A Theory of Complexity Aversion*

Xavier Gabaix

March 17, 2025. Preliminary and incomplete. Please do not circulate.

Abstract

I propose a tractable model of "complexity aversion". The key ingredient is "first order complexity aversion": when people know they're making a mistake (because the situation is complex) they experience some dread, which is a utility loss proportional to the absolute value of the expected error. I show how complexity aversion leads to optimally simple mechanisms. I illustrate this in five examples complexity aversion makes a large difference. (i) If complexity aversion is high enough, the price of a good will be constant over time, even though the marginal cost might be variable, to avoid annoying the consumer with a complex price system. (ii) In the theory of optimal taxation, if complexity aversion is high enough, the optimal tax system is "simple", e.g. just features a uniform tax rate rather than a different tax rate for each good, as recommended by the traditional Ramsey model. (iii) Whereas the traditional model predicts that contracts should be indexed aggregate factors (e.g. on inflation, GDP, or the stock market), with enough complexity aversion, contracts are non-indexed, "simple". (iv) Complexity aversion leads to a model of a non-traditional (first order) cost of inflation, which in calibration is quite important: as different sources of income do not react equally to inflation, higher inflation leads to a more complex planning process. (v) This in turn changes optimal monetary policy, which will de facto target a zero inflation (or, more generally, zero deviation from the inflation target), to the exclusion of other goals, except in rare extreme circumstances such as an extreme recession. I finally discuss how using this model of complexity aversion will lead to a useful "behavioral mechanism design" theory, and more realistic—simpler—mechanisms.

*xgabaix@fas.harvard.edu. For helpful comments, I thank Ben Enke, Ed Glaeser, Thomas Graeber, Brian Hall, Olivier Hart, Louis Kaplow, David Laibson, Shengwu Li, Florencia Marotta-Wurgler, Eric Maskin, Emi Nakamura, Ricardo Reis, Josh Schwarzstein, Kathy Spiers, Jeremy Stein, Jón Steinsson, Dmitry Taubinsky, Dhara Yu, and seminar participants at various seminars. I gratefully acknowledge financial support from Ferrante Economics Research Fund.

1 Introduction

I propose a tractable notion of "complexity aversion", and explore its implication for "simple mechanisms", in several settings: simple price systems, simple taxes, simple contracts (that are not indexed), and the cost of inflation as increasing choice complexity, and the resultant optimal monetary policy.

Prices systems are typically much "simpler" than a frictionless rational model would predict. 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.[1] Likewise, the barber typically has prices that don't depend on the time of day, even though the 5pm slot is much more popular than the 2pm slot, so that (generically at least) their prices should differ.

I show how complexity aversion leads to optimally simple mechanisms. For instance, 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. Two key properties yields full simplicity.

First, the consumer 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 cost, 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 his dread. This is a sense in which cognitive costs, and complexity dread are distinct, though they interact with each other: cognitive costs generate mistakes, which in terms generate complexity aversion.

Second, this complexity aversion cost is (in the main specification) first order in the size of the mistake rather than second order: if the expected size of the mistake is e (in absolute value), the dread cost is proportional to e, rather than e^2. This "first order complexity aversion" is a close cousin to the "first order risk aversion" found necessary to explain the behavior under risk (Kahneman and Tversky (1979), Rabin (2000)). Traditional model

[1] 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.

ambiguity aversion would generate something akin to second order complexity aversion. But to robustly generate simplicity at the optimum, we need first order complexity aversion. Why? This is because making mechanism more "rationally efficient" by a small factor b will create extra allocative gains proportional to b, typically. With first order complexity aversion, the dread costs will also be proportional to b, and so they might overcome the "rational benefits" of order b. However, if the complexity aversion is proportional to b^2, then it will be much lower than the the first order rational allocative gains (of order b); so at the optimum, we shall want to have b away from 0: the mechanism will be a complex, as b is not 0. So second order complexity aversion does not (most of the time) generate fully simple mechanism. The logic is actually more rich, as the size of the mistakes, hence the dread, depends on attention which is endogenous, but this is the essence of the reasoning.

Complexity aversion creates "aversion to deciding". If an "active decision" (e.g., rebalance one's portfolio) entails complexity aversion, while a "passive course" (e.g. do nothing), when complexity aversion is high enough, and the decision is hard enough, people will stay passive.[2]

As a second example of "simple mechanisms", this paper revisits 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 appears 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.

A third application is aversion to indexation. In the basic economic theory, most contracts should be indexed in aggregate risks, such as inflation, GDP, or the stock market (see e.g. Holmström (1979)), as that reduces the risk faced by consumers. For instance, debt contracts ought to be indexed on inflation, and GDP; mortgage contracts "ought" to be indexed on inflation, or on GDP. But that's typically not the case in practice. Intuitively, this is because "this would be too complicated". The model formalizes that, and, sensibly, delivers that contracts will generally not be indexed, except if the risk-reduction effects are large enough.

A fourth application is to the cost of inflation. One often mentioned cost is that inflation makes things complicated, as people need to juggle between prices that follow inflation (such as the price of most goods), and other streams that do follow it only infrequently (e.g. the wage, being adjusted to inflation only once a year, or a fixed-interest rate mortgage, that is

[2] Years ago, I describe this to Robert Lucas. He told me that this what he did—he had never rebalanced his retirement portfolio, and just let it drift.

not indexed). The theory leads to first order losses of inflation $|\pi|$, rather than, as in the traditional model (Woodford (2003)), second order losses π^2. In calibration, this is much bigger, and indeed much more in line with people's intuition.

The fifth application is to monetary policy. We just saw that complexity aversion leads to "first order" rather than "second order" losses from inflation, and, under some conditions, of the output gap. Hence, the central bank's objective function has terms in $|\pi_t|$ rather than π_t^2. This implies that exactly zero inflation is optimal in a large range of circumstances, even when this is at the cost of an output gap (this is in the spirit of the Lasso).

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 (i.e. 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.

The reader might be a bit bewildered by all those examples. Are they really needed? I found it useful to have several examples, for two 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. So it is useful to analyze simplicity across a domain of situations.

Related literature There is relatively little work on explaining why contracts are simple. One paper is Tirole (2009). Another literature, on the lack of indexation, emphasizes asymmetric information: the person who proposes to index the contract may have more information than the one would has to accept or reject the indexation (Spier (1992), Hartman-Glaser and Hébert (2020)). 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 worse-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 of complexity. Some are conceptual 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 sort of model is meant to be reasonable compromise between naive bounded rationality and fully Bayesian information processing. A very behavioral approach, where attention is insensitive to cost-benefit analysis, is developed in a series of papers by Bordalo et al. (2013). 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), Woodford (2020), Khaw et al. (2021). The analysis we propose here could be applied to those styles of models. A cumulative body of papers show inattention (in particular in its behavioral flavor) is a good candidate for one unifying force in behavioral economics (see Gabaix (2019) for a manifesto, and Enke et al. (2024) for new systematic evidence).

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 (Abeler and Jäger (2015), Martínez-Marquina et al. (2019), Oprea (2020), Dean and Neligh (2023), Enke et al. (2023)). Huck and Weizsäcker (1999), Sonsino et al. (2002), Fudenberg and Puri (2023), Enke and Shubatt (2023), Puri (2024) find lotteries are less tempting when they have more outcomes (so are more "complex"). Oprea (2024b) finds that this behavior may be linked to complexity rather risk per se. 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.

It is clear intuitively a high complexity is a key determinants in mistakes, and this has been directly or indirectly shown in a number of papers, in all parts of economics. For instance, in macro, losses from lack of consumption planning can be very high (Augenblick et al. (2024)), and where the non-indexation of debt contract amplify financial crises (Gertler and Karadi (2011)).[3] 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)). 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 (orth); Ben-David and Chinco (2024)). This paper could be useful in those parts of economics, as a way model the taste for simplicity, and its equilibrium consequences.

Outline The paper first lays out the basic model, then explores successively simple prices, the fact that people avoid deciding at all, simple tax systems, aversion to indexation, the complexity costs of inflation, and optimal monetary policy. Most proofs are in the appendix.

[3] If debt contracts were indexed on aggregate economic activity, typically contracts would automatically channel funds to financial firms in crises, as in contingent convertible bonds—but we scarcely see that in practice, arguably because investors want simple, standard bonds.

2 How complexity aversion leads to simple mechanisms

2.1 Motivation: many institutions are "simple"

Many institutions 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.[4] This sort of situation is extremely prevalent: for instance, restaurant typically (not always) charge the same price for meals are different hours of the evening; bars do not implement "congestion pricing".

I show how this model of complexity does generate "complexity aversion", which leads to "simple mechanisms" – for instance rigid prices.

2.2 Model

I first review an existing model of thinking cost, then add the key new ingredient: complexity aversion.

2.2.1 A model of thinking cost: Background material

I first review model laid out in Gabaix (2014) and Gabaix and Graeber (2024), that does not feature complexity aversion yet. The task is to maximize over a continuous action a an objective function $u(a,x)$, where u is smooth (three times continuously differentiable) and concave in a. The vector of disturbances x is drawn from a distribution with mean normalized to 0, and its component 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 default action a^d the optimal action "at the default", i.e. when all x are equal to 0, $a^d = \text{argmax}_a\, u(a,0)$, and I normalize it to 0; a_{x_i} is the partial derivative at a default point. The rational answer is thus (after linearization, so up to second order terms in x_i, assuming that the deviations x_i are "small")

$$a^r = \sum_i a_{x_i} x_i = \sum_i y_i, \qquad y_i := a_{x_i} x_i \tag{1}$$

The agent's objective is to cognitively construct those $y_i := a_{x_i} x_i$, which indicates by how much dimension i of the problem should change the rational action.

[4] 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 modelling would be similar as the one we proposed here, and might apply to the situations documented in DellaVigna and Gentzkow (2019). It couldn't 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.

People receive noisy signals y_i^s about y_i:[5]

$$y_i^s = m_i y_i + (1 - m_i) y_i^d + \sqrt{m_i (1 - m_i)} \varepsilon_i \tag{2}$$

where $m_i \in [0, 1]$ is the precision of the signal, y_i^d is a default value, equal to 0 when the mean of x_i is 0, and ε_i a mean-zero noise with variance $\sigma_{\varepsilon_i}^2 = \sigma_{y_i}^2$. If all the shocks are jointly Gaussian, with the prior of y_i equal to y_i^d, we have $\mathbb{E}[y_i | y_i^s] = y_i^s$. The rational case corresponds to $m_i = 1$, and the extremely inattentive case, to $m_i = 0$.[6] Accordingly, I posit that if the decision maker sees y_i^s, she takes the decision

$$a = \sum_i y_i^s.$$

The traditional utility losses from an imperfect decision are then:

$$\mathbb{E}[u(a^s, x)] - \mathbb{E}[u(a^r, x)] = \frac{1}{2} \left| u'' \left(a^d \right) \right| \sum_i \sigma_{y_i}^2 (1 - m_i) \tag{3}$$

Production function of thought

Typically, in calculations, we will assume a production function of thought (or precision) as follows: an effort L_i leads to a precision

$$m_i (L_i) = \min \left(\left(\frac{L_i}{c_i} \right)^{1-\alpha}, 1 \right), \qquad \alpha \in (0, 1) \tag{4}$$

where $\alpha \in (0, 1)$ and c_i parametrizes the basic "cost" of thinking, and is taken here as a primitive (it can be endogenized, as in Gabaix and Graeber (2024)). The cognitive cost $C_i (m_i)$ is then $w L_i$, where L_i is the effort needed to reach precision m_i according to (4), and w is a cost of effort, which can be taken as a constant in applications, so is

$$C_i (m_i) = w c_i m_i^{\frac{1}{1-\alpha}}, \qquad 0 < \alpha < 1 \tag{5}$$

We could also have the production function of thought with $\alpha > 1$:

$$m_i (L_i) = \max \left(1 - \left(\frac{L_i}{c_i} + 1 \right)^{1-\alpha}, 0 \right), \qquad \alpha > 1 \tag{6}$$

[5] The noisiness is not central — the fact that signals are imperfect is.

[6] 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 $\mathbb{E}[y_1 | y_1^s, y_2^s]$ to infer y_1, but instead we model the agent as being a "limited Bayesian", who simply performs $\mathbb{E}[y_1 | y_1^s]$.

This model induces sparsity, i.e. when benefits are small enough, the optimal effort and attention are exactly 0. Then, the cost is

$$C_i(m_i) = wc_i\left((1-m_i)^{\frac{1}{1-\alpha}} - 1\right), \qquad \alpha > 1 \tag{7}$$

Yet another tractable format is a linear-quadratic cost of cognition.

The cost $C(m)$ is convex in m, $C(0) = 0$, and we normalize $w = 1$.

2.2.2 Complexity aversion: The innovation in this paper

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

$$V(m) = \mathbb{E}[u(a^s, x)] - C(m) - C^{\text{CA}}(m) \tag{8}$$

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 $C(m)$ is a cognitive cost function, which is the effort needed to achieve precision m (which comes from (4)).

The key new term in this paper is the "complexity aversion" term $C^{\text{CA}}(m)$. The leading version is *first order* complexity aversion:

$$C^{\text{CA},1} = \xi \sum_i \mathbb{E}\left[\left|\sigma_a u_{aa} a^r_{x_i} x_i\right|^2\right]^{1/2} (1-m_i) \tag{9}$$

where $\xi \geq 0$ is a "complexity aversion" parameter, which is unitless, σ_a is the natural scale of a, explained below. The first order penalty is proportional to σ_{x_i} rather than $\sigma^2_{x_i}$. This is in the spirit of first order risk aversion (where the risk penalty is proportional to the standard deviation), which has been found many times to describe people better than rational model's prediction of second order risk aversion (where the risk penalty is proportional to the squared standard deviation), see Kahneman and Tversky (1979), Rabin (2000).

The crucial part in (9) is that, when $m = 0$ (when the agent doesn't think), the CA term is strictly positive (even though the cognition cost is $C(0) = 0$). To build intuition, consider a very boundedly rational agent that has $m = 0$. Then, the cognition cost is 0, but the dread cost is strictly positive, as the agent is conscious enough that he is making a mistake (as in Enke and Graeber (2023)). Hence, there is an incentive for the planner to create a "simple scheme", that will not create too much "dread", i.e. complexity aversion.

On the other hand, when the consumer understands fully the situation (so $m_i = 1$), there is no room mistake, and the dread cost becomes 0 in (9).

A variant that study in the appendix (and then ultimately discard) is "second order complexity aversion", which which is proportional to the expected loss from mis-optimization:

$$C^{\mathrm{CA},2}=\xi\mathbb{E}\left[u\left(a^{r},x\right)-u\left(a^{s},x\right)\right]=\xi\sum_{i}\mathbb{E}\left[-\frac{1}{2}u_{aa}\left(a_{x_i}^{r}x_i\right)^{2}\right]\left(1-m_i\right) \tag{10}$$

It is in the tradition of regret, e.g. Loomes and Sugden (1982), Sarver (2008). It is more natural given the tradition, but ultimately I will discard, as it typically does not generate full simplicity, e.g. rigid prices, or no indexation.

The key difference between "complexity aversion" (proposed here) and "regret aversion" (from the tradition) are:

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)

2. CA applies to the mistakes made by the agents in their reasoning, not to the ex post regret because of nature's stochasticity. Hence, it does not change the utility of a rational agent, only of one that makes mistakes.

3. CA applies features 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, so the loss is $\xi\sum_i w_i\left(1-m_i\right)$ for some weight w_i, rather than in a "unified manner". To see why this matters, consider the alternative, e.g. having e.g. $CA'=\xi\sqrt{b_i^2\sigma_{x_i}^2\left(1-m_i\right)}$. Then, the partial impact of each b_i in b_i^2,not $|b_i|$: so de facto we have second order complexity aversion at the feature level, and we will not get full simplicity.

Multidimensional actions The following is advanced material, and best skipped at a first reading. If $\tilde{a}$ represents the randomness of the action across situations, the complexity aversion is:

$$C^{\mathrm{CA},1}=\xi\sum_{i}\left(\mathbb{E}\left|\tilde{a}u_{aa}a_{x_i}x_i\right|^{2}\right)^{1/2}\left(1-m_i\right) \tag{11}$$

If there are several dates, and utility is separable, it may be useful to sum separately across periods:

$$C^{\mathrm{CA},1}=\xi\sum_{i,t}\left(\mathbb{E}\left[\left|\tilde{a}_t u_{a_t a_t}a_{t,x_i}x_i\right|^{2}\right]\right)^{1/2}\left(1-m_i\right). \tag{12}$$

3 A simple example: rigid pricing schemes

I present in detail the application of the above general model to a simple situation: the pricing of one good. This is a very simple context, that will provide many of the key economic lessons, valid in later applications.

3.1 Setup and basic analysis

A competitive firm (maximizing social welfare) sells to a behavioral consumer, who can buy a continuous quantity of its good.[7] The 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 unit is is $p(x) = 1 + bx$, where b is chosen by the benevolent firm, 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 firm might choose another b. For instance, if $b = 0$, the price is constant over time, which "simplifies" the consumer's choice.

The consumer chooses to consume a quantity $C = 1 + a$, where a will be a deviation of consumption from a value of 1, and will have a mean 0. Her total utility, net of the cost of good, is:

$$u(a, x) = U(1 + a) - p(x)(1 + a) \tag{13}$$

where the gross utility from consuming a quantity $1 + a$ of the good is:

$$U(1 + a) = 1 + a - \frac{1}{2\psi}a^2 \tag{14}$$

where $\psi > 0$ is the elasticity of demand (when the price is centered at 1). The rational action is $a^r = \operatorname{argmax}_a u(a, x)$ i.e.

$$a^r = -\psi bx \tag{15}$$

3.1.1 The behavioral consumer's utility

Imperfect cognition The consumer perceives the price as $p^s = 1 + bx^s$ with

$$x^s = mx + \sqrt{m(1 - m)}\varepsilon \tag{16}$$

with $\sigma_\varepsilon = \sigma_x$ and $m \in [0, 1]$ is the precision of the perception (exo- or endogenous). She takes the action $a = -\psi bx^s$, see (15). The utility gains from taking the right action (i.e. $a^r = -\psi bx$) is:

$$\chi = -\frac{1}{2}u_{aa}\mathbb{E}\left[(a_x^r x)^2\right] = \frac{1}{2}\psi\sigma_x^2 b^2 \tag{17}$$

[7]Here, I take the benchmark of firms forced to maximize social welfare by competitive forces, hence to be "benevolent". Section A.1 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 (2018a)), because there is a "shrouded attribute" (e.g. the spare part to be bought later) in addition to a "visible base good" (the car). Considering that those issues are now well-understood, I focus complexity with a benevolent firms. Of course, complexity aversion with an exploiting firm would be an interesting next step.

Traditional consumption utility is:[8]

$$U^{\text{cons}} = \mathbb{E}\left[u\left(a,x\right)\right] = \chi m$$

so is increasing in attention.[9]

Complexity aversion With first order complexity aversion (9), we have the dread cost, of complexity aversion cost:

$$C^{\text{CA},1} = \xi \mathbb{E}\left[\left(\sigma_a u_{aa} a_x^r x\right)\right]^{1/2}\left(1-m\right) = \xi \sigma_a \frac{1}{\psi}\psi\left|b\right|\sigma_x\left(1-m\right)$$

hence

$$C^{\text{CA},1} = \xi \sigma_a \sigma_x \left|b\right|\left(1-m\right) \tag{18}$$

The crucial part is that, when $m=0$ (when the agent doesn't think), the dread term $C^{\text{CA},1}$ is strictly positive (even though cognition cost is zero, $C\left(0\right)=0$). Specifically, it is proportional to the expected absolute mistake from not thinking.[10] Also, in (18), if the consumer makes no mistake, $m=1$, the dread term is 0: there is no cognitive dread when all is well-understood.

With second order CA, the CA penalty (10) is the expected loss from mis-optimization, times the coefficient of complexity aversion ξ:

$$C^{\text{CA},2} = \xi \frac{1}{2}\psi \sigma_x^2 b^2 \left(1-m\right) \tag{19}$$

The interpretation is similar to that of the first order CA (18, but with a term proportional to b^2 rather than $|b|$.

The consumer's full utility is, from (8):

$$V^{\text{cons}}\left(m\right) = \chi m - C\left(m\right) - C^{\text{CA}}\left(m\right). \tag{20}$$

[8]The derivation is:

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

[9]It may be surprising at first that a more volatile price makes the consumer better off, even though the consumer is risk-averse (U^{cons} increases with $\sigma_x^2 b^2$). This is always true with quasi-linear utility $u\left(c_0,\ldots,c_n\right) = c_0 + U\left(c_1,\ldots,c_n\right)$, with U concave, which is the case here. This is not true for general non-quasi linear 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.

[10]To anticipate why this is important, consider a very cognitively limited consumer that exerts no cognition in equilibrium ($m=0$): the dread terms increases the deadweight loss from the mechanism, as it makes the agent "feel anxious", in addition to making him lose utils from poor decision-making.

In the case of endogenous attention, I posit that the consumer optimizes cognition, i.e. does

$$\max_m V^{\text{cons}}(m). \tag{21}$$

3.1.2 Social welfare

Social welfare is:

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

i,.e. the consumer's full utility (including dread and cognition cost), plus profits from consumption $1 + a(x^s)$, which also reflects that the social cost of the good is c, not p:

So, the social planner's (or benevolent firm's) wants to maximize total social surplus, while taking into account that the variability of prices (indexed by b) affects the endogenous attention m. Mathematically:

$$\max_b W(m,b) \text{ subject to } m \in \operatorname*{argmax}_m V^{\text{cons}}(m,b) \tag{22}$$

where $V^{\text{cons}}(m,b)$ is (20). We next calculate that social welfare.

Lemma 1. *Social welfare is*

$$W(m,b) = \left(b - \frac{b^2}{2}\right)\psi\sigma_x^2 m - C^{CA} - C(m) \tag{23}$$

We verify that when there is no cognition cost nor complexity aversion ($m = 1$, $\xi = 0$), social welfare is $W = \left(b - \frac{b^2}{2}\right)\psi\sigma_x^2$, so is maximized when $b = 1$, i.e. when the consumer pays the marginal cost.

3.2 When is the optimal price system "simple"?

We now arrive at the meat of the analysis. We can derive the optimum price complexity, particularly focusing on when the price is "perfectly simple" and constant ($b = 0$).

First order complexity aversion reliably generates simplicity Let us start with a transparent case, with exogenous attention m.

Proposition 1. (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:*

$$b = \max\left(1 - \frac{\xi\sigma_a\sigma_x(1-m) - W_m\frac{dm}{db}}{\psi\sigma_x^2 m}, 0\right) \tag{24}$$

So, with exogenous attention ($\frac{dm}{db} = 0$), the optimum entails rigid prices if and only if:

$$\xi \sigma_a \sigma_x (1 - m) \geq \psi \sigma_x^2 m \tag{25}$$

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

Proof. With first order CA,

$$W = \left(b - \frac{1}{2} b^2\right) \psi \sigma_x^2 m - \xi \sigma_a \sigma_x |b| (1 - m) - C(m)$$

So,

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

Solving for b gives the announced expression (24). □

Proposition 1, especially (25) also shows that contracts are more likely to be "simple" ($b = 0$) when (i) complexity aversion ξ is high; (ii) when 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)); when (iii) volatility of the demand would be low ($\psi \sigma_x$ low) even under a frictionless model; (iv) when the degree of understanding is low (low m).

The important point is that we get a simple price system if the cost of cognition 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 goods is much more elastic (e.g. raspberries vs blueberries, or raspberries at Whole Foods vs another store), and the marginal cost of perishable goods is very variable, depending on crops for instance. Likewise, in very competitive markets (e.g. airlines), the elasticity of demand is very high, and as a result the schemes are "complex".

The role of complexity aversion cost The next proposition shows that the CA is essential to obtain simple mechanism. Without CA, we do not get outright simplicity ($b = 0$), and instead we get $b = 1$ as the optimum.

Proposition 2. (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.*

We conclude that, with first order CA, a high CA parameter ξ generates simplicity, $b = 0$. Let us next verify that this works with endogenous attention.

Deepening the role of the thinking cost with endogenous attention The following gives sufficient conditions for full simplicity.

Proposition 3. (Optimal price simplicity, first order complexity aversion, endogenous attention) *Suppose first order complexity aversion, and endogenous attention m. 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 (25) holds. For a perfectly simple contract to be* globally *optimal, the following are sufficient conditions: $C(1) > \psi\sigma_x^2$ (reaching full precision is very costly), and the complexity aversion cost ξ is large enough: $\xi \geq \xi^*$, with:*

$$\xi^* := \max_m \frac{\psi\sigma_x^2 - C(m)}{\sigma_a\sigma_x(1-m)} \tag{26}$$

The intuition is the following: if it's 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 (18) 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 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 are low enough, compared to the social surplus gains of having a complex contract.

Proposition 4. (When full rationality is easy enough, contracts shouldn't be maximally simple*) Suppose that $C'(1) < \frac{1}{2}\psi\sigma_x^2$. 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 $C'(1)$ is less than the surplus generated at the optimum, which is $\frac{1}{2}\psi\sigma_x^2$. Hence, some complexity is important if the task is "easy" ($C'(1)$ is low, which implies that $C(1)$ is low, as by convexity $C(1) \leq C'(1)$), or a "very important" (it yields a high social surplus $\frac{1}{2}\psi\sigma_x^2$). Hence, we have simplicity 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 1 and 3). At the same time, complexity aversion is necessary to account for full simplicity (Proposition 2).

The appendix A.2 analyzes second order complexity aversion, and finds that it can sometimes generates simplicity, but in a non-robust way.

4 When people prefer not to decide

Very often, people prefer not to decide, at all. This typically is linked to hyperbolic discounting (Laibson (1997)), as the thinking cost is paid now and the benefits are received later. Another reason, is that the act of choosing is itself aversive (see in particular the evidence in Bernheim et al. (2024)), or that it is about loss aversion (Kőszegi and Rabin (2009), Andries and Haddad (2020)), as choice involves looking up information, which can lead to bad news. Complexity aversion offers another a natural hypothesis for this: thinking is aversive.

Suppose that the active decision is the traditional utility:

$$U^{\text{active}} = \mathbb{E}\left[u\left(a^s\left(x^s\right), x\right)\right] - C^{\text{cognition}}\left(m\right) - C^{\text{CA}}\left(m\right) \tag{27}$$

where a^s is the noisy decision given the imperfect perception x^s. But there is a "passive" decision a^d, evaluated as:[11]

$$U^{\text{passive}} = \mathbb{E}\left[u\left(a^d, x\right)\right] \tag{28}$$

Hence, we can hypothesize that people will stay passive, and do nothing, when $U^{\text{passive}} \geq U^{\text{active}}$. The main psychological difference is that the passive decision does not bear any complexity aversion.

To complete this simple psychology, we can add "motivated cognition", e.g. "ostrich effect" (Karlsson et al. (2009); Sicherman et al. (2016)), with[12]

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

where $\mu > 0$ is weight on motivated attention. This means that reacting to "bad news" (news x_i that lower utility if $u_{x_i} x_i^s < 0$) is penalized.

To see the effects, take the model $\tilde{U}\left(a, x\right) = -\frac{1}{2}\left(a - x\right)^2 + vx$, so that the passive decision at default action a is $U^{\text{passive}} = -\frac{1}{2}\left[\left(a^d\right)^2 + \sigma_x^2\right]$. Take the case $C\left(m\right) = \kappa m$ for simplicity.

Proposition 5. (When the agent fails to act, and passively follows the default) *The agent takes passive course of action ($\tilde{U}^{passive} \geq \max_m U^{active}\left(m\right)$) if and only if:*

$$\min\left(\xi \sigma_a \sigma_x, \kappa - \frac{\sigma_x^2}{2} - \mu v x\right) \geq \frac{1}{2}\left(a^d\right)^2 \tag{29}$$

Condition (29) 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" ($vx < 0$ high), and if the passive allocation isn't too far off the optimum ($\left|a^d\right|$ small). If there is no complexity aversion, the

[11] An interesting variant would be: $U^{\text{passive}} = \mathbb{E}\left[u\left(a^d, x^d\right)\right]$, which gives an even higher utility.

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

whole effect disappears.[13]

Take the lack of rebalancing of one's portfolio. The active decision involves thinking about the correct determinants for portfolio choice, a complicated vector x, involving e.g. the expected returns of US vs International stocks, or long vs short term bonds. Hence, many people just choose not rebalance at all.[14] They will do so only when the portfolio is very off, and when they have a no complexity aversion.

5 Simple tax systems: Ramsey problem with optimally simple taxes

The Ramsey (1927) taxation problem is the most basic taxation problem, and still leads to complex optimal taxes: generically, each good has its own specific tax, proportional to the inverse elasticity of demand. This is even true when agents are behavioral, in the sense that they misperceive the tax, but are not complexity averse (Mullainathan et al. (2012), Farhi and Gabaix (2020), Rees-Jones and Taubinsky (2020)). Let us see how complexity aversion lead to an optimal simple tax system.

5.1 Traditional, rational version

We 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 the a residual good with price 1 (which is utility from residual income), untaxed. Hence the rational demand is $c_g = 1 - \psi_g(p_g - 1)$, so that ψ_g is a 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 $\tilde{p}_g := p_g(1+\tau_g)$. 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:

$$W = U(c) + (1+\lambda)\sum_g \tau_g c_g \tag{30}$$

where $\lambda > 0$ is the excess marginal value of public funds, which are the collected taxes, $\sum_g \tau_g c_g$.

[13] In this model with linear cost, if the agent is active, then $m = 1$, so that she fully optimizes. With costly attention, then even under "active" decision, the attention wouldn't be full.

[14] Calvet et al. (2009) finds some rebalancing at the level of individual securities. Gabaix et al. (2025) finds very low rebalancing at the of aggregate risk-taking.

Following the tradition, we take the limit of a small λ, so that taxes are small. Taking a Taylor expansion up to second order terms, rational utility is

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

where W^0 is the utility with 0 taxes. With rational consumers, optimizing over τ_g gives the traditional Ramsey (1927) rule

$$\tau_g = \frac{\lambda}{\psi_g} \tag{31}$$

Each good has a different tax: the tax system is very "complex" in that sense.

5.2 Complexity averse agents

The tax is: $\tau_g = \bar{\tau} + \hat{\tau}_g$, where $\bar{\tau}$ is the average tax and $\hat{\tau}_g$ is a deviation. I assume that the consumer sees a tax

$$\tau_g^s = \bar{\tau} + \hat{\tau}_g^s$$

where the subjective perception $\hat{\tau}_g^s$ is an inattentive version of the true $\hat{\tau}_g$: the consumer sees average tax $\bar{\tau}$ perfectly (the idea is that if the tax is on average 10%, consumers will know it and take it into account), but sees the deviation $\hat{\tau}_g$ only partially.

The limited capacity to react to taxes is intuitive. Chetty et al. (2009), Taubinsky and Rees-Jones (2017) and Rees-Jones and Taubinsky (2020) provide compelling experimental evidence that people do not react optimally to taxes, and indeed typically under-react to them, consistent with limited attention or processing capacity. 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.

This model applies directly to the US, where it's customary for firms to show the pre-tax price p_g, but not the price inclusive of the tax, $\tilde{p}_g := p_g (1 + \tau_g)$. 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, intermediary production is not distorted, so that firms do not pay the VAT on intermediary inputs. Then, the model still applies, but the complexity is for firms, not consumers. However, the economics are the same.[15]

[15]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 intermediary inputs X_g^h of good g, $\max_X p_h F^h \left(X_g^h \right) - \sum_g \frac{\tilde{p}_h}{1+\tau_g} X_g^h$. So, it's the firm that now faces the complexity cost of knowing each τ_g.

The consumer's demand is:

$$c_g = 1 - \psi_g \tau_g^s \tag{32}$$

Calling m_g the attention to the special tax $\hat{\tau}_g$ to good g, the (first order) complexity aversion cost associated with good g is:

$$C_g^{\mathrm{CA}} = \xi \sigma_g \left|\hat{\tau}_g\right| (1 - m_g) \tag{33}$$

where σ_g is the typical variation in the consumption of good g. This is derived as in (18). Social welfare is:

$$W = \mathbb{E}\left[\sum_g u^g (c_g)\right] - C_g^{\mathrm{CA}} - C_g + (1+\lambda) \sum_g \tau_g c_g \tag{34}$$

i.e. the consumer's utility (inclusive of the complexity aversion, and the cognition cost C_g^{CA}, plus the government revenues (with a weight $1+\lambda$, where $\lambda \geq 0$).

Lemma 2. (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:*

$$W \quad = W^0 + \quad \sum_g \left[\frac{-1}{2}\psi_g \mathbb{E}\left[\left(\bar{\tau} + \check{\tau}_g^s\right)^2\right] - \xi \sigma_{c_g} \left|\check{\tau}_g\right| (1 - m_g) - C_g + \lambda \left(\bar{\tau} + \check{\tau}_g\right)\right] \tag{35}$$

where W^0 social when all taxes are zero. Also,

$$\mathbb{E}\left[\left(\bar{\tau} + \check{\tau}_g^s\right)^2\right] = \bar{\tau}^2 + m_g \left(2\bar{\tau}\check{\tau}_g + \check{\tau}_g^2\right)$$

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

If all consumers are rational ($\xi = 0$, $\check{\tau}_g^s = \check{\tau}_g$), optimizing social welfare (35), we obtain $\tau_g = \frac{\lambda}{\psi_g}$, the traditional Ramsey inverse elasticity rule: the tax system is complex, as each good has a different price.

With complexity-averse consumers, we obtain the following very different result (the proof is in the Appendix).

Proposition 6. (Optimal tax simplicity, with a concave cost of cognition) *Suppose first order complexity aversion, and exogenous attention. Define the average elasticity $\bar{\psi} = \frac{1}{G}\sum_g \psi_g$.*

As decision-makers in firms are people, then if prices inclusive of taxes are shown, those people will pay the complexity cost instead of the consumer.Their attention may be higher than for consumers, but (at least for non-giant firms), likely to be imperfect.

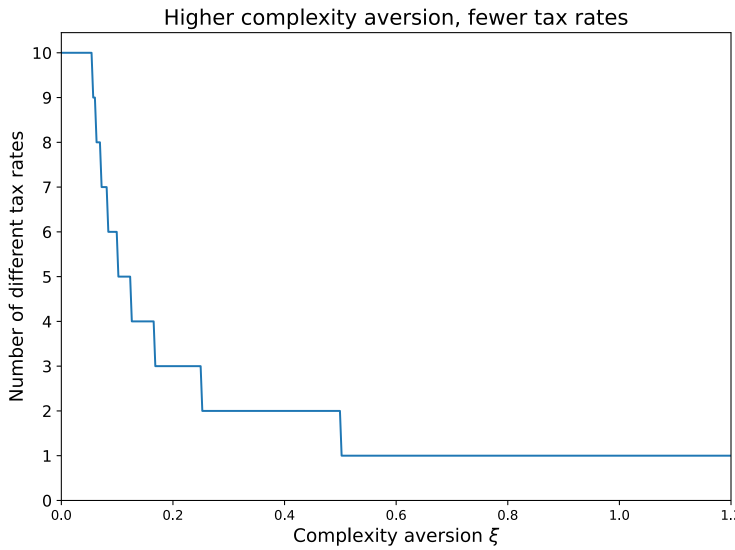

Figure 1: As the complexity aversion increases, the optimal number of different tax rates decreases. *Notes*. In this illustration there are $N = 10$ goods. When complexity aversion is very small, there are 10 different tax rates, as in Ramsey. However, when it is large enough, there is just one uniform tax rate.

Suppose for all goods, the average attention to the tax is imperfect, $m_g < 1$. Then if complexity aversion $\xi > 0$ is large enough, then 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 for this simple tax system is that ξ is large enough to satisfy is:

$$\max_g \lambda m_g \frac{\psi_g}{\bar{\psi}} - \xi \sigma_g (1 - m_g) \leq \min_g \lambda m_g \frac{\psi_g}{\bar{\psi}} + \xi \sigma_g (1 - m_g) \tag{36}$$

When m_g and σ_g are the same across goods, this necessary and sufficient condition becomes:

$$\lambda \frac{\max_g \psi_g - \min_g \psi_g}{\bar{\psi}} \leq 2\xi\sigma \frac{1-m}{m} \tag{37}$$

Condition (37) means that we have 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 (36) has the same message, in a more general way.[16]

This makes intuitive sense: when cognition costs are very high, it's best to have a uniform tax. However, finding the conditions (namely, first order complexity aversion) for that is not trivial.[17]

[16] There is a value $\xi_* \geq 0$ such that (36) holds if and only if $\xi \geq \xi_*$. Indeed, observe that the difference between the right hand size of (36) and its left-hand size increases in ξ, and goes to ξ as $\xi \to \infty$.

[17] We started Farhi and Gabaix (2020) hoping to obtain simple tax systems, but we didn't. Optimal taxes

Extending the proposition, one could imagine that as the complexity aversion cost $\xi > 0$ becomes lower, a more and more complex tax system is warranted. Indeed, it is clear that more generally, when complexity aversion ξ increases, there are fewer tax buckets. The situation will like Figure 1.

More generally, this approach yields a principled way to do "behavioral mechanism design", i.e. design optimal institutions (i.e. price or tax systems) that aim both at high allocative efficiency and not too high complexity (see Gonczarowski et al. (2023) for advances in behavioral mechanism design). This is a fruitful research avenue.

6 Aversion to indexation

When a risk-averse agents 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 other macroeconomic variables (e.g., Holmström (1979)). However, the stylized facts on indexation are the following. (i) Most 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.

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

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

Income at time 2 is $2(\bar{y} + \hat{y})$, where $\hat{y}$ is a mean 0 random variable, and there is no income at times 0 and 1. The interest rate is 0.

At time 0, the consumer can choose to "index", i.e. to enter into a contract adding $-2\beta x$ to his future time 2 income, where x is an variable, correlated with $\hat{y}$. So, time 2 income is, inclusive of the indexation,

$$y_2 = 2(\bar{y} + \hat{y} - \beta x) \tag{38}$$

where β is the consumer's choice of indexation.

I call a the time-1 consumption. So utility is $u(a) = v(a) + v(y_2 - a)$, i.e.

$$u(a, \beta) = v(a) + v(2(\bar{y} + \hat{y} - \beta x) - a)$$

in the Ramsey problem were modulated by attention, but where not simple. Inattention, and information economics was not enough as an ingredient. But now new ingredient of complexity aversion allows to obtain simple taxes.

The rational decision at time 1 is

$$a = \bar{y} + \hat{y} - \beta x \tag{39}$$

The value of $\hat{y}$ and x is announced at time 1. To minimize variance (i.e. maximize utility $\mathbb{E}_0 [u(a,\beta)]$), from the point of view of time 0, a rational consumer would choose $\beta^* = \text{argmin}_b \, var(\hat{y} - \beta x)$, i.e. $\beta^* = \frac{cov(\hat{y},x)}{\sigma_x^2}$.

But intuitively, a behavioral consumer will dislike the additional complexity of having to think through βx at time 1 (the values of r, x are realized at time 1). To model this, let's say that only x is hard to see. Then, the perceived value is $x^s = mx + (1-m)\varepsilon$. The consumer sees $a = \bar{y} + \hat{y} - \beta x^s$. So, the objective utility is

$$V = V^* - \frac{1}{2}\Gamma\beta^2\sigma_x^2 (1-m) - \xi\Gamma\sigma_a\sigma_x |\beta| (1-m)$$

So, if ξ is large enough, it is best to choose $\beta = 0$.

Proposition 7. (How complexity dampens or annuls the incentive to index) *Suppose first order complexity aversion. Normalize $cov(\hat{y},x) \geq 0$ (otherwise, replace x by $-x$). Suppose exogenous attention m. The optimal indexation level is*

$$\beta = \max\left(cov(\hat{y},x) - \xi\sigma_a\sigma_x(1-m), 0\right)\frac{1}{\sigma_x^2} \tag{40}$$

In particular, if the marginal benefit of indexation ($cov(\hat{y},x)$) is less than marginal complexity cost of indexation ($\xi\sigma_x$) (i.e., if $cov(\hat{y},x) < \xi\sigma_a\sigma_x(1-m)$), then the optimal indexation for the consumer is 0.

Proposition 7 is qualitatively consistent with the main facts. When stakes are small, there is no indexation. However, when stakes increase (e.g. inflation is higher), then there is some indexation. This may explain why most contracts (e.g. debt contracts) are not indexed.[18]

This example made the point that complexity-averse 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), but with a more explicit cognitive structure, which would lend itself to measurement. I defer this exploration to a future iteration of this project.

[18]Proposition 18 in the online appendix shows that with second-order complexity aversion, there is always a bit of indexation.

7 The complexity costs of inflation, its implications for monetary policy

7.1 The complexity cost of inflation are first order, so much larger than the usual second order costs

There is a large literature on the welfare cost of inflation, which includes many traditional determinants. But typically, they are very small. For instance, a prominent cost (Woodford (2003)) is the fact that inflation adds dispersion between firm prices: for instance, if Coca-Cola last reset its nominal price τ period before Pepsi, and inflation is π, the relative price between Coca-Cola and Pepsi is off by $\pi\tau$; that distortion yields a welfare loss (Harberger triangle) order $(\pi\tau)^2$. This is typically very small (less than 0.1% per year).

However, one very plausible cost of inflation is that it adds "complexity" for the decision makers (Binetti et al. (2024)). Let us model it precisely.

Time t is continuous. The log price level is $p_t = \ln P_t$. I write $t = n + h$ with n an integer represents is "year" and continuous variable $h \in [0,1)$ is the "calendar date within the year). The log price increase continuously within the year, as $p_{n+h} = p_n + \pi_{n+1}h$, where inflation in year n is π_{n+1}. The agent gets a raise at the beginning of the year. Income is paid continuously as

$$y^{\$}_{n+h} = p_n + \frac{1}{2}\pi_{n+1} \tag{41}$$

as $\frac{1}{2}\pi_{n+1}$ compensates for the lack of raise within the year ($t \in [n, n+1)$). Inflation has expected value $\bar{\pi}$.

For starkness, let us assume that inflation is can known in advance: for instance, it is constant at $\pi_n = \bar{\pi}$. Here, a rational agent would face no cost of inflation.[19]

Figure 2 plots prices (for a constant inflation), and income. The infrequent adjustment of income creates an extra complexity, as income and prices do not move in sync — inflation makes real income less regular. This is also true even if inflation is constant. In that sense, inflation adds extra complexity.

Optimal (rational) consumption is (in logs)

$$c^{\$}_{n+h} = y^{\$}_{n+h} - \left(\frac{1}{2} - h\right)\pi_{n+1}$$

Hence, to the problem is added an extra complication: the π_{n+1} term. It captures the following: early in the year ($h \simeq 0$), agent should save: as he just got a raise, his real wage

[19] It would be easy to make inflation uncertain – then in the income process (41) we'd replace π_n by its expectation.

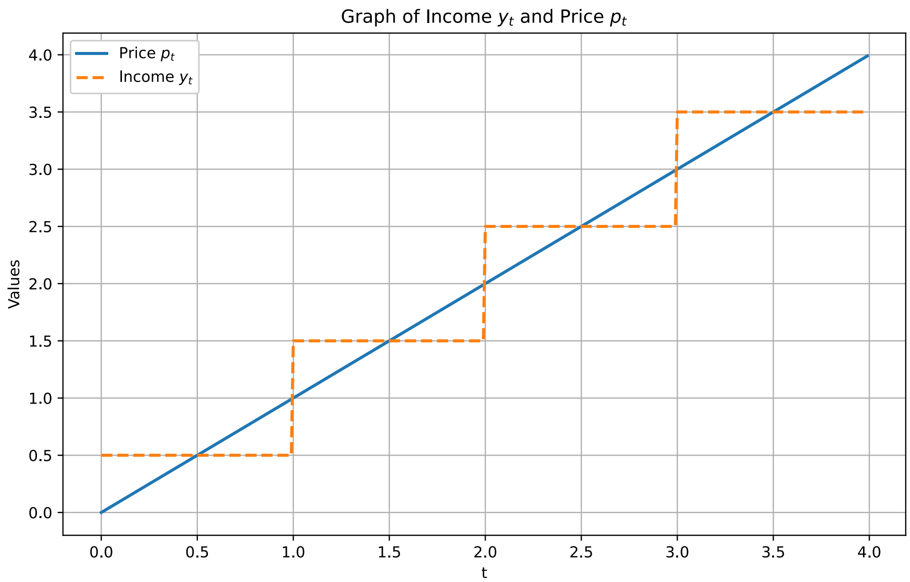

Figure 2: The price increase continuously, but the agents's nominal income increases only once a year (those quantities are in logs). That makes the planning process more complex, as the theory develops.

is high; later in the year ($h \simeq 1$), agent should dis-save: his real wage has been eroded by the inflation since the beginning of the year, so his real income is low.

The cost of inflation are the following. For simplicity, we suppose here that inflation is constant, at π.

Proposition 8. (Losses from inflation) *Suppose that nominal wages are readjusted only between periods of length τ. Call m the attention to inflation between wages increases, and assume first order complexity aversion. Then, the utility losses (expressed as a fraction of losses in permanent consumption) are:*

$$\ell^{CA} = \frac{1}{4}\gamma\xi\left|\tau\pi\right|(1-m) \tag{42}$$

while the "variance" costs are $\ell^{var} = \frac{1}{24}\gamma\tau^2\pi^2(1-m)$.

In particular, the complexity costs are higher than the variance costs iff $\pi < \frac{6\xi}{\tau} = \frac{6\times\frac{1}{2}}{1\text{yr}} =$ 300%/yr.

Because of first order complexity aversion, the loss is proportional to π, not π^2 as in the traditional model. So, for a moderately boundedly rational consumer ($m = 1/2$), then the loss is quite high, about ξ times $\tau\pi$. With $\xi = 0.5$ and $\gamma = 4$, and reset of wages every $\tau = 1$ year, the looses are $\ell^{\mathrm{CA}} \simeq \frac{1}{2}|\pi|$.[20] If annual inflation is $\pi = 4\%$, the loss is 2% — something sizable, while at $\pi = 20\%$,the loss is a large 10%. This is qualitatively consistent with empirical evidence that shows that people very averse to inflation even at low inflation rates in surveys, (Shiller (1997), Stantcheva (2024), Georgarakos et al. (2025)) and in elections (Di Tella et al. (2001)). This loss is much larger than the quadratic loss π^2 of the New Keynesian model, with both rational agents (Woodford (2003)), and behavioral but not complexity averse agents (Gabaix (2020)).

7.2 The sparsity of outcomes principle

With complexity costs, we have a first-order costs of having disturbances. We next present a simple result on their optimization.

Proposition 9. (Sparsity in outcomes principle) *Suppose that we minimize, over action a, the function:*

$$L = w_1 |x_1| + w_2 |x_2| \tag{43}$$

with $w_i > 0$, and $x_i = b_i a + \xi_i$, for some b_i, and ξ_i. Suppose that x_1 have higher "social welfare impact" than x_2, $|w_1 b_1| > |w_2 b_2|$. Then, at the optimum, $x_1 = 0$, while, generically, $x_2 \neq 0$.

This means that the planner wishes to set one of the objective variables to 0. We call this the "sparsity of outcomes" principle: the optimal policy makes one of the two outcomes sparse.[21] This result is mathematically extremely simple, coming from simple algebra and in the spirit of the Lasso (Tibshirani (1996)). Nonetheless, it is useful to keep in mind when thinking about optimal policy.

Let us see how that affects optimal monetary policy.

7.3 Complexity-aware welfare in the New Keynesian model

7.3.1 Economic environment

We take a simple New Keynesian model. Intuitively, given Proposition 9, the planner will wish to always set either inflation or the output gap to 0: this helps the consumer, this way. This is in contrast with the "quadratic loss" case, where generically both are non-zero.

[20] Georgarakos et al. (2025) reports that people are on average willing for forgo 5% per year in annual consumption to eliminate inflation, which they perceive to be around 5%. This can be matched with a more aggressively behavioral calibration, say with $m = 0$, $\xi = 1$ and $\pi = 5\%$, which is the inflation often perceived by survey respondents. Then, the loss is $\ell^{\mathrm{CA}} = 5\%$.

[21] The principle generalizes to high-dimensional settings. With V variables and J policy instruments (so a loss $L = \sum_{i=1}^{V} w_i |x_i|$, and an action vector a of dimension J), it will be often (i.e., in an open set of parameters) the case that J of the variables are exactly equal to 0.

Let us see this in some detail. We take the model in Gabaix (2020), which generalizes Woodford (2003) and Galí (2015), to allow for partially myopic agents, which is relevant for monetary policy. It microfounds the following the behavior of the output gap x_t and inflation π_t:

$$x_t = M\mathbb{E}_t [x_{t+1}] - \sigma (i_t - \mathbb{E}_t \pi_{t+1} - r_t^n) \text{ (IS curve)}, \tag{44}$$

$$\pi_t = \beta M^f \mathbb{E}_t [\pi_{t+1}] + \kappa x_t + \nu_t \text{ (Phillips curve)}, \tag{45}$$

where $M, M^f \in [0, 1]$ are the aggregate-level behavioral "cognitive discounting" parameters of consumers and firms, respectively. The rational case corresponds to $M = M^f = 1$.

Aggregate production is $Y_t = e^{\zeta_t} N_t$, where TFP level ζ_t, follows an AR(1), and N_t is labor supply. Equation (45) also allows for a "cost-push" shock ν_t, which could be microfounded as a markup shock (Galí (2015), Section 5.2), or an expectational error.

7.3.2 Welfare with complexity costs

In the traditional models (Woodford (2003)), the welfare function is:

$$W^{\text{trad}} = -K^{\text{trad}} \mathbb{E}_0 \sum_{t=0}^{\infty} \frac{1}{2} \beta^t \left(\pi_t^2 + \vartheta^{\text{trad}} (x_t - x_*)^2 \right) + W_-, \tag{46}$$

for positive coefficients K^{trad}, ϑ^{trad}.[22] The term x_* is zero is there are no distortions at the natural rate of interest (i.e., when removing monetary frictions). The losses come from misallocation costs, due to inflation π_t (the relative price of firms is off), and output gap x_t (people don't work exactly the correct amount). In contrast, complexity aversion adds one more term.

Proposition 10. *In the New Keynesian model with complexity aversion costs, welfare is*

$$W^{tot} = -\sum_t \beta^t C_t^{CA} + W^{trad}$$

where the complexity aversion costs at time t is:

$$C_t^{CA} = w_\pi |\pi_t| . \tag{47}$$

for positive coefficients w, and W^{trad} is the traditional welfare loss (46), with quadratic loses in inflation and output.

[22] Here $\vartheta^{\text{trad}} = \frac{\bar{\kappa}}{\varepsilon}$, $K^{\text{trad}} = u_c c (\gamma + \phi) \frac{\varepsilon}{\bar{\kappa}}$, and W_- is a constant, $\bar{\kappa}$ is the Phillips curve coefficient with rational firms, and ε is the elasticity of demand.

But for small inflations and output gap, first order costs dwarf second-order costs. Hence, the welfare is, to the leading order, only made of first order costs, $\sum_t \beta^t C_t^{CA}$.[23]

The planner wants to maximize $\sum_t \beta^t W_t$, where the loss at time t is

$$W_t = K\left(-\left|\pi_t\right| + g\left(x_t\right)\right), \qquad g\left(x_t\right) = \gamma x_t - \frac{1}{2}\theta x_t^2 \tag{48}$$

where the first order costs come from complexity aversion, and $\theta = \frac{\vartheta^{\text{trad}}}{w_\pi} > 0$ is the relative weight on output gaps, and $\gamma = -x_* \vartheta^{\text{trad}}$. Empirically, almost everyone "feels" that booms are better than recessions, so $\gamma > 0$ seems the empirically relevant case.[24]

So now we have a someone different problem: optimal policy with a absolute value penalty on inflation, rather than a quadratic one. This apparently minor and technocratic difference leads to quite deep changes in the optimal monetary policy, as we next see.

7.4 Optimal monetary policy with complexity costs

We can now analyze policy – here with commitment, which is arguably the relevant case (the appendix deals with the no-commitment case). When there are no cost-push shocks, the optimal monetary policy is very simple, and the same as the traditional New Keynesian model: ensure zero output gap and zero inflation, by adjusting the nominal rate to be the natural rate ($i_t = r_t^n$ in (44), so that $x_t = 0$ and $\pi_t = 0$ are solutions).[25]

However, when there is cost-push shock, the $\nu_t \neq 0$ in (45), one cannot get the first best: $x_t = \pi_t = 0$ is not consistent with (45). In the traditional model, a positive cost-push shocks is best handles by a bit of recession and a bit of inflation. With complexity aversion, things change.

In what follows, the sign (π) is meant to be the generalized derivative of the function $|\pi|$. So it is 1 if $\pi > 0$, -1 if $\pi < 0$, and some number between -1 and 1 if $\pi = 0$.[26]

Proposition 11. (Optimal response to a cost-push shocks, with commitment: conditions to have 0 inflation as a state by state target). *With welfare (48), the optimal commitment policy is characterized by*

$$\kappa \operatorname{sign}\left(\pi_t\right) = g'\left(x_t\right) - M^f g'\left(x_{t-1}\right) 1_{t>0} \tag{49}$$

[23] For simplicity, only consumers are complexity averse, not firms.

[24] In the traditional NK model, this is dues to the fact that there positive markups at the steady state, which this microfoundation is not necessary

[25] I omit here the issue of the zero-lower bound, i.e. the constraint $i_t \geq 0$. The first best solution is then use fiscal policy.

[26] This could be made more rigorous with the machinery of subgradients, but that would make the analysis less transparent for most readers.

Call $x_t^* := \frac{-\nu_t}{\kappa}$ *the output gap that ensures zero inflation. Suppose that:*

$$\kappa \geq \left|\gamma\left(1-M^f\right)-\theta\left(x_t^*-M^f x_{t-1}^*\right)\right| \tag{50}$$

then the optimal central bank policy under commitment is simply to keep inflation at 0 at all dates,

$$\pi_t = 0 \tag{51}$$

and adjust the output gap x_t *so that* $x_t = x_t^*$. *In turn, the nominal interest rate* i_t *is set so that (44) holds.*

First, take the case $\theta = 0$, so that the welfare impact of recessions is simply linear (rather than quadratic). Condition (50) is simply

$$\kappa \geq \left|\gamma\left(1-M^f\right)\right| \tag{52}$$

i.e. the slope κ of the Phillips curve must be high enough. Then, we have a starkly simple policy: if welfare increases linearly with the output gap, you want to always target inflation, and set is to 0. This is the "sparsity principle" at work: given the absolute value penalty in $-|\pi_t|$, in an open zone of parameters, one wants to have exactly zero inflation.

When $\theta > 0$, the intuition is similar: we set inflation to 0, if the recession is not too deep (i.e. x_t is not too negative), and the shocks ν_t do not mean-revert too fast: only long lasting inflation shocks are worth purging via a shock to the business cycle.

The case where (50) is violated is more involved. Let us a consider a special, but illustrative case.

Proposition 12. (Optimal response to a large, persistent cost-push shocks, with commitment) *Suppose that we have a permanent cost-push shock* $\nu_t = \nu$. *Define* $\nu_- := \frac{1}{\theta}\left(\gamma - \frac{\kappa}{1-M^f}\right)$, $\nu_+ := \frac{1}{\theta}\left(\gamma + \frac{\kappa}{1-M^f}\right)$, *and* $x_\pm := -\frac{\nu_\pm}{\kappa}$. *Then, the steady state commitment optimal policy is the following. (i) For moderate cost-push shocks,* $\nu \in [\nu_-, v_+]$, *we set inflation at 0,* $\pi = 0$ *and* $x = -\frac{\nu}{\kappa}$. *(ii) For very inflationary pressures* ($\nu > \nu_+$), *the central banks sets* $x = x_-$ *(there is a recession floor), and* $\pi = \frac{\nu-\nu_+}{1-\beta M^f} > 0$*: inflation is positive and increasing in the cost-push shock. (iii) Similarly, if deflationary pressures are very high* ($\nu < \nu_-$), *the central bank sets* $x = x_+$ *(i.e. a maximum-allowed boom to fight deflation), and set have* $\pi = \frac{\nu-\nu_-}{1-\beta M^f} < 0$.

As in Proposition 11, moderate cost-push pressures call for exactly zero inflation. However, if there is a very high cost-push shock ($\nu > \nu_+$), then there is a recession floor: $x = x_-$, while inflation bears the brunt of the cost-push shocks, $\pi = \frac{\nu-\nu_+}{1-\beta M^f} > 0$.

7.5 Extension: Non-zero inflation target, what what it should be

So far, the optimal inflation target was 0. I now study what happens with non-zero inflation target $\bar{\pi}$. The traditional reason for that is that workers exhibit an aversion to nominal (not just real) wage cuts, so a bit of inflation helps the "morale" in the labor market (Akerlof et al. (1996); Bewley (1999)). Then, I derive the optimal target—providing a newish simple model for this, that perhaps calibrates better given the new complexity costs of inflation. Then, I derive optimal monetary policy. The earlier conclusions broadly remain the same, replacing "inflation" by "inflation minus target".

7.5.1 Modelling workers' happiness from a higher nominal wage

So far, the optimal inflation target was zero. Now, the real world brings one more complication: workers have some nominal illusion, and in particular do not like nominal wage cuts. So, if the neoclassical real wage was to fall, with zero inflation, firms would not cut wages, and rather fire some workers. To avoid this, it is useful to have positive inflation.

To model this, let us say that workers morale affects raw log productivity ζ_t, to change it into effective log productivity:

$$\zeta_t^{\text{eff}} = \zeta_t + f_t$$

where the new term f_t captures the efficiency wage idea that disgruntled work less well, and indeed may sabotage a firm. We take the following formulation for their morale:

$$f_t = f\left(w_t - \zeta_{t-1} + \alpha_\pi \pi_t\right) \tag{53}$$

for some increasing function f, so that workers compare the wage to the "fair" wage from past productivity, but have a bit illusion, index by α_π. This is in the spirit previous formulations (e.g. Benigno and Antonio Ricci (2011); Schmitt-Grohé and Uribe (2016)), which typically use a "hard" constraint such as a rule that the nominal wage can never fall. The formulation here likely would yield rather similar conclusion as those earlier models, but is more tractable.[27]

[27] A tempting formulation would be:

$$f_t^0 = f\left(w_t^\$ - w_{t-1}^\$\right) = f\left(w_t - w_{t-1} + \pi_t\right) \tag{54}$$

where w_t (resp. $w_t^\$$) is the log real (resp. nominal) wage, for some function f that may be a sigmoid, that saturates with a range $\pm 5\%$. But this formulation makes the problem a little more complicated, because of the backward looking term. The function f_t we choose in (53) has similar properties to the "natural" one f_t^0: a bit of inflation increases total efficiency; and that's particularly true if the real wage is low, compared to past wages.

7.5.2 Optimal nonzero steady state inflation

Given we have inflation has both a first order cost (complexity) and a first order benefit (worker morale linked to the nominal wage), we can discuss an optimal steady state inflation. Total per period log consumption is

$$C_t = e^{\zeta_t + f_t} L_t$$

in this model, without complexity costs of inflation. With complexity costs, log utility becomes:

$$W_t = -\omega_\pi |\pi_t| + f_t$$

as in the steady state, there is no trend productivity growth, and we normalize average log productivity to be 0. So optimal steady state inflation balances the complexity costs of inflation with the efficiency inflation benefits: $\max_\pi W$, with:

$$W = -\omega_\pi |\pi| + f(\pi)$$

To calibrate, let's imagine $\alpha_\pi = 1$, and take $f(x) = a\left(\pi - \frac{1}{2\Pi}\pi^2\right)$ around 0, with $a = 1\text{yr}$ (so 1% deflation would great a 1% productivity loss), where $\Pi = 4\%$ gives the order of magnitude of the potential maximum effect. Those are taken here as psychological primitives, though in turn they might be microfounded. Then, the optimal inflation π satisfies $\max_a -\omega_\pi |\pi| + a\pi - \frac{a}{\Pi}\pi^2$, i.e. is

$$\pi = \left(1 - \frac{\omega_\pi}{a}\right)^+ \Pi$$

This optimum inflation target balances the complexity costs. Given $\omega_\pi = \frac{1}{2}\text{yr}$ (from above), $a = 1\text{yr}$, we get an optimal target of 2%. Of course, this is all quite a guesstimate, but at least, we have roughly plausible order of magnitude, and a framework to do a more systematic investigation.

7.5.3 Welfare and optimal policy with a non-zero inflation target

Call $\bar{\pi}$ the target inflation, and $\hat{\pi}_t = \pi_t - \bar{\pi}$ the deviation from the target. I now use the enriched model in Gabaix (2020, section V.6), that allows for non-zero target inflation: then, the basic model (44)-(45) remains the same, replacing π_t by $\hat{\pi}_t$ in the Phillips curve (45).

Proposition 13. *Suppose that we morale reasons to have positive steady inflation, which is set optimally. Then, up to second order terms in inflation, welfare is:*

$$W_t = \bar{W} - h(\hat{\pi}_t) + g(x_t) \tag{55}$$

with g as in (48), and

$$h(\hat{\pi}_t) = \omega_{\hat{\pi}}^{-} |\hat{\pi}_t| + \omega_{\pi} \hat{\pi}_t = \omega_{\hat{\pi}}^{-} |\hat{\pi}_t| 1_{\hat{\pi}_t < 0} + \omega_{\hat{\pi}}^{+} |\hat{\pi}_t| 1_{\hat{\pi}_t > 0} \tag{56}$$

where $0 \le \omega_{\hat{\pi}}^{+} \le \omega_{\hat{\pi}}^{-}$ (the values are in the proof) so that a temporary increase in inflation is less harmful than a temporary decrease.

Hence, temporary inflation $\hat{\pi}_t$ hurts welfare, much like $-|\hat{\pi}_t|$, but with a $\omega_{\hat{\pi}}^{+}$ weight for positive temporary inflation, and a higher weight $\omega_{\hat{\pi}}^{-}$ for negative temporary inflation. A negative deviation of inflation is worse than a positive one, because the complexity cost are the same in both directions, but the labor market gains are positive for positive inflation.

Intuitively, this implies the central bank policy is again to set $\hat{\pi}_t = 0$, in a broad set of circumstances, but with some asymmetry. We make this more precise.

Proposition 14. *Make the same assumption as Proposition 13. The first order condition for optimal monetary policy becomes*

$$\kappa h'(\hat{\pi}_t) = g'(x_t) - M^f g'(x_{t-1}) \tag{57}$$

So, it is "more acceptable" to have positive deviations (as $|h'(0^+)| \le |h'(0^-)|$) than negative deviations of temporary inflation, as negative ones are costlier. This may be a reason why empirically inflation seems to overshoot its target more often than undershoot it.

Otherwise, the earlier analysis goes through, with the caveat that the previous $\text{sign}(\pi_t)$ changes into $h'(\hat{\pi}_t)$. For instance, the condition (50) to have always $\hat{\pi}_t = 0$ becomes:

$$g'(x_t) - M^f g'(x_{t-1}) \in \left[\kappa h'\left(0^-\right), \kappa h'\left(0^+\right)\right]. \tag{58}$$

We finished our tour of the complexity costs of inflation. It leads to much higher cost, a different monetary policy (targeting exactly zero inflation, or zero deviation from a target), and indeed allows to think about the optimal steady state level of inflation.

8 Discussion

This section discusses some remaining high-level points.

Why is it hard to handle tasks requiring to juggle with 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 computer is often easy for people and vice-versa). *Why is it hard to handle tasks requiring to juggle with 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 quite easier. In the model, this is because the cost of processing c_i in (4) is high when dealing with mathematical operations on numbers, and and low when processing visual information: e.g., there's 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 (9) is very low. Under a heavy rain, so that seeing the road is harder, imprecision is higher, and driving is harder.

Are people aware that they're making a mistake? In this model, people are completely aware that they're making a mistake. This is clearly just a benchmark, probably right for more decisions. However, it is useful to keep in mind that there are some problems where a "tempting" answer seems "obvious", and is actually wrong. One famous one is the Monty Hall three door problems (which, amazingly, the great mathematician Erdős got wrong initially, and was convinced only after several days, see Vazsonyi (1999)), and behavioral economics has provided other (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 (2018a)), those consumers are not aware that they're making a mistakes. It would be nice to extend the model for that, with an "attention" to the mistake.

9 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 model with no complexity costs.

Along the way, the model shed light on a number of issues of basic modelling. One, cognition costs or information cost per se are not enough to generate complexity: one needs to have a form of "complexity aversion" that is non-zero even when the agent pays no cognition cost. Another, we saw that first order complexity aversion is the robust way to generate simplicity, much more than second order complexity aversion (or regret from regret theory)

Many interesting issues seem now within reach. In particular, designing optimal mechanisms, taking into account the complexity costs by consumers, seem particularly useful from a practical point of view. The present paper's analysis of simple tax systems for a taxation of goods is a beginning, but clearly much more can be done.

References

Abeler, J. and S. Jäger (2015). Complex tax incentives. *American Economic Journal: Economic Policy 7*(3), 1–28.

Aghion, P., U. Akcigit, M. Gravoueille, M. Lequien, and S. Stantcheva (2024). Tax simplicity or simplicity of evasion? evidence from self-employment taxes in france. Working paper, National Bureau of Economic Research.

Akerlof, G. A., W. T. Dickens, and G. L. Perry (1996). The macroeconomics of low inflation. *Brookings papers on economic activity 1996*(1), 1–76.

Andries, M. and V. Haddad (2020). Information aversion. *Journal of Political Economy 128*(5), 1901–1939.

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

Augenblick, N., B. K. Jack, S. Kaur, F. Masiye, and N. Swanson (2024). Retrieval failures and consumption smoothing: A field experiment on seasonal poverty. Technical report, Technical Report, University of California, Berkeley.

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

Ben-David, I. and A. Chinco (2024). Expected EPS × Trailing P/E. Technical report, National Bureau of Economic Research.

Benigno, P. and L. Antonio Ricci (2011). The inflation-output trade-off with downward wage rigidities. *American Economic Review 101*(4), 1436–1466.

Bernheim, B. D., K. Kim, and D. Taubinsky (2024). Welfare and the act of choosing. Technical report, National Bureau of Economic Research.

Bewley, T. F. (1999). *Why wages don't fall during a recession*. Harvard university press.

Binetti, A., F. Nuzzi, and S. Stantcheva (2024). People's understanding of inflation. Technical report, National Bureau of Economic Research.

Bordalo, P., N. Gennaioli, and A. Shleifer (2013). Salience and consumer choice. *Journal of Political Economy 121*(5), 803–843.

Calvet, L. E., J. Y. Campbell, and P. Sodini (2009). Fight or flight? Portfolio rebalancing by individual investors. *Quarterly Journal of Economics 124*(1), 301–348.

Caplin, A., D. Csaba, J. Leahy, and O. Nov (2020). Rational inattention, competitive supply, and psychometrics. *The Quarterly Journal of Economics 135*(3), 1681–1724.

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

Célérier, C. and B. Vallée (2017). Catering to investors through security design: Headline rate and complexity. *Quarterly Journal of Economics 132*(3), 1469–1508.

Chetty, R., A. Looney, and K. Kroft (2009). Salience and taxation: Theory and evidence. *American economic review 99*(4), 1145–1177.

Dari-Mattiacci, G. and F. Marotta-Wurgler (2022). Learning in standard-form contracts: Theory and evidence. *Journal of Legal Analysis 14*(1), 244–314.

Dean, M. and N. L. Neligh (2023). Experimental tests of rational inattention. *Journal of Political Economy*.

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

Di Tella, R., R. J. MacCulloch, and A. J. Oswald (2001). Preferences over inflation and unemployment: Evidence from surveys of happiness. *American economic review 91*(1), 335–341.

Enke, B. and T. Graeber (2023). Cognitive Uncertainty. *Quarterly Journal of Economics 138*(4), 2021–2067.

Enke, B., T. Graeber, and R. Oprea (2023). Complexity and time. Working paper, National Bureau of Economic Research.

Enke, B., T. Graeber, R. Oprea, and J. Yang (2024). Behavioral attenuation. Working paper, National Bureau of Economic Research.

Enke, B. and C. Shubatt (2023). Quantifying lottery choice complexity. Technical report, National Bureau of Economic Research.

Enke, B. and F. Zimmermann (2019). Correlation neglect in belief formation. *The Review of Economic Studies 86*(1), 313–332.

Eyster, E., K. Madarász, and P. Michaillat (2020, 08). Pricing Under Fairness Concerns. *Journal of the European Economic Association 19*(3), 1853–1898.

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

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic perspectives 19*(4), 25–42.

Fudenberg, D. and I. Puri (2023). Evaluating and extending theories of choice under risk. Working paper, MIT Economics.

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

Gabaix, X. (2019). Behavioral inattention. *Handbook of Behavioral Economics 2*, 261–344.

Gabaix, X. (2020). A behavioral new keynesian model. *American Economic Review 110*(8), 2271–2327.

Gabaix, X. and T. Graeber (2024). The complexity of economic decisions. Working Paper 33109, National Bureau of Economic Research.

Gabaix, X., R. S. Koijen, F. Mainardi, S. S. Oh, and M. Yogo (2025). Limited risk transfer between investors: A new benchmark for macro-finance models. Technical report, National Bureau of Economic Research.

Gabaix, X. and D. Laibson (2006). Shrouded attributes, consumer myopia, and information suppression in competitive markets. *Quarterly Journal of Economics 121*(2), 505–540.

Gabaix, X., D. Laibson, D. Li, H. Li, S. Resnick, and C. G. de Vries (2016). The impact of competition on prices with numerous firms. *Journal of Economic Theory 165*, 1–24.

Gabaix, X., D. Laibson, G. Moloche, and S. Weinberg (2006, September). Costly information acquisition: Experimental analysis of a boundedly rational model. *American Economic Review 96*(4), 1043–1068.

Galí, J. (2015). *Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework and its Applications*. Princeton University Press.

Georgarakos, D., K. H. Kim, O. Coibion, M. Shim, M. A. Lee, Y. Gorodnichenko, G. Kenny, S. Han, and M. Weber (2025). How costly are business cycle volatility and inflation? a vox populi approach. Technical report, National Bureau of Economic Research.

Gertler, M. and P. Karadi (2011). A model of unconventional monetary policy. *Journal of monetary Economics 58*(1), 17–34.

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

Gonczarowski, Y. A., O. Heffetz, and C. Thomas (2023). Strategyproofness-exposing mechanism descriptions. Technical report, National Bureau of Economic Research.

Gormsen, N. J. and K. Huber (forth.). Corporate discount rates. Technical report.

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

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

Hart, O. (1995). *Firms, contracts, and financial structure.* Clarendon press.

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

Heidhues, P. and B. Kőszegi (2018a). Behavioral industrial organization. *Handbook of Behavioral Economics: Applications and Foundations 1 1*, 517–612.

Holmström, B. (1979). Moral hazard and observability. *The Bell journal of economics*, 74–91.

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

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

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

Karlsson, N., G. Loewenstein, and D. Seppi (2009). The ostrich effect: Selective attention to information. *Journal of Risk and uncertainty 38*(2), 95–115.

Khaw, M. W., Z. Li, and M. Woodford (2021). Cognitive imprecision and small-stakes risk aversion. *The Review of Economic Studies 88*(4), 1979–2013.

Kőszegi, B. and M. Rabin (2009). Reference-dependent consumption plans. *The American Economic Review 99*(3), 909–36.

Laibson, D. (1997). Golden eggs and hyperbolic discounting. *Quarterly Journal of Economics 112*(2), 443–478.

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

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

Matějka, F. and A. McKay (2015). Rational inattention to discrete choices: A new foundation for the multinomial logit model. *American Economic Review 105*(1), 272–298.

Mullainathan, S., J. Schwartzstein, and W. J. Congdon (2012). A reduced-form approach to behavioral public finance. *Annual Review of Economics 4*(1), 511–540.

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

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

Puri, I. (2024). Simplicity and risk. *Journal of Finance, Forthcoming*.

Rabin, M. (2000). Risk aversion and expected-utility theory: A calibration theorem. *Econometrica 68*(5), 1281–1292.

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

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

Rubinstein, A. (1998). *Modeling Bounded Rationality.* MIT press.

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

Schmitt-Grohé, S. and M. Uribe (2016). Downward nominal wage rigidity, currency pegs, and involuntary unemployment. *Journal of Political Economy 124*(5), 1466–1514.

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

Shiller, R. J. (1997). Why do people dislike inflation? In *Reducing inflation: Motivation and strategy*, pp. 13–70. University of Chicago Press.

Sicherman, N., G. Loewenstein, D. J. Seppi, and S. P. Utkus (2016). Financial attention. *The Review of Financial Studies 29*(4), 863–897.

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

Sonsino, D., U. Benzion, and G. Mador (2002). The complexity effects on choice with uncertainty–experimental evidence. *Economic Journal 112*(482), 936–965.

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

Stantcheva, S. (2024). Why do We Dislike Inflation? *Brookings Papers on Economic Activity 55*(1).

Taubinsky, D. and A. Rees-Jones (2017). Attention variation and welfare: Theory and evidence from a tax salience experiment. *Review of Economic Studies 85*(4), 2462–2496.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological) 58*(1), 267–288.

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

Vazsonyi, A. (1999). Which door has the cadillac? *Decision Line 30*(1), 17–19.

Woodford, M. (2003). *Interest and Prices: Foundations of a Theory of Monetary Policy.* Princeton University Press.

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

A Theory complements

A.1 A microfoundation for the fact that competitive firms want maximize social surplus

Section 3 assumes that the firm is "benevolent" – wants 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 2 firms, and N consumers. On day t, the marginal costs 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 is chooses the average price $\bar{p}_f$ and the slope b_f) and also "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).[28]

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).[29]

The consumer is fully conscious her expected utility, including and cognition and complexity aversion costs.[30] She chooses a firm f that maximizes his expected utility surplus, from this pricing scheme, i.e.

$$V_f = -F_f + TV_f^{\text{cons}} \tag{59}$$

The first part if 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 of consumers choosing the firms, and the last term is the average

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

[29] This is important, to prevent arbitrage: otherwise, take the case 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 $1 + \frac{1}{2}x_t < 1$, explain to consumers that they save money by coming to this deviating firm, and make a profit.

[30] This is where the paper deviates from the literature on shrouded attributes and firms exploiting naive consumers (see footnote 7): 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.

profit per consumer.

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

Lemma 3. *In the competition game above, the only Nash equilibrium is: $\bar{p}_f = 1$, $F_f = 0$, and b_f as in the paper, maximizing total welfare. 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 right should be the average marginal cost, so $\bar{p}_f = 1$. 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. □

A.2 Second order complexity aversion sometimes generates simplicity, but in a non-robust way

This appendix probes second-order complexity aversion, and finds that it generally fails to generate full simplicity, e.g. rigid prices or zero indexation.

A.2.1 The price scheme problem of Section 3

Proposition 15. (Optimal price simplicity, with second order complexity aversion) *Suppose second order complexity aversion. Suppose that attention m is exogenous. Then the optimum slope is*

$$b = \frac{m}{m + (1 + \xi)(1 - m)} \in [0, 1]$$

Hence, contracts are typically non-simple, as $b > 0$ whenever $m > 0$.

Proof. With second order CA, we use the analogue of (62)-(63), with $\bar{\xi} = 1 + \xi$

$$W = \left(b - \frac{1}{2}b^2\right)\psi\sigma_x^2 m - \bar{\xi}\frac{1}{2}\psi\sigma_x^2 b^2 (1 - m) - C(m)$$

so

$$W_b = (1 - b)\psi\sigma_x^2 m - \bar{\xi}\psi\sigma_x^2 b (1 - m) + W_m \frac{dm}{db}$$

First, with exogenous attention, so $\frac{dm}{db} = 0$. Then, the optimum is $b = \frac{m}{m + \bar{\xi}(1-m)}$. In general, $b = \frac{m + \frac{1}{\psi\sigma_x^2} W_m \frac{dm}{db}}{m + \bar{\xi}(1-m)}$. □

So, if there is a "default" level of attention that's positive, the some positive slope $b > 0$ is optimum: contracts are not "perfectly simple". They are not the optimum contract with rational agents, which would be $b = 1$.

Let us next explore a different variant, to probe the robustness of the second order complexity aversion. I suppose that there is a non-zero fraction $\varepsilon \in (0,1)$ of rational consumers, with zero cognition costs, and that the same contract must be offered to all consumers.

Proposition 16. (Optimal contract when there is a small fraction of rational consumers) *Suppose that there is a fraction $\varepsilon \in (0,1)$ of rational consumers. Then, with first order complexity aversion: with either exogenous attention $m < 1$, or endogenous attention with the conditions of Proposition 3, then if ξ is large enough one still wants a simple contract ($b = 0$). However, with second order complexity aversion, a simple contract is never optimal.*

The moral is that first order complexity aversion is a much more robust way to get simplicity than second order CA.

Now, let us turn to endogenous attention, which gives another way to appreciate the same phenomenon.

Proposition 17. (Optimal price simplicity, with second order complexity aversion, endogenous attention) *Suppose second order complexity aversion, and endogenous attention. Then, we have the following: with $\alpha < \frac{2}{3}$ or $\alpha > 1$ (where α parametrizes the curvature of the production function of thought, (4)-(7)), then if complexity aversion $\xi > 0$ is large enough, then the optimum entails rigid prices ($b = 0$). However, if $\alpha \in \left(\frac{2}{3}, 1\right)$, then the optimum entails flexible prices ($b > 0$).*

We see that both first and second order complexity aversion generate simplicity. However, the first order complexity works with fewer condition. Both tools achieve complexity, but first order CA is very directly effective, like sword, whereas second order CA is more delicate, like a scalpel. It is useful to have both tools in the economist's kit, though the sword will be more powerful most of the time.

A.2.2 The indexation problem of Section 6

First order complexity aversion predicted (in a range of parameters) no indexation (Proposition 7). In contrast, the next proposition shows that second order complexity aversion will fail to do so.

Proposition 18. (With second order complexity aversion, there is always some indexation) *In contrast to Proposition 7, with second order complexity aversion, there is always some indexation.*

Proof. Utility is:

$$V = -\frac{\Gamma}{2}\left(\sigma_r^2 - 2\beta\sigma_{xr} + \beta^2\sigma_x^2\right) - \xi\Gamma\beta^2\sigma_x^2\left(1-m\right) - C$$

so $V_\beta \neq 0$ at $\beta = 0$, hence some $\beta \neq 0$ is optimal. □

B Omitted proofs

Proof of Lemma 1 We have:

$$\mathbb{E}\left[\left(p\left(x\right)-c\left(x\right)\right)a\left(x^s\right)\right] = \mathbb{E}\left[\left(b-1\right)x\left(-\psi b x^s\right)\right] = \left(1-b\right)b\psi\mathbb{E}\left[xx^s\right] = \left(b-b^2\right)\psi\sigma_x^2 m$$

and as $\mathbb{E}\left[\left(p\left(x\right)-c\left(x\right)\right)\right] = 0$, the total producer surplus is:

$$\mathbb{E}\left[\left(p\left(x\right)-c\left(x\right)\right)\left(1+a\left(x^s\right)\right)\right] = \left(b-b^2\right)\psi\sigma_x^2 m \tag{60}$$

This implies:

$$\begin{aligned} W &= V^{\text{cons}}\left(m\right) + \mathbb{E}\left[\left(p\left(x\right)-c\left(x\right)\right)\left(1+a\left(x^s\right)\right)\right] \\ &= \frac{1}{2}\psi\sigma_x^2 b^2 - C^{\text{CA}} - C\left(m\right) + \psi\sigma_x^2\left(b-b^2\right)m \\ &= \left(b-\frac{b^2}{2}\right)\psi\sigma_x^2 m - C^{\text{CA}} - C\left(m\right) \end{aligned}$$

Proof of Proposition 2 Recall that with $\chi = \frac{1}{2}\psi\sigma_x^2$, social welfare is:

$$W = \Pi + V^{\text{cons}}, \qquad \Pi = 2\left(b-b^2\right)m\chi, \qquad V^{\text{cons}} = b^2 m\chi - C\left(m\right)$$

where Π is the average profit (60), 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

$$W_b = 2\left(1-b\right)m\chi + 2\left(b-b^2\right)m_b\chi \tag{61}$$

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 $C'\left(m\right) = b^2\chi$.

If $C'(0) < \chi$, when $b \in [0,1]$ is such that $b^2\chi > C'(0)$ we have $m > 0$, and $b = 1$ is a strict global optimum.

If $C'(0) \leq \chi$, 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 with a minuscule fraction of rational consumers (as in Proposition 16) shifts the social optimum to $b = 1$, strictly.

Proof of Proposition 3 Suppose now that we have endogenous attention. Then, as $b = 0$, $W_m = 0$, so

$$\frac{d}{db} W_{|b=0} = \psi \sigma_x^2 m^d - \xi \sigma_a \sigma_x sign(b) \left(1 - m^d\right)$$

so locally, $\frac{d}{db} W_{|b=0} \leq 0$ iff the same condition (25) 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 useful to define with

$$V(m) := V^{\text{cons}}(m) - \frac{1}{2}\psi \sigma_x^2 b^2$$

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

$$V(m) := V^{\text{cons}}(m) - \frac{1}{2}\psi \sigma_x^2 b^2 = \frac{1}{2}\psi \sigma_x^2 b^2 m - \xi \sigma_a |b| \sigma_x (1 - m) - C(m) - \frac{1}{2}\psi \sigma_x^2 b^2$$

hence

$$V(m) = -B(1 - m) - C(m), \qquad B := \frac{1}{2}\psi \sigma_x^2 b^2 + \xi \sigma_a |b| \sigma_x, \tag{62}$$

which gives $m^* = \operatorname{argmax}_m V(m)$. . By (60) social welfare is: $W = V^{\text{cons}}(m) + (b - b^2)\psi \sigma_x^2 m$, i.e.

$$W = V(m) + \left(b - \frac{b^2}{2}\right)\psi \sigma_x^2 m \tag{63}$$

Suppose that $C(1) > \psi \sigma_x^2$ 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'll 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.

$$\min_m \left(\frac{1}{2}\psi\sigma_x^2 b^2 + \xi\sigma_a b\sigma_x\right)(1-m) + C(m) \geq b\psi\sigma_x^2$$

i.e. (after division by b), for all b,

$$\min_m \left(\frac{1}{2}\psi\sigma_x^2 b + \xi\sigma_a\sigma_x\right)(1-m) + \frac{C(m)}{b} \geq \psi\sigma_x^2$$

A sufficient condition for that is that

$$\xi\sigma_a\sigma_x(1-m) + C(m) \geq \psi\sigma_x^2, \qquad \forall m \in [0,1]$$

i.e. $\xi \geq \xi^* := \max_m \frac{\psi\sigma_x^2 - C(m)}{\sigma_a\sigma_x(1-m)}$.

Proof of Proposition 4 If $b=0$, social welfare is $W(0)=0$. Now suppose that the firm sets $b=1$. Then, optimal attention allocation follows $\max_m Bm - C(m)$ with $B = \frac{1}{2}\psi\sigma_x^2 + \xi\sigma_a\sigma_x$. So, by the proposition's assumption, we have $C'(1) < \frac{1}{2}\psi\sigma_x^2 \leq B$, so consumer attention is $m=1$. Hence, welfare is $W(1) = \frac{1}{2}\psi\sigma_x^2 - C(1)$. As $C(1) \leq C'(1)$ by convexity, we have $C(1) < \frac{1}{2}\psi\sigma_x^2$ and $W(1) > 0 = W(0)$. Choosing $b=1$ yields greater social welfare better than $b=0$.

Proof of Proposition 5. For this utility $\tilde{U}(a) = -\frac{1}{2}(a-x)^2 + vx$, we have

$$\tilde{U}^{\text{active}}(m) = -B(1-m) - C(m), \qquad B = \frac{1}{2}\sigma_x^2 + \xi\sigma_a\sigma_x \tag{64}$$

while the passive decision is $U^{\text{passive}} = -\frac{1}{2}\left[\left(a^d\right)^2 + \sigma_x^2\right]$. We have

$$\begin{aligned}\tilde{U}^{\text{active}} &= -B(1-m) - C(m) + \mu vx = -B(1-m) - C(m) - \mu vx(1-m) + \mu vx \\ &= -\tilde{B}(1-m) - C(m) + \mu vx\end{aligned}$$

with $\tilde{B} = B + \mu vx$. We observe that

$$\min_{m\in[0,1]} \tilde{B}(1-m) + \kappa m = \min\left(\tilde{B}, \kappa\right) \tag{65}$$

as indeed the optimum attention is $m = 1_{\tilde{B}\leq\kappa}$. So, optimizing on attention m, with $C(m) = \kappa m$, we get:

$$\tilde{U}^{\text{active}} = -\min\left(\tilde{B}, \kappa\right) + \mu vx = -\min\left(\tilde{B} - \mu vx, \kappa - \mu vx\right) = -\min(B, \kappa - \mu vx)$$

Hence,

$$\tilde{U}^{\text{active}} = -\min\left(B, \kappa - \mu v x\right)$$

Hence, we have an passive decision iff $\tilde{U}^{\text{active}} \leq U^{\text{passive}}$, i.e. iff

$$\min\left(B, \kappa - \mu v x\right) \geq \frac{1}{2}\left[\left(a^d\right)^2 + \sigma_x^2\right]$$

i.e.

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

Proof of Proposition 7. Calling V^r the utility of a fully rational consumer, utility is V where, calling $\bar{\xi} = \xi\sigma_a\sigma_x\left(1-m\right)$

$$\begin{aligned} V - V^r &= -\frac{\Gamma}{2}\mathbb{E}\left[\left(\hat{y} - \beta x\right)^2\right] - \Gamma\bar{\xi}\left|\beta\right| - C \\ &= -\frac{\Gamma}{2}\left(\sigma_{\hat{y}}^2 - 2\beta\sigma_{\hat{y}x} + \beta^2\sigma_x^2 + \bar{\xi}\left|\beta\right|\right) - C \end{aligned}$$

We see that V is a concave function of β. At the optimum, we have:

$$0 = \frac{V_\beta}{\Gamma} = \sigma_{\hat{y}x} - \beta\sigma_x^2 - \bar{\xi} sign\left(\beta\right)$$

If the optimal β is non-zero, then must be positive, and $\beta = \frac{\sigma_{\hat{y}x} - \bar{\xi}}{\sigma_x^2} = \frac{\sigma_{\hat{y}x} - \xi\sigma_a\sigma_x(1-m)}{\sigma_x^2}$. The optimal β is 0 if $\sigma_{\hat{y}x} - \xi\sigma_a\sigma_x\left(1-m\right) \leq 0$.

Proof of Proposition 8 The complexity costs are, in general:

$$C^{\text{CA}} = \xi\sum_{i,t}\mathbb{E}\left|\tilde{a}_t u_{a_t a_t} a_{t,x_i} x_i\right|\left(1-m_i\right) \tag{66}$$

In our context, this gives:

$$C^{\text{CA}}/\xi = \sum_t \mathbb{E}\left|\gamma\beta^t c_t^{1-\gamma}\frac{\partial \ln c_t}{\partial x_i}\sigma_{x_i}\right|\left(1-m_i\right)$$

In the limit $r\tau \to 0$,

$$\begin{aligned}
C^{\text{CA},1}/\xi &= \gamma c_0^{1-\gamma}(1-m)\int_0^\infty e^{-rt}\left|\frac{\partial \ln c_t}{\partial x_i}\sigma_{x_i}\right| dt \\
&= \gamma c_0^{1-\gamma}(1-m)\,\sigma_{x_i}\sum_{n\geq 0} e^{-rn}\int_0^1 \left|\frac{1}{2}-h\right| dt \\
&= \gamma c_0^{1-\gamma}(1-m)\,\sigma_{x_i}\sum_{n\geq 0} e^{-rn}\frac{1}{4} \\
&= \frac{1}{r}c_0^{1-\gamma}\frac{1}{4}\xi\gamma(1-m)\,\sigma_{x_i}
\end{aligned}$$

Next, the variance term is

$$V = \frac{1}{2}\gamma c_0^{1-\gamma}(1-m)\int_0^\infty e^{-rt}\left|\frac{\partial \ln c_t}{\partial x_i}\sigma_{x_i}\right|^2 dt$$

By the same reasoning, we have $\int_{h=0}^1 \left(\frac{\partial \ln c_{t+h}}{\partial \pi_t}\right)^2 dh = \int_{h=0}^1 \left(\frac{1}{2}-h\right)^2 dh = \left[-\frac{1}{3}\left(\frac{1}{2}-h\right)^3\right]_0^1 = \frac{1}{12}$

In the end,

$$L = \sum_t \beta^t\left(\frac{1}{4}\gamma\xi\left|\pi_t\right|\tau + \frac{1}{24}\gamma\pi_t^2\tau^2\right)(1-m)$$

Proof of Proposition 9 This is very simple. Write $L = \sum_{i=1}^2 w_i\left|b_i a + \xi_i\right|$, so

$$L_a = \sum_{i=1}^2 w_i b_i sign\left(x_i\right).$$

If none of the x_i is 0, then $L_a = \sum_i \pm w_i b_i$ is non-zero, as $|w_1 b_1| > |w_2 b_2|$. Hence, one of them is 0. Suppose that it's x_2 that is 0. Then, $L_a = \pm w_1 b_1 + s_2 w_2 b_2$, where $s_2 \in [-1,1]$ is the "generalized sign". So, we must have $L_a \neq 0$, as $|w_1 b_1| > |w_2 b_2|$. Hence, we have $x_1 = 0$. Indeed, we verify that $L_a = s_1 w_1 b_1 \pm w_2 b_2$ can be equal to 0, with $|s_1| = \frac{|w_2 b_2|}{|w_1 b_1|}$.

Proof of Proposition 11 Let us call $g(x) = \gamma x - \frac{1}{2}\theta x^2$ the welfare impact of output gap x. The lagrangian is:

$$L = \mathbb{E}_0\sum_{t=0}^\infty \beta^t\left[-\left|\pi_t\right| + g\left(x_t\right) - \Xi_t\left(-\pi_t + \beta M^f\pi_{t+1} + \kappa x_t + \nu_t\right)\right], \tag{67}$$

where Ξ_t are Lagrange multipliers. The first order conditions are: $L_{x_t} = 0$ and $L_{\pi_t} = 0$, i.e.:

$$\begin{aligned} g'(x_t) &= \Xi_t \kappa && (68) \\ \operatorname{sign}(\pi_t) &= \Xi_t - M^f \Xi_{t-1} 1_{t>0} && (69) \end{aligned}$$

hence

$$\operatorname{sign}(\pi_t) = g'(x_t) - M^f g'(x_{t-1}) 1_{t>0} \tag{70}$$

which gives (49).

We have solution with $\pi_t = 0$ at all dates iff (i) $\kappa x_t + \nu_t = 0$ from the Phillips curve, so $x_t = x_t^* := \frac{-\nu_t}{\kappa}$ and (ii) (49) holds with $|\operatorname{sign}(\pi_t)| \leq 1$, i.e.

$$\kappa \geq \left| g'(x_t) - M^f g'(x_{t-1}) \right| = \left| \gamma \left(1 - M^f\right) - \theta \left(x_t^* - M^f x_{t-1}^*\right) \right|$$

which is the announced expression (50).

If this inequality is violated, then the optimum is remains by (49).

Proof of Proposition 12 From the Phillips curve, the steady state values π, x satisfy

$$\pi = \frac{\kappa x + \nu}{1 - \beta M^f} \tag{71}$$

and the optimality condition (49) gives: $\kappa \operatorname{sign}(\pi) = \left(1 - M^f\right) g'(x)$ i.e.

$$\kappa \operatorname{sign}(\pi) = \left(1 - M^f\right)(\gamma - \theta x) \tag{72}$$

We have $\pi = 0$ iff $x = \frac{-\nu}{\kappa}$ and (72) holds, i.e. $\kappa \geq \left| \left(1 - M^f\right)(\gamma - \theta x) \right|$. Hence, the boundary condition is $\pm\kappa = \left(1 - M^f\right)(\gamma - \theta x)$, i.e. $x = \frac{\gamma}{\theta} - \frac{\pm\kappa}{\theta\left(1 - M^f\right)}$, as announced. The corresponding value of ν is $\nu_\pm := -\kappa x_\pm$.

When $\nu > \nu_+$, (so inflationary pressures are very large), we have $\pi > 0$, so $\kappa = \left(1 - M^f\right)(\gamma - \theta x)$, and $x = x_+$. Then, (71) gives $\pi = \frac{\nu - \nu_+}{1 - \beta M^f}$ Likewise, when $\nu < \nu_-$, we have $\pi < 0$ and $x = x_-$.

Proof of Proposition 13 Welfare is:

$$W_t = -\omega_{\bar{\pi}} |\bar{\pi}| - \omega_{\hat{\pi}} |\hat{\pi}_t| + g(x_t) + f_t$$

where f_t is the "efficiency" from term. As temporarily inflation is less well perceived than steady state one, we have $\omega_{\hat{\pi}} \geq \omega_{\bar{\pi}} \geq 0$. Around the steady state, $f_t = f(\bar{\pi} + \hat{\pi}_t)$. Taking the Taylor expansion around $\hat{\pi}_t = 0$, we get:

$$f_t - f(\bar{\pi}) = f'(\bar{\pi})\,\hat{\pi}_t = \omega_{\bar{\pi}}\hat{\pi}_t$$

so $W_t = \bar{W} - \omega_{\hat{\pi}}|\hat{\pi}_t| + \omega_{\bar{\pi}}\hat{\pi}_t + g(x_t)$ i.e.

$$W_t = \bar{W} - h(\hat{\pi}_t) + g(x_t) \tag{73}$$

with

$$h(\hat{\pi}_t) = \omega_{\hat{\pi}}|\hat{\pi}_t| - \omega_{\bar{\pi}}\hat{\pi}_t = \omega_{\hat{\pi}}^{-}|\hat{\pi}_t|\,1_{\hat{\pi}_t<0} + \omega_{\hat{\pi}}^{+}|\hat{\pi}_t|\,1_{\hat{\pi}_t>0}$$

with $\omega_{\hat{\pi}}^{+} := \omega_{\hat{\pi}} - \omega_{\bar{\pi}}, \omega_{\hat{\pi}}^{-} := \omega_{\hat{\pi}} + \omega_{\bar{\pi}}$. so, we get an asymmetric kink in $\hat{\pi}_t$.

Proof of Proposition 16 With a fraction ε of rational consumers, the welfare function has now a term:

$$W = W^0(1-\varepsilon) + \varepsilon\frac{1}{2}\psi\sigma_x^2\left(b - \frac{b^2}{2}\right)$$

where W^0 is the social welfare with behavioral aversion, and $\frac{1}{2}\psi\sigma_x^2\left(b - \frac{b^2}{2}\right)$ is the social surplus linked to rational consumers.

With first order CA, we had (**??**). Now, this gives:

$$W \leq \underline{W}(b) := -(1-\varepsilon)\min\left(kd^{1-\gamma}, d\right) + b.$$

So, the exact same reasoning as in the proof of Proof of Proposition 14 shows that, if ξ is large enough, $\underline{W}(b) < 0$ for $b > 0$. So, $b = 0$ is preferable.

With second order CA, (74) implies that $W^0 \geq -Eb^2$ for a positive constant E. So, we have

$$W \geq -(1-\varepsilon)Eb^2 + \varepsilon b$$

hence social surplus W is strictly positive for a small b. As a result, some $b > 0$ is preferable to $b = 0$.

Proof of Proposition 17 I build on the proof of Proposition 15. I use the attention function in (4). Let us analyze the incentive to go from $b = 0$ to a small $b > 0$. First, we have $\chi = \frac{1}{2}\psi\sigma_x^2 b^2$, $m = \left(\bar{\xi}\frac{\chi}{wc}\right)^{\frac{1}{\alpha}-1} = \bar{k}b^{\frac{2}{\alpha}-2}$, for a positive constant $\bar{k}$ independent of b, and one can show that the cost is: $C = (1-\alpha)m\bar{\xi}\chi = O(b^2)m$, so that, using (23), in the neighborhood of $b = 0$, we have

$$\begin{aligned} W &= \left(b - \frac{1}{2}b^2\right)\psi\sigma_x^2 m + m\bar{\xi}\frac{1}{2}\psi\sigma_x^2 b^2 - C(m) - \bar{\xi}\frac{1}{2}\psi\sigma_x^2 b^2 \\ &= mb\left(D + o(1)\right) - Eb^2, \qquad \text{where } D, E \text{ are positive constants} \\ W &= \bar{k}b^{\frac{2}{\alpha}-1}\left(D + o(1)\right) - Eb^2 \end{aligned} \tag{74}$$

Hence, if $\alpha \in \left(\frac{2}{3}, 1\right)$ we have $\frac{2}{\alpha} - 1 < 2$, so $W(b) > 0$ for b in a neighborhood of 0: locally $b = 0$ is worse than $b > 0$.

If $\alpha < \frac{2}{3}$, then the reverse is true then, $W(b) < 0$ for small b's, and $b = 0$ is a local optimum. One can also show that, if ξ is large enough, this is a global optimum, as in the proof of Proposition 3.