measure is a useful proxy in some cases.

Second-order complexity aversion: Traditional regret theory A variant would be "second-order complexity aversion", which is proportional to the expected loss from mis-optimization:

$$C^{\text{CA},2} = \xi \mathbb{E}\left[u\left(a^r, x\right) - u\left(a^s, x\right)\right] = \xi \sum_i \mathbb{E}\left[-\frac{1}{2} u_{aa} \left(a^r_{x_i} x_i\right)^2\right] (1 - m_i) \tag{65}$$

It is in the tradition of regret, e.g. Loomes and Sugden (1982), Sarver (2008). The problem is that it violates the "Generation of full simplicity" criterion.

It may be useful to record the main differences between "complexity aversion" (proposed here) and "regret aversion" (from the tradition):

1. CA is first order (the loss is proportional to the error e, not e^2): otherwise, second-order losses will not robustly lead to full simplicity (e.g. no indexation); while regret aversion is second order. This is crucial, as second-order regret aversion violates the "Generation of full simplicity" criterion.

2. CA applies to the mistakes made by the agents in their mind, not in nature. It does not apply to ex post regret over nature's moves. Hence, CA does not change the utility of a rational agent (who makes no mistakes), unlike ambiguity aversion.

3. CA applies feature by feature: if there are several features (e.g. the decision entails looking at a tax rate, and an expected return, and a price), each of them is subject to CA, and the penalties add up. To see why this matters, consider the "square root" variant to regret aversion, e.g. having $C^{\text{CA},2\prime} = \bar{\xi}\sqrt{\sum_i \left(a^r_{x_i}\right)^2 \sigma^2_{x_i} (1 - m_i)}$. This will still violate the "Generation of full simplicity" criterion. Indeed, as soon as at least two terms under the square root are non-zero, the incremental penalty from each $a^r_{x_i}$ is proportional to $\left(a^r_{x_i}\right)^2$, not $\left|a^r_{x_i}\right|$: so de facto we have second-order complexity aversion at the feature level, and we will not get full simplicity.

7.3.3 Some other formulations

Finally, here are two proposals that can satisfy the criteria, but that I do not use for reasons of psychological applicability. Then, my CA proposal is just an alternative, applying more naturally perhaps than those proposals.

Maximin and taste for robustness. That has been a productive enterprise (Gilboa and Schmeidler, 1989; Hansen and Sargent, 2008; Cerreia-Vioglio et al., 2015; Carroll, 2015; Ilut and Schneider, 2023; Fudenberg and Mudekereza, 2026). Its psychology is more one of "pessimism" than of "taste for simplicity". Also, though mathematically possible, application to robustness seems, in a number of situations, awkward. For instance, in the context of Section 4, if the price of bread varies a bit, is there a great rational loss in some terrible scenario? In the Whole Foods example, should I be worried if the cost of delivery at 10am is one dollar lower? The complexity aversion I present here seems more plausible than a maximin behavior. In any case, the present model is clearly different from this robustness model; its tractability makes it possible to handle new applications; and it is good for the ecosystem to have different proposals for complexity.

Loss aversion for risk in the world. I am closest to it, as first-order complexity aversion is "loss aversion for randomness in one's mind", rather than randomness in the objective world. First, it might work for some cases, e.g. to buy a good with an expected price, as in Section 4. However, it does not seem to apply well to the other examples, e.g. lack of indexation: to use loss aversion, one would need to segregate the two variables ($\hat{y}$ and x), and apply loss aversion to them (Kőszegi and Rabin, 2006). Also, it seems awkward to apply to the Whole Foods case, where many prices are visible with full information—arguably the consumer seems not to like that because of complexity aversion, not because there is a loss. In any case, it is again useful to have a materially different model, and defer to future research some empirical assessment of the merit of the models.

8 Conclusion

This paper proposed a tractable theory of complexity aversion. It explains fairly easily a number of "obvious" features of the real world (that contracts and prices are "simple", e.g. not indexed)—obvious to common sense, but puzzling to the traditional economic model with no complexity costs. Because the penalty scales with the incentives of the problem, it is governed by a single unitless parameter ξ, and the same value $\xi \simeq \frac{2}{3}$ gives sensible orders of magnitude across contexts, such as indexation, price rigidity, taxes, inertia and choice overload.

Along the way, the model shed light on a number of issues of basic modeling. First-order complexity aversion is a robust way to generate simplicity. But cognition costs or information costs alone do not generate simplicity: they typically generate a dampened, but non-zero response to incentives.

Many interesting issues seem now within reach. A companion paper (in progress) explores the consequences for macro, starting from the fact that inflation makes planning more complex, hence has first order costs. Designing optimal mechanisms, taking into account the complexity costs borne by consumers, seems particularly useful from a practical point of view. The present paper's analysis of simple tax systems for the taxation of goods is a beginning, and much more can be done.

References

Aghion, Philippe, Ufuk Akcigit, Maxime Gravoueille, Matthieu Lequien, and Stefanie Stantcheva, "Tax Simplicity or Simplicity of Evasion? Evidence from Self-Employment Taxes in France," Working Paper, NBER 2024.

Agnew, Julie, Pierluigi Balduzzi, and Annika Sundén, "Portfolio Choice and Trading in a Large 401(k) Plan," *American Economic Review*, 2003, *93* (1), 193–215.

Ameriks, John and Stephen P. Zeldes, "How Do Household Portfolio Shares Vary with Age?," 2004. TIAA-CREF Working Paper, Columbia University.

Andriès, Marianne and Valentin Haddad, "Information aversion," *Journal of Political Economy*, 2020, *128* (5), 1901–1939.

Atkeson, Andrew and Ariel Burstein, "Pricing-to-market, trade costs, and international relative prices," *American Economic Review*, 2008, *98* (5), 1998–2031.

Banki, Daniel, Uri Simonsohn, Robert Walatka, and George Wu, "Decisions under Risk Are Decisions under Complexity: Comment," BFI Working Paper 2025-29, 2025.

Bar-Gill, Oren, *Seduction by Contract: Law, Economics, and Psychology in Consumer Markets*, Oxford University Press, 2012.

Ben-David, Itzhak and Alex Chinco, "Expected EPS × Trailing P/E," Technical Report, NBER 2024.

Bernheim, B Douglas and Dmitry Taubinsky, "Behavioral public economics," *Handbook of Behavioral Economics: Applications and Foundations 1*, 2018, *1*, 381–516.

_ , **Kristy Kim, and Dmitry Taubinsky**, "Welfare and the Act of Choosing," *Journal of Political Economy*, forthcoming.

Bordalo, Pedro, Nicola Gennaioli, Giacomo Lanzani, and Andrei Shleifer, "A cognitive theory of reasoning and choice," *The Quarterly Journal of Economics*, 2027.

Brown, Alexander L., Taisuke Imai, Ferdinand M. Vieider, and Colin F. Camerer, "Meta-analysis of Empirical Estimates of Loss Aversion," *Journal of Economic Literature*, June 2024, *62* (2), 485–516.

Brunnermeier, Markus K. and Stefan Nagel, "Do Wealth Fluctuations Generate Time-Varying Risk Aversion? Micro-Evidence on Individuals' Asset Allocation," *American Economic Review*, 2008, *98* (3), 713–736.

Buturak, Gökhan and Özgür Evren, "Choice Overload and Asymmetric Regret," *Theoretical Economics*, 2017, *12* (3), 1029–1056.

Caplin, Andrew, Dániel Csaba, John Leahy, and Oded Nov, "Rational inattention, competitive supply, and psychometrics," *The Quarterly Journal of Economics*, 2020, *135* (3), 1681–1724.

_ , **Mark Dean, and Daniel Martin**, "Search and Satisficing," *The American Economic Review*, 2011, *101* (7), 2899–2922.

Carroll, Gabriel, "Robustness and linear contracts," *American Economic Review*, 2015, *105* (2), 536–563.

Cerreia-Vioglio, Simone, David Dillenberger, and Pietro Ortoleva, "Cautious Expected Utility and the Certainty Effect," *Econometrica*, 2015, *83* (2), 693–728.

Chernev, Alexander, Ulf Böckenholt, and Joseph Goodman, "Choice Overload: A Conceptual Review and Meta-Analysis," *Journal of Consumer Psychology*, 2015, *25*, 333–358.

Chetty, Raj, Adam Looney, and Kory Kroft, "Salience and taxation: Theory and evidence," *American economic review*, 2009, *99* (4), 1145–1177.

Colliard, Jean-Edouard and Co-Pierre Georg, "Measuring regulatory complexity," *Journal of Financial Economics*, 2025, *174*, 104186.

Dari-Mattiacci, Giuseppe and Florencia Marotta-Wurgler, "Learning in Standard-Form Contracts: Theory and Evidence," *Journal of Legal Analysis*, 2022, *14* (1), 244–314.

de Clippel, Geoffroy, Paola Moscariello, Pietro Ortoleva, and Kareen Rozen, "Caution in the Face of Complexity," Technical Report, Mimeo 2025.

DellaVigna, Stefano and Matthew Gentzkow, "Uniform pricing in us retail chains," *The Quarterly Journal of Economics*, 2019, *134* (4), 2011–2084.

Enke, Benjamin and Cassidy Shubatt, "Quantifying lottery choice complexity," Technical Report, NBER 2023.

_ and Thomas Graeber, "Cognitive Uncertainty," *Quarterly Journal of Economics*, 2023, *138* (4), 2021–2067.

_ , _ , Ryan Oprea, and Jeffrey Yang, "Behavioral Attenuation," *Journal of Political Economy*, forthcoming.

Ergin, Haluk and Todd Sarver, "A unique costly contemplation representation," *Econometrica*, 2010, *78* (4), 1285–1339.

Eyster, Erik, Kristóf Madarász, and Pascal Michaillat, "Pricing Under Fairness Concerns," *Journal of the European Economic Association*, 08 2020, *19* (3), 1853–1898.

Farhi, Emmanuel and Xavier Gabaix, "Optimal Taxation with Behavioral Agents," *American Economic Review*, 2020, *110* (1), 298–336.

Fudenberg, Drew and Florian Mudekereza, "Complexity and Misspecification," 2026.

_ and Indira Puri, "Simplicity and Probability Weighting in Choice under Risk," *AEA Papers and Proceedings*, 2022, *112*, 421–425.

Gabaix, Xavier, "A sparsity-based model of bounded rationality," *Quarterly Journal of Economics*, 2014, *129* (4), 1661–1710.

_ , "Behavioral Inattention," *Handbook of Behavioral Economics*, 2019, *2*, 261–344.

_ , "A behavioral New Keynesian model," *American Economic Review*, 2020, *110* (8), 2271–2327.

_ and David Laibson, "Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets," *Quarterly Journal of Economics*, 2006, *121* (2), 505–540.

_ and Thomas Graeber, "The Complexity of Economic Decisions," Working Paper 33109, NBER 2024.

_ , David Laibson, Deyuan Li, Hongyi Li, Sidney Resnick, and Casper G de Vries, "The impact of competition on prices with numerous firms," *Journal of Economic Theory*, 2016, *165*, 1–24.

_ , _ , Guillermo Moloche, and Stephen Weinberg, "Costly Information Acquisition: Experimental Analysis of a Boundedly Rational Model," *American Economic Review*, September 2006, *96* (4), 1043–1068.

_ , Ralph SJ Koijen, Federico Mainardi, Sangmin Oh, and Motohiro Yogo, "Asset demand of US households," Technical Report, National Bureau of Economic Research 2024.

Gagnon-Bartsch, Tristan, Matthew Rabin, and Joshua Schwartzstein, "Channeled Attention and Stable Errors," Working Paper, Harvard Business School 2026.

Giglio, Stefano, Matteo Maggiori, Johannes Stroebel, and Stephen Utkus, "Five facts about beliefs and portfolios," *American Economic Review*, 2021, *111* (5), 1481–1522.

Gilboa, Itzhak and David Schmeidler, "Maxmin expected utility with non-unique prior," *Journal of mathematical economics*, 1989, *18* (2), 141–153.

Glaeser, Edward L and Andrei Shleifer, "A reason for quantity regulation," *American Economic Review, Papers and Proceedings*, 2001, *91* (2), 431–435.

Gonczarowski, Yannai A, Ori Heffetz, and Clayton Thomas, "Strategyproofness-exposing mechanism descriptions," Technical Report, NBER 2023.

Goodman, Aaron and Indira Puri, "Overvaluing simple bets: Evidence from the options market," *Journal of Financial Economics*, 2025, *172*, 104140.

Gormsen, Niels Joachim and Kilian Huber, "Corporate Discount Rates," *American Economic Review*, 2025, *115* (6), 2001–2049.

Graham, John R, "Presidential address: Corporate finance and reality," *The Journal of Finance*, 2022, *77* (4), 1975–2049.

Gulati, Mitu and Robert E Scott, *The three and a half minute transaction: Boilerplate and the limits of contract design*, University of Chicago Press, 2012.

Hansen, Lars Peter and Thomas J Sargent, *Robustness*, Princeton university press, 2008.

Hart, Oliver, *Firms, contracts, and financial structure*, Clarendon press, 1995.

Hartman-Glaser, Barney and Benjamin Hébert, "The insurance is the lemon: Failing to index contracts," *The Journal of Finance*, 2020, *75* (1), 463–506.

Heidhues, Paul and Botond Kőszegi, "Behavioral industrial organization," *Handbook of Behavioral Economics: Applications and Foundations 1*, 2018, *1*, 517–612.

Huck, Steffen and Georg Weizsäcker, "Risk, complexity, and deviations from expected-value maximization: Results of a lottery choice experiment," *Journal of Economic Psychology*, 1999, *20* (6), 699–715.

Ilut, Cosmin and Martin Schneider, "Ambiguity," in "Handbook of economic expectations," Elsevier, 2023, pp. 749–777.

Iyengar, Sheena S and Emir Kamenica, "Choice proliferation, simplicity seeking, and asset allocation," *Journal of Public Economics*, 2010, *94* (7-8), 530–539.

_ **and Mark R Lepper**, "When choice is demotivating: Can one desire too much of a good thing?," *Journal of Personality and Social Psychology*, 2000, *79* (6), 995.

Kahneman, Daniel and Amos Tversky, "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, 1979, *47* (2), 263–291.

Kamenica, Emir, "Contextual Inference in Markets: On the Informational Content of Product Lines," *The American Economic Review*, 2008, *98* (5), 2127–2149.

Karlsson, Niklas, George Loewenstein, and Duane Seppi, "The ostrich effect: Selective attention to information," *Journal of Risk and uncertainty*, 2009, *38* (2), 95–115.

Khaw, Mel Win, Ziang Li, and Michael Woodford, "Cognitive imprecision and small-stakes risk aversion," *The Review of Economic Studies*, 2021, *88* (4), 1979–2013.

Kőszegi, Botond and Matthew Rabin, "A model of reference-dependent preferences," *Quarterly Journal of Economics*, 2006, *121* (4), 1133–1165.

Laibson, David, "Golden Eggs and Hyperbolic Discounting," *Quarterly Journal of Economics*, 1997, *112* (2), 443–478.

Li, Shengwu, "Obviously strategy-proof mechanisms," *American Economic Review*, 2017, *107* (11), 3257–87.

Loomes, Graham and Robert Sugden, "Regret Theory: An Alternative Theory of Rational Choice Under Uncertainty," *The Economic Journal*, 12 1982, *92* (368), 805–824.

Manzini, Paola and Marco Mariotti, "Stochastic choice and consideration sets," *Econometrica*, 2014, *82* (3), 1153–1176.

Martínez-Marquina, Alejandro, Muriel Niederle, and Emanuel Vespa, "Failures in contingent reasoning: The role of uncertainty," *American Economic Review*, 2019, *109* (10), 3437–3474.

Matějka, Filip and Alisdair McKay, "Rational Inattention to Discrete Choices: A New Foundation for the Multinomial Logit Model," *American Economic Review*, 2015, *105* (1), 272–298.

Nagel, Lea and Roberto Saitto, "A Measure of Complexity for Strategy-Proof Mechanisms," *Econometrica*, 2027. Forthcoming.

Oprea, Ryan, "What makes a rule complex?," *American Economic Review*, 2020, *110* (12), 3913–51.

_ , "Decisions under risk are decisions under complexity," *American Economic Review*, 2024, *114* (12), 3789–3811.

_ , "Initial Reply to Banki, Simonsohn, Walatka and Wu (2025)," 2025. Available at SSRN 5182750.

Ortoleva, Pietro, "The price of flexibility: Towards a theory of Thinking Aversion," *Journal of Economic Theory*, 2013, *148* (3), 903–934.

Puri, Indira, "Preference for simplicity," 2018. Available at SSRN 3253494.

_ , "Simplicity and risk," *The Journal of Finance*, 2025, *80* (2), 1029–1080.

Rabin, Matthew, "Risk Aversion and Expected-utility Theory: A Calibration Theorem," *Econometrica*, 2000, *68* (5), 1281–1292.

Ramsey, Frank P., "A Contribution to the Theory of Taxation," *Economic Journal*, 1927, *37* (145), 47–61.

Rees-Jones, Alex and Dmitry Taubinsky, "Measuring 'schmeduling'," *The Review of Economic Studies*, 2020, *87* (5), 2399–2438.

Rubinstein, Ariel, *Modeling Bounded Rationality*, MIT press, 1998.

Sarver, Todd, "Anticipating regret: Why fewer options may be better," *Econometrica*, 2008, *76* (2), 263–305.

Scheibehenne, Benjamin, Rainer Greifeneder, and Peter M. Todd, "Can There Ever Be Too Many Options? A Meta-Analytic Review of Choice Overload," *Journal of Consumer Research*, 2010, *37* (3), 409–425.

Schwartzstein, Joshua and Adi Sunderam, "Using models to persuade," *American Economic Review*, 2021, *111* (1), 276–323.

Segal, Ilya, "Complexity and renegotiation: A foundation for incomplete contracts," *The Review of Economic Studies*, 1999, *66* (1), 57–82.

Shubatt, Cassidy and Jeffrey Yang, "Tradeoffs and comparison complexity," *arXiv preprint arXiv:2401.17578*, 2024.

Sims, Christopher A, "Implications of rational inattention," *Journal of Monetary Economics*, 2003, *50* (3), 665–690.

Spier, Kathryn E, "Incomplete contracts and signalling," *The RAND Journal of Economics*, 1992, *23* (3), 432–443.

Taubinsky, Dmitry and Alex Rees-Jones, "Attention Variation and Welfare: Theory and Evidence from a Tax Salience Experiment," *Review of Economic Studies*, 2017, *85* (4), 2462–2496.

Tibshirani, Robert, "Regression Shrinkage and Selection via the Lasso," *Journal of the Royal Statistical Society. Series B (Methodological)*, 1996, *58* (1), 267–288.

Tirole, Jean, "Cognition and incomplete contracts," *The American Economic Review*, 2009, *99* (1), 265–294.

Woodford, Michael, "Modeling imprecision in perception, valuation, and choice," *Annual Review of Economics*, 2020, *12*, 579–601.

Young, H Peyton and Mary A Burke, "Competition and custom in economic contracts: a case study of Illinois agriculture," *American Economic Review*, 2001, *91* (3), 559–573.

Online Appendix for "A Theory of Complexity Aversion"

Xavier Gabaix

August 21, 2026

Contents

A Theory complements **51**

A.1 Some simple cognitive microfoundations 51

A.2 How to recover frames, attention and default from data 52

A.3 Extensions to multidimensional actions and problems with budget constraints 55

A.4 Slightly different specifications for the basic penalty 55

A.5 Complexity aversion in discrete choice 57

A.6 A closed form for endogenous attention in the indexation problem 59

A.7 A microfoundation for the fact that competitive firms want to maximize social surplus 60

A.8 Complements on the role of thinking with endogenous attention 62

A.9 Shrouded attributes with complexity averse consumers: Complexity aversion as a partial force for consumer protection 63

A.10 Fixed price vs Cost plus contract: why the slope in a contract is often exactly 0 or 1 68

B Complements to the Ramsey problem **69**

B.1 More on simple taxes 69

B.2 Optimal tax system with costly endogenous attention 70

B.3 Ramsey problem: the regime with low complexity aversion, and "stealth taxes" 74

B.4 In the Ramsey problem, cognition costs alone (without complexity aversion) do not yield full simplicity 77

B.5 Proof of Proposition 10 78

B.6 Proof of Proposition 11 84

B.7 Proofs of Propositions 12 and 13 85

B.8 Proof of Proposition 32 85

C An analysis of de Clippel, Moscariello, Ortoleva and Rozen, based on complexity aversion **86**
C.1 Compound risk 89
C.2 Bayesian updating 91

D Omitted proofs **96**
D.1 Proof of Theorem 1 96
D.2 Proof of Proposition 5 98
D.3 Proof of Proposition 6 98
D.4 Proof of Proposition 8 99
D.5 Proof of Proposition 14 99
D.6 Proof of Proposition 16 101
D.7 Proof of Proposition 17 101
D.8 Proof of Proposition 21 101
D.9 Proof of Proposition 22 103

Notation. In the proofs, I abbreviate the cognition cost $C^{\text{Cognition}}$ as C^{Cog}.

A Theory complements

A.1 Some simple cognitive microfoundations

Derivation of (3) One derivation of (3), and the functional form for the noise, is the following. People receive noisy "raw" signals $\tilde{x}_i^s$ about x_i: $\tilde{x}_i^s = x_i + \tilde{\eta}_i$, where $\tilde{\eta}_i$ is a mean-zero noise uncorrelated with x_i. Then, if x_i and $\tilde{\eta}_i$ are jointly Gaussian, we have $\mathbb{E}\left[x_i|\tilde{x}_i^s\right] = x_i^s \coloneqq m_i\tilde{x}_i^s$, with $m_i = \frac{\sigma_{x_i}^2}{\sigma_{x_i}^2+\sigma_{\tilde{\eta}_i}^2}$. If noises are not Gaussian, then $x_i^s = m_i\tilde{x}_i^s$ is still the linear least-squares projection of x_i onto $\tilde{x}_i^s$, i.e. $m_i = \text{argmin}_{m_i}\,\mathbb{E}\left[(x_i - m_i\tilde{x}_i^s)^2\right]$, which yields $m_i = \frac{\mathbb{E}\left[x_i\tilde{x}_i^s\right]}{\mathbb{E}\left[\left(\tilde{x}_i^s\right)^2\right]} = \frac{\sigma_{x_i}^2}{\sigma_{x_i}^2+\sigma_{\tilde{\eta}_i}^2}$. We then define the "processed" signal as $x_i^s \coloneqq m_i\tilde{x}_i^s$, and set $\eta_i = m_i\tilde{\eta}_i$, which is uncorrelated with x_i. This yields (3), with:

$$\sigma_{\eta_i}^2 = m_i^2\sigma_{\tilde{\eta}_i}^2 = m_i^2\sigma_{x_i}^2\left(\frac{1}{m_i} - 1\right) = \sigma_{x_i}^2 m_i\left(1 - m_i\right)$$

Cost functions Typically, in calculations, we will assume a production function of thought (or precision), which leads to a cost function as follows:

$$C_i\left(m_i\right) = c_i m_i^{\frac{1}{1-\alpha}}, \qquad 0 < \alpha < 1 \tag{66}$$

where the cost c_i can be endogenized, as in Gabaix and Graeber (2024).

Why is it hard to handle tasks that require juggling numbers, and relatively "easy" to drive—does the model see that? We conclude with a remark on complexity across domains (linked to Moravec's Paradox: what's hard for computers is often easy for people and vice-versa). *Why is it hard to handle tasks that require juggling numbers, and relatively "easy" to drive* (at least, after two dozen hours of lessons)*—does the model see that?* Intuitively, this is because processing actual numbers is hard, while processing visual information is much easier. In the model, this is because the cost of processing parameter c_i in (66) is high when dealing with mathematical operations on numbers, and low when processing visual information: e.g., there is a fork on the road, and I should turn right. Hence, inside the model, when driving, the imprecision $1 - m_i$ is very small, and the complexity dread term (6) is very low. In heavy rain, so that seeing the road is harder, imprecision is higher, and driving is harder.

When people do not understand that they are making a mistake In this model, people are completely aware that they are making a mistake. This is clearly just a benchmark, probably right for most decisions. However, it is useful to keep in mind that there are some problems where a "tempting" answer seems "obvious", and is actually wrong. A famous one is the Monty Hall three-door problem[64], and behavioral economics has provided others (Frederick, 2005; Enke and Zimmermann, 2019). Those people are not "aware" of the mistake. This matters for economics: for instance, firms can exploit consumers' naïveté (Gabaix and Laibson, 2006; Heidhues and Kőszegi, 2018), as those consumers are not aware that they are making a mistake. Appendix A.9 builds a bridge with partial awareness: complexity aversion makes it hard for firms to exploit consumers.

A.2 How to recover frames, attention and default from data

In this research, the "attention", the "default", and the "frame" are taken as knowable. Here, I discuss how to elicit them from data.[65] In practice, there is often an "obvious" frame (e.g. prices of apples and bananas are processed separately), though sometimes the choice is a bit delicate (e.g. do people use nominal or inflation-adjusted prices).[66]

[64] Amazingly, the great mathematician Erdős got it wrong initially, and he was convinced only after several days (Vazsonyi, 1999)

[65] I already discussed how to elicit attention in Gabaix (2014) and Gabaix (2019); but here I discuss how to elicit the frame.

[66] One could also endogenize them, e.g. by using the "most useful" frame for instance.

First, let us define terms. We consider problems $\max_a u(a,x)$, where $a \in \mathbb{R}^A$ and $x \in \mathbb{R}^N$. Linearizing, and normalizing the default a to be 0, the rational action is:

$$a^r = a^r_x x$$

where a^r_x is an $A \times N$ matrix. Here 0 is only the origin for the expansion of the rational policy; the cognitive default x^d introduced next is a distinct object, and the associated default action $a^d = a^r_x x^d$ is generally non-zero. With inattention, we get, for the average action, an attention matrix M, which we take to be diagonal, and the default vector x^d:

$$a^s = a^r_x x^s, \qquad x^s = Mx + (I - M)\, x^d$$

But one could imagine that people use a transformation $\tilde{x} = Fx$, where $F = \frac{\partial \tilde{x}}{\partial x} = \tilde{x}_x$ is an $N \times N$ matrix (we keep the transformation to be linear; or F could be the derivative for small x of a non-linear transformation). For instance, the frame could be nominal, or real, assuming the inflation is somewhere in x. We can call the choice of F (or equivalently, the representation $x \mapsto \tilde{x}$) the "frame". Under that frame

$$\tilde{x}^s = M\tilde{x} + (I - M)\, \tilde{x}^d = MFx + (I - M)\, Fx^d$$

and the action is $a^s = a^r_{\tilde{x}} \tilde{x}^s = a^r_x x^s$ with $a^r_{\tilde{x}} = a^r_x F^{-1}$ and $x^s = F^{-1}\tilde{x}^s$, so

$$a^s = a^r_x x^s, \qquad x^s = F^{-1}MFx + \left(I - F^{-1}MF\right) x^d \tag{67}$$

Suppose that we have data on x, $a^s(x)$, where $a^s(x)$ is the average choice of the behavioral agent (with s meaning subjective, or sparse), and perhaps $x^s(x)$—where we elicit the perceived value of a variable (say, after the choice). Can we recover attention M and the frame F? Throughout, we suppose that we know a^r_x—perhaps, this is elicited under clear and distinct conditions, with the help of a guide.

Proposition 18. (Recovery of attention, and the frame) *Suppose that M's eigenvalues are all different, and that we know a^r_x. Suppose that we know $D_x = G_a a^s_x + G_x x^s_x \in \mathbb{R}^{d_D \times N}$, where the selection matrices G_a and G_x have dimensions $d_D \times A$ and $d_D \times N$, and that $G_a a^r_x + G_x$ has rank N; a sufficient condition, directly checkable in the data, is that D_x have rank N. For instance we know x^s_x, or we know a^s_x and a^r_x has rank N (which requires $A \geq N$). Then, up to reordering and rescaling, we can recover the attention M and the frame F. That is, we can recover some M_0 and F_0 that account for the data; and all the other solutions have the form $M = \Pi M_0 \Pi^{-1}$ and $F = \Pi S F_0$, where Π is a permutation matrix, and S is a diagonal*

invertible "scaling" matrix. If we know the frame F, the requirement is much weaker than rank N: attention m_i is recovered if and only if the i-th column of $(G_a a_x^r + G_x) F^{-1}$ is non-zero, i.e. if and only if dimension i moves the observed data. Hence M is recovered if and only if that matrix has no zero column—which a matrix of rank 1 can satisfy.

Proof. Let us start with a special case, where we have x_x^s from data. In the model, this is $x_x^s = F^{-1}MF$. Hence, given M is diagonal, x_x^s must be diagonalizable, and M is (up to permutation) the matrix of its eigenvalues. The rest is well-known diagonalization algebra: any solution for the diagonal has the form $M = \Pi M_0 \Pi^{-1}$, where Π is a permutation matrix (there are $N!$ of them). Also, a solution can be rescaled, as in $F = \Pi S F_0$ where S is a diagonal invertible "scaling" matrix. Indeed $F^{-1}MF = F_0^{-1} M_0 F_0 = x_x^s$, and this describes fully the set of solutions.

In the more general case, we have the data $D_x = G_a a_x^s + G_x x_x^s = (G_a a_x^r + G_x) F^{-1}MF$, for some (selection) matrices G_a and G_x. The proposition assumes that $G_a a_x^r + G_x$ has rank N. Hence, we can left-invert it and recover $x_x^s = F^{-1}MF$ from data, and we apply the special case we start with.

If we know the frame F (for instance, it is the identity matrix, a sort of natural frame), we no longer need to recover $F^{-1}MF$ in full. Write $B \coloneqq G_a a_x^r + G_x$, and post-multiply the data by F^{-1}: as M is diagonal, $D_x F^{-1} = (BF^{-1})\, M$, which reads, column by column,

$$\left(D_x F^{-1}\right)_{\cdot i} = m_i \left(BF^{-1}\right)_{\cdot i}.$$

Hence m_i is identified if and only if the i-th column of BF^{-1} is non-zero, and M is identified if and only if BF^{-1} has no zero column. This is much weaker than rank N: a rank 1 matrix, e.g. $BF^{-1} = (1, \ldots, 1)$, will do. But $D_x \neq 0$ does not suffice: if a column of BF^{-1} vanishes, the corresponding dimension never moves the observed data, and its attention cannot be recovered.[67] □

Economically, the "up to reordering and rescaling" makes sense: we do not care about the order of the perception vector (i.e., does the agent order the price vector as $(p_{\text{apple}}, p_{\text{banana}}, p_{\text{cherries}})$ or with another order), or scale (i.e. does the agent count in grams or kilograms). The hard data requirement is that to recover the frame we need enough data on actions, or on perceptions: we need N^2 pieces of independent data, enough to pin down the whole matrix $F^{-1}MF$.[68]

[67] For instance, take $N = 2$, $F = I$, and data on the action only, with $a_x^r = (1, 0)$. Then $D_x = (m_1, 0) \neq 0$, but the action never responds to x_2, so m_2 is not identified.

[68] The special case where we only have a_x^s is also interesting. We can recover $a_x^s = a_x^r x_x^s = a_x^r F^{-1}MF$. This is an $A \times N$ matrix. However, the unknowns live in a manifold of dimension N^2: N for M; plus $N^2 - N$ for F, up to the diagonal scaling matrix S. Hence, a necessary condition is $A \times N \geq N^2$, i.e. $A \geq N$. That

It is comparatively easier to recover the default x^d, as I record here.

Proposition 19. (Recovery of the default) *Suppose that we observe $D^d = G_a a^d + G_x x^d$, some transform of the default action and default parameter, where $G_a a_x^r + G_x$ has rank N. Then, we can recover the default x^d.*

Proof. We observe $D^d = (G_a a_x^r + G_x)\, x^d$. Then, as $G_a a_x^r + G_x$ has rank N, we recover x^d. □

In conclusion, we can in principle recover all the essential elements: the attention, the default, and the frame.

A.3 Extensions to multidimensional actions and problems with budget constraints

Here are natural extensions of the basic model. Suppose that the action is multidimensional, $a = (a_1, \ldots, a_K)$. Then, the complexity aversion is:

$$C^{\text{CA}} = \xi \sum_k \sum_i |u_{a_k a_k}| s_{a_k} |a_{k,x_i}^r| \sigma_{x_i} \left(1 - m_i^k\right) \tag{68}$$

where m_i^k is the attention to the i-th dimension when taking the k-th action; a_{k,x_i}^r is the derivative of the action a_k with respect to the information x_i; all derivatives being evaluated at the default action a^d and the default information $x^d = 0$. This specification sums own-curvature terms $|u_{a_k a_k}|$, so it is valid when the local curvature is (approximately) diagonal, or when the action components are behaviorally primitive; the general case would require the full Hessian.

If we have a budget constraint, then one can proceed in the same way, using the Lagrangian L (utility + Lagrange multiplier times constraints), replacing u_{aa} by L_{aa}. To obtain a value that is invariant to non-linear transformations of the utility function, one can use the "compensated" derivative $\bar{a}_x$, defined exactly as in Gabaix (2014), Section V.D of the main text and Section X of the online appendix.

A.4 Slightly different specifications for the basic penalty

In the main paper, I took the specification:

$$a^\delta := \mathbb{E}\left[a^r - a^s \mid x\right] = a^r \left(1 - m\right) \tag{69}$$

count is not sufficient by itself: as in the proposition, a_x^r must have rank N, so that it can be inverted on the left—which requires $A \geq N$, and holds for a generic a_x^r.

and for the error: $e \coloneqq \sigma_{a^\delta}$, which led to the core CA penalty:

$$e = |a_x^r|\, \sigma_x\, (1 - m) \tag{70}$$

The charm is that the penalty is affine in m. Given traditional losses from imperfect cognition are also proportional to $1 - m$ (see e.g. (4)), this leads to lots of tractability.

But of course, one could imagine lots of small variants, that would lead to slightly different penalty: as in Theorem 1, one could conceive of any $h(m)$, not just $h(m) = 1 - m$.[69] In particular, a tempting formulation might be:

$$e_2 \coloneqq \left(\mathbb{E}\left[(a^r - a^s)^2\right]\right)^{1/2} = |a_x^r|\, \sigma_x\, (1 - m)^{1/2} \tag{71}$$

The problem is that this is nonlinear in m. Conceptually, this is innocuous—essentially nothing very important would change, but it would make the analysis messier. Hence, I did not use it.

If someone absolutely wanted a measure based on the expected squared loss, they might take:

$$e_3 \coloneqq \frac{\mathbb{E}\left[(a^r - a^s)^2\right]}{\left(\mathbb{E}\left[(a^r - a^d)^2\right]\right)^{1/2}} = \frac{|a_x^r|^2\, \sigma_x^2\, (1 - m)}{\left(|a_x^r|^2\, \sigma_x^2\right)^{1/2}} = |a_x^r|\, \sigma_x\, (1 - m) = e \tag{72}$$

so that it is identical to the basic penalty e in (70) that I use in the main paper. There, the penalty is first order in the stakes ($|a_x^r|\, \sigma_x$), which is what matters.

Nonlinear expressions for the CA penalty One could also have a nonlinear expression for the CA penalty, such as the following (where $a^r(x_i, 0)$ means that all the entries other than x_i are set to their default value of 0):

$$C^{\text{CA,nonlin}} \coloneqq \bar{\xi} \sum_i \mathbb{E}[|\mathbb{E}[a^r(x_i, 0) - a^r(x_i^s, 0)|x_i]|^2]^{1/2} \tag{73}$$

Then, by Taylor expansion, for small σ_{x_i}, we recover the expression in the main text, as

$$\mathbb{E}[|\mathbb{E}[a^r(x_i, 0) - a^r(x_i^s, 0)|x_i]|^2]^{1/2} = \left|a_{x_i}^r\right|\, \sigma_{x_i}\, (1 - m) + O(\sigma_{x_i}^2)$$

[69] The main economics is that $h(m)$ is (weakly) decreasing in m, and $h(1) = 0$: more confusion (lower m) leads to more dread.

A.5 Complexity aversion in discrete choice

Most of this paper is for a continuous action (perhaps multidimensional). Here I show a way to handle discrete choice.

A.5.1 Prior results on discrete choice, without complexity aversion

I draw from Gabaix and Graeber (2024), which I first summarize here. We consider the discrete choice

$$\max_{k:k\in K\subseteq\mathbb{N}} v\left(k,x\right) \tag{74}$$

where the choice k is in a discrete set K, and yields rational utility $v\left(k,x\right)$, where for each k, $v\left(k,x\right)$ is twice continuously differentiable in x.

We define as $a\left(k\right)$ the probability of playing action k, so $\sum_k a\left(k\right)=1$, $a\in\Delta\left(K\right)$. We then define the extended utility function, concave in a, as

$$u\left(a,x\right)\coloneqq\sum_k a\left(k\right)v\left(k,x\right)-\beta D\left(a,a^d\right), \tag{75}$$

where $D\left(a,a^d\right)$ is a divergence measure that is convex in a, $a^d\in\Delta\left(K\right)$ is a default action probability, and $\beta>0$ is discussed later. We shall take the Kullback-Leibler divergence between a and a^d:

$$D\left(a,a^d\right)=1+\sum_k a\left(k\right)\left(\ln\frac{a\left(k\right)}{a^d\left(k\right)}-1\right) \tag{76}$$

The main analytical results are the following. The solution of $\max_a u\left(a,x\right)$ is $a=a^F\left(a^d,v\left(x\right),\beta\right)$, where the probability of playing action k is:

$$a^{F,k}\left(a^d,v\left(x\right),\beta\right)=\frac{a^{k,d}e^{\frac{v^k}{\beta}}}{\sum_{k'}a^{k',d}e^{\frac{v^{k'}}{\beta}}} \tag{77}$$

A change in information dx_i creates a change in the probability of choosing action k equal to

$$da^k=\frac{1}{\beta}a^k\sum_i\left(v^k_{x_i}-\bar{v}_{x_i}\right)dx_i \tag{78}$$

where $\bar{v}_{x_i}$ is the average value of valuation changes created by the change in variable x_i, evaluated at the default, where $a=a^d$:

$$\bar{v}_{x_i}=\sum_k a^{k,d}v^k_{x_i}. \tag{79}$$

In addition, $u_{aa} = -\beta \operatorname{Diag}\left(\frac{1}{a^k}\right)$.

Hence, the expected utility after acquiring precision m_i is, up to higher order terms in $\sigma_{x_i}^2$:

$$U(m) = U(0) + \sum_i V_i m_i, \qquad U(0) = \sum_k a^{d,k} \mu^k \tag{80}$$

where V_i is the value of information in dimension i, with:

$$V_i = \frac{V_i^0}{2\beta}, \qquad V_i^0(a) = \sum_k a^k \left(v_{x_i}^k - \bar{v}_{x_i}\right)^2 \sigma_{x_i}^2 \tag{81}$$

If information in dimension i only affects an option $k(i)$ (i.e., if $v_{x_i}^{k'} = 0$ for $k' \neq k$), V_i^0 reduces to:

$$V_i^0(a) = a^{k(i)} \left(1 - a^{k(i)}\right) \left(v_{x_i}^{k(i)}\right)^2 \sigma_{x_i}^2 \tag{82}$$

Finally, we set a^d and β in the following manner. First, we define $a^{d,0}(k) \coloneqq \frac{1}{|K|}$ the "ignorance prior" on the actions, assigning equal probability to all actions; and compute an initial value of β, $\beta^0 \coloneqq \beta^F\left(a^{d,0}\right)$,

$$\beta^F\left(a^d\right) \coloneqq \bar{\beta}(\sum_i V_i^0\left(a^d\right))^{1/2}, \tag{83}$$

where $\bar{\beta}$ is a unitless parameter and V_i^0 is in (81). Then, we obtain the default action probability, $a^d = a^F\left(a^{d,0}, v\left(x^d\right), \beta^0\right)$, where a^F is the function in (77), and its associated $\beta = \beta^F\left(a^d\right)$.

A.5.2 Complexity aversion in discrete choice

The agent adds a complexity penalty to each option k:

$$\tilde{v}(k,x) \coloneqq v(k,x) - C^{\mathrm{CA},k} \tag{84}$$

$$C^{\mathrm{CA},k} = \xi \sum_i |v_{x_i}^k| \sigma_{x_i} (1 - m_i) \tag{85}$$

Hence, the CA penalty for option k is the sum of the "errors" $|v_{x_i}^k|\sigma_{x_i}(1-m_i)$, times the unitless CA parameter ξ.[70]

[70]Here, I choose a complexity penalty proportional to $|v_{x_i}^k|$. Hence, an option with zero features has no complexity penalty. The alternative would have been to use $|v_{x_i}^k - \bar{v}_{x_i}|$ rather than $|v_{x_i}^k|$, where $\bar{v}_{x_i}$ is the average impact of x_i on valuations, as in (79). This might have been more in line with the information-theoretic perspective of (81), but, I think, psychologically implausible, as it would give the 0 option a non-zero complexity penalty. In this inductive step of mathematizing the psychology, it is better to follow the paths suggested by psychology, rather than the most obvious paths suggested by mathematics. There

Total complexity aversion is

$$C^{\text{CA}} = \sum_k a^{k,d} C^{\text{CA},k} \tag{86}$$

Otherwise, the model is the same. The optimization of attention does take into account the CA term. So, the full utility is:

$$U(m) = U^r - \sum_i V_i (1 - m_i) - C^{\text{CA}}$$

As a side note, we can expand total CA as:

$$C^{\text{CA}} = \sum_k a^{k,d} C^{\text{CA},k} = \xi \sum_i \left(\sum_k a^{k,d} \left| v_{x_i}^k \right| \right) \sigma_{x_i} (1 - m_i) = \xi \sum_i D_{x_i}^u \sigma_{x_i} (1 - m_i) \tag{87}$$

$$D_{x_i}^u \equiv \sum_k a^{k,d} \left| v_{x_i}^k \right| \tag{88}$$

This $D_{x_i}^u$ is the discrete-action counterpart of what we had in the basic continuous action case: $D_{x_i}^u = \left| s_a u_{aa} \frac{\partial a^r}{\partial x_i} \right|$. It captures how much x_i changes the action, times the utility consequences of that. It shows how the discrete-choice penalty maps back into the continuous-action one.

I developed this machinery for general applicability. It is applied in Section 6.2 to analyze choice overload.

A.6 A closed form for endogenous attention in the indexation problem

Solving for the optimum m can be a bit ugly. One closed form obtains with

$$C^{\text{Cognition}}(m) = \kappa \left(m - m^d \right)^+ \tag{89}$$

for some "default" attention $m^d \in [0, 1)$. This captures that there is a "free" minimum of attention m^d, but that paying higher attention is effortful.

Proposition 20. (Optimal attention and indexation with linear cognitive cost function). *Suppose the cognition cost (89). Then, the full utility function of the consumer (22) is*

could be advantages to this path, e.g. it simplifies decision-making (at a meta-cognitive level, as in Gabaix and Graeber (2024), Section 6 on imperfect metacognition), but I will not explore that extra microfoundation layer here.

maximized at either $m = m^d$ or $m = 1$. We have $m = 1$ iff $\mathcal{V}(1) \geq \mathcal{V}\left(m^d\right)$, i.e. iff:

$$b(1)^2 - b\left(m^d\right)^2\left(2 - m^d\right) \geq \kappa\left(1 - m^d\right) \tag{90}$$

and otherwise $m = m^d$ and $b = b\left(m^d\right)$.

In particular, when $b\left(m^d\right) = 0$, the optimum indexation level is either no indexation ($b = 0$), or full indexation ($b = b(1) = b^r$).

Proof. Using (21) and the envelope theorem, with $\mathcal{V}(m) \coloneqq V(b(m), m)$

$$\mathcal{V}'(m) = \frac{d}{dm}V(b(m), m) = \frac{\partial V(b, m)}{\partial m} = b^2 + 2\xi|b| - \kappa$$

Observe that $b(m)$ is weakly increasing. So, if for some $m_0 \in \left[m^d, 1\right]$, $\mathcal{V}'(m_0) \geq 0$, then for all $m \geq m_0$, we will have $\mathcal{V}'(m) \geq \mathcal{V}'(m_0) \geq 0$ hence $\mathcal{V}(1) \geq \mathcal{V}(m)$. Likewise, if $\mathcal{V}'(m_0) \leq 0$, $\mathcal{V}\left(m^d\right) \geq \mathcal{V}(m)$. Hence, the two candidate solutions are $m = m^d$ or $m = 1$. Hence, we have $m = 1$ iff $\mathcal{V}(1) \geq \mathcal{V}\left(m^d\right)$. This is true if the cognition costs κ are not too large. In the case $m^d = 0$, we have $m = 1$ iff $\mathcal{V}(1) \geq \mathcal{V}(0)$, i.e., as $b(1) = b^r$, iff $\left(b^r\right)^2 - 2b(0)^2 \geq \kappa$. This simplifies to $\left(b^r\right)^2 \geq \kappa$ when $b(0) = 0$, i.e. when $b^r \leq \xi$, so that the consumer does not index at zero attention. □

This "extremal property" (that the optimal attention is either m^d or 1) works also for "nice" cost functions in a range $[m_d, 1]$. Indeed, observe that (calling $C^{\text{Cognition}}(m)$ simply $C^{\text{Cog}}(m)$):

$$\mathcal{V}'(m) = \bar{c}(m) - C^{\text{Cog}\prime}(m), \qquad \bar{c}(m) \coloneqq b^2 + 2\xi|b| \tag{91}$$

So, if $C^{\text{Cog}\prime}(m) \leq \bar{c}(m)$ for all m, then $\mathcal{V}'(m) \geq 0$ for all m, and the optimum is $m = 1$. If $C^{\text{Cog}\prime}(m) \geq \bar{c}(m)$ for all m, then $\mathcal{V}'(m) \leq 0$ for all m, and the optimum is $m = m^d$.

A.7 A microfoundation for the fact that competitive firms want to maximize social surplus

Section 4 assumes that the firm is "benevolent" —wants to maximize social surplus. Here is a specific microfoundation. It is quite traditional: as in Adam Smith and Arrow-Debreu, competitive firms, by their actions, maximize social surplus (at least, when consumers appreciate enough their expected utility). Still, for completeness, it is useful to fill in the details.

There are at least two firms, and a mass N of consumers. On day t, the marginal cost is $1 + x_t$, the same across firms (this is not essential). At time 0 (at the beginning of the year),

each firm f proposes a pricing scheme $p_f(x_t) = \bar{p}_f + b_f x_t$ (where it chooses the average price $\bar{p}_f$ and the slope b_f) and also a "membership fee" that the consumer has to pay, a lump-sum payment F_f (this is not essential either, but makes the analysis more general, and simpler).[71]

At the beginning of the year, each consumer sees the menus offered by the firms, and chooses which firm to sign up with—which firms to go to every day $t = 1, \ldots, T$ in that year. For instance, the consumer chooses which bakery to go to every day, in a way that maximizes her utility over the year. Formally, the firm gets an exclusive contract within the year (but has committed to the pricing scheme above).[72]

The consumer is fully conscious of her expected utility, including cognition and complexity aversion costs.[73] She chooses a firm f that maximizes her expected utility surplus, from this pricing scheme, i.e.

$$V_f = -F_f + TV_f^{\text{cons}} \tag{92}$$

The first part is the lump-sum payment. The second part is the expected utility of the consumer, net of the cost. The consumption is chosen as in the main body of the paper, as $1 + a_f(x^s) = \text{argmax}_{1+a} U(1+a) - p_f(x^s)(1+a)$, based on the perceived price $p_f(x^s)$.

The firm's profit is

$$\Pi_f = N_f \left(F_f + T\mathbb{E}\left[\left(p_f(x) - c(x)\right)\left(1 + a_f(x^s)\right)\right]\right)$$

where N_f is the number (mass) of consumers choosing the firm, and the last term is the average profit per consumer.

The timing is: each firm f proposes simultaneously $(F_f, \bar{p}_f, b_f)$, to maximize its expected profit Π_f; then each consumer decides which firm to choose. Firms play Nash.

Lemma 2. *In the competition game above, the only Nash equilibrium is: $\bar{p}_f = 1$, b_f as in the paper, maximizing total welfare, and the fee F_f set to rebate the operating margin, $F_f = -T\psi b_f(1 - b_f) m\sigma_x^2$. Firms' profits are zero.*

Proof. The reasoning is completely conventional. Because we allow for fixed payments F_f, firms have zero profits, and offer the maximum consumer's surplus given zero profits—hence, they maximize total social surplus. Given that, the average price should be the

[71] I assume that the firm can commit to the scheme. This is obviously easy to do when the outcome is a constant price $\bar{p}_f$.

[72] This is important, to prevent arbitrage: otherwise, take the case where the outcome is a constant price equal to 1. Then, on a day with a negative x_t, a deviating firm could undercut its competitors (which offer a price of 1), by offering a price of $1 + \frac{1}{2}x_t < 1$, explaining to consumers that they save money by coming to this deviating firm, and make a profit.

[73] This is where the paper deviates from the literature on shrouded attributes and firms exploiting naive consumers (see footnote 29): in that benchmark, the consumer is not conscious of all the future elements of the situations, for instance does not take into account "surprise" fees.

average marginal cost, so $\bar{p}_f = 1$.[74] Then, the derivation of the optimal b_f is the one that maximizes social surplus (which is the consumer's utility, as equilibrium profits are zero) as in the main body of this paper. Finally, at $\bar{p}_f = 1$ the expected operating profit per consumer per period is $\mathbb{E}\left[(p-c)(1+a)\right] = \psi b_f (1-b_f) m\sigma_x^2$ by (148), which is strictly positive for an interior slope $b_f \in (0,1)$; so competition drives the fee down to the announced rebate, and $F_f = 0$ only in the corner cases $b_f \in \{0,1\}$ or $m = 0$. □

A.8 Complements on the role of thinking with endogenous attention

The following gives sufficient conditions for full simplicity, in the context of the model of section 4.

Proposition 21. (Optimal price simplicity, first-order complexity aversion, endogenous attention) *Suppose first-order complexity aversion, and endogenous attention m, with attention strictly costly beyond its free default, $C^{Cognition\prime}\left(m^{d+}\right) > 0$. A perfectly simple contract ($b = 0$) is* locally *optimal (i.e. the social welfare $W(b)$ has a local optimum at 0) if and only if (31) holds. For a perfectly simple contract to be* globally *optimal, the following are sufficient conditions: $C^{Cognition}(1) > \psi\sigma_x^2$ (reaching full precision is very costly), and the complexity aversion cost ξ is large enough: $\xi \geq \xi^*$, with:*

$$\xi^* := \max_{m<1} \frac{\psi\sigma_x^2 - C^{Cognition}(m)}{s_a\sigma_x(1-m)} \tag{93}$$

The intuition is the following: if it is mentally costly enough to fully understand the situation, people will have a limited understanding of it, so $m < 1$. But then, if $b > 0$ and the complexity aversion ξ is high enough, the "dread" term (26) is high enough to wipe out the allocative benefits, $\frac{1}{2}\psi\sigma_x^2$.

The cutoff ξ^* is higher when cognition is cheaper ($C^{\text{Cognition}}$ is lower), and when the traditional benefits from a rational allocation ($\frac{1}{2}\psi\sigma_x^2$) are higher (when the rational elasticity of demand ψ and the volatility of prices σ_x are higher).

To complete the picture, let us consider a case where we need not have perfect simplicity: intuitively, this is the case when the marginal cost of thinking is low enough, compared to the social surplus gains of having a complex contract.

Proposition 22. (When full rationality is easy enough, contracts should not be maximally simple) *Suppose that $C^{Cognition}$ is convex and differentiable with $C^{Cognition}(0) = 0$, and that*

[74] One can also check that directly.

$\frac{dC^{Cognition}}{dm}(1) < \frac{1}{2}\psi\sigma_x^2$. *Then, perfect simplicity ($b = 0$) is non-optimal (and indeed, less desirable than $b = 1$, the optimum slope with traditional agents), even with a high complexity aversion parameter ξ.*

The economic interpretation of the condition is that the (marginal) cost of thinking $\frac{dC^{\text{Cognition}}}{dm}(1)$ is less than the surplus generated at the optimum, which is $\frac{1}{2}\psi\sigma_x^2$. So, some complexity is important if the task is "easy", or "very important" (it yields a high social surplus $\frac{1}{2}\psi\sigma_x^2$). Hence, we have complexity when the cost of getting it exactly right is not too large.

This concludes the analysis of first-order complexity aversion. We see that when it is large enough (ξ large enough), it generates full simplicity, at least if it is hard enough for the agents to completely understand the situation (Propositions 7 and 21). At the same time, complexity aversion is necessary to account for full simplicity (Proposition 8).

A.9 Shrouded attributes with complexity averse consumers: Complexity aversion as a partial force for consumer protection

How does complexity aversion modify firms' ability to manipulate consumers? Complexity (adding features) has a new effect, (i) a source of suspicion: it lowers utility by making the product feel more "suspicious"—in addition to two effects already understood, (ii) a source of profit: the pitfall for naive consumers, and (iii) a source of inefficiency: the effort required of rational consumers. To see how they balance, I revisit Gabaix and Laibson (2006), but with complexity averse consumers. We will see that complexity aversion helps make it less likely that the market is shrouded, see mostly (97) and Proposition 24.

A.9.1 Environment

I revisit Gabaix and Laibson (2006), henceforth GL, with complexity-averse naive consumers. The environment is theirs, with one change, isolated in Assumption A.9.1 below.

Assumptions.

(A1) A continuum of identical firms competes with free entry. Each posts a contract $(p, \hat{p})$: a visible base price $p \in \mathbb{R}$ and an add-on price $\hat{p} \in [0, \bar{p}]$, where $\bar{p}$ is a legal cap. Marginal costs are zero. Consumers observe the menu of posted contracts and each buys one unit of the base good from one firm.

(A2) A fraction $\alpha \in (0,1)$ of consumers is naive and $1-\alpha$ is rational. Rational consumers see $\hat{p}$ and may avoid the add-on at effort cost e, so

$$U^r(p,\hat{p}) = -p - \min(\hat{p}, e).$$

(A3) $0 < e < \bar{p}$. (If $e \geq \bar{p}$, avoidance never binds and the shrouding problem is trivial.)

(A4) Naive consumers do not see $\hat{p}$ and pay it, but feel complexity aversion in proportion to it. Writing

$$\bar{\xi} \coloneqq M\xi \tag{94}$$

for the awareness-adjusted complexity aversion, where $M \in [0,1]$ is the naives' awareness of the shrouded attribute, their *perceived* surplus is

$$V^n(p,\hat{p}) = -p - \bar{\xi}\hat{p},$$

while their true surplus is $-p-\hat{p}$. Choices are made on V^n. GL is the case $\bar{\xi}=0$, and I focus on $\bar{\xi} \in (0,1)$.

Profit per customer is $\pi^n = p + \hat{p}$ on a naive, and $\pi^r = p + \hat{p}\mathbf{1}_{\hat{p}\leq e}$ on a rational, since a rational buys the add-on exactly when $\hat{p} \leq e$. Where indifferent, I adopt the convention that a rational buys the add-on; all statements below are otherwise the $\varepsilon \to 0$ limits of contracts with $\hat{p} = e - \varepsilon$.

A.9.2 Sorting wedge

The new term $-\bar{\xi}\hat{p}$ makes the two types value the add-on price differently, and so creates a screening instrument absent from GL. Define the sorting wedge

$$w(\hat{p}) \coloneqq \min(\hat{p}, e) - \bar{\xi}\hat{p}, \tag{95}$$

the amount by which a rational's disutility from the add-on exceeds a naive's. It is tent-shaped: increasing at rate $1-\bar{\xi}$ on $[0,e]$, decreasing at rate $\bar{\xi}$ on $[e,\bar{p}]$, with peak $w(e) = e\left(1-\bar{\xi}\right)$.

Lemma 3. (Cream-skimming) *Let a market configuration deliver equilibrium utilities U^n (perceived) to naives and U^r to rationals. A contract attracting naives but not rationals exists if and only if $w(\hat{p}) \geq U^n - U^r$ for some $\hat{p}$, and the largest profit any such contract earns, per naive, is*

$$\sup\left\{\hat{p}\left(1-\bar{\xi}\right) - U^n \;:\; w(\hat{p}) \geq U^n - U^r\right\}. \tag{96}$$

Proof. A contract $(p, \hat{p})$ attracts naives and repels rationals iff $-p - \bar{\xi}\hat{p} \geq U^n$ and $-p - \min(\hat{p}, e) \leq U^r$, i.e. $-U^r - \min(\hat{p}, e) \leq p \leq -U^n - \bar{\xi}\hat{p}$. This interval is non-empty iff $w(\hat{p}) \geq U^n - U^r$. Profit per naive is $p + \hat{p}$, maximized at the upper limit, giving $\hat{p}\left(1 - \bar{\xi}\right) - U^n$. □

Lemma 3 already isolates the role of complexity aversion. In the GL shrouded configuration $(p, \hat{p}) = (-\alpha\bar{p}, \bar{p})$ we have $U^n = \bar{p}\left(\alpha - \bar{\xi}\right)$ and $U^r = \alpha\bar{p} - e$, so $U^n - U^r = e - \bar{\xi}\bar{p} = w(\bar{p})$: the wedge binds exactly at the incumbent contract. When $\bar{\xi} = 0$ the wedge is flat on $[e, \bar{p}]$ and there is no slack anywhere, which is why GL's equilibrium cannot be skimmed. When $\bar{\xi} > 0$ the wedge is strictly decreasing there, so every $\hat{p} < \bar{p}$ has slack $\bar{\xi}(\bar{p} - \hat{p}) > 0$, and (96) approaches $\bar{p}(1 - \alpha) > 0$ as $\hat{p} \uparrow \bar{p}$.

A.9.3 Threshold

Lemma 4. *The transparent contract $T \coloneqq (-e, e)$ breaks even on both types, gives $U^r = 0$ and $U^n = e\left(1 - \bar{\xi}\right)$, and admits no profitable cream-skimming deviation.*

Proof. Both types buy the add-on at $\hat{p} = e$, so revenue per customer is $p + \hat{p} = 0$. Then $U^n - U^r = e\left(1 - \bar{\xi}\right) = \max_{\hat{p}} w(\hat{p})$, attained only at $\hat{p} = e$. By Lemma 3 the only admissible skim is $\hat{p} = e$ itself, whose profit is $e\left(1 - \bar{\xi}\right) - U^n = 0$. □

The economics: at transparent contract T the firm recovers its base-price subsidy through a fee set just low enough that rationals pay it rather than avoid it. No type cross-subsidizes the other, so there is nothing to skim.

Proposition 23. (Shrouding threshold with complexity averse naive consumers) *Define*

$$\alpha^* \coloneqq \frac{e}{\bar{p}} + \bar{\xi}\left(1 - \frac{e}{\bar{p}}\right). \tag{97}$$

Both types strictly prefer the shrouded contract $S \coloneqq (-\alpha\bar{p}, \bar{p})$ to the transparent contract $T = (-e, e)$ if and only if the proportion of naive consumers is $\alpha > \alpha^*$.

Proof. $U^r(S) - U^r(T) = \alpha\bar{p} - e$, positive iff $\alpha > e/\bar{p}$. $U^n(S) - U^n(T) = \bar{p}\left(\alpha - \bar{\xi}\right) - e\left(1 - \bar{\xi}\right)$, positive iff $\alpha > \alpha^*$. Since α^* is a convex combination of 1 and $e/\bar{p}$, it exceeds $e/\bar{p}$, so the naive constraint is the binding one. □

Note that $\alpha^* > \max\left(\bar{\xi}, e/\bar{p}\right)$ strictly whenever $\bar{\xi} \in (0, 1)$ and $e \in (0, \bar{p})$. The relevant outside option for a naive consumer is not the marginal-cost offer $(0, 0)$—which gives her zero—but the break-even low-add-on offer T, which gives her $e\left(1 - \bar{\xi}\right) > 0$ because she discounts by $\bar{\xi}$ a fee she only half-perceives. Shrouding is therefore harder to sustain than a comparison against marginal-cost entry would suggest.

A.9.4 Anticipatory equilibrium

Under the standard Nash concept, no equilibrium exists above the threshold: the shrouded contract is vulnerable to the cream-skimming deviation of Lemma 3. I therefore use the following anticipatory concept, under which an equilibrium does exist.[75]

Definition 1. *(Wilson, 1977)* A menu is an anticipatory equilibrium if no firm can add a contract that earns positive profit after those existing contracts that become unprofitable in its presence are withdrawn, and consumers re-optimize over what remains.

Proposition 24. (Complexity aversion makes shrouding less likely) *Under Definition 1, and with the pivotal value for the fraction of naive consumers α^* in* (97)*:*

1. *If $\alpha \geq \alpha^*$, the shrouded contract $S = (-\alpha\bar{p}, \bar{p})$ is an equilibrium.*

2. *If $\alpha \leq \alpha^*$, the transparent contract $T = (-e, e)$ is an equilibrium.*

The shrouding region shrinks as complexity aversion rises: $\partial\alpha^*/\partial\bar{\xi} = 1 - e/\bar{p} > 0$. With no complexity aversion ($\bar{\xi} = 0$) the condition is simply $\alpha \geq e/\bar{p}$, as in GL. Hence complexity aversion helps make it less likely that the market is shrouded. The reason is that the naives "feel bad" about the product (via complexity aversion), as they sense that there is a snag, and defect to the simpler scheme, with a smaller shrouded attribute. Hence complexity aversion is a partial force for consumer protection from firms' complexity.

A.9.5 Welfare

I next calculate welfare. As in the rest of the paper, I count CA as a welfare loss.

Proposition 25. *Aggregate true surplus, relative to the frictionless benchmark, is*

$$W_S = -e\,(1-\alpha) - \alpha\bar{\xi}\bar{p}$$

under the shrouded equilibrium $S = (-\alpha\bar{p}, \bar{p})$, and

$$W_T = -\alpha\bar{\xi}e$$

under the transparent equilibrium $T = (-e, e)$. The transparent equilibrium dominates at every parameter value,

$$W_T - W_S = e\,(1-\alpha) + \alpha\bar{\xi}\,(\bar{p} - e) > 0,$$

[75] The anticipatory concept is the natural one here because the skimmer's profit is parasitic: it depends on the continued presence of the contract it is skimming, which is precisely the case Definition 1 was designed for.

both because no effort is wasted avoiding the add-on and because the fee that naives dread is smaller. Moreover, under the shrouded equilibrium naives lose $\bar{p}(1-\alpha)$ in material terms and rationals gain $\alpha\bar{p}-e$, whereas under the transparent equilibrium both types obtain zero material surplus: complexity aversion removes the cross-subsidy from naive to sophisticated consumers, and not merely the inefficiency.

A.9.6 Proofs for shrouded attributes with complexity averse naive consumers

Proof of Proposition 24. *Part 1.* Consider deviations against S, at which $U^r = \alpha\bar{p} - e$ and $U^n = \bar{p}\left(\alpha - \bar{\xi}\right)$. A pooling deviation $(p, \hat{p})$ must satisfy $p \leq -U^r - \min(\hat{p}, e)$ and $p \leq -U^n - \bar{\xi}\hat{p}$. If $\hat{p} \leq e$, its profit is $p + \hat{p} \leq -U^r = e - \alpha\bar{p} < 0$, since $\alpha > e/\bar{p}$. If $\hat{p} > e$, the first constraint gives $p \leq -\alpha\bar{p}$, and it binds because $-U^n - \bar{\xi}\hat{p} + \alpha\bar{p} = \bar{\xi}(\bar{p} - \hat{p}) \geq 0$, so the profit is $p + \alpha\hat{p} \leq \alpha(\hat{p} - \bar{p}) \leq 0$. Rational-only deviations require $p < -U^r - \min(\hat{p}, e)$ and so lose money. There remains the cream-skim D at $\hat{p} = \bar{p} - \varepsilon$: by Lemma 3 it earns $\approx \bar{p}(1-\alpha) > 0$ while S remains. But once D takes the naives, S retains only rationals and earns $-\alpha\bar{p} < 0$ per customer, so S is withdrawn. The rationals then move to D, whose profit becomes

$$p_D + \alpha\hat{p}_D = \left(-\alpha\bar{p} + \bar{\xi}\varepsilon\right) + \alpha(\bar{p} - \varepsilon) = \varepsilon\left(\bar{\xi} - \alpha\right) < 0,$$

since $\alpha \geq \alpha^* > \bar{\xi}$. So D is not profitable in the anticipatory sense.

Part 2. At T we have $U^r = 0$ and $U^n = e\left(1 - \bar{\xi}\right)$. Cream-skims are blocked by Lemma 4, and rational-only deviations lose money as above. A pooling deviation with $\hat{p} \leq e$ earns at most $\min\left(0, \left(1 - \bar{\xi}\right)(\hat{p} - e)\right) \leq 0$. One with $\hat{p} > e$ faces $p \leq -e - \bar{\xi}(\hat{p} - e)$, and so earns at most $-e(1-\alpha) + (\hat{p} - e)\left(\alpha - \bar{\xi}\right)$, maximized at $\hat{p} = \bar{p}$, which is non-positive precisely when $\alpha \leq \alpha^*$. Since Definition 1 admits a subset of these deviations, T is an anticipatory equilibrium.

Proof of Proposition 25. Under S the add-on price is $\hat{p} = \bar{p} > e$, so rational consumers avoid it at cost e each while naives pay it. With $p = -\alpha\bar{p}$, a rational's material surplus is $\alpha\bar{p} - e$ and a naive's is $\alpha\bar{p} - \bar{p} = -\bar{p}(1-\alpha)$; the transfers cancel in the aggregate, leaving $(1-\alpha)(\alpha\bar{p} - e) + \alpha(\alpha\bar{p} - \bar{p}) = -e(1-\alpha)$, the avoidance effort. Each naive additionally bears dread $\bar{\xi}\bar{p}$, giving W_S.

Under T the add-on price is $\hat{p} = e$, so no consumer avoids it and both types pay $p + \hat{p} = 0$ in total: material surplus is zero for everyone. Each naive bears dread $\bar{\xi}e$, giving W_T. Subtracting gives $W_T - W_S$, positive since $\bar{p} > e$.

A.10 Fixed price vs Cost plus contract: why the slope in a contract is often exactly 0 or 1

To build a house, most contracts are fixed price: the buyer pays a pre-determined price $\bar{p}$ to the seller; but otherwise, they tend to be "cost plus" (C+).[76] So the slope of the payment is either $b = 0$ or $b = 1$, rather than some intermediate solution. How to explain that? A key model (Bajari and Tadelis, 2001) models whether people will choose Fixed or Cost+, but they need to posit that those are the only two possibilities; they acknowledge that, logically, intermediate values of $b \in (0, 1)$ would be optimal, and they exclude them by fiat (adding some intuitive reasons why that might be possible).

Here I show how complexity aversion can be a natural force that can be added to their analysis, and make natural the fact that (often) people will use only $b = 0$ or $b = 1$. To show that, I take the "chemically pure" case where complexity-sharing and risk-sharing are the only forces, without incentives or incomplete contracts.

Suppose that the cost is

$$c = \bar{c} + x$$

where x is random with mean 0. The buyer pays

$$p = \bar{p} + bx$$

to the seller, for some $\bar{p}$ and the slope b to be determined. The utilities of the buyer and seller are:

$$v^B = u^B(-p) \qquad v^S = u^S(p - c)$$

Assume linear-quadratic utility, absolute risk aversion Γ_B, Γ_S. The full utilities are, removing the constants:

$$V_B = -\bar{p} - \frac{1}{2}\Gamma_B b^2 \sigma_x^2 - \xi_B |b| \sigma_x, \qquad V_S = \bar{p} - \bar{c} - \frac{1}{2}\Gamma_S (1-b)^2 \sigma_x^2 - \xi_S |1-b| \sigma_x$$

where $\xi_B = \xi\kappa(1 - m)$ is the effective complexity aversion of the buyer, with κ the unit converter (and symmetrically for the seller). So, the optimal contract will maximize (Nash bargaining would yield something very similar):

$$\max_b V, \qquad V = V_B + V_S$$

First, without complexity aversion (or first-order risk aversion), we get an interior so-

[76] The "plus" refers to e.g. a fixed cost, or some extra incentive e.g. on-time performance, or an extra 10% "administrative" cost, which I fold into the total cost.

lution, $b = \frac{\Gamma_S}{\Gamma_B+\Gamma_S}$, so it is typically strictly interior — contradicting the empirical facts — except for a risk-neutral firm, which does lead to $b = 0$.

Next, with complexity aversion, we have the following result:[77]

Proposition 26. (Optimality of Fixed vs Cost+ contracts) *We have $b = 0$ iff $\xi_B - \xi_S \geq \Gamma_S \sigma_x$. We have $b = 1$ iff $\xi_B - \xi_S \leq -\Gamma_B \sigma_x$. In between, if $\xi_B - \xi_S \in (-\Gamma_B \sigma_x, \Gamma_S \sigma_x)$ we have $b = \frac{\Gamma_S - \frac{\xi_B - \xi_S}{\sigma_x}}{\Gamma_B+\Gamma_S} \in (0, 1)$.*

When risk aversions become negligible ($\Gamma_B, \Gamma_S \to 0$), only $b = 0$ and $b = 1$ are possible. More generally, in an open zone of parameters, we have $b = 0$ or $b = 1$. In most cases with a "small" buyer and a large firm, $\xi_B \gg \xi_S$, so $b = 0$ and we have a fixed contract. In contrast, when the "government" is large, perhaps $\xi_B \ll \xi_S$, and $b = 1$.

So, we find that complexity aversion, or its elder cousin first-order risk aversion, explains easily why contracts will have a slope of 0 or 1. This means that complexity aversion is an ingredient that can be added to previous analysis (e.g. Bajari and Tadelis, 2001), which helps guide the analysis towards simpler schemes.

We have "complexity sharing"—rather than "risk sharing". But very often, in "complexity sharing", just one party bears all the complexity.

B Complements to the Ramsey problem

This section contains complements to the Ramsey problem of Section 5. I will use the following transform of ξ:

$$\bar{\xi} \coloneqq \frac{\xi}{\lambda} \frac{(1-m)}{m}, \qquad m > 0. \tag{98}$$

B.1 More on simple taxes

I ask when there is a fully uniform tax system with a finite number of goods.

Proposition 27. (Optimal tax simplicity) *Suppose exogenous attention. Define the average elasticity $\bar{\psi} = \frac{1}{G}\sum_g \psi_g$. Suppose that for all goods, the average attention to the tax is imperfect, $m_g < 1$. Then, if complexity aversion $\xi > 0$ is large enough, the optimum entails a simple tax system: the tax rate is the same for all goods, $\tau_g = \bar{\tau} = \frac{\lambda}{\bar{\psi}}$. More specifically, a necessary and sufficient condition for this simple tax system is that ξ is large enough to satisfy:*

$$\max_g \left[\lambda m_g \frac{\psi_g}{\bar{\psi}} - \xi s_{c_g}(1 - m_g)\right] \leq \min_g \left[\lambda m_g \frac{\psi_g}{\bar{\psi}} + \xi s_{c_g}(1 - m_g)\right] \tag{99}$$

[77] The proof involves simply taking the derivative. E.g., taking the case $\sigma_x = 1$, $V_{b|b=0} = -\xi_B \operatorname{sign}(b) + \Gamma_S + \xi_S$, so $b = 0$ is optimal iff $\xi_B \geq \Gamma_S + \xi_S$.

When m_g and s_{c_g} are the same across goods, this necessary and sufficient condition becomes:

$$\lambda \frac{\max_g \psi_g - \min_g \psi_g}{\bar{\psi}} \leq 2\xi s_c \frac{1-m}{m} \tag{100}$$

In particular, in the case of uniformly distributed elasticities, a uniform tax system is optimal if $s_c\bar{\xi} \geq v_\psi$, where $\bar{\xi}$ is defined in (98) and v_ψ is the half-width of the range of elasticities. This is $\bar{\xi} \geq v_\psi$ under the normalization $s_c = 1$ of the main text, which this section relaxes.

Condition (100) means that there is a simple tax system if complexity aversion ξ is large enough compared to the relative dispersion of elasticities ($\frac{\max_g \psi_g - \min_g \psi_g}{\bar{\psi}}$), and if attention m is low enough. Condition (99) has the same message, in a more general way.[78]

Proof of Proposition 27 First, observe that if at the optimum all the $\hat{\tau}_g = 0$, then the optimum $\bar{\tau} = \frac{\lambda}{\bar{\psi}}$.

Next, I examine the optimum $\hat{\tau}_g$. I maximize W, subject to the constraint that the average deviation $\hat{\tau}_g$ is 0, so the Lagrangian is $L = W + \mu \sum_g \hat{\tau}_g$. The FOC for $\hat{\tau}_g$ is:

$$0 = L_{\hat{\tau}_g} = -\psi_g m_g \left(\bar{\tau} + \hat{\tau}_g\right) - \xi s_{c_g} \left(1 - m_g\right) \text{sign}\left(\hat{\tau}_g\right) + \lambda + \mu$$

So $\hat{\tau}_g = 0$ iff:

$$v_g^- \coloneqq \psi_g m_g \bar{\tau} - \xi s_{c_g} \left(1 - m_g\right) \leq \lambda + \mu \leq \psi_g m_g \bar{\tau} + \xi s_{c_g} \left(1 - m_g\right) =: v_g^+$$

So all $\hat{\tau}_g = 0$ iff one can find a μ for which this inequality holds for all g, i.e. iff:

$$\max_g v_g^- \leq \min_g v_g^+$$

This is the expression in (99), as $\bar{\tau} = \frac{\lambda}{\bar{\psi}}$.

B.2 Optimal tax system with costly endogenous attention

I complement the picture with a study of the optimal tax system for intermediate values of the complexity aversion parameter ξ, as opposed to large values of ξ in the main text (which I think should remain the leading case).

Assumptions and notation. As in the proof of Proposition 10, write $t_{G_k} \coloneqq \frac{\tau_{G_k} - \bar{\tau}}{\bar{\tau}}$ for the tax deviation of bracket G_k, and s_{G_k} for its mass of goods. Two objects recur in the results:

[78] There is a value $\xi_* \geq 0$ such that (99) holds if and only if $\xi \geq \xi_*$. Indeed, observe that the difference between the right-hand side of (99) and its left-hand side increases in ξ, and goes to $+\infty$ as $\xi \to \infty$.

the shrinkage parameter (98), evaluated at an attention level m, and the compression factor it induces,

$$\bar{\xi}(m) \coloneqq \frac{\xi(1-m)}{\lambda m}, \qquad M(m) \coloneqq \max\left(1 - \frac{\bar{\xi}(m)}{v_\psi}, 0\right). \tag{101}$$

So the factor multiplying $\left(\frac{\lambda \bar{\tau} v_\psi^2}{3\kappa}\right)^{1/3}$ in (43) is $M(\bar{m})$.

The consumer chooses a separate attention m_{G_k} for each bracket deviation, with a free floor and a hard ceiling:

$$m_{G_k} \in [\underline{m}, \bar{m}], \qquad C^{\text{Cognition}} = \kappa \sum_k (m_{G_k} - \underline{m}), \qquad 0 \leq \underline{m} < \bar{m} < 1. \tag{102}$$

The floor $\underline{m}$ is a free, default minimum of attention (the specification (89), used for the indexation problem): every line item of the tax code is at least dimly perceived. The ceiling $\bar{m} < 1$ caps precision: even unbounded effort does not fully decode a tax system. Raising attention on an item from the floor to m_{G_k} costs $\kappa(m_{G_k} - \underline{m})$.

The consumer's surplus is separable across brackets and linear in each m_{G_k}: attention to bracket k avoids the misoptimization loss and relieves the dread on the bracket's goods, so its marginal value is, to the leading order,[79]

$$\theta_{G_k} \coloneqq s_{G_k} \bar{\tau} \left(\frac{\lambda}{2} t_{G_k}^2 + \xi |t_{G_k}|\right). \tag{103}$$

Against the linear cost κ, attention is bang-bang between the corners:

$$m_{G_k} = \bar{m} \text{ if } \theta_{G_k} \geq \kappa, \qquad m_{G_k} = \underline{m} \text{ otherwise.} \tag{104}$$

A bracket is paid attention to only if it is large enough (mass s_{G_k}) and its tax distinctive enough ($|t_{G_k}|$). A bracket with $t_{G_k} = 0$ has $\theta_{G_k} = 0$: there is nothing to learn, so it stays at the floor, at no cost. The planner sets the brackets and each deviation, anticipating the rule (104)—so, when optimizing each tax, he takes the consumer's cognition into account.

The no-stealth condition. The new possibility, relative to exogenous attention, is that the planner may want to tax *through* inattention: place a surcharge on a bracket left at the floor, which is only dimly seen (at rate $\underline{m}$), hence barely distorts demand. The temptation is real: relative to raising the transparent average tax, a floor surcharge saves the distortion margin $(\bar{m} - \underline{m})\lambda$ per unit; but it inflicts dread $\xi(1 - \underline{m})$ per unit, *because first-order com-*

[79] Per good, attention avoids the misoptimization loss $\frac{\psi_g}{2}\hat{\tau}_G^2(1-m)$ (the analogue of (4)) and relieves the dread $\xi|\hat{\tau}_G|(1-m)$; both are linear in m, so the marginal value of attention is independent of its level. Averaging over the bracket, and using $\hat{\tau}_G = \bar{\tau} t_G$ and $\bar{\psi}\bar{\tau}^2 = \lambda\bar{\tau}$, gives (103).

plexity aversion is felt however dimly the tax is seen. Hence the condition for "no stealth taxation":

$$\xi \geq \frac{(\bar{m} - \underline{m})\,\lambda}{1 - \underline{m}}. \tag{105}$$

The floor itself protects: the right-hand side decreases in $\underline{m}$, and vanishes as $\underline{m} \to \bar{m}$—there is nothing to hide from a consumer whose free attention already matches her maximal one.

Proposition 28. (Optimal tax system with costly endogenous attention) *Suppose* (102)*, and the no-stealth condition* (105) *strengthened by an* $O(v_\psi)$ *margin:*

$$\xi \geq \frac{\lambda\,(\bar{m} - \underline{m} + \underline{m} v_\psi)}{1 - \underline{m}}. \tag{106}$$

Then, to the leading order, the optimal tax system is as follows, with the shrinkage $\bar{\xi}(\cdot)$ *and the compression factor* $M(\cdot)$ *as defined in* (101)*.*

(i) (Attention) *The central band of goods is taxed at exactly* $\bar{\tau}$*, and rests at the free floor of attention; every outer bracket is studied at the ceiling,* $m_{G_k} = \bar{m}$*.*

(ii) (Taxes) *The outer brackets carry the soft-thresholded taxes of Proposition 9 evaluated at attention* $\bar{m}$*:* $t_{G_k} = -T\left(\hat{\psi}_{G_k}, \bar{\xi}(\bar{m})\right)$*; and the bracket geometry (the central band* b^* *of* (124)*, read with* $\tilde{\xi} = \bar{\xi}(\bar{m})/v_\psi$*, and equal-sized outer brackets) is as with exogenous attention* $\bar{m}$*.*

Proof. Attention enters welfare bracket by bracket. For a bracket at the floor, the Lagrangian terms are those of (117) evaluated at $m = \underline{m}$: per unit mass,

$$-\frac{1}{2}\left(1 + \hat{\psi}_G\right)\left(1 + 2\underline{m}t + \underline{m}t^2\right) - \underline{m}\bar{\xi}(\underline{m})\,|t| + 1 + \mu t.$$

The subgradient condition for $t = 0$ to be optimal is $\left|\mu - \underline{m}\left(1 + \hat{\psi}_G\right)\right| \leq \underline{m}\bar{\xi}(\underline{m}) = \frac{\xi(1-\underline{m})}{\lambda}$. The multiplier is pinned by the studied brackets, whose first-order conditions are those of the proof of Proposition 10 with $m = \bar{m}$; averaging them as in (119) gives $\mu = \bar{m} + O\left(v_\psi^2\right)$. Hence floor brackets are optimally null if and only if $\bar{m} - \underline{m} + \underline{m}\left|\hat{\psi}_G\right| \leq \frac{\xi(1-\underline{m})}{\lambda}$ for the relevant goods, which the no-stealth condition with margin guarantees for all $\left|\hat{\psi}_G\right| \leq v_\psi$. Null floor brackets pool into the central band. The studied brackets all share $m_{G_k} = \bar{m}$, so their taxes and geometry solve the same problem as with exogenous attention $\bar{m}$—the cognition cost $\kappa(\bar{m} - \underline{m})$ per studied bracket depends only on their number—giving (ii). The consumer's rule (104) then implements (i): the central band has $\theta = 0$, so it stays at the floor for free, and the outer brackets clear the attention threshold, $\theta_{G_k} \geq \kappa$ (Proposition 29 quantifies the slack). □

The number of brackets. Since the central band rests at the free floor, cognition costs are paid only on the $n-1$ studied brackets, at $\kappa(\bar{m}-\underline{m})$ each. Using the welfare gain of Proposition 10 evaluated at $m=\bar{m}$, and writing $\Omega \coloneqq \frac{\lambda\bar{\tau}v_\psi^2 M(\bar{m})^3}{6}$, the planner solves

$$\max_n \; W(n) - W(1) = \bar{m}\,\Omega\left(1-\frac{1}{n^2}\right) - \kappa(\bar{m}-\underline{m})(n-1). \tag{107}$$

A remark on why this maximization is the naive one. In general, endogenous attention drives a wedge into the planner's calculus: the consumer picks attention to maximize her own surplus, ignoring that it moves demand, hence tax revenue—so that $\frac{dW}{dn}$ acquires an uninternalized *fiscal externality* term, $R_m\frac{dm}{dn}$, which complexity aversion makes non-zero. Here that term vanishes, because attention sits at a corner: every studied bracket is at the ceiling $\bar{m}$ and every other at the floor $\underline{m}$, so a marginal change in n moves no attention level, $\frac{dm_{G_k}}{dn}=0$. The wedge exists in potential—$R_m>0$ still—but is inoperative, and (107) may be differentiated holding $\bar{m}$ fixed. This requires only that the attention constraints be slack, so that a small change in n does not flip a bracket's attention decision; the next proposition verifies that they are, and identifies the exception.

Proposition 29. (Optimal number of brackets) *With $\bar{\xi}(\cdot)$ and $M(\cdot)$ as in* (101)*, the optimal (continuous) number of brackets is*

$$n^* = \max\left(M(\bar{m})\left(\frac{\lambda\bar{\tau}v_\psi^2}{3\kappa}\cdot\frac{\bar{m}}{\bar{m}-\underline{m}}\right)^{1/3}, 1\right), \tag{108}$$

and the welfare gain from bracketing is

$$W(n^*) - W(1) = \bar{m}\,\Omega\,\frac{(n^*-1)^2(n^*+2)}{n^{*3}} \geq 0: \tag{109}$$

bracketing (weakly) always pays, and n^* falls to 1 continuously *as the cognition cost κ rises. Moreover, each studied bracket clears its attention threshold with slack $\frac{\theta_{G_k}}{\kappa} \geq 6\,\frac{\bar{m}-\underline{m}}{\bar{m}}\left(1+\frac{\xi n^*}{\lambda v_\psi M(\bar{m})}\right)$, which exceeds* 1 *whenever $\underline{m} < \frac{5}{6}\bar{m}$; for a floor closer to the ceiling, the attention threshold, rather than the upgrade cost, caps the number of brackets.*

Proof. Attention is at the corner $m_{G_k}=\bar{m}$ on every studied bracket, independently of n, so (107) is differentiated at fixed $\bar{m}$, with no attention-response term. Its first-order condition is $\frac{2\bar{m}\Omega}{n^3} = \kappa(\bar{m}-\underline{m})$, giving (108) when the solution exceeds 1, and $n^*=1$ otherwise, continuously. Substituting back, $1-\frac{1}{n^2}-\frac{2(n-1)}{n^3} = \frac{n^3-3n+2}{n^3} = \frac{(n-1)^2(n+2)}{n^3}$, giving (109). For the slack: the innermost studied bracket has mass $\frac{M}{n}$ and deviation $|t_1| = \frac{2v_\psi M}{n}$ (from the geometry of the proof of Proposition 10: width $\delta_u = \frac{2M}{n}$, midpoint at distance δ_u from $\tilde{\xi}$),

so $\theta_1 = \frac{2\lambda\bar{\tau}v_\psi^2 M^3}{n^3}\left(1+\frac{z\xi n}{M}\right)$ with $z = \frac{1}{\lambda v_\psi}$, while $\kappa = \frac{\lambda\bar{\tau}v_\psi^2 M^3}{3n^{*3}} \cdot \frac{\bar{m}}{\bar{m}-\underline{m}}$ at an interior optimum. □

In sum, the optimal tax system with costly endogenous attention reads: post the average tax $\bar{\tau}$; keep a central band of goods exactly at $\bar{\tau}$, resting at the consumer's free attention; partition the tail goods into $n^* - 1$ equal brackets, each large and distinctive enough to earn the consumer's scrutiny, and tax each at its shrunk Ramsey level. Complexity aversion shapes the system thrice: it compresses the taxes within brackets, it limits the number of brackets, and—when strong enough—it forecloses the taxation-through-inattention that would otherwise be the planner's temptation. Section B.3 studies that last regime, when complexity aversion is too weak to foreclose it.

B.3 Ramsey problem: the regime with low complexity aversion, and "stealth taxes"

What if complexity aversion is weak—below the no-stealth condition (105) that Section B.2 assumed? Throughout this subsection the cognition cost is (102), the no-stealth condition fails, and I adopt the tie-breaking convention that an indifferent consumer does not upgrade her attention.

Proposition 30. (Taxes in the stealth regime) *Suppose* (102) *and that the no-stealth condition fails outright:*

$$\xi < \frac{(\bar{m}-\underline{m})\,\lambda}{1-\underline{m}}. \tag{110}$$

Then, to the leading order:

(i) (The stealth surcharge) *An unattended bracket of mass s carries the surcharge $\bar{\tau}\,t^{stealth}(s)$, with*

$$t^{stealth}(s) = \min\left\{t^{int}, t(s)\right\}, \qquad t^{int} := \frac{(\bar{m}-\underline{m})\,\lambda - \xi\,(1-\underline{m})}{\lambda\underline{m}}, \tag{111}$$

where the attention bound $t(s)$ *is the positive root of $s\bar{\tau}\left(\frac{\lambda}{2}t^2+\xi t\right) = \kappa$:*

$$t(s) = \frac{\xi}{\lambda}\left(\sqrt{1+\frac{2\lambda\kappa}{s\bar{\tau}\xi^2}}-1\right), \tag{112}$$

valid for $\xi > 0$; at $\xi = 0$ the positive-root definition still gives the continuous extension $t(s) = \sqrt{2\kappa/(\lambda s\bar{\tau})}$. This is the largest surcharge that leaves the consumer's stake at her attention threshold. For small items the interior cap t^{int} binds: fragmenting a stealth set strictly raises revenue only down to the scale $s^\dagger$ at which $t\left(s^\dagger\right) = t^{int}$, and is neutral beyond—the planner's problem attains its maximum, even with a continuum of goods.

(ii) (The optimal system) *Goods with* $\left|\hat{\psi}_g\right| \leq \rho^*$ *are floor items carrying the common surcharge* t^{int} *(in items of size at most* $s^\dagger$*); goods with* $\left|\hat{\psi}_g\right| > \rho^*$ *sit in studied brackets, taxed* $t_{G_k} = c - T\left(\hat{\psi}_{G_k}, \bar{\xi}\left(\bar{m}\right)\right)$ *with the uniform shading* $c = -\frac{s_S}{1-s_S} t^{int} < 0$ *(where* s_S *is the stealth mass): the advertised structure is shaded down, and the revenue is recouped through the dimly seen surcharges—as with loss leaders and shrouded add-ons (Gabaix and Laibson, 2006). The cutoff equalizes the two uses of the marginal good:*

$$\rho^* = \bar{\xi}\left(\bar{m}\right) + \sqrt{\frac{\underline{m}}{\bar{m}}}\, t^{int}. \tag{113}$$

(iii) (Caps) *The aggregate take is capped by the shading feedback: with* $\mu = \bar{m}\left(1+c\right)$*, surcharges pay only while* $\mu > \underline{m}\left(1 + \bar{\xi}\left(\underline{m}\right)\right)$*, i.e. while* $|c| < 1 - \frac{\underline{m}+\xi(1-\underline{m})/\lambda}{\bar{m}}$*.*

Proof. (i) A floor bracket's objective, per unit mass, is $-\frac{1}{2}\left(1+\hat{\psi}_G\right)\left(1 + 2\underline{m}t + \underline{m}t^2\right) - \underline{m}\bar{\xi}\left(\underline{m}\right)|t| + 1 + \mu t$: strictly concave in t, as the demand response operates at rate $\underline{m}$. Its unconstrained maximum, from the first-order condition $-\left(1+\hat{\psi}_G\right)\underline{m}\left(1+t\right) - \underline{m}\bar{\xi}\left(\underline{m}\right) + \mu = 0$ with $\mu = \bar{m}\left(1+c\right)$, is t^{int} as announced, to the leading order (the formula is otherwise exact in t: with linear demand, welfare is exactly quadratic in the taxes). The item must also stay below the consumer's attention threshold, $\theta \leq \kappa$, whence the bound $t\left(s\right)$, binding for $s > s^\dagger$ and slack below; implicit differentiation gives $t'\left(s\right) < 0$, so splitting items below the scale $s^\dagger$ leaves all surcharges at t^{int}, and revenue unchanged. The value of a floor item at the interior surcharge is $\frac{\underline{m}}{2} t^{\text{int}2}$ per unit mass, in the units of (117).

(ii) Per unit mass, a good contributes $\frac{\bar{m}}{2}\max\left(\left|\hat{\psi}_g\right| - \bar{\xi}\left(\bar{m}\right), 0\right)^2$ in a studied bracket, and $\frac{\underline{m}}{2} t^{\text{int}2}$ as a floor item; the latter is flat in $\left|\hat{\psi}_g\right|$ and the former increasing, whence the cutoff (113). The shading: deviations average to zero, $s_S t^{\text{int}} + \left(1 - s_S\right)c = 0$ (the thresholded terms average to zero by the symmetry of the studied set); the studied brackets' first-order conditions then hold with multiplier $\mu = \bar{m}\left(1+c\right)$.

(iii) Surcharges pay while the floor first-order condition gives $t > 0$: $\mu > \underline{m}\left(1 + \bar{\xi}\left(\underline{m}\right)\right)$. □

The economics deserve note. First, rearranged, the floor item's tax satisfies $1 + t = \frac{\mu - \underline{m}\bar{\xi}(\underline{m})}{\underline{m}\left(1+\hat{\psi}_G\right)}$: it is the *same* tax rule as any bracket, evaluated at attention $\underline{m}$. To the planner, a dimly seen good is simply a good with a low *perceived* elasticity $\underline{m}\psi_g$, and the "stealth tax" is ordinary Ramsey pricing on the attention-deflated elasticity, net of dread; it is finite precisely because the floor keeps the perceived elasticity positive. Second, the regime predicts a recognizable pathology: a shaded advertised structure, plus a proliferation of small opaque line items—"junk fees"—each individually not worth the consumer's scrutiny. Without

complexity aversion ($\xi = 0$), this exploitation of inattention would be the dominant use of tax complexity; it is the dread term—paid however dimly the tax is seen—that makes hidden taxes unprofitable when (105) holds. As with shrouded attributes (Section A.9), complexity aversion acts as a form of taxpayer protection.

One might worry that stealth taxes contradict the requirement that deviations average to zero: hidden surcharges seem to raise the average tax. They do not: $\bar{\tau}$ is *defined* as the average tax, and is assumed perfectly seen (a salience assumption), so there is no stealth on the aggregate margin. A stealth tax is a *cross-sectional* reshuffle at a fixed, visible average: surcharges on floor items, financed by shading the studied brackets' taxes. The consumer is not fooled about the average burden, which she sees exactly; what she partly misses is the *incidence*—which goods carry the load. A direct perturbation displays the logic: raising the actual tax of a floor bracket of mass s_0 by $d\tau$ changes welfare by $s_0 (1 - s_0) [(\bar{m} - \underline{m}) \lambda - \xi (1 - \underline{m})] d\tau$: only the deviation component $(1 - s_0) d\tau$ is (partially) hidden, and at $s_0 = 1$, a "stealth surcharge" is just a visible increase in the average tax, with no gain.[80]

In between: a fringe of stealth. Conditions (106) and (110) leave a window of width $\frac{2\lambda \underline{m} v_\psi}{1-\underline{m}}$ in ξ unaccounted for. Nothing dramatic happens there: the two regimes are the ends of a continuous interpolation, which a single statistic organizes.

Proposition 31. (Partial stealth: the intermediate regime) *Suppose* (102). *Keeping the dependence on the good's elasticity gap that Propositions 28 and 30 treat as higher order, a floor good at gap $\hat{\psi}$ carries, to the leading order, the surcharge*

$$t\left(\hat{\psi}\right) = \max\left(t^{int} - \hat{\psi}, 0\right), \qquad t^{int} = \frac{(\bar{m} - \underline{m}) \lambda - \xi (1 - \underline{m})}{\lambda \underline{m}}, \tag{114}$$

another soft threshold: as $\hat{\psi}$ ranges over $[-v_\psi, v_\psi]$, the stealth set is $\left\{\hat{\psi} < t^{int}\right\}$. Everything turns on the position of t^{int} relative to $\pm v_\psi$, i.e. on ξ:

(i) $t^{int} \leq -v_\psi$, which is exactly (106)*: no good is surcharged—Proposition 28.*

(ii) $t^{int} \geq v_\psi$: every floor good is surcharged, and as t^{int} grows beyond v_ψ the surcharge becomes common across them—Proposition 30, whose condition (110) *($t^{int} > 0$) marks where stealth first appears, and whose uniform-surcharge description is exact for $t^{int} \gg v_\psi$.*

(iii) $-v_\psi < t^{int} < v_\psi$, the intermediate window: partial stealth. *Only the goods more inelastic than the cutoff, $\hat{\psi} < t^{int}$, carry a surcharge—sensibly, since these are the ones the*

[80] If part of the average itself could be shrouded, the planner would command a form of hidden lump-sum taxation—a much stronger force; the salience of the average is what disciplines stealth to this cross-sectional form.

Ramsey rule wants taxed most—and the surcharges are $O(v_\psi)$, of the same order as ordinary tax deviations.

Proof. For a floor good at gap $\hat{\psi}$, the first-order condition in the proof of Proposition 30 reads $(1+t)\left(1+\hat{\psi}\right) = \frac{\mu - \underline{m}\bar{\xi}(\underline{m})}{\underline{m}} = 1 + t^{\text{int}}$, with $\mu = \bar{m}$ to the leading order. Hence $t = \left(1+t^{\text{int}}\right)\left(1-\hat{\psi}\right) - 1 + O\left(v_\psi^2\right) = t^{\text{int}} - \hat{\psi} + O\left(v_\psi^2\right)$ in the window where $t^{\text{int}} = O(v_\psi)$; and the subgradient condition at $t = 0$ (proof of Proposition 28) truncates it at zero, giving (114). Case (i) restates that subgradient condition at the worst good $\hat{\psi} = -v_\psi$, which is (106). In case (ii), $t\left(\hat{\psi}\right) > 0$ for all $\hat{\psi} \leq v_\psi$, and $\frac{t\left(\hat{\psi}\right)}{t^{\text{int}}} = 1 + O\left(\frac{v_\psi}{t^{\text{int}}}\right)$ is common across goods when $t^{\text{int}} \gg v_\psi$. □

So the intermediate case is a mild one: a fringe of the least elastic central goods is taxed slightly above the average, with no "junk fee" proliferation—that pathology needs $t^{\text{int}} \gg v_\psi$, i.e. ξ well below the threshold. The stealth mass and the surcharges both shrink continuously to zero as ξ rises to (106). Two caveats. The cutoff is determined jointly with the shading: t^{int} depends on the multiplier $\mu = \bar{m}(1+c)$, and c on the stealth mass and surcharges, so (114) is a fixed point in this window (where $c = O(v_\psi)$ shifts t^{int} within the window, but does not move it out). Helpfully, the feedback is self-limiting: more stealth means more shading, hence a lower μ, hence a lower t^{int} and less stealth.

Orders of magnitude. In terms of the maintained small parameter $\bar{\xi}(\bar{m})$, the no-stealth condition (105) reads $\bar{\xi}(\bar{m}) \geq \frac{(1-\bar{m})(\bar{m}-\underline{m})}{\bar{m}(1-\underline{m})}$. It is thus compatible with the several-brackets regime $\bar{\xi}(\bar{m}) = O(v_\psi)$ if and only if $(1-\bar{m})(\bar{m}-\underline{m}) = O(v_\psi)$: a high ceiling ($1-\bar{m}$ small), or a high floor ($\bar{m}-\underline{m}$ small). Away from those, with weak complexity aversion, the stealth motive is first order—each unattended category is worth up to $\kappa\frac{(\bar{m}-\underline{m})\lambda - \xi(1-\underline{m})}{\xi}$, and the planner multiplies such items. In calibrations, however, the sensible corner is the protected one: with $\xi \simeq \frac{2}{3}$ (Section 7.1) and a modest value of public funds λ, condition (105) holds easily—and then $\bar{\xi}(\bar{m})$ is sizable, so the optimal system has one or very few brackets: strong complexity aversion delivers the simplicity and the taxpayer protection jointly.

B.4 In the Ramsey problem, cognition costs alone (without complexity aversion) do not yield full simplicity

I finally show how with zero complexity aversion, the optimal tax system is always complex: it has strictly more than one tax bracket. To prevent indeterminacies, I suppose that a small fraction $\epsilon \in (0,1)$ of the agents are rational: they see the taxes perfectly, at no cognition cost, and have no complexity aversion. The remaining $1-\epsilon$ are as above. Their presence gives an incentive for the planner to have complex tax systems.

Proposition 32 (Complexity aversion is necessary to obtain a simple tax system)**.** *Consider the tax problem above, with a fraction $\epsilon \in (0,1)$ of rational agents.*

(i) Suppose there is no complexity aversion ($\xi = 0$). Then the optimum has strictly more than one tax bracket. This holds for any cognition cost function $C^{Cog}(m)$.

(ii) Suppose complexity aversion ξ is high enough. Then the optimum is a uniform tax, with just one bracket.

Part (i) is the counterpart, with endogenous attention, of the fact that pure information costs do not generate full simplicity (Proposition 3). Without complexity aversion, the cost of a complex tax system is a cost that the information-theoretic consumer (but not complexity-averse) can decline to pay: she sets $m_{G_k} = 0$, pays nothing, and the rational agents still react to the brackets, so starting from full simplicity, a bit of extra complexity always gives a first-order gain in welfare. This is true for any cognition cost function $C^{\text{Cog}}(m)$, including the case where cognition is infinitely costly. In contrast, the complexity-averse consumer will always suffer (provided $m < 1$ and $\xi > 0$) from any extra complexity.

B.5 Proof of Proposition 10

I call $\bar{\psi}$ the average elasticity (so $\bar{\psi} = \frac{\psi_{\min}+\psi_{\max}}{2}$); $v_\psi = \frac{\psi_{\max}-\bar{\psi}}{\bar{\psi}}$ the normalized half-width of the range of elasticities; and $\hat{\psi}_g \coloneqq \frac{\psi_g-\bar{\psi}}{\bar{\psi}}$, so that $\frac{\hat{\psi}_g}{v_\psi}$ is uniformly distributed in $[-1,1]$. Likewise, for a bracket G with tax τ_G, I call $t_G \coloneqq \frac{\tau_G-\bar{\tau}}{\bar{\tau}}$ the (multiplicative) special tax for τ_G, so that $\tau_G = \bar{\tau}(1+t_G)$ and the additive deviation is $\hat{\tau}_G = \bar{\tau}t_G$. The proof works mostly with t_G.

Recall the thresholding function T defined in (38).

As said in the main text, I take the limit where v_ψ is "small", and of the same order of magnitude as $\bar{\xi}$ defined in (98). I look for an odd number of brackets, $n = 2N+1$ for an integer N. This is without loss of generality, as the central bracket could have a mass of 0.

The shape of the boundaries of the optimal tax brackets, before computing the optimal taxes. I start with a simple result on the shape of optimal tax brackets, which is useful to compute the optimal taxes numerically.

Lemma 5. *The tax brackets, expressed in terms of $\frac{\hat{\psi}_g}{v_\psi}$, are of the form: a central one in $[-b,b]$ for some $b \in (0,1]$; and otherwise, $2N$ equal-sized brackets partitioning $[-1,1]\setminus[-b,b]$ into $2N$ intervals of size $\delta_{\hat{\psi}} \coloneqq \frac{1-b}{N}$, so intervals $I_k = (b+(k-1)\delta_{\hat{\psi}}, b+k\delta_{\hat{\psi}}]$ for $k = 1,\dots,N$ and symmetrically in the negatives.*

Proof. This proof is mostly mathematical, and the reader is encouraged to skip it, and go to the derivation of the actual taxes. Write $u_g \coloneqq \hat{\psi}_g/v_\psi \in [-1,1]$ (uniformly distributed)

and $\tilde{\xi} \coloneqq \bar{\xi}/v_\psi$. A "simple" tax system is a partition of the goods into n brackets together with one common tax per bracket. The proof rests on the fact, established below, that the problem is one of *optimal piecewise-constant approximation*: the welfare contributed by a bracket depends on it only through a convex function of its mean elasticity gap.

Step 1 (a bracket's value depends only on its mass and mean gap, via a convex even function). Take an arbitrary bracket G (a measurable set of goods) of mass s_G and mean gap $\hat{\psi}_G \coloneqq \frac{1}{s_G}\int_G \hat{\psi}_g$, on which the planner sets a common tax t. Per unit mass, a good's contribution to the Lagrangian is the summand of (117) with $\hat{\psi}_g$ in place of $\hat{\psi}_{G_k}$; it is *exactly affine* in the good's characteristic $\hat{\psi}_g$, which enters only through $-\frac{1}{2}\hat{\psi}_g\,(1+2mt+mt^2)$. Averaging over G therefore replaces $\hat{\psi}_g$ by $\hat{\psi}_G$ exactly, and the bracket contributes (dropping t-independent constants, and the third-order term $\hat{\psi}_G t^2$)

$$s_G\Big[-\tfrac{m}{2}t^2+\Big(\mu-m-m\hat{\psi}_G\Big)\,t-m\bar{\xi}\,|t|\,\Big].$$

Optimizing over the common tax t gives the soft-thresholded value $t_G = T\left(-\hat{\psi}_G, \bar{\xi}\right)$ (using that the multiplier satisfies $\mu - m = O(v_\psi^2)$, of higher order, as derived below in (119)) and the maximized per-unit value

$$g\left(\hat{\psi}_G\right) \coloneqq \tfrac{1}{2}mt_G^2 = \tfrac{m}{2}\max\left(\left|\hat{\psi}_G\right|-\bar{\xi},0\right)^2, \tag{115}$$

so the bracket contributes $s_G\, g\left(\hat{\psi}_G\right)$. The function g is even and convex; it vanishes on $\left\{\left|\hat{\psi}\right| \le \bar{\xi}\right\}$ and is a shifted quadratic on $\left\{\left|\hat{\psi}\right| \ge \bar{\xi}\right\}$.

Step 2 (brackets are intervals). Total welfare is $\sum_G s_G g\left(\hat{\psi}_G\right)$. Take an optimal partition and move an infinitesimal mass of goods located at gap $\hat{\psi}$ into a bracket G. Adding mass ϵ at $\hat{\psi}$ to G changes $s_G g\left(\hat{\psi}_G\right)$ by $\epsilon\big[g\left(\hat{\psi}_G\right)+g'\left(\hat{\psi}_G\right)\left(\hat{\psi}-\hat{\psi}_G\right)\big]$, so the marginal value of assigning a good at $\hat{\psi}$ to G is the value at $\hat{\psi}$ of the tangent to g at $\hat{\psi}_G$,

$$L_G\left(\hat{\psi}\right) = g\left(\hat{\psi}_G\right)+g'\left(\hat{\psi}_G\right)\left(\hat{\psi}-\hat{\psi}_G\right).$$

At the optimum each good must lie in the bracket with the highest $L_G\left(\hat{\psi}\right)$ (otherwise a profitable move exists). Because g is convex, the L_G are its tangent lines, with slopes $g'\left(\hat{\psi}_G\right)$ increasing in $\hat{\psi}_G$; their upper envelope $\max_G L_G$ is convex and piecewise linear, and the set of $\hat{\psi}$ on which a given L_G is highest is an interval. Hence every bracket is an interval in $\hat{\psi}$, equivalently in u.

Step 3 (a symmetric central band). Since the density of u is symmetric and g is even,

the reflection about 0 of an optimal partition is optimal; so there is an optimal partition symmetric about 0. By the tangent-line argument of Step 2, a good at $\hat{\psi}$ with $\left|\hat{\psi}\right| < \bar{\xi}$ has $L_G\left(\hat{\psi}\right) \leq g\left(\hat{\psi}\right) = 0$ for every bracket G (convexity places the tangent below g), so it is weakly better placed in a zero-tax bracket; all such goods can be pooled into a single central bracket, which by Step 2 is an interval $[-b, b]$ with $b \geq \tilde{\xi}$. The remaining $2N$ brackets tile $[-1, -b] \cup [b, 1]$ and, by symmetry, come in mirror pairs.

Step 4 (the outer brackets are equal-sized). Consider the N outer brackets in $[b, 1]$, of lengths $\ell_1, \ldots, \ell_N$ with $\sum_k \ell_k = 1 - b$. There, by (115), $g(u) = \frac{m}{2} v_\psi^2 \left(u - \tilde{\xi}\right)^2$ is quadratic, so for an interval I with (mid-point) mean $\bar{u}_I$, under the uniform measure ν,

$$s_I \, g(\bar{u}_I) = \tfrac{m}{2} v_\psi^2 \Big[\int_I \left(u - \tilde{\xi}\right)^2 d\nu - \int_I (u - \bar{u}_I)^2 \, d\nu \Big].$$

Summed over the outer intervals, the first term equals $\int_b^1 \left(u - \tilde{\xi}\right)^2 d\nu$ and is independent of the partition, so maximizing welfare is equivalent to minimizing the total within-interval dispersion $\sum_k \int_{I_k} (u - \bar{u}_{I_k})^2 \, d\nu = \frac{1}{24} \sum_k \ell_k^3$. Since $t \mapsto t^3$ is convex, $\sum_k \ell_k^3$ subject to $\sum_k \ell_k = 1 - b$ is minimized when all ℓ_k are equal, i.e. $\ell_k = \delta_{\hat{\psi}} = \frac{1-b}{N}$. By symmetry the same holds on $[-1, -b]$, which yields the stated intervals I_k. □

Next, I determine the optimal tax for bracket G_k. I call $\hat{\psi}_{G_k}$ the average elasticity gap in bracket k, so

$$\frac{\hat{\psi}_{G_k}}{v_\psi} = b + \left(k - \frac{1}{2}\right) \delta, \tag{116}$$

where $\delta \coloneqq \delta_{\hat{\psi}} = \frac{1-b}{N}$ is the common width of the outer brackets from the Lemma above.

The average welfare coming from bracket k (with special tax t_{G_k}) is:

$$V_{G_k} \coloneqq \frac{-1}{2} \bar{\psi} \left(1 + \hat{\psi}_{G_k}\right) \mathbb{E}\left[\bar{\tau}^2 \left(1 + 2m t_{G_k} + m t_{G_k}^2\right)\right] - \xi \bar{\tau} \left|t_{G_k}\right| (1 - m_{G_k}) + \lambda \bar{\tau} \left(1 + t_{G_k}\right)$$

The planner's problem (given attentions, for now taken as exogenous) is (where s_{G_k} is the mass in each bracket $k \neq 0$)

$$\max_{t_{G_k}} \sum_k s_{G_k} V_{G_k} \text{ s.t. } \sum_k s_{G_k} t_{G_k} = 0$$

The last condition means that the average tax deviation is 0. I call $\mu - 1$ the multiplier for

that last condition. Note that $\bar{\tau} = \frac{\lambda}{\bar{\psi}}$, so the Lagrangian is

$$\frac{L}{\lambda\bar{\tau}} = \frac{1}{\lambda\bar{\tau}}\left(\sum_k s_{G_k} V_{G_k}\right) + (\mu - 1)\sum_k s_{G_k} t_{G_k}$$

So, with $\bar{\xi}$ defined in (98), the Lagrangian is

$$\frac{L}{\lambda\bar{\tau}} = \sum_k \left[\frac{-1}{2}\left(1 + \hat{\psi}_{G_k}\right)\left(1 + 2mt_{G_k} + mt_{G_k}^2\right) - m\bar{\xi}\left|t_{G_k}\right| + 1 + \mu t_{G_k}\right] s_{G_k} \tag{117}$$

Orders of magnitude. Before solving, I make the bookkeeping explicit. There are two small parameters, λ and v_ψ. Taxes are small because λ is: $\bar{\tau} = \frac{\lambda}{\bar{\psi}}\left(1 + O(v_\psi^2)\right) = O(\lambda)$,[81] and the Lagrangian (117) is written in units of $\lambda\bar{\tau} = O(\lambda^2)$, the scale of welfare. In this model the Lagrangian is moreover *exact* in λ: demand is linear, so welfare is exactly quadratic in the taxes—there is no $O(\lambda^3)$ truncation, unlike for a general non-quadratic demand. Within a bracket the deviations are small because v_ψ is: $\hat{\psi}_{G_k} = O(v_\psi)$, $t_{G_k} = O(v_\psi)$, and I take $\bar{\xi}$ of the same order, $\bar{\xi} = O(v_\psi)$—the interesting regime, since $\bar{\xi} \gg v_\psi$ forces a uniform tax and $\bar{\xi} \ll v_\psi$ makes complexity aversion negligible. I therefore solve the FOC to first order in v_ψ, carrying the remainder explicitly.

The FOC of (117) is exact:

$$-\left(1 + \hat{\psi}_{G_k}\right) m\left(1 + t_{G_k}\right) - m\bar{\xi}\,\mathrm{sign}\left(t_{G_k}\right) + \mu = 0.$$

Assuming an interior solution with $t_{G_k} \neq 0$,

$$1 + t_{G_k} = \frac{\mu - m\bar{\xi}\,\mathrm{sign}\left(t_{G_k}\right)}{m\left(1 + \hat{\psi}_{G_k}\right)}. \tag{118}$$

The multiplier. Averaging the exact relation (118) over the non-central brackets (mass-weighted), with $\langle t_{G_k}\rangle = 0$: the left side is 1; on the right, $\left\langle \frac{1}{1+\hat{\psi}_{G_k}} \right\rangle = 1 + O(v_\psi^2)$ and $\bar{\xi}\left\langle \frac{\mathrm{sign}\left(t_{G_k}\right)}{1+\hat{\psi}_{G_k}} \right\rangle = \bar{\xi}\left\langle \left|\hat{\psi}_{G_k}\right| \right\rangle + O(v_\psi^3) = O(v_\psi^2)$ (using $\langle \mathrm{sign}\left(t_{G_k}\right)\rangle = 0$ by the symmetry of the construction, and $\mathrm{sign}\left(t_{G_k}\right) = -\,\mathrm{sign}\left(\hat{\psi}_{G_k}\right)$). Hence

$$\mu = m + O(v_\psi^2). \tag{119}$$

Its precise second-order value is immaterial for what follows, as μ enters only through $\frac{\mu}{m} - 1 =$

[81] The correction is $O(v_\psi^2)$ because $\langle 1/\psi_g\rangle = \frac{1}{\bar{\psi}}\left\langle 1 - \hat{\psi}_g + \hat{\psi}_g^2 - \cdots \right\rangle = \frac{1}{\bar{\psi}}\left(1 + O(v_\psi^2)\right)$.

$O(v_\psi^2)$.

The tax. Expanding (118) in v_ψ, with $\frac{1}{1+\hat{\psi}_{G_k}} = 1 - \hat{\psi}_{G_k} + O(v_\psi^2)$, $\frac{\mu}{m} = 1 + O(v_\psi^2)$, and $\bar{\xi}\,\mathrm{sign}\,(t_{G_k})\,\hat{\psi}_{G_k} = O(v_\psi^2)$:

$$\begin{aligned} t_{G_k} &= \left(\frac{\mu}{m} - \bar{\xi}\,\mathrm{sign}\,(t_{G_k})\right)\left(1 - \hat{\psi}_{G_k}\right) - 1 + O(v_\psi^2) \\ &= -\hat{\psi}_{G_k} - \bar{\xi}\,\mathrm{sign}\,(t_{G_k}) + O(v_\psi^2). \end{aligned}$$

I record this. The optimal tax deviation for a bracket k is, to first order in v_ψ:

$$t_{G_k} = T\left(-\hat{\psi}_{G_k}, \bar{\xi}\right) + O(v_\psi^2) = -T\left(\hat{\psi}_{G_k}, \bar{\xi}\right) + O(v_\psi^2) \tag{120}$$

So, when t_{G_k} is not zero, it is

$$t_{G_k} = -\left(\hat{\psi}_{G_k} - \bar{\xi}\,\mathrm{sign}\left(\hat{\psi}_{G_k}\right)\right) + O(v_\psi^2). \tag{121}$$

I next express it in terms of tax rates, which gives Proposition 9.

There is a dampening by the term $\bar{\xi}$, which signifies that there is a Lasso-like shrinkage towards 0, to mitigate complexity aversion.

I next calculate welfare, given the optimized tax rate for each bracket. Let me next calculate the gain V_{G_k} from moving t_{G_k} from zero to its optimal value. Observe that in the problem $\max_x -\frac{1}{2}Ax^2 + Bx$, the maximum is at $x = \frac{B}{A}$, and the value is $\frac{B^2}{2A} = \frac{A}{2}x^2$. Applying this to (117), with $x = t_{G_k}$, $A = \left(1 + \hat{\psi}_{G_k}\right) m = m\,(1 + O(v_\psi))$, and $t_{G_k}^2 = O(v_\psi^2)$, this gives (the $O(v_\psi)$ relative error absorbing the factor $1 + \hat{\psi}_{G_k}$):

$$\frac{1}{\lambda\bar{\tau}} V_{G_k} = \frac{1}{2} m\,(t_{G_k})^2\, s_{G_k}\,(1 + O(v_\psi)) \tag{122}$$

This is, summing over the $2N$ outer brackets (the factor 2 from symmetry; each outer bracket has mass $s_{G_k} = \frac{\delta}{2}$, as the density of u on $[-1, 1]$ is $\frac{1}{2}$),

$$\frac{1}{\lambda\bar{\tau}} V \;=\; 2 \cdot \frac{1}{2} m \sum_{k=1}^{N} (t_{G_k})^2 \frac{\delta}{2} = \frac{m\delta}{2} \sum_{k=1}^{N} \left(v_\psi\left(b + \left(k - \frac{1}{2}\right)\delta\right) - \bar{\xi}\right)^2$$

with $\delta = \frac{1-b}{N}$. Normalizing $v_\psi = 1$, the last sum is a composite midpoint rule for $\int_b^1 \left(x - \bar{\xi}\right)^2 dx = \frac{(1-\bar{\xi})^3 - (b-\bar{\xi})^3}{3}$: the N outer brackets have equal width δ and midpoints $b + \left(k - \frac{1}{2}\right)\delta$, with $b + N\delta = 1$. As the integrand is quadratic (constant second derivative $f'' = 2$), the midpoint rule is exact up to the standard error term $\frac{(1-b)\delta^2}{24} f'' = \frac{(1-b)\delta^2}{12} = \frac{(1-b)^3}{12N^2}$, by which the sum

underestimates the convex integral. Hence

$$\tilde{V} \coloneqq \frac{2}{\lambda\bar{\tau}m}V = \frac{(1-\bar{\xi})^3 - (b-\bar{\xi})^3}{3} - \frac{(1-b)^3}{12N^2}.$$

So

$$\frac{d\tilde{V}}{db} = -(b-\bar{\xi})^2 + \frac{(1-b)^2}{4N^2}$$

which is 0 for $b = \bar{\xi} \pm \frac{\delta}{2}$. When b goes from $-\infty$ to $+\infty$, $\frac{d\tilde{V}}{db}$ is negative, then positive in $\left(\bar{\xi} - \frac{\delta}{2}, \bar{\xi} + \frac{\delta}{2}\right)$, then negative. Hence, the maximum of $\tilde{V}$ is reached for:

$$b^* = \bar{\xi} + \frac{\delta}{2} \tag{123}$$

which is the value of b for which the optimal tax in the central band, $t_0 = b - \frac{\delta}{2} - \bar{\xi}$, is just 0. Also, as $b^* = \bar{\xi} + \frac{1-b^*}{2N}$, this gives, more explicitly:

$$b^* \coloneqq \frac{\bar{\xi} + \frac{1}{2N}}{1 + \frac{1}{2N}} \tag{124}$$

When $b = b^*$, this is

$$\frac{2}{\lambda\bar{\tau}m}V = \frac{(1-\bar{\xi})^3}{3}\left(1 - \frac{1}{(2N+1)^2}\right) = \frac{1}{3}(1-\bar{\xi})^3\left(1 - \frac{1}{n^2}\right) \tag{125}$$

Recall from (121) that the optimal special tax for bracket k is the soft-thresholded Ramsey tax

$$t_{G_k} = -T\left(\hat{\psi}_{G_k}, \bar{\xi}\right).$$

I record it, together with the welfare gain, here, now expressed in terms of the tax $\tau_{G_k} = \bar{\tau}\left(1 + t_{G_k}\right)$ rather than t_{G_k}:

Lemma 6. *With n tax brackets (with optimally chosen bounds), the optimal tax in bracket k is in* (39)*, and the welfare gain from moving from one to n tax brackets (so $V(1) = 0$) is*

$$V(n) = \frac{1}{6}\lambda\bar{\tau}v_\psi^2 m \max\left(1 - \frac{\bar{\xi}}{v_\psi}, 0\right)^3 \left(1 - \frac{1}{n^2}\right)(1 + O(v_\psi)) \tag{126}$$

where $\bar{\xi}$ is in (98). The displayed term is the leading one as $v_\psi \to 0$; it is $O(\lambda^2)$ overall (since $\lambda\bar{\tau} = O(\lambda^2)$), and—demand being linear—exact in λ, so the $O(v_\psi)$ remainder is the only approximation.

Moments of the optimal tax deviations. These hold in the differentiated-tax regime $\tilde{\xi} < 1$; when $\tilde{\xi} \geq 1$ soft thresholding sets every deviation to zero, and both moments are zero.

With $b = b^*$, and to the leading order in v_ψ (the deviations t_{G_k} being given by (121)), the outer brackets have $|t_{G_k}| = v_\psi \left(u_k - \tilde{\xi}\right)$ at the midpoints $u_k = b^* + \left(k - \frac{1}{2}\right)\delta$, each with mass $\frac{\delta}{2}$, while the central bracket has $t_0 = 0$. The midpoint rule is *exact* for affine integrands, so

$$\langle |t_G| \rangle = 2 \sum_{k=1}^{N} \frac{\delta}{2} v_\psi \left(u_k - \tilde{\xi}\right) = v_\psi \int_{b^*}^{1} \left(u - \tilde{\xi}\right) du = \frac{v_\psi}{2} \left[\left(1 - \tilde{\xi}\right)^2 - \left(b^* - \tilde{\xi}\right)^2\right]$$

and since $b^* - \tilde{\xi} = \frac{\delta}{2} = \frac{1-\tilde{\xi}}{2N+1} = \frac{1-\tilde{\xi}}{n}$ (from (124)),

$$\langle |t_G| \rangle = \frac{v_\psi}{2} \left(1 - \tilde{\xi}\right)^2 \left(1 - \frac{1}{n^2}\right). \tag{127}$$

For the second moment, the display above reads $V = \frac{\lambda \bar{\tau} m}{2} \langle t_G^2 \rangle$, so the welfare formula gives

$$\langle t_G^2 \rangle = \frac{v_\psi^2}{3} \left(1 - \tilde{\xi}\right)^3 \left(1 - \frac{1}{n^2}\right). \tag{128}$$

where $\tilde{\xi} = \bar{\xi}/v_\psi$, so that $1 - \tilde{\xi} = M(m)$ when the threshold binds.

B.6 Proof of Proposition 11

(i) Immediate from Proposition 10.

(ii) I use $\bar{V} \coloneqq V^{1/3}$. I write $\zeta \coloneqq \left(\frac{1}{6}\lambda \bar{\tau} v_\psi^2\right)^{1/3} > 0$ for the scale of the welfare gains, and $z \coloneqq \frac{1}{\lambda v_\psi}$, so that $z\xi\frac{1-m}{m} = \frac{\bar{\xi}}{v_\psi}$ is the compression term appearing in (40). In the interior case (otherwise $n^* = 1$), (40) gives, with $m = m(n)$,

$$\bar{V} = \zeta\, m^{1/3} \left(1 - \tfrac{1}{n^2}\right)^{1/3} \left(1 - z\xi\tfrac{1-m}{m}\right) = A(n) - \xi B(n),$$

where $A(n) = \zeta\, m^{1/3} \left(1 - \frac{1}{n^2}\right)^{1/3}$ and $B(n) = \zeta z (1-m)\, m^{-2/3} \left(1 - \frac{1}{n^2}\right)^{1/3}$. Thus $\bar{V}_\xi = -B(n)$. As $m(n)$ decreases in n, each factor of B—namely $(1-m)$, $m^{-2/3}$, and $\left(1 - \frac{1}{n^2}\right)^{1/3}$—is positive and weakly increasing in n, and the last one is increasing, so $B(n)$ is weakly increasing and

$$\bar{V}_{\xi n} = -B'(n) \leq 0,$$

with equality only where $m(n) = 1$, so that $B(n) = 0$. Hence by monotone comparative statics the optimum $n^*(\xi)$ is decreasing in ξ. Finally, for ξ large enough, $V(n) = 0$ and $n^* = 1$—this is also the message of Proposition 9.

(iii) comes from the fact that $V(n)$ increases in n, strictly so as $m > 0$.

(iv) comes from the fact that $V(2) > V(1) = 0$, provided people can pay some attention to two tax brackets, i.e. $m(2) > 0$.

B.7 Proofs of Propositions 12 and 13

Both are specializations to $\underline{m} = 0$ of the analysis with a free floor of attention. Proposition 12 is Proposition 28 (optimal tax system with costly endogenous attention) at $\underline{m} = 0$: there, the no-stealth condition (106) becomes simply $\xi \geq \lambda \bar{m}$, as its $O(v_\psi)$ margin carries a factor $\underline{m}$. Proposition 13 is Proposition 29 at $\underline{m} = 0$, where the cost-correction factor $\frac{\bar{m}}{\bar{m}-\underline{m}}$ in (108) equals one. The attention constraints are then slack by a factor of at least 6, so the corner $m_{G_k} = \bar{m}$ is robust to a marginal change in n.

B.8 Proof of Proposition 32

Call $\bar{m}^a$ the average attention to a bracket deviation across the two groups, so $\bar{m}^a = \epsilon + (1-\epsilon) m_G$ when the behavioral agents pay attention m_G. In the units of (117), and using $\sum_k s_{G_k} t_{G_k} = 0$, the terms of welfare that depend on the tax deviations are

$$\sum_k s_{G_k} \left[-\bar{m}^a \hat{\psi}_{G_k} t_{G_k} - \frac{\bar{m}^a}{2} t_{G_k}^2 - (1-\epsilon) \frac{\xi (1 - m_{G_k})}{\lambda} |t_{G_k}| \right] - \frac{(1-\epsilon)\kappa}{\lambda \bar{\tau}} \sum_k m_{G_k} \qquad (129)$$

where the rational agents contribute no dread and no cognition cost.

(i) At the uniform tax all deviations are zero, so attention has no value, the behavioral agents set $m_G = 0$, and no cognition cost is paid. Fix $\alpha \in (0,1)$ and take two brackets, symmetric around the average elasticity, with $t_G = -\alpha \hat{\psi}_G$, where $\hat{\psi}_G$ is the bracket's average elasticity gap; this satisfies $\langle t_G \rangle = 0$, and by symmetry the two brackets carry the same attention.

Suppose first that the behavioral agents still choose $m_G = 0$. Then $\bar{m}^a = \epsilon$, no cognition cost is paid, and, as $\xi = 0$, (129) reduces to its first two terms, giving a welfare gain of

$$\epsilon \alpha \left(1 - \frac{\alpha}{2}\right) \left\langle \hat{\psi}_G^2 \right\rangle > 0.$$

Suppose instead that they choose some $m_G = m > 0$. Their own surplus is then weakly higher than at $m_G = 0$, which was available to them; the rational agents are unaffected; and tax revenue is higher, as

$$\frac{\partial R}{\partial m} = (1-\epsilon) \bar{\psi} \bar{\tau}^2 \alpha (1-\alpha) \left\langle \hat{\psi}_G^2 \right\rangle > 0 \qquad \text{for } \alpha < 1,$$

attention making demand respond to deviations that are still short of their Ramsey values. Welfare is therefore at least what it is in the previous case, hence again strictly above the uniform tax. Neither step uses any property of C^{Cog} beyond $C^{\text{Cog}}(0) = 0$, so the conclusion holds for every cognition cost function.

(ii) I show that the uniform tax is optimal when

$$\xi \geq \xi^*(\epsilon) := \frac{\lambda v_\psi (\epsilon + (1-\epsilon)\bar{m})}{(1-\epsilon)(1-\bar{m})}, \tag{130}$$

which is the sense of "high enough". The threshold $\xi^*(\epsilon)$ increases with ϵ, and goes to infinity as $\epsilon \to 1$: a larger rational fringe calls for more complexity aversion. For $\epsilon = 0$ it is $\frac{\lambda v_\psi \bar{m}}{1-\bar{m}}$, i.e. $M(\bar{m}) = 0$, which is the condition of Proposition 9.

Take any deviations t_{G_k}, with the behavioral agents responding as in (104). Consider first the brackets they attend, so $m_{G_k} = \bar{m}$ and $\bar{m}^a = \epsilon + (1-\epsilon)\bar{m}$. As $\left|\hat{\psi}_{G_k}\right| \leq v_\psi$, and the quadratic and cognition terms are negative, their contribution to (129) is at most

$$\sum_k s_{G_k} |t_{G_k}| \left[(\epsilon + (1-\epsilon)\bar{m}) v_\psi - (1-\epsilon) \frac{\xi(1-\bar{m})}{\lambda} \right],$$

and the bracket is non-positive exactly when (130) holds. A bracket they do not attend has $m_{G_k} = 0$, hence a lower $\bar{m}^a$ and the full dread $(1-\epsilon)\frac{\xi}{\lambda}|t_{G_k}|$: the same bound applies, and is slacker, provided $\xi \geq \lambda\bar{m}$ holds with a margin of order v_ψ, as in the no-stealth condition—otherwise such a bracket would be a stealth tax. So no set of deviations improves on the uniform tax.

C An analysis of de Clippel, Moscariello, Ortoleva and Rozen, based on complexity aversion

de Clippel et al. (2025) run three large experiments showing that people put a lower dollar value on an option because they find it complex, in a way no expected-utility model with beliefs can produce. They conclude that their evidence calls for a model that subtracts a first-order complexity cost from expected utility, and they leave the construction of that model open. This section takes up the invitation: I apply (6) to two of their three experiments—compound risk (their Experiment C) in Section C.1, and Bayesian updating (their Experiment A) in Section C.2—and confront it with their numbers.[82] The exercise

[82] I leave aside their Experiment B (visual perception). It is the one experiment where the object the subject is uncertain about is a state of the world rather than her own computation.

is quantitatively demanding, because the parameter ξ was calibrated in Section 7.1 from an unrelated source—the rejection of small gambles—and I hold it fixed at $\xi \simeq \frac{2}{3}$ throughout. In each case the rest of the calculation is fixed by the design, so what follows is an a priori calibration, not an estimation.

Common apparatus: complexity aversion in a valuation task. All three experiments share a format. A bet pays a prize $Z = \$30$ if a drawn ball or card is purple, \$0 otherwise, and subjects state a dollar value for it through a multiple price list (MPL): for each sure amount $s \in \{\$1, \ldots, \$30\}$ they choose between s for sure and the bet, with a single switching point. Each complex bet is compared to a plain 50/50 lottery on the same prize, whose value I write $V^{50/50}$; measured complexity aversion is $\text{CA} \coloneqq V^{50/50} - V^{\text{bet}}$. Subjects also report, for each scenario, their belief about the exact chance of purple and their "cognitive uncertainty" (CU): 100% minus their stated confidence in that belief, following Enke and Graeber (2023). Throughout I read CU, which lies in $[0, 1]$, as the imprecision $1 - m$ of the model.

The adaptation of the model needed is small: I apply the discrete-choice version of the complexity-aversion cost (Section A.5) to a valuation task. Over these stakes utility is essentially linear in money, so I take utils to be dollars.[83] The MPL is a binary menu. The sure amount s is a riskless number on the screen: $v^{\text{sure}} = s$ does not depend on anything the agent must work out, so $v^{\text{sure}}_{x_i} = 0$ and, by (85), it carries no complexity penalty. The bet is worth $v^{\text{bet}} = \pi(x)\, Z$, where π is the winning probability and x collects whatever the agent must process to arrive at it. Hence, by (85),

$$C^{\text{CA}}_{\text{bet}} = \xi \sum_i \left| v^{\text{bet}}_{x_i} \right| \sigma_{x_i} (1 - m_i) = \xi\, Z\, \Sigma_\pi\, (1 - m)\,, \qquad \Sigma_\pi \coloneqq \sum_i |\pi_{x_i}|\, \sigma_{x_i}, \tag{131}$$

taking a common m across the features of a given scenario. I call Σ_π the "probability stakes of thinking": the typical size, in probability units, of the error the agent would make if she did not think at all. It is the object that plays the role of $|a^r_x|\, \sigma_x$ in (6). The agent switches on the MPL where $v^{\text{bet}} - C^{\text{CA}}_{\text{bet}} = s$, and the 50/50 lottery has $\pi_x = 0$, hence no penalty. So the model's prediction is exactly the object these experiments measure.

Proposition 33. (Complexity aversion in a valuation task) *In an MPL valuation of a bet paying Z on an event of probability $\pi(x)$, the measured complexity aversion is*

$$CA = V^{50/50} - V^{bet} = \xi\, Z\, \Sigma_\pi\, (1 - m) = \xi\, Z\, \Sigma_\pi \cdot CU, \tag{132}$$

[83] The 50/50 lottery on \$30 is valued at \$14.26, so risk aversion is small; and in any case CA is a difference between two bets with the same prize, so curvature nets out to second order.

reading the imprecision $1 - m$ off measured cognitive uncertainty. It is proportional to the prize, to the probability stakes, and to cognitive uncertainty, with the unitless ξ as the only free parameter.

Three features of (132) do the work below. It is first order in the stakes; it is proportional to cognitive uncertainty, which delivers parameter-free cross-treatment and cross-subject tests; and it is proportional to Σ_π, so that an object can be complicated to look at and still carry no penalty, if nothing the agent might get wrong would change what she should do.

One reading convention deserves to be stated once. In the model σ_{x_i} is the modulus—the typical size—of the disturbance x_i, evaluated at the default and not at its realization. In the compound-risk experiment of Section C.1 this is literally a standard deviation, since the deck is genuinely drawn at random. In the updating experiment of Section C.2 the numbers on the screen are fixed, and I read σ_{x_i} as the size of the piece of information the subject is asked to process—the "size modulus" usage of Section 5. This is the one judgment call in what follows, and I flag it rather than bury it.

The Ellsberg bet, and why complexity aversion should correlate with ambiguity aversion. Their mechanism result is that undervaluation is driven by the interaction of cognitive uncertainty with ambiguity aversion, measured in a separate Ellsberg task as $\mathrm{AA} = V^{50/50} - V^{\mathrm{Ellsberg}}$. The model delivers that structure, if one grants one extension. In their Ellsberg task an urn holds 100 balls of unknown composition θ, "all purple, all green, or any combination." The agent's action is her valuation; the right action depends on θ, which no amount of thinking will reveal. This is the corner $m = 0$, with a modulus given by the prior over compositions: with θ uniform on $[0, 1]$, $\sigma_\theta = \frac{1}{\sqrt{12}}$. Then (131) gives[84]

$$\mathrm{AA}_i = \xi_i \, Z \, \sigma_\theta = \frac{Z}{\sqrt{12}} \, \xi_i : \tag{133}$$

the Ellsberg task is a measurement of the individual's ξ_i, up to a known constant. Substituting into (132),

$$\mathrm{CA}_i = \frac{\Sigma_\pi}{\sigma_\theta} \, \mathrm{AA}_i \times \mathrm{CU}_i. \tag{134}$$

So the model predicts that complexity aversion enters as a pure product of ambiguity aversion and cognitive uncertainty, with no main effect of either, and with a coefficient equal to

[84]This is a genuine extension rather than a restatement: in Section 7.3 I noted that complexity aversion, unlike ambiguity aversion, does not change the utility of a rational agent. That is right when the imprecision is the agent's own; here it is irreducible, so $m = 0$ even for a fully rational agent, and the two coincide. Read this way, first-order complexity aversion contains Ellsberg-style ambiguity aversion as its $m = 0$ corner. It is a satisfying place for the two literatures to meet, and it is why the correlation that de Clippel et al. (2025) document should exist at all.

the ratio of the two moduli—$\Sigma_\pi\sqrt{12}$, which is 1.04 in their Experiment C and 0.70 in their Experiment A. Their Table 1 is the textbook version of the first two predictions: the interaction is 0.43 ($p < .01$) while ambiguity aversion alone (0.09) and cognitive uncertainty alone (-0.76) are both insignificant. The magnitude is the weak leg: the estimated interactions (0.26 to 0.60 across their three experiments) run below the predicted ratios, as one should expect, since AA and CU are each differences of noisy elicitations, so classical measurement error both attenuates the interaction and shifts part of it onto the main effect of AA. I read the structure as the informative prediction and the coefficient as a lower bound.

A remark on dominance. de Clippel et al. (2025) sketch a representation $V(f) = \mathbb{E}_\pi[u(f)] - \theta k(f)$, with k a measure of perceived complexity, and note that "models of this type may generate violations of first-order stochastic dominance unless properly disciplined." The criteria of Theorem 1 are that discipline. Because the penalty in (131) is proportional to the stakes $|v_{x_i}|\,\sigma_{x_i}$ rather than to a free-standing index of how complicated the option looks, a riskless act is never penalized, and an act's penalty is bounded by ξ times its own dispersion. For the binary bets used here—prizes $\{0, Z\}$ with probability π—the value is $\pi Z - \xi Z \Sigma_\pi (1-m)$, which exceeds the dominated sure amount 0 whenever $\xi\Sigma_\pi(1-m) \leq \pi$: at $\xi = \frac{2}{3}$ the left-hand side is below 0.1 in every scenario below, against $\pi = \frac{1}{2}$. More generally, a two-point act paying c or $c + \Delta$ with equal chances is never valued below c provided $\xi(1-m) \leq 1$, which the calibrated ξ satisfies with room to spare. The unitlessness and moderate size of ξ are thus not just conveniences for portability: they are what keeps a first-order penalty well behaved.[85]

C.1 Compound risk

I begin with their experiment on compound risk, where the probability stakes are read off the design most directly. The prize is $Z = \$30$ throughout.

The experiment. Three objects are valued, all of which reduce to a 50/50 bet on \$30.

- The 50/50 lottery: the deck is 50% purple. This is the benchmark.
- The non-trivial compound lottery: there are two decks, one 20% purple and one 80% purple; the computer picks a deck at random, then draws a card. This reduces to 50%, but not transparently.

[85] The discipline is not unconditional. For a very skewed act—a small chance of a large prize—the dispersion can exceed the mean gain over the dominated sure amount, and a first-order penalty can then conflict with dominance. This is a known cost of first-order models, shared with first-order risk aversion, and it does not arise in the near-symmetric bets studied here.

- The trivial compound lottery: the same two-stage structure, but both decks are 50% purple. This also reduces to 50%, and transparently so.

A One-Stage Mirror treatment replaces "pick a deck at random, then draw" by "combine the two decks, then draw"—the same two numbers to process, but only one stage of risk.

The design's virtue is that the null is airtight. Under expected utility, with any beliefs, Bayesian or not, and any risk aversion, a subject who holds some belief about the chance of purple must value the compound lottery at least as much as some 50/50 lottery.[86]

The facts. In the Main treatment, CA = \$2.56 for the non-trivial compound (with an average CU = 42%), and CA = \$0.19—essentially zero—for the trivial compound (with CU = 26%). In the One-Stage Mirror, aversion to the non-trivial object falls to \$1.40, and cognitive uncertainty falls to 23%. The undervaluation is large: \$2.56 is 17% of the bet's expected value of \$15.

The probability stakes. The design hands us Σ_π for each object.

- Non-trivial compound. There is exactly one thing to figure out: which deck was selected. Writing $\pi = \frac{1}{2} + x$, we have $x = \pm 0.3$ with equal probabilities, so $\pi_x = 1$, $\sigma_x = 0.3$, and $\Sigma_\pi = 0.3$.

- Trivial compound. The deck identity does not move the right answer: $\pi_x = 0$, so $\Sigma_\pi = 0$. The stage is there, the numbers are there, the agent is not fully sure of herself (CU = 26%)—but there is no mistake to be made, so there is nothing to dread.

- One-Stage Mirror, non-trivial. The two decks are pooled, so $\pi = \frac{1}{2}(\pi_1 + \pi_2)$ with $\pi_1 = 0.2$, $\pi_2 = 0.8$. Now $\pi_{x_j} = \frac{1}{2}$ for each of the two deck compositions, each with modulus $\sigma_{x_j} = 0.3$, so $\Sigma_\pi = \frac{1}{2}(0.3) + \frac{1}{2}(0.3) = 0.3$ again. The treatment removes a stage of risk, but not one iota of the arithmetic: the model says it should change m, and only m.

The predictions. Plugging $\xi = \frac{2}{3}$, $Z = 30$ and the measured cognitive uncertainties into (132) gives Table 1. There is no fitted parameter: ξ comes from the calibration of Section 7.1, Z and Σ_π from the design, and CU is data.

Three things are worth emphasizing.

[86] This remains true under most non-expected-utility models with well-defined beliefs, including cumulative prospect theory with the usual weighting functions; see de Clippel et al. (2025), footnote 5.

	Σ_π	CU	Predicted CA	Observed CA
Non-trivial compound (Main)	0.3	42%	\$2.52	\$2.56
Non-trivial, One-Stage Mirror	0.3	23%	\$1.38	\$1.40
Trivial compound (Main)	0	26%	\$0.00	\$0.19

Table 1: Complexity aversion in the compound-risk experiment of de Clippel et al. (2025). Predicted CA is $\xi Z \Sigma_\pi \text{CU}$ from (132), with $\xi = \frac{2}{3}$ and $Z = \$30$.

First, the level is right: \$2.52 predicted against \$2.56 observed, with ξ imposed at $\frac{2}{3}$ rather than fitted here. I did not expect a match this close, and would not read too much into the last digit; but it is the same parameter, in a domain the calibration never saw.

Second, the comparative static across treatments is right, and it is the sharper test. Since the One-Stage Mirror leaves Σ_π untouched, (132) says the ratio CA/CU should be the same across the two treatments, and equal to $\xi Z \Sigma_\pi = 6.0$. In the data it is $2.56/0.42 = 6.10$ and $1.40/0.23 = 6.09$. Equivalently: aversion falls by 45% across the treatments, and cognitive uncertainty falls by 45%. de Clippel et al. (2025) note the co-movement; the model says the elasticity should be exactly one, and it is.

Third, and most diagnostically, the trivial compound. It has the same number of stages, the same numbers on the screen, the same support $\{\$0, \$30\}$, the same first-stage entropy, and a strictly positive measured cognitive uncertainty of 26%—and essentially zero aversion. In (131) this is the factor $|\pi_{x_i}|$ doing its work: the Lasso-like term of (6) prices the complexity of the policy, not of the environment. A penalty in information space, a counting measure of the number of parameters (Section 7.3.2), a support-size measure (Puri, 2025), or an entropy measure would all charge the two compound lotteries the same amount. The data separate them by a factor of thirteen. Relatedly, the model predicts the "relative complexity aversion" that de Clippel et al. (2025) construct—the value gap between the trivial and the non-trivial compound—to be $\xi Z \Sigma_\pi \text{CU} = \$2.52$, against \$2.37 observed.

C.2 Bayesian updating

Another of their experiments is about updating. Its probability stakes are also fixed by the design, once one works in the frame in which Bayes' rule is additive, so it too admits an a priori calibration; and it is the setting in which the failure of expected utility is starkest. The prize is again $Z = \$30$.

The experiment. Subjects see: a bag with 20 green and 30 purple balls; $\frac{1}{2}$ of the green balls and $\frac{1}{3}$ of the purple balls are marked with an X; a ball has been drawn and is marked

with an X. The Updating Bet pays \$30 if that ball is purple. Bayes' rule gives

$$\pi = \frac{30 \cdot \frac{1}{3}}{30 \cdot \frac{1}{3} + 20 \cdot \frac{1}{2}} = \frac{10}{20} = \frac{1}{2}, \tag{135}$$

so the bet is worth exactly what the 50/50 lottery is worth. A Complement Updating Bet inverts the colors, which is equivalent to betting on green in the same scenario. Subjects value both, and separately report their belief and their cognitive uncertainty. A Mirror treatment presents an arithmetically identical task stripped of its conceptual difficulty: "$\frac{1}{2}$ of the green balls and $\frac{1}{3}$ of the purple balls were put in the bag"—so the bag visibly holds 10 green and 10 purple, and no updating is involved.

The null is again airtight, and here it bites against the entire non-Bayesian updating literature. Whatever posterior a subject holds—Bayesian, base-rate neglecting, or otherwise—at least one of the two complementary events has probability at least $\frac{1}{2}$, so at least one of the two bets must be worth at least as much as the 50/50 lottery.

The facts. The two updating bets are valued at \$12.68 and \$12.75, against \$14.26 for the 50/50 lottery: both are first-order stochastically dominated by it. Average complexity aversion is CA = \$1.54, at an average CU = 35%. The 35% of subjects who report the correct Bayesian posterior have CU = 17%, against 45% for everyone else, and they still display CA = \$0.79. In the Mirror treatment cognitive uncertainty falls to 15% and complexity aversion disappears.

The probability stakes, from the design. The frame in which to compute Σ_π is the one in which the agent's problem separates into features that enter additively—which, for updating, is log-odds, since that is the frame in which Bayes' rule is an addition:[87]

$$\operatorname{logit} \pi = \underbrace{\ln \frac{N_P}{N_G}}_{x_1\text{: base rate}} + \underbrace{\ln \frac{\ell_P}{\ell_G}}_{x_2\text{: signal}}. \tag{136}$$

There are exactly two features, and they enter linearly, as in the model's $a^r = \sum_i a^r_{x_i} x_i$. The default $x^d = (0, 0)$—a flat base rate and an uninformative signal—gives $\pi^d = \frac{1}{2}$, so the model's default action is $a^d = \frac{Z}{2}$: the experiment's benchmark object, the 50/50 lottery, is the model's default. That is a fortunate alignment of their design with the theory.

The design then fixes both the slopes and the moduli. With $N_P = 30$, $N_G = 20$, $\ell_P = \frac{1}{3}$

[87] Log-odds are also the normal way to write perceptions of uncertainty, as they live in $(-\infty, +\infty)$ rather than in a bounded interval such as $[0, 1]$; see Gabaix (2019).

and $\ell_G = \frac{1}{2}$,

$$x_1 = \ln \frac{3}{2} = +0.406, \qquad x_2 = \ln \frac{2}{3} = -0.406, \tag{137}$$

equal in size and opposite in sign—which is precisely why the answer in (135) is exactly one half. The slopes are $\frac{\partial \pi}{\partial x_i} = \pi\,(1-\pi) = \frac{1}{4}$. Hence

$$\Sigma_\pi = \frac{1}{4}\,|x_1| + \frac{1}{4}\,|x_2| = \frac{1}{2} \ln \frac{3}{2} = 0.203. \tag{138}$$

The answer barely depends on the frame, which is what makes this a calibration rather than a choice.[88]

This is a clean illustration of why the penalty must be a sum over features rather than a norm of them. The net evidence here is exactly zero, $x_1 + x_2 = 0$. Collapse the two features into one and $\Sigma_\pi = 0$, predicting no complexity aversion at all; use the "square root" variant discussed in Section 7.3 and one gets $\frac{1}{4}\sqrt{x_1^2 + x_2^2} = 0.143$, hence \$1.00. The subject must handle two pieces of evidence of size $\ln \frac{3}{2}$ each, and dreads both, even though they cancel. This is the "CA applies feature by feature" property at work.

The prediction: symmetry, plus a common shift. The two bets are mirror images: $\pi^{\text{green}} = 1 - \pi^{\text{purple}}$, so $|\pi_{x_i}|$ and hence Σ_π are identical for the two, and perception is unbiased, $\mathbb{E}\,[\pi^s] = \pi$. So (132) predicts that the two bets are valued equally, and both displaced below the benchmark by the same amount:

$$V^{\text{purple}} = V^{\text{green}} = V^{50/50} - \xi Z \Sigma_\pi\,\text{CU} = 14.26 - \frac{2}{3} \cdot 30 \cdot 0.203 \cdot 0.35 = \$12.84, \tag{139}$$

against \$12.68 and \$12.75 observed.[89] Equivalently, in the units the experiment reports, predicted CA = \$1.42 against \$1.54 observed.

The shape of this prediction is worth as much as its magnitude. A model in which subjects hold a distorted posterior $\hat{\pi}$ and maximize expected utility can move the two values apart, in opposite directions, but can never move them down together: that requires a term which does not come from beliefs at all.

[88] Redo it in probability space: writing p for the share of purple balls and $s := \frac{\ell_P}{\ell_P + \ell_G}$ for the normalized strength of the signal, one has $\text{logit}\,\pi = \text{logit}\,p + \text{logit}\,s$ exactly, so both slopes equal 1 at the default, and with $p = 0.6$ and $s = 0.4$ the deviations are ± 0.1, giving $\Sigma_\pi = 0.1 + 0.1 = 0.200$. Log-odds give 0.203, probabilities 0.200: the deviations are small enough that the coordinate system is immaterial.

[89] Note that the benchmark in (139) is the measured $V^{50/50} = \$14.26$, not the risk-neutral $\frac{Z}{2} = \$15$. The 74-cent gap between them is risk aversion and the discreteness of the MPL grid; it is common to all the bets, so it drops out of CA, but it does not drop out of the levels.

A test that uses no parameter at all. The probability stakes Σ_π are a property of the problem, which is the same for everyone; cognitive uncertainty is a property of the subject. So (132) implies that, across any partition of the sample facing the same task,

$$\frac{\mathrm{CA}^{\text{group 1}}}{\mathrm{CA}^{\text{group 2}}} = \frac{\mathrm{CU}^{\text{group 1}}}{\mathrm{CU}^{\text{group 2}}}, \tag{140}$$

which requires knowing neither ξ nor Σ_π. Split on whether the subject computed the correct posterior. The measured cognitive uncertainties are 17% and 45%, a ratio of 0.38. The measured complexity aversions are \$0.79 and \$1.94, a ratio of 0.41.[90] Equivalently, in levels: $\$1.54 \times \frac{17}{35} = \$0.75$ predicted for the correct-posterior group, against \$0.79 observed.

A subject who has worked out that the answer is one half, and has said so, still pays about eighty cents to avoid the bet—and pays it in proportion to how unsure she remains. That is precisely the psychology (6) is built to capture: the dread is not a consequence of holding a wrong belief, it is the felt possibility of having erred, and it survives being right. It also disposes of a natural alternative reading of the whole phenomenon, namely that undervaluation is just miscalculation in disguise.

The Mirror treatment, and how the two experiments identify m and Σ_π separately. In the Mirror, cognitive uncertainty falls from 35% to 15%. Were only m changing, (132) would predict aversion of $\$1.54 \times \frac{15}{35} = \$0.66$; in fact it disappears. That is as it should be, because the Mirror does not merely make the same problem clearer: by displaying the bag composition directly it removes the deviation from 50/50 that the subject had to construct—in the language of (136), both x_1 and x_2 are set to zero and the subject reads 10 against 10—so it drives Σ_π to zero as well as raising m. The attractive feature of having two experiments is that they separate the two channels, as summarized in Table 2.

Manipulation	Main channel	Σ_π	CU	CA
Exp. C: trivial vs non-trivial compound	stakes	$0.30 \to 0$	$42\% \to 26\%$	$\$2.56 \to \$0.19$
Exp. C: One-Stage Mirror	attention	$0.30 \to 0.30$	$42\% \to 23\%$	$\$2.56 \to \$1.40$
Exp. A: Mirror	both	$0.203 \to 0$	$35\% \to 15\%$	$\$1.54 \to \approx 0$

Table 2: The three "switch-off" treatments of de Clippel et al. (2025), read through (132). Killing the probability stakes kills complexity aversion outright; raising attention lowers it proportionately.

[90] The paper reports \$0.79 for the correct-posterior group, which is 35% of the sample, and \$1.54 overall; the residual group therefore averages \$1.94. As a consistency check on the two reported splits, $0.35 \times 17\% + 0.65 \times 45\% = 35\%$, the reported overall average.

An over-identifying restriction. Because Σ_π here is predicted rather than fitted, the model has a degree of freedom left over, and it can be spent on a second moment. Every subject faces the same numbers, so the dispersion of their stated beliefs comes from the perception noise alone: with $x_i^s = mx_i + \eta_i$ and $\sigma_{\eta_i}^2 = m\,(1-m)\,\sigma_{x_i}^2$,

$$\operatorname{var}(\pi^s) = m\,(1-m)\sum_i \left(\pi_{x_i}\sigma_{x_i}\right)^2. \tag{141}$$

What enters is the sum of the squares, not the square of the sum Σ_π; the two coincide only when there is a single feature. This is the feature-by-feature structure of (138) showing up again, from the other side: dread adds the features linearly, while their perception errors add in quadrature. With the two equal features here the sum of squares is $\frac{1}{2}\Sigma_\pi^2$, so the model predicts a cross-sectional standard deviation of reported posteriors of

$$\operatorname{sd}(\pi^s) = \Sigma_\pi\sqrt{\frac{m\,(1-m)}{2}} = 0.203 \times \sqrt{\frac{0.65 \times 0.35}{2}} \simeq 0.068, \tag{142}$$

around a mean of exactly $\frac{1}{2}$, since $x_1 + x_2 = 0$. The elicited beliefs are in their data, so this is a genuine over-identifying restriction, and one I would like to see checked—bearing in mind that it is a floor rather than a point prediction, since the model gives every subject the same attention and the same frame, while their errors are plainly heterogeneous: 35% give exactly 0.50, whereas the salient mistakes sit further out, at 0.60 using the base rate alone, 0.40 using the likelihoods alone, and 0.33 reading off "one third".

Assessment. de Clippel et al. (2025) conclude that their evidence calls for a model that subtracts a first-order complexity cost from expected utility. That is (6). Table 3 collects what it delivers. The single unitless parameter is held at the value calibrated on gambles, everything else is design or data, and nothing is estimated.

		Σ_π	CU	Predicted CA	Observed CA
Exp. C	Non-trivial compound (Main)	0.30	42%	\$2.52	\$2.56
Exp. C	Non-trivial, One-Stage Mirror	0.30	23%	\$1.38	\$1.40
Exp. C	Trivial compound (Main)	0	26%	\$0.00	\$0.19
Exp. A	Updating bet (Main)	0.203	35%	\$1.42	\$1.54

Table 3: All four cells, at $\xi = \frac{2}{3}$ and $Z = \$30$. Predicted CA is $\xi Z \Sigma_\pi \mathrm{CU}$ from (132); Σ_π is computed from the design in each case, and CU is the measured cognitive uncertainty.

What I find encouraging is not only that the sign and the mechanism come out right, but that the magnitudes do. Four cells, two experiments, one parameter fixed in advance at

$\xi = \frac{2}{3}$: every prediction lands within about ten percent of the observation, and the fourth cell, where the design sets the stakes to zero, is predicted to show nothing and shows nothing. On top of that the model gets the cross-subject gradient of the updating experiment without using ξ at all. It also suggests that their design—value a complex object and a simple object with the same reduced payoff distribution, and measure cognitive uncertainty—is a portable instrument for estimating ξ in other domains. Given how few unitless, transportable parameters behavioral economics has, that seems worth having.

D Omitted proofs

D.1 Proof of Theorem 1

Throughout, f and g are C^2, and f is finite. Since attention m ranges over the compact set $[0,1]$, the function h defined in Step 1 below is then continuous and finite, hence bounded: $\bar{H} \coloneqq \sup_{m\in[0,1]} h(m) < \infty$.

The logic of the proof is as follows: the Separable homogeneity criterion alone forces C^{CA} to be a multiple of an ℓ^β norm of $(b_i h(m_i))_i$, for some exponent β; the Generation of full simplicity criterion requires that $\beta \leq 1$. But the Finiteness criterion requires $\beta \geq 1$. Hence only $\beta = 1$ remains possible, and indeed satisfies both criteria.

Step 0: Preliminaries on g. By Unit invariance and Separable homogeneity, $C^{\mathrm{CA}} = g(S)$ with $S \coloneqq \sum_i f(b_i, m_i)$. Homogeneity of degree 1 applied at $b = 0$ gives $C^{\mathrm{CA}}(0,m) = \lambda C^{\mathrm{CA}}(0,m)$ for all $\lambda > 0$, hence

$$C^{\mathrm{CA}}(0,m) = 0. \tag{143}$$

Next, $\partial_{b_i} C^{\mathrm{CA}} = g'(S) f_i'(b_i)$, so Generation of full simplicity forces $g' \neq 0$; as g is C^2 and S ranges over an interval, g' has a constant sign. Since $C^{\mathrm{CA}} \geq 0$ with (143), that sign is positive (otherwise replace (f,g) by $(-f, g(-\cdot))$, which leaves C^{CA} unchanged). So g is increasing, and $f_i' > 0$ whenever $m_i < 1$.

Step 1: Separable homogeneity yields an ℓ^β norm. Write $f_i'(b_i) \coloneqq \partial_{b_i} f(b_i, m_i)$ and $P \coloneqq \sum_i b_i f_i'(b_i)$. Euler's identity for homogeneity of degree 1 reads

$$g'(S)\, P = g(S). \tag{144}$$

Differentiating (144) with respect to a single b_k, using $\partial_{b_k} S = f_k'(b_k)$ and $\partial_{b_k} P = f_k'(b_k) + b_k f_k''(b_k)$, gives

$$g''(S) f_k'(b_k) P + g'(S) [f_k'(b_k) + b_k f_k''(b_k)] = g'(S) f_k'(b_k).$$

The term $g'(S) f_k'(b_k)$ appears on both sides and cancels, leaving $g''(S) f_k'(b_k) P + g'(S) b_k f_k''(b_k) = 0$. Substituting $P = g(S)/g'(S)$ from (144) and dividing by $g'(S) f_k'(b_k)$:

$$\frac{b_k f_k''(b_k)}{f_k'(b_k)} = -\frac{g''(S) g(S)}{g'(S)^2}. \tag{145}$$

As there are at least two variables, the right-hand side of (145) can be moved (by perturbing some b_ℓ, $\ell \neq k$) while the left-hand side is held fixed. Hence both sides equal a constant, which I call $\beta - 1$, and that constant is common to all k. Integrating the two resulting ordinary differential equations yields $f(b, m) = a(m) b^\beta + c(m)$ and $g(S) = A(S - S_0)^{1/\beta}$, with $A > 0$ and $a \geq 0$ by Step 0, and $a(m) > 0$ for $m < 1$.

Two constants remain to be disposed of. First, $\beta > 0$: for $\beta < 0$ (resp. $\beta = 0$, where the ODE gives $f = a(m) \ln b + c(m)$) we would have $f(0, m) = \pm\infty$ whenever $a(m) \neq 0$, contradicting the finiteness of $f(0, \cdot)$. Second, c is an irrelevant constant: by (143) and the strict monotonicity of g, $\sum_i c(m_i) = S_0$ for every m, and varying a single m_i shows that c is constant; the constants then cancel from $S - S_0 = \sum_i a(m_i) b_i^\beta$. Setting $h \coloneqq a^{1/\beta}$,

$$C^{\mathrm{CA}}(b, m) = A \left(\sum_i (b_i h(m_i))^\beta \right)^{1/\beta}. \tag{146}$$

Step 2: Full simplicity imposes $\beta \leq 1$. Differentiating (146) at $b_i = 0^+$, with some other $b_\ell > 0$, gives $+\infty$ if $\beta < 1$, $Ah(m_i) > 0$ if $\beta = 1$, and 0 if $\beta > 1$. So the criterion is satisfied iff $\beta \leq 1$.

Step 3: Finiteness imposes $\beta \geq 1$. Suppose that $\beta < 1$. We will see that Finiteness is violated. Define $b_i \coloneqq i^{-\delta}$ for some $\delta \in (1, 1/\beta)$. Then, $\sum_i b_i < \infty$, but $\sum_i b_i^\beta = \infty$, so CA is infinite.

However, if $\beta \geq 1$, Finiteness is satisfied. This is well-known, but let us verify it in a self-contained way. Define $\tilde{b}_i \coloneqq b_i h(m_i)$. As $h \leq \bar{H}$, we have $T \coloneqq \sum_i \tilde{b}_i \leq \bar{H} \sum_i b_i < \infty$. Write $\hat{b}_i \coloneqq \tilde{b}_i / T \in [0, 1]$; the pointwise bound $\hat{b}_i^\beta \leq \hat{b}_i$ gives

$$S \coloneqq \sum_i \tilde{b}_i^\beta = T^\beta \sum_i \hat{b}_i^\beta \leq T^\beta \sum_i \hat{b}_i = T^\beta < \infty.$$

Conclusion. Steps 2 and 3 give $\beta = 1$, so that (146) yields

$$C^{\mathrm{CA}}(b, m) = \sum_i b_i \bar{h}(m_i), \qquad \bar{h} \coloneqq Ah.$$

The $\beta = 1$ case of Step 2 gives $\bar{h}(m) > 0$ for $m < 1$, and Finiteness makes $\bar{h}$ bounded: if $\bar{h}\left(m^{(i)}\right) \to \infty$ along some sequence, choosing b_i with $\sum_i b_i < \infty$ but $\sum_i b_i \bar{h}\left(m^{(i)}\right) = \infty$ (e.g. $b_i = 2^{-i}$ along a subsequence with $\bar{h}\left(m^{(i)}\right) \geq 4^i$) violates the criterion. The second display of the theorem then follows from the Unit invariance for utility and actions criterion by dimensional analysis, which does not interact with β: that criterion fixes the prefactor at $A = |u_{aa} s_a|$, up to a function h that does not depend on u.

Conversely, this functional form satisfies the four criteria.

D.2 Proof of Proposition 5

Then, the value function has a term (with $b \geq 0$): $Ab - Bb^2$ with $A = 2b^r - 2\xi(1 - m)$ and $B = 2 - m$. The optimum (subject to $b \geq 0$) is $b = \frac{A}{2B}$, so the value function is $\frac{A^2}{4B} = Bb^2$. Hence,

$$V(m) = b(m)^2 (2 - m) - C^{\mathrm{Cog}}(m) \tag{147}$$

Hence, if $C^{\mathrm{Cog}}(m) \geq b(m)^2 (2 - m)$, then $V(m) \leq 0 = V(0)$. So $m = b = 0$ is optimal (a strict inequality would make it the unique optimum). That is, the optimum has m and b equal to 0.

D.3 Proof of Proposition 6

We have:

$$\mathbb{E}\left[(p(x) - c(x)) a(x^s)\right] = \mathbb{E}\left[(b - 1) x(-\psi b x^s)\right] = (1 - b) b\psi \mathbb{E}\left[x x^s\right] = \left(b - b^2\right) \psi \sigma_x^2 m$$

and as $\mathbb{E}\left[(p(x) - c(x))\right] = 0$, the total producer surplus is:

$$\mathbb{E}\left[(p(x) - c(x))(1 + a(x^s))\right] = \left(b - b^2\right) \psi \sigma_x^2 m \tag{148}$$

This implies:

$$\begin{aligned} W &= V^{\text{cons}}(m) + \mathbb{E}\left[\left(p(x) - c(x)\right)\left(1 + a\left(x^s\right)\right)\right] \\ &= \frac{1}{2}\psi\sigma_x^2 b^2 m - C^{\text{CA}} - C^{\text{Cog}}(m) + \psi\sigma_x^2\left(b - b^2\right)m \\ &= \left(b - \frac{b^2}{2}\right)\psi\sigma_x^2 m - C^{\text{CA}} - C^{\text{Cog}}(m) \end{aligned}$$

D.4 Proof of Proposition 8

Recall that with $\chi = \frac{1}{2}\psi\sigma_x^2$, social welfare is:

$$W = \Pi + V^{\text{cons}}, \qquad \Pi = 2\left(b - b^2\right)m\chi, \qquad V^{\text{cons}} = b^2 m\chi - C^{\text{Cog}}(m)$$

where Π is the average profit (148), and V^{cons} is the consumer's expected utility. Hence,

$$W_b = \Pi_b + V_b^{\text{cons}} + \left(\Pi_m + V_m^{\text{cons}}\right)m_b$$

As the consumer optimizes m to maximize V^{cons}, we have $V_m^{\text{cons}} = 0$, so

$$W_b = 2(1 - b)m\chi + 2\left(b - b^2\right)m_b\chi \tag{149}$$

For all $b \in [0, 1)$, as $m_b \geq 0$, we have $W_b \geq 0$; and whenever $m > 0$, we have $W_b > 0$. Hence, $b = 1$ is the social optimum.

We next study when we have some $m > 0$. We observe that the first-order condition for m is $C^{\text{Cog}\prime}(m) = b^2\chi$.

If $C^{\text{Cog}\prime}(0) < \chi$, when $b \in [0, 1]$ is such that $b^2\chi > C^{\text{Cog}\prime}(0)$ we have $m > 0$, and $b = 1$ is a strict global optimum.

If $C^{\text{Cog}\prime}(0) \geq \chi$, then naive consumers always give $m = 0$ (they "give up" on paying attention), and any $b \in [0, 1]$ gives the same social surplus. But a minuscule fraction of rational consumers (as in Proposition 32) shifts the social optimum to $b = 1$, strictly.

D.5 Proof of Proposition 14

The utility is $u(a, x) = -\frac{1}{2}(a - x)^2 + vx$. I consider the median agent, i.e. the agent who receives the median signal x^s (which is mx) when evaluating his "motivated cognition" part; but, with respect to the other consequences of his actions, the agent still behaves like the regular agent in this paper, i.e. we will get terms in σ_x^2 for the terms that do not involve motivated cognition, i.e. are not proportional to vx. This allows us to get a closed-form

solution, without losing much economics. Note that in the choice of attention m, the agent is allowed to get a "glimpse" of the value of x. This is a convenient modeling assumption, and is defensible: we have a rough sense of the value of x before we pay attention to it (this is the "ex post" attention determination discussed in Gabaix (2014)).[91] The utility of this median agent is:

$$\tilde{U}^{\text{active}}(m) = -B(1-m) - C^{\text{Cog}}(m) + \mu v m x - P, \qquad B = \frac{1}{2}\sigma_x^2 + \xi s_a \sigma_x \tag{150}$$

where $P = \mathbb{E}\left[u\left(a^d, x\right)\right] = -\frac{1}{2}\left(\left(a^d\right)^2 + \sigma_x^2\right)$ is the traditional expected utility from the passive action a^d—recall that U^{active} is measured net of it, so that $U^{\text{passive}} = 0$.

We have

$$\begin{aligned}
\tilde{U}^{\text{active}}(m) &= -B(1-m) - C^{\text{Cog}}(m) + \mu v m x - P \\
&= -B(1-m) - C^{\text{Cog}}(m) - \mu v x(1-m) + \mu v x - P \\
&= -\tilde{B}(1-m) - C^{\text{Cog}}(m) + \mu v x - P
\end{aligned}$$

with $\tilde{B} = B + \mu v x$. We observe that

$$\min_{m\in[0,1]}\left[\tilde{B}(1-m) + \kappa m\right] = \min\left(\tilde{B}, \kappa\right) \tag{151}$$

as indeed the optimal attention is $m = 1_{\kappa \leq \tilde{B}}$. So, optimizing on attention m, with $C^{\text{Cog}}(m) = \kappa m$, we get:

$$\begin{aligned}
\max_m \tilde{U}^{\text{active}}(m) &= -\min\left(\tilde{B}, \kappa\right) + \mu v x - P \\
&= -\min\left(\tilde{B} - \mu v x, \kappa - \mu v x\right) - P = -\min\left(B, \kappa - \mu v x\right) - P
\end{aligned}$$

Hence, we have a passive decision iff $\max_m \tilde{U}^{\text{active}}(m) \leq 0$, i.e. iff

$$\min\left(B, \kappa - \mu v x\right) \geq -P = \frac{1}{2}\left[\left(a^d\right)^2 + \sigma_x^2\right]$$

i.e.

$$\min\left(\xi s_a \sigma_x, \kappa - \frac{\sigma_x^2}{2} - \mu v x\right) \geq \frac{1}{2}\left(a^d\right)^2$$

[91] The reader who does not like this modeling assumption can set $v = 0$, or replace σ_x by $|x|$.

D.6 Proof of Proposition 16

Write $z(m) = \frac{\xi\sqrt{p}\,(1-m)}{\sqrt{m}}$, so that $P_0 = \Phi(z(m))^n$ by Proposition 15.

For the first part, fix n and $m < 1$: then $z(m) \to \infty$ as $\xi \to \infty$, so $\Phi(z(m)) \to 1$ and $P_0 \to 1$.

For the second part, fix $\xi > 0$ and let $z_n = z(m(n))$. Since $m(n)\ln n \to 0$ and $\ln n \to \infty$, we have $m(n) \to 0$, so for n large enough $(1-m(n))^2 \geq \frac{1}{2}$ and $z_n \geq 1$. Mills' inequality $1 - \Phi(z) \leq \varphi(z)/z \leq e^{-z^2/2}$ for $z \geq 1$ then gives

$$n\big(1-\Phi(z_n)\big) \;\leq\; \exp\Big(\ln n - \tfrac{1}{2}z_n^2\Big) \;\leq\; \exp\Big(\ln n - \frac{\xi^2 p}{4m(n)}\Big) \longrightarrow 0, \tag{152}$$

because $\frac{\xi^2 p}{4m(n)} \big/ \ln n = \frac{\xi^2 p}{4\,m(n)\ln n} \to \infty$. Finally, using $\ln(1-u) \geq -2u$ for $u \in [0, \frac{1}{2}]$,

$$0 \;\geq\; \ln P_0 = n\ln\big(1-(1-\Phi(z_n))\big) \;\geq\; -2n\big(1-\Phi(z_n)\big) \longrightarrow 0, \tag{153}$$

so $P_0 \to 1$.

The result is easy to extend with extreme value theory techniques, e.g. if the CDF of X_k^s had exponential rather than Gaussian tails, the condition would be the very slightly stronger $\lim_{n\to\infty} m(n)(\ln n)^2 = 0$.

D.7 Proof of Proposition 17

This is very simple. By Proposition 15, $P_0(n) = \Phi(z_n)^n$ with $z_n = \xi\sqrt{p}\,(1-m(n))/\sqrt{m(n)}$. As $m(1) = 1$, $z_1 = 0$ and $P_0(1) = \Phi(0) = \frac{1}{2}$. Hence $P_0(2) > P_0(1)$ iff $\Phi(z_2)^2 > \frac{1}{2}$, i.e. $\Phi(z_2) > 2^{-1/2}$, i.e. $z_2 > \Phi^{-1}(2^{-1/2})$, which is (49).

D.8 Proof of Proposition 21

Suppose now that we have endogenous attention, and call m^d the attention that the consumer chooses at $b = 0$. At $b = 0$ her surplus (27) is just $-C^{\mathrm{Cog}}(m)$: attention brings her nothing, so she takes the highest level that is available at no marginal cost. Hence $m^d = 0$ if attention is costly throughout, and m^d is the free default of (89) when there is one. I assume that attention is strictly costly beyond it, $C^{\mathrm{Cog}\prime}\left(m^{d+}\right) > 0$.

Then attention does not move around $b = 0$. Indeed, by (27) the marginal benefit of raising attention above m^d is $\frac{1}{2}\psi\sigma_x^2 b^2 + \xi s_a \sigma_x |b| = O(|b|)$, which is below $C^{\mathrm{Cog}\prime}\left(m^{d+}\right)$ for $|b|$ small enough. So $\frac{dm}{db} = 0$ at $b = 0$, and the total derivative of welfare reduces to its partial

derivative:

$$\frac{d}{db}W_{|b=0} = W_{b|b=0} = \psi\sigma_x^2 m^d - \xi s_a \sigma_x \operatorname{sign}(b)\left(1-m^d\right)$$

It is $\frac{dm}{db} = 0$, and not $W_m = 0$, that removes the attention-response term here: at $b = 0$ we have $W_m = -C^{\text{Cog}\prime}(m)$, which need not vanish. So locally, $\frac{d}{db}W_{|b=0} \leq 0$ iff the same condition (31), evaluated at $m = m^d$, holds. We note that a simple way to think about this is that the allocative benefits are $\sim bm$, while CA costs are ξb, so that around $b = 0$, rigidity is better if ξ is large enough.

The consumer's attention is: $\max_m V^{\text{cons}}(m)$. It is useful to define

$$V(m) \coloneqq V^{\text{cons}}(m) - \frac{1}{2}\psi\sigma_x^2 b^2$$

where $\frac{1}{2}\psi\sigma_x^2 b^2$ is the surplus with a rational agent. Hence, $V(m) \leq 0$ and it is easier to reason with it. We have

$$V(m) \coloneqq V^{\text{cons}}(m) - \frac{1}{2}\psi\sigma_x^2 b^2 = \frac{1}{2}\psi\sigma_x^2 b^2 m - \xi s_a |b| \sigma_x (1-m) - C^{\text{Cog}}(m) - \frac{1}{2}\psi\sigma_x^2 b^2$$

hence

$$V(m) = -B(1-m) - C^{\text{Cog}}(m), \qquad B \coloneqq \frac{1}{2}\psi\sigma_x^2 b^2 + \xi s_a |b| \sigma_x, \tag{154}$$

which gives $m^* = \operatorname{argmax}_m V(m)$. By (148) social welfare is: $W = V^{\text{cons}}(m) + (b - b^2)\psi\sigma_x^2 m$, i.e.

$$W = V(m) + \left(b - \frac{b^2}{2}\right)\psi\sigma_x^2 m \tag{155}$$

Suppose that $C^{\text{Cog}}(1) > \psi\sigma_x^2$ and ξ is large enough. Let us show that with endogenous attention, a perfectly simple contract is optimal.

When $b = 0$, welfare W is 0. So, a sufficient condition is: for all $b \in (0, 1]$ (later we will say "all b"), $W \leq 0$, i.e.

$$-V(m^*) \geq \left(b - \frac{b^2}{2}\right)\psi\sigma_x^2 m^*$$

A sufficient condition for that is: for all b

$$-V(m^*) \geq b\psi\sigma_x^2$$

i.e.

$$\min_m \left(\frac{1}{2}\psi\sigma_x^2 b^2 + \xi s_a b\sigma_x\right)(1-m) + C^{\text{Cog}}(m) \geq b\psi\sigma_x^2$$

i.e. (after division by b), for all b,

$$\min_m \left(\frac{1}{2}\psi\sigma_x^2 b + \xi s_a \sigma_x\right)(1-m) + \frac{C^{\text{Cog}}(m)}{b} \geq \psi\sigma_x^2$$

A sufficient condition for that is that

$$\xi s_a \sigma_x (1-m) + C^{\text{Cog}}(m) \geq \psi\sigma_x^2, \qquad \forall m \in [0,1]$$

i.e. $\xi \geq \xi^* \coloneqq \max_{m<1} \frac{\psi\sigma_x^2 - C^{\text{Cog}}(m)}{s_a\sigma_x(1-m)}$. Hence, for $m=1$, we need to have $C^{\text{Cog}}(1) > \psi\sigma_x^2$.

D.9 Proof of Proposition 22

If $b=0$, social welfare is $W(0)=0$. Now suppose that the firm sets $b=1$. Then, optimal attention allocation follows $\max_m Bm - C^{\text{Cog}}(m)$ with $B = \frac{1}{2}\psi\sigma_x^2 + \xi s_a \sigma_x$. So, by the proposition's assumption, we have $C^{\text{Cog}\prime}(1) < \frac{1}{2}\psi\sigma_x^2 \leq B$, so consumer attention is $m=1$. Hence, welfare is $W(1) = \frac{1}{2}\psi\sigma_x^2 - C^{\text{Cog}}(1)$. As $C^{\text{Cog}}(1) \leq C^{\text{Cog}\prime}(1)$ by convexity, we have $C^{\text{Cog}}(1) < \frac{1}{2}\psi\sigma_x^2$ and $W(1) > 0 = W(0)$. Choosing $b=1$ yields greater social welfare than $b=0$.

References for the Online Appendix

Bajari, Patrick and Steven Tadelis, "Incentives versus transaction costs: A theory of procurement contracts," *RAND Journal of Economics*, 2001, *32* (3), 387–407.

de Clippel, Geoffroy, Paola Moscariello, Pietro Ortoleva, and Kareen Rozen, "Caution in the Face of Complexity," Technical Report, Mimeo 2025.

Enke, Benjamin and Florian Zimmermann, "Correlation neglect in belief formation," *The Review of Economic Studies*, 2019, *86* (1), 313–332.

_ **and Thomas Graeber**, "Cognitive Uncertainty," *Quarterly Journal of Economics*, 2023, *138* (4), 2021–2067.

Frederick, Shane, "Cognitive reflection and decision making," *Journal of Economic perspectives*, 2005, *19* (4), 25–42.

Gabaix, Xavier, "A sparsity-based model of bounded rationality," *Quarterly Journal of Economics*, 2014, *129* (4), 1661–1710.

_ , "Behavioral Inattention," *Handbook of Behavioral Economics*, 2019, *2*, 261–344.

_ **and David Laibson**, "Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets," *Quarterly Journal of Economics*, 2006, *121* (2), 505–540.

_ **and Thomas Graeber**, "The Complexity of Economic Decisions," Working Paper 33109, NBER 2024.

Heidhues, Paul and Botond Kőszegi, "Behavioral industrial organization," *Handbook of Behavioral Economics: Applications and Foundations 1*, 2018, *1*, 517–612.

Puri, Indira, "Simplicity and risk," *The Journal of Finance*, 2025, *80* (2), 1029–1080.

Vazsonyi, Andrew, "Which door has the Cadillac?," *Decision Line*, 1999, *30* (1), 17–19.

Wilson, Charles, "A model of insurance markets with incomplete information," *Journal of Economic Theory*, 1977, *16*, 167–207.