An Attention-Based Theory of Mental Accounting*

Botond Kőszegi† and Filip Matějka‡

October 7, 2019

Abstract

We analyze how an agent with costly attention optimally attends and responds to taste, consumption-opportunity, and price shocks in basic multi-product consumption problems, explaining several types of mental accounting and making other predictions. If the problem is to choose the consumption levels of many goods with different degrees of substitutability, the agent may create budgets for the more substitutable products (e.g., entertainment). In some situations, it is optimal to specify budgets in terms of consumption quantities, but when most products have an abundance of substitutes, specifying budgets in terms of nominal spending tends to be optimal. If the goods are complements, in contrast, the agent — consistent with naive diversification — may choose a fixed, unconsidered mix of products. And if the agent's problem is to choose one of multiple products to fulfill a given consumption need (e.g., for gasoline or a bed), it is often optimal for her to allocate a fixed sum for the need.

Keywords: mental accounting, naive diversification, rational inattention.
JEL Codes: D01, D11, D14

* We thank Yuriy Gorodnichenko, Paul Heidhues, Rupal Kamdar, Marc Kaufmann, David Laibson, Andrei Shleifer, Philipp Strack, and seminar participants at BI Norwegian Business School, Carnegie Mellon University, Cornell University, D-TEA, Harvard University, LSE, Oxford University, Princeton University, Queen Mary University, UC Berkeley, UC Davis, University of Copenhagen, SITE, Washington University, and Wharton for excellent comments. This project has received funding from the European Research Council (grant agreement numbers 678081 and 788918) and the UNCE project (UNCE/HUM/035).
†Central European University. botondkoszegi@gmail.com
‡CERGE-EI, a joint workplace of Charles University in Prague and the Economics Institute of the Czech Academy of Sciences. filip.matejka@cerge-ei.cz

1 Introduction

Individuals and households must make a myriad decisions on how to allocate money in the face of many competing uses and a barrage of relevant information. Thaler (1985, 1999), and the literature following him, propose that to help solve such allocation problems, individuals create different virtual "mental accounts" for different purposes (entertainment, clothing, etc.), and treat these accounts as separate when responding to changes in circumstances. Researchers have interpreted many empirical, experimental, and anecdotal observations as signs of mental accounting, making the concept one of the most commonly invoked ideas in behavioral economics. Yet there is no theory that explains how a person creates separate mental accounts from fungible finances, and how this process interacts with her reactions to shocks.

In this paper, we formulate a theory of expenditure allocation based on the premise that attention is costly, and therefore individuals expend it only on types of information that are sufficiently valuable. In consumption problems, this means that the information a person pays attention to depends systematically on her preferences and the economic situation, allowing us to explain mental accounting, to connect mental accounting to naive diversification — a phenomenon that has hitherto been treated separately in the literature — and to make other predictions.

After illustrating the logic of our results in a simple example in Section 2, in Section 3 we develop tools for analyzing the effects of costly attention on decision-making when both a person's action and her information are multidimensional. Our general methods, which extend the rational-inattention approach of Sims (2003) using the water-filling algorithm from information theory (Telatar, 1999, Cover and Thomas, 2006), are also of independent interest as they are likely to apply to many economic situations.

In Section 4, we turn our main topic, consumption decisions with costly attention. We first consider how a person allocates expenditure when she faces uncertainty about her preferences or consumption opportunities (but not prices), and she can reduce any aspect of that uncertainty through costly attention. We assume that goods can be grouped into nested consumption categories, and they are more substitutable within than between categories. For instance, a restaurant dinner and a play could both be in the category "entertainment" under the larger category "discretionary spending," with the two being more substitutable with each other than either is with products outside the entertainment category.

Our main result says that the agent often behaves as if she had separate mental accounts for separate categories: (i) consumption in a category is independent of shocks to other categories, and (ii) total consumption is unresponsive, but individual consumption levels are smoothly responsive, to shocks within the category. In a classical consumption problem, (i) holds only if utility is separable across categories — which our model does not assume — and (ii) does not hold for any natural utility function. Intuitively, the most relevant consideration for the agent to think about is which of multiple highly substitutable products are worth buying, so if she has sufficiently costly attention, she thinks only about this consideration. As a result, she does not think about shocks to the optimal level of consumption, and hence her budget is fixed. Even when she does not have such a hard budget, her spending in a category varies less than with full information, so she can be interpreted as having a soft budget.

Our budgeting result helps explain evidence that many individuals and households separate expenditures into budgetary categories (Rainwater et al., 1959, Kahneman and Tversky, 1984, Lave, 1995, Ameriks et al., 2003, Antonides et al., 2011), and makes the novel prediction that products are grouped into mental accounts according to their substitutability. And through a simple reinterpretation, our theory predicts that individuals may use budgeting strategies for other types of decisions, for instance allocating separate time budgets for substitute tasks.

We demonstrate in our simple example that mental budgeting can interact in an economically interesting way with budget constraints. Much like an unconstrained agent, a budget-constrained agent may prefer not to think about how much to consume in total. This implies that if her budget constraint is relatively tight, she always exhausts her budget — despite lower consumption being optimal with some probability. For budget-constrained individuals, therefore, costly attention increases consumption as well as the marginal propensity to consume out of increases in the budget.

An entirely different prediction emerges when we assume that the products are complements, and (similarly to the case of substitutes) they are more complementary within than between categories. Because the optimal consumption levels of complementary products tend to move together, the agent may now not think about her relative value for products at all, only about how much she should consume in total. Hence, she may choose a fixed, unconsidered mix of products. We argue that this prediction is consistent with the phenomenon of naive diversification in financial (Benartzi and Thaler, 2001, 2007) and consumption (Simonson, 1990) decisions. For instance, if the products are different funds in an employer-based retirement program, and the agent knows nothing about the funds to start with, then she may follow the $1/N$ rule, investing equal amounts in the available funds. Mental budgeting and naive diversification can therefore be viewed as solutions to the same type of decision-making problem that apply in different circumstances.

In Section 5, we ask whether the agent still wants to set budgets for substitute products when there are price shocks, and whether she prefers budgets expressed in *quantities of consumption* or *amounts of spending*. Accordingly, we allow the agent to make and execute plans in two different ways: she can choose the quantity of consumption for each product, or she can choose the amount of spending on each product. While these two ways of thinking are equivalent in a classical consumer problem in which prices are known, in our framework and with price uncertainty — in which the agent may not fully learn prices before making decisions — they are not equivalent. We establish that thinking in terms of spending is optimal whenever optimal total consumption is sufficiently price sensitive, or there are sufficiently many substitutable products in a category. Intuitively, fixing the amount to be spent on a product means that consumption responds to unforeseen changes in the product's price, and this is optimal if an average of the relevant optimal price elasticities (both the elasticities of substitution and elasticities of total consumption) is sufficiently high. Furthermore, we show that a consumer who thinks in terms of spending often sets spending budgets. These results explain the prevalence of spending budgets as well as the greater prevalence of spending budgets among (generally more price-sensitive) lower-income households, but they also predict consumption budgets in some plausible circumstances. For instance, a rich time-constrained consumer who is not price-sensitive may set an entertainment budget in consumption quantities, such as the number of nights out per month.

In Section 6, we consider a variant of our model in which the agent has unit demand for each product — e.g., she needs a single mattress or computer to replace her old one or a given amount of gasoline to drive that month — but has multiple versions of each product to choose from. Similarly to above, we ask whether deciding the version of the product (e.g., the grade of gasoline) to buy or the amount to spend is optimal for a consumer who does not process all price information before making decisions. We show that thinking in terms of spending is optimal if and only if product prices are on average sufficiently positively correlated with premiums for better products. And thinking in terms of spending implies, in line with evidence by Hastings and Shapiro (2013) on gasoline purchases, that when prices for all varieties of a product rise, the agent switches to a lower-priced variety. While in Hastings and Shapiro's setting the price and price premium are not positively correlated, our explanation applies if such situations are sufficiently uncommon, and consumers do not think about the correlation separately for gasoline.

In Section 7, we discuss how our model relates to existing theories and the many distinct phenomena labeled mental accounting. Our paper is about mental accounting as a decision-making aid when multiple uses of money compete for finances. While previous work explores another central aspect of this issue, self-control problems (Shefrin and Thaler, 1988, Galperti, forthcoming), we are the first to explain how a person creates budgets from fungible finances and how this affects her reactions to shocks, as well as to connect mental accounting formally to naive diversification. Our paper does not study mental-accounting phenomena that pertain to the framing and evaluation of individual transactions when tradeoffs with other products are not explicit.

We conclude in Section 8 by mentioning mental-accounting phenomena that our current model cannot explain, but (we argue) closely related attention-based models can. We also add, however, that it would be fruitful to study the interaction between our mental-accounting framework and others, especially self-control problems and loss aversion.

2 Example

In this section, we illustrate the logic of our budgeting result, and its relationship with naive diversification, using a simple example. We substantially generalize this example, and derive other predictions, in Section 4. The agent chooses the consumption levels of two goods, y_1 and y_2, to maximize the expectation of

$$(\overline{x}+x_1)y_1+(\overline{x}+x_2)y_2-\frac{y_1^2}{2}-\frac{y_2^2}{2}-\theta y_1 y_2-(y_1+y_2), \tag{1}$$

where $\overline{x}$ is her average taste for the goods, x_1 and x_2 are independent taste shocks drawn from $N(0,1)$, and $\theta \in (-1,1)$ is a substitutability parameter, with the goods being substitutes for $\theta > 0$ and complements for $\theta < 0$. The price of both goods is 1, and the disutility of spending \$1 is also 1, so $y_1 + y_2$ is the total disutility of spending. That prices are deterministic (and equal) means that consumption budgets and spending budgets are equivalent; in Section 5, we consider which type of budget the agent prefers.

Before choosing y_1 and y_2, the agent can observe exactly one of x_1, x_2, x_1+x_2, and x_1-x_2: she can think about her taste for one of the goods or her total or relative taste for the two goods. We ask: what does she optimally choose to think about, and how does this affect her consumption?

To facilitate an answer, we put the problem in a different form. Instead of working with the tastes x_1 and x_2, we work with the relative and total tastes, $x_- = x_1 - x_2$ and $x_+ = x_1 + x_2$; and instead of solving for the consumption levels y_1 and y_2, we solve for the relative and total consumption levels, $y_- = y_1 - y_2$ and $y_+ = y_1 + y_2$. Up to a function of x_1 and x_2 — which the agent cannot influence — the objective (1) can then be written as

$$-\frac{\Big(x_- - (1-\theta)y_-\Big)^2}{2(1-\theta)} - \frac{\Big(x_+ - (1+\theta)y_+\Big)^2}{2(1+\theta)} + (\overline{x}-1)y_+. \tag{2}$$

To maximize her expected utility conditional on her information, the agent therefore chooses

$$y_- = \frac{E[x_-|\text{info}]}{1-\theta} \quad \text{and} \quad y_+ = \frac{(\overline{x}-1)+E[x_+|\text{info}]}{1+\theta}. \tag{3}$$

The agent's maximization problem is equivalent to minimizing the expected loss relative to perfect information (knowing x_- and x_+). Plugging the optimal y_- and y_+ from (3) into (2), this is

$$\frac{var[x_-|\text{info}]}{2(1-\theta)} + \frac{var[x_+|\text{info}]}{2(1+\theta)}. \tag{4}$$

Optimal information acquisition is now obvious from how information affects the variances of x_- and x_+. If the products are substitutes (i.e., $\theta > 0$, and therefore $1/(1-\theta) > 1/(1+\theta)$), then the agent chooses to observe x_-. Since x_- and x_+ are independent, observing x_- provides no information about x_+, so $y_+ = ((\overline{x}-1)+E[x_+])/(1+\theta) = (\overline{x}-1)/(1+\theta)$. This means that y_1+y_2 is constant: the agent has a fixed budget determined by her average taste $\overline{x}$ for the products. Since $y_- = x_-/(1-\theta)$, however, the consumption levels y_1 and y_2 are not fixed — the agent does respond to changes in circumstances, but not by changing her total budget.

If the products are complements (i.e., $\theta < 0$, and therefore $1/(1-\theta) < 1/(1+\theta)$), then the agent chooses to observe x_+. As a result, she learns nothing about x_-, so $y_- = E[x_-]/(1-\theta) = 0$. This means that $y_1 = y_2$: the agent naively diversifies, always choosing the goods in equal proportion.

Since $y_+ = x_+/(1+\theta)$, however, the consumption levels y_1 and y_2 are not fixed — the agent does think about the problem, but not by changing the ratio in which she buys the products.

We use variants of our simple model to make a few further points. First, our model assumes that x_1 and x_2 are independent. If x_1 and x_2 are positively correlated, then $var[x_+] > var[x_-]$, which by Equation (4) increases the value of observing x_+. Hence, in this case naive diversification is more likely to occur. Intuitively, if the tastes for two products are highly positively correlated, then the consumer is unlikely to learn much from thinking about which one she likes. Conversely, a negative correlation between x_1 and x_2 increases the value of observing x_-, increasing the tendency toward budgeting.

In the case of consumer products, it is difficult to know what correlation between tastes or consumption opportunities is most realistic. But in the case of retirement investments (the primary example of naive diversification), it is likely that preferences are positively correlated. That one fund is a good investment reflects in part that employer-sponsored retirement savings is a good investment in general, and therefore other funds in the program are good investments as well. Hence, the motive for naive diversification is even stronger than for independent preferences.

Second, by treating the disutility of spending money as a constant, we have implicitly assumed that the agent knows it or does not want to think about it. Uncertainty in the value of saving affects the disutility of spending on both products equally, so — if the agent can lower the uncertainty through thinking — it is equivalent to a positive correlation between x_1 and x_2.[1] If the value of saving is highly uncertain, therefore, our budgeting result fails. In this sense, figuring out one's value of saving to a point where one no longer wants to think about it much is a precursor to budgeting. Since the value of saving tends to change slowly, this precondition is likely to hold in most periods of one's life. For investments, in contrast, uncertainty in the value of money only strengthens the tendency toward naive diversification.

Third, consider also what happens when the goods are substitutes, and the agent has a relatively tight budget constraint $y_+ \leq y_+^{max} \leq (\overline{x}-1)/(1+\theta)$. This means that without information, the constraint would be binding, with the agent choosing $y_+ = y_+^{max}$ and $y_- = 0$. Since observing x_- is only useful for choosing y_-, the constraint — which does not restrict y_- — leaves the value of

[1] To see this formally, let the disutility of spending be $1+\mu$, with μ being the uncertainty in the value of money, and let x'_1 and x'_2 be the independent taste shocks. The agent's utility is then $(\overline{x}+x'_1)y_1 + (\overline{x}+x'_2)y_2 - y_1^2/2 - y_2^2/2 - \theta y_1 y_2 - (1+\mu)y_1 - (1+\mu)y_2$. Setting $x_1 = x'_1 - \mu$ and $x_2 = x'_2 - \mu$ gives Expression (1), where x_1 and x_2 are now positively correlated.

observing x_- unchanged. In contrast, since observing x_+, x_1, or x_2 is useful for choosing y_+, the constraint — which prevents increases in y_+ in response to news — decreases the value of observing any of these variables. Hence, the agent still prefers to observe x_-, and her total consumption is $y_+ = y_+^{max}$, i.e., she always exhausts her spendable funds. Intuitively, while consuming less might be optimal, thinking about this is less valuable than thinking about how to split her spendable funds between the goods. When the budget constraint is relatively tight, therefore, costly attention increases consumption. Furthermore, because the agent's marginal propensity to consume out of increases in available funds equals the probability with which her budget constraint binds, costly attention also increases the marginal propensity to consume from as low as 1/2 to 1.

3 Theoretical Tools

In this section, we develop a methodology for analyzing rational-inattention models in which — as with mental accounting — the agent's information and action are multi-dimensional.[2] Since these tools are potentially applicable to many economic settings, we present them in a general form. We lay out our results on mental accounting in a self-contained way, so readers not interested in the general tools can skip to Section 4.

3.1 Multi-Dimensional Rational Inattention

The agent maximizes the expectation of the utility function $U(\mathbf{y}, \mathbf{x})$, which depends on an exogenous random vector of states $\mathbf{x} \in \mathbb{R}^J$ and her chosen vector of actions $\mathbf{y} \in \mathbb{R}^N$, less the cost of information processing. U takes the form

$$U(\mathbf{y}, \mathbf{x}) = -\mathbf{y}'C\mathbf{y} + \mathbf{x}'B\mathbf{y},$$

where $B \in \mathbb{R}^{N\times J}$, $C \in \mathbb{R}^{N\times N}$, and C is symmetric and positive definite. The matrix C summarizes interactions between actions, while B summarizes interactions between states and actions. We assume that the prior uncertainty about $\mathbf{x}$ is multivariate Gaussian with the variance-covariance matrix ψ. To focus on the allocation of attention driven by preferences only, we let $\psi = \sigma_0^2 I$.

[2] For some previous applications of rational inattention, see Veldkamp (2006), Mackowiak and Wiederholt (2009), Woodford (2009), Luo and Young (2014), Caplin and Dean (2015), Matějka and McKay (2015), and Matějka (2016). See Mackowiak et al. (2018) for a review. Recent papers by Miao et al. (2019) and Verstyuk (2019) also solve models of multi-dimensional rational inattention.

Before choosing $\mathbf{y}$, the agent can obtain any Gaussian signal about $\mathbf{x}$. The resulting posterior beliefs are also Gaussian, with the agent being able to choose the posterior variance-covariance matrix Σ subject to the constraint that $\psi - \Sigma$ is positive definite — i.e., that the posterior is more precise than the prior. Denoting by $|\cdot|$ the determinant of a matrix, we posit that the cost of information is $(\lambda/2) \cdot (\log|\psi| - \log|\Sigma|)$, where $\lambda \geq 0$ is the agent's attention cost. This specification of decision-making with costly attention is the reduced form of a general rational-inattention model in which the agent can obtain not just Gaussian signals, but any signal at a cost equal to the reduction in the entropy of her beliefs.[3]

As in the previous literature, there are three main reasons for using the entropy-based functional form for attention costs. First, it is highly tractable. Second, it has the basic property that information is costly (if the agent learns $\mathbf{x}$ more precisely, then $|\Sigma|$ is lower, and therefore $(\lambda/2) \cdot (\log|\psi| - \log|\Sigma|)$ is higher). Third, it implies that all information has the same cost — what matters is the amount of uncertainty reduction, not what the uncertainty is about — so it can be viewed as ideal for studying information acquisition based on endogenous considerations about the benefits of information, and not based on exogenous assumptions about the costs of information.

At the same time, researchers have raised various concerns about specifying attention costs to be linear in entropy reduction. Woodford (2012) points out that the entropy-based cost function fails to predict the finding from perceptual experiments that subjects make smaller errors in more likely states. Dean and Neligh (2017) find that experimental subjects' behavior is consistent with a cost function that is convex in entropy reduction. Similarly, Morris and Strack (2017) establish that a constant marginal cost of signals in sequential information-acquisition problems corresponds to a convex entropy-based cost function. Accordingly, theoretical work generalizes the entropy-based cost function to allow for differences in comparison costs across versus within nests of products (Fosgerau et al., 2017), on different dimensions of the state space (Pomatto et al., 2019), and for nearby versus distant states (Morris and Yang, 2016), and Caplin and Dean (2015) study a broader class of cost functions called posterior separable. With alternatives going beyond entropy-based costs, our decision problem would be difficult or impossible to analyze. It seems clear, however, that such extensions would not affect the logic of our results, but would merely add the consideration

[3] Sims (2003) shows that in a rational-inattention model with entropy costs, it is optimal for an agent with our linear-quadratic consumption utility to collect Gaussian signals; hence, we simply assume that the agent does so. In addition, the entropy of a Gaussian distribution with variance-covariance matrix Σ is a constant plus $\log|\Sigma|/2$.

that the agent is more likely to obtain less costly information.

3.2 Optimal Information Acquisition and Actions

In the proof of Proposition 1 in Appendix A, we describe the following steps in detail. We first show that the agent's objective, expected utility less the cost of information, can be written as

$$-E\left[(\tilde{\mathbf{x}}-\mathbf{x})'\Omega(\tilde{\mathbf{x}}-\mathbf{x})\right]+\frac{\lambda}{2}\log|\Sigma|, \tag{5}$$

where $\Omega = BC^{-1}B'/4$ and $\tilde{\mathbf{x}}$ is the random mean of the posterior beliefs about $\mathbf{x}$, which depends on the realization of noise in signals. The first term in (5) is the expected loss from misperceptions $(\tilde{\mathbf{x}}-\mathbf{x})$, which are distributed according to $N(0,\Sigma)$ and translated into losses by Ω. The second term is the cost of information, with the constant $(\lambda/2)\cdot\log|\psi|$ dropped.

Decomposition into One-Dimensional Problems. Let $v^1,\ldots,v^J$ be an orthonormal basis of eigenvectors of the loss matrix Ω (which is symmetric), with the eigenvalue corresponding to v^i denoted by Λ_i. The utility term in (5) can be conveniently expressed using the transformation of coordinates to this basis. Letting $(\tilde{\mathbf{x}}-\mathbf{x})=\sum_i \tilde{\eta}_i v^i$, we have

$$(\tilde{\mathbf{x}}-\mathbf{x})'\Omega(\tilde{\mathbf{x}}-\mathbf{x})=\Big(\sum_i \tilde{\eta}_i v^i\Big)'\Omega\Big(\sum_i \tilde{\eta}_i v^i\Big)=\sum_i \Lambda_i\tilde{\eta}_i^2.$$

The eigenvalue Λ_i is thus a scaling parameter for how uncertainty about the linear combination $(v^i\cdot\mathbf{x})$ translates into losses. Now the expectation of $\tilde{\eta}_i^2$ is by definition the posterior variance of $v^i\cdot(\tilde{\mathbf{x}}-\mathbf{x})$. Since the x_i are i.i.d. with prior variance σ_0^2, the random variables $(v^i\cdot\mathbf{x})$ are also i.i.d. with prior variance σ_0^2. Let us denote the posterior variance of $(v^i\cdot\mathbf{x})$ by $\sigma_i^2\le\sigma_0^2$. In the proof we show that Σ must be diagonal in the basis of the eigenvectors, and thus $\log|\Sigma|=\sum_i\log\sigma_i^2$. The agent's problem therefore reduces to

$$\max_{\sigma_i^2\le\sigma_0^2}\Big(\sum_i -\Lambda_i\sigma_i^2+\frac{\lambda\log\sigma_i^2}{2}\Big). \tag{6}$$

This can now be solved separately for each i, yielding a simple information-acquisition strategy:

Proposition 1 (Information Acquisition)**.** *The optimal information-acquisition strategy is to acquire independent signals of $v^i\cdot\boldsymbol{x}$ such that the posterior variance of $v^i\cdot\boldsymbol{x}$ is* $\min\{\sigma_0^2,\lambda/(2\Lambda_i)\}$.

Intuitively, the agent processes more information about vectors in the space of $\mathbf{x}$ that are more costly to misestimate. Specifically, if $\sigma_0^2 \leq \lambda/(2\Lambda_i)$, then the agent acquires no information about $v^i \cdot \mathbf{x}$; and if $\sigma_0^2 > \lambda/(2\Lambda_i)$, then she observes a signal about $v^i \cdot \mathbf{x}$ with precision chosen to bring the posterior variance of $v^i \cdot \mathbf{x}$ down to $\lambda/(2\Lambda_i)$. Hence, when the cost of information λ is high ($\lambda/(2\Lambda_i) > \sigma_0^2$ for all i), then the agent does not process any information. If the cost is somewhat lower, then the agent processes information about the $v^i \cdot \mathbf{x}$ with the highest Λ_i, but she processes no other information. At even lower costs, the agent processes information about more $v^i \cdot \mathbf{x}$, etc.

Responsiveness of Actions. Next, we discuss implications for actions. We show in the appendix that $\mathbf{y} = H\tilde{\mathbf{x}}$, where $H = C^{-1}B'/2$. We define ε_i^λ as the average change in the action $\mathbf{y}$ when $\mathbf{x}$ changes in direction v^i by 1. We can think of it as the average responsiveness of the agent's behavior to shocks along v^i. Using this notation, the responsiveness under perfect information — when the agent has no attention costs — is ε_i^0.

Proposition 2 (Optimal Actions)**.**

1. *The space of actions is spanned by* $\{Hv^i | \lambda/(2\Lambda_i) < \sigma_0^2\}$.
2. *The agent underresponds to shocks relative to the perfect-information case* ($\varepsilon_i^\lambda < \varepsilon_i^0$)*, with*

$$\frac{\varepsilon_i^0 - \varepsilon_i^\lambda}{\varepsilon_i^0} = \min\left(1, \lambda/(2\sigma_0^2\Lambda_i)\right). \tag{7}$$

3. *In the range* $\Lambda_i > \Lambda_j > \lambda/(2\sigma_0^2)$*, the relative responsiveness* $\varepsilon_i^\lambda/\varepsilon_j^\lambda$ *is strictly increasing in* λ.

Part 1 says that the agent's action moves only along directions that are sufficiently important to pay attention to — that is, along the directions in which losses are highest. Part 2 says that the agent underresponds to shocks. Because the agent pays only partial attention to information, on average she does not notice the extent of shocks, so she does not respond as much as an agent with zero attention costs. More interestingly, Part 3 says that with costly attention, optimal behavior calls for concentrating reactions to shocks in directions that are the most important. As a result, the responsiveness to shocks along v^i relative to v^j is higher than under perfect information if and only if $\Lambda_i > \Lambda_j$.

4 Consumption Patterns

We now apply the tools from Section 3 to analyze how a person attends and responds to taste or consumption-opportunity shocks when choosing a consumption basket from many products with different degrees of substitutability or complementarity. We analyze price shocks in the next section.

There are N goods, each of which has price equal to 1. The agent's utility from consumption or, equivalently, spending levels $y_1, \ldots, y_N \in \mathbb{R}$ is

$$-\sum_m y_m^2 - \sum_{m \neq n} \Theta_{mn} y_m y_n + \sum_m (\overline{x}_m + x_m) y_m - \sum_m y_m, \tag{8}$$

where $\Theta \in \mathbb{R}^N \times \mathbb{R}^N$ with $\Theta_{mm} = 1$ is a symmetric positive definite matrix that captures the substitutability patterns between the goods, $\overline{x}_m$ is the baseline marginal utility of consuming good m, and x_m is a shock to this marginal utility. Uncertainty in x_m could arise from uncertainty about taste — the agent does not know what combination of restaurant dinners, laptops, housing amenities, etc. maximizes her well-being — or from shocks to consumption opportunities — e.g., if better bands happen to be in town, then the marginal utility of going to concerts is higher. Finally, $\sum_m y_m$ is the disutility of spending money.

To be able to analytically solve and economically interpret our model, we posit a specific structure for Θ. In particular, the goods can be grouped into $L \geq 1$ levels of categories. The level $l = L$ is the largest category (e.g., discretionary spending), which includes all N goods; the level $l = L - 1$ is the set of second-largest categories (e.g., entertainment), and so on, with the smallest ($l = 1$) categories being individual consumption goods (e.g., a dinner out). We denote by $R^{k,l} \subset \{1, \ldots, N\}$ the consumption category k at level l. We assume that all categories at level l are of the same size ($|R^{k,l}| = |R^{k',l}|$ for all k, k', l), and that each category at level $l < L$ is a subset of a higher category (for each $l < L, k$, there is a k' such that $R^{k,l} \subset R^{k',l+1}$). The substitutability of two goods is determined by the smallest category to which they both belong. For two goods m and n, let l be the smallest l' such that there is a k with $m, n \in R^{k,l'}$. Then, $\Theta_{mn} = \gamma^l$, where γ^2 through γ^L are constants.[4]

We posit that the x_m are i.i.d. normal random variables with mean zero and variance σ_0^2, and

[4] Note that we have introduced the notion of categories merely to facilitate the definition of the substitutability matrix Θ and the statement of our results; we do not presume that the agent thinks of goods in the same category separately from other goods.

the agent can obtain any multivariate normal signal about $(x_1, \ldots, x_N)$. Part of this thinking could, for instance, involve mentally simulating future consumption (as in Gabaix and Laibson, 2017), or searching for information about consumption opportunities. The agent's cost of attention is the same as in Section 3, so that she maximizes the sum of her expected utility given her posterior beliefs plus $\lambda \log |\Sigma|/2$, where $\lambda \geq 0$ is her cost of attention, Σ is the variance-covariance matrix of her posterior, and $|\Sigma|$ is the determinant of Σ.

The assumption that the agent can think about the vector $(x_1, \ldots, x_N)$ in a fully flexible way is of course unrealistic. For instance, it is unlikely that one can obtain a noisy signal of an arbitrary linear combination of this month's entertainment programs. At the same time, there is clearly flexibility in what a person thinks about or focuses on, and our framework captures such flexibility without making potentially ad-hoc assumptions on its limits. Fortunately, the optimal solution we identify below involves highly plausible and intuitive ways of thinking. Hence, if we allowed only plausible ways of thinking, the same solutions would obtain. Furthermore, note that the attention cost in our framework can be reinterpreted as a calculation cost when the agent knows her tastes or consumption opportunities (or, in Section 5, prices), but without thinking does not know what they imply for optimal consumption. It is plausible that one can perform such optimization calculations flexibly.

As a benchmark, we identify how the agent behaves if she has costless attention, and how she responds to ex-ante known changes (i.e., changes in the $\overline{x}_m$). For instance, the agent's average taste may evolve over time. To state the result, let $\mathbf{y} = (y_1, \ldots, y_N)', \overline{\mathbf{x}} = (\overline{x}_1, \ldots, \overline{x}_N)', \mathbf{x} = (x_1, \ldots, x_N)'$.

Fact 1. *If* $\lambda = 0$*, then* $\mathbf{y} = \Theta^{-1}(\overline{\mathbf{x}} + \mathbf{x})/2$*. For any* $\lambda \geq 0$*,* $E[\mathbf{y}] = \Theta^{-1}\overline{\mathbf{x}}/2$*.*

The agent's average behavior responds to ex-ante known changes in exactly the same way as with perfect information. This also means that her utility function (i.e., the matrix Θ) can be extracted from her responses to ex-ante known changes. As we show below, her responses to ex-post shocks she needs to think about are often markedly different, and by implication do not accurately reflect her true preferences over consumption. Nevertheless, these responses can be predicted from her (from ex-ante known shocks measurable) true preferences.

4.1 Substitutes: Mental Budgeting

First, we consider substitute goods, assuming that $0 < \gamma^L < \cdots < \gamma^2 < 1$. This captures the idea that a good is a better substitute for other goods in its category than for goods in a different category. For instance, a French dinner is a closer substitute to a Chinese dinner than to a movie. Then:

Proposition 3 (Hard Budgeting of Substitute Products)**.** *There are $\lambda_1, \ldots, \lambda_L$ satisfying $\lambda_L < \cdots < \lambda_1$ such that*

$$\lambda \geq \lambda_l \Leftrightarrow \sum_{m \in R^{k,l}} y_m = \textit{constant for all } k. \tag{9}$$

Proposition 3 says that if (and only if) her attention cost is sufficiently high, the agent has a fixed mental budget — a constant total expenditure — for each l-category of products. Accordingly, the higher is her cost of attention — e.g., because she has lower cognitive ability or is busy with other things — the more likely she is to budget, and the narrower are her budgets. To appreciate ways in which such behavior differs from that of a classical decision-maker, suppose that $\lambda_2 < \lambda < \lambda_1$, and one category at level 2 is entertainment. Denoting the entertainment category by R:

Corollary 1. *(i) For any $m \in R$ and $n \notin R$, y_m does not depend on x_n; (ii) $\sum_{m \in R} y_m$ is constant; and (iii) for any $m \in R$, $E[y_m|\mathbf{x}]$ is a function of the vector $(x_m - x_n)_{n \in R \setminus \{m\}}$ that is strictly increasing in each component.*

Corollary 1 implies two related phenomena. First, Part (i) says that the agent's consumption decisions regarding entertainment are independent of other shocks. In a classical consumption problem, this occurs only if the utility from entertainment is separable from the rest of the utility function. We do not impose such separability; in fact, with full information $\partial y_m / \partial x_n < 0$ for all $n \neq m$.[5] Second, Parts (ii) and (iii) imply that the agent's total consumption of entertainment is independent of shocks, but her consumption within the category responds smoothly to within-category shocks. This is in general not the case in any classical model.

Intuitively, knowing about a shock to the relative marginal utility of movies and theater is very valuable, as it allows for substantial readjustment of both consumption levels through substitution. Knowing about a shock to the relative marginal utility of movies and clothing is less valuable, since

[5] This observation follows from Fact 1 applied to the current problem. Given the structure we have imposed on Θ, the off-diagonal entries of Θ^{-1} are all negative.

the scope for substitution between these goods is lower. And knowing about a shock to the marginal utility of movies is also less valuable, as it leads mainly to the adjustment of movies consumption. With the agent's attention being costly, she thinks only about the most important thing, the relative utility of movies and theater. As a result, she fixes total entertainment consumption.

Proposition 3 explains evidence that many consumers have category-specific budgets. As a stark manifestation of this phenomenon, many households used to place budgets allocated for different purposes into different envelopes or tin cans (Rainwater et al., 1959, Lave, 1995). More recently, Ameriks et al. (2003) and Antonides et al. (2011) document that the mental budgeting (if not physical separation) of expenses is still common. Indeed, most of the many online financial management tools seem to presume that users want to set budgets for separate categories. To go further, Proposition 3 makes the novel prediction that the most substitutable goods go into the same budget.

The logic applies in other domains as well. For instance, there is evidence that some individuals have mental budgets for time allocation, such as hours per day devoted to studying (Rajagopal and Rha, 2009). This follows from our model by reinterpreting y_m as the time allocated to task m, and x_m as a shock to the return of working on task m. Furthermore, our theory predicts that a person creates budgets for substitute tasks, for instance different ways of studying for an exam.

Our model is static in the sense that the agent solves a single optimization problem over what information to obtain and what to consume. But she does not have to make all choices at the same time. When choosing budgets, she can leave her plans incomplete, and obtain information about shocks only when relevant consumption opportunities start arising, even making decisions separately for separate categories of products. This piecemeal execution is facilitated by the separable nature of the optimal plan, and is in fact optimal if obtaining the same information or mentally simulating consumption at the earlier budgeting stage is costlier.[6]

Having budgets leads to specific patterns in how a person reacts to shocks. Suppose, for instance, that the x_m in the entertainment category all increase by the same amount — i.e., unusually

[6] Technically speaking, at the budgeting stage it is necessary for the agent to understand exactly what she will do at the execution stage. Interpreted more broadly, it is sufficient for her to have (perhaps based on experience) a reasonable understanding of the average value of increasing her budget.

Relatedly, when the agent acquires information piecemeal, the question arises how costly each piece of information is. A simple assumption consistent with our formulation is that at each stage, the cost of information equals $(\lambda/2)(\log|\Sigma_0| - \log|\Sigma_1|)$, where Σ_0 are Σ_1 are the variance-covariance matrices of her previous and new beliefs, respectively.

fun entertainment opportunities present themselves across the board. Then, the agent's average consumption of entertainment as well as other goods remains unchanged. Since she evaluates entertainment goods only relative to each other, on average she does not see a reason to change her behavior. If she had unlimited attention, in contrast, she would respond to such positive shocks by increasing entertainment consumption and decreasing other consumption. Similarly, if a single x_m increases, that leads the agent to increase y_m. If she had full information, she would also decrease the consumption of all other goods. Because she has a budget, however, she concentrates the substitution to within the category.

Proposition 3 identifies a stark form of budgeting, in which the budget is completely fixed: if $\lambda \geq \lambda_l$, then the correlation between the consumption of a good and the total consumption of other goods in its l-category is -1. Beyond this extreme result:

Proposition 4 (Soft Budgeting of Substitute Products)**.** *Suppose $\lambda < \lambda_l$. For any k and $m \in R^{k,l}$, the correlation between y_m and $\sum_{n \in R^{k,l} \setminus \{m\}} y_n$ is strictly decreasing in λ.*

Proposition 4 says that the higher is the agent's attention cost, the more she restricts consumption adjustments to substitutions within a category. As a result, although her total consumption is not completely fixed, it varies less than one would expect based on her preferences. In this sense, she can be viewed as having a soft budget for l-categories.

Figure 1 illustrates Propositions 3 and 4 in an example. We consider four goods grouped into categories $\{1, 2\}$ and $\{3, 4\}$, and draw the joint distribution of y_1 and y_2 for different levels of λ. For costless attention ($\lambda = 0$), the distribution of possible consumption pairs is quite dispersed. At the other extreme, for very high attention cost ($\lambda = 1$), the consumption amounts are fixed. For lower, but relatively high attention costs ($\lambda = 0.75, 0.5$), the agent sets a budget for the two products, so her consumption is always on the same budget line. These situations correspond to Proposition 3. For even lower positive attention costs ($\lambda = 0.48, 0.45$), the agent starts substituting goods 1 and 2 with goods 3 and 4, but not as much as with costless attention, so the distribution of y_1 and y_2 is closer to a budget line than with costless attention. These situations correspond to Proposition 4.

Asymmetries in the prior variances of x_m or prices also lead to a kind of soft budgeting. To illustrate, suppose that $L = 2$ and $N = 4$ — there is a single category of four products — and the prior variances $\sigma^2_{0,m}$ satisfy $\sigma^2_{0,1} \neq \sigma^2_{0,2} = \sigma^2_{0,3} = \sigma^2_{0,4}$. We show in Appendix C that there are λ_1 and α such that if $\lambda \geq \lambda_1$, then $\alpha y_1 + y_2 + y_3 + y_4$ is constant, with numerical simulations

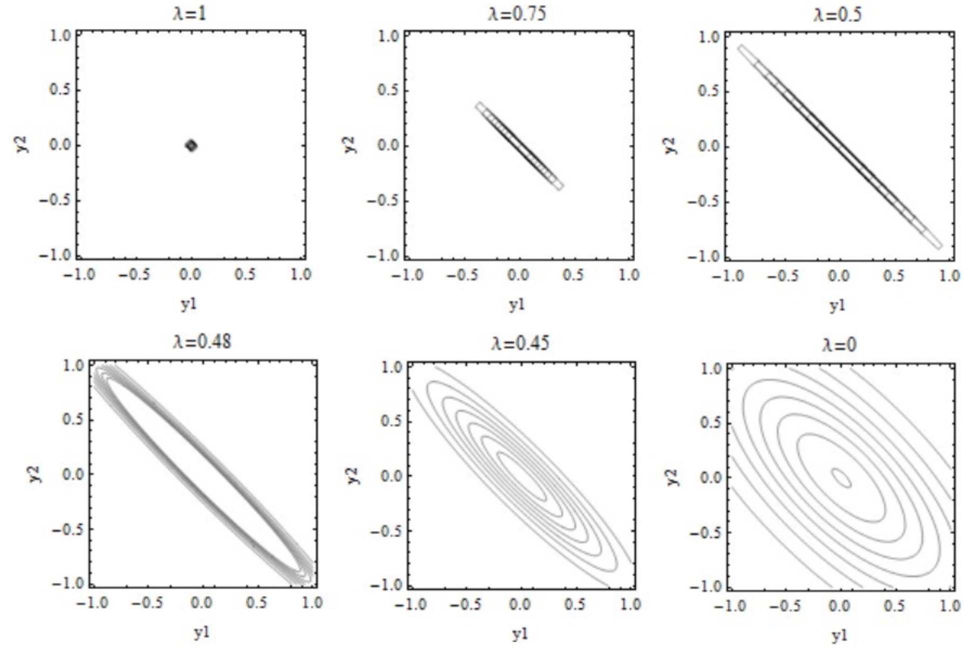

Figure 1: Joint distributions of consumption of goods 1 and 2 for $\sigma_0^2 = 1$ and $\gamma^2 = 1/2, \gamma^3 = 1/4$ under different costs of attention. Iso-density curves are shown.

indicating that $\alpha < 1$ if and only if $\sigma_{0,1}^2$ is greater than the other $\sigma_{0,m}^2$. Hence, total spending equals $y_1 + y_2 + y_3 + y_4 = \text{constant} + (1 - \alpha)y_1$. Furthermore, simulations show that unless the asymmetry is very large, an increase in y_1 is associated with a decrease in $y_2 + y_3 + y_4$ much more than with full information. This can be interpreted as saying that the agent has a soft target budget, allowing herself to go over the target if she happens to have a high value for a good with more volatile value. Relatedly, if good 1 has price $p_1 \neq 1$, then total spending is $p_1 y_1 + y_2 + y_3 + y_4 = \text{constant} + (p_1 - \alpha)y_1$: now the agent also allows herself to go over the target if she has a high value for a more expensive product. If choosing between cheaper chicken and more expensive beef, for instance, she allows herself to splurge when especially nice beef is available.

4.2 Complements: Naive Diversification

We turn to complementary products, assuming that $\gamma^2 < \cdots < \gamma^L < 0$. This means that products are arranged in a nested fashion into categories, with products belonging to smaller categories being stronger complements in consumption. For instance, different features of a car (e.g., driving experience, seats, sound system) might be highly complementary to each other, but not to one's furniture. To simplify our statement as well as to capture situations in which the products are ex ante equally desirable, we also assume that the $\overline{x}_m$ are equal. Then:

Proposition 5 (Naive Diversification)**.** *There are $\lambda_2, \ldots, \lambda_L$ satisfying $\lambda_2 > \cdots > \lambda_L$ such that*

$$\lambda \geq \lambda_l \Leftrightarrow \textit{for any } k \textit{ and any } m, n \in R^{k,l},\ y_m = y_n. \tag{10}$$

Proposition 5 says that if the agent's attention cost is sufficiently high, then she chooses a fixed mix of products in category l. This contrasts with the case of substitute products, where it was not the *mix*, but the *budget* that was fixed. Intuitively, because the optimal consumption levels of complementary products tend to move together, the agent does not think about their optimal relative consumption at all, only about how much she should consume overall. Continuing with the example of cars, the agent does not think separately about the quality of the engine, seats, sound system, etc. she wants — she only thinks about whether she wants an economy or luxury car.

An important application of the above result is naive diversification in financial decisions, whereby a person chooses a simple mix of investments that is unlikely to be fully optimal. For instance, Benartzi and Thaler (2001) document that many employees in employer-based retirement savings plans divide their investments equally across available funds, and relatedly, employees invest more in stocks if there are more stock funds available. Huberman and Jiang (2006) find a similar pattern for plans offering 10 or fewer funds, although not for plans offering more funds. To see how our model can account for this phenomenon in an example, suppose that an investor with mean-variance preferences decides the amounts y_1 and y_2 to invest into two assets. There are two equally likely states, with asset 1's net return being $x_1 + 1$ in state 1 and $x_1 - 1$ in state 2, and asset 2's net return being $x_2 - 1$ in state 1 and $x_2 + 1$ in state 2. It is easy to check that the mean of the investor's wealth is $x_1 y_1 + x_2 y_2$ and the variance is $(y_1 - y_2)^2$, so the utility function can be written in the form (8) with $\Theta_{12} = \gamma_2 = -1$. Hence, Proposition 5 predicts that an investor with sufficiently costly attention splits her investment equally between the two assets. More generally, because diversification is desirable, different investments are often complements, so Proposition 5 predicts that investors may diversify naively.[7]

[7] In the illustrative example above, the complementarity of the two investments relies on the asset returns being negatively correlated. Even for uncorrelated or somewhat positively correlated asset returns, investments are complements if the investor's disutility from variance is strictly concave. Furthermore, with a precautionary savings motive, risky and safe investments are often complements.

Investigating a completely different domain, Simonson (1990) finds that individuals naively diversify when choosing items to consume at different future dates.[8] Our model explains this finding if individuals both have a taste for variety — which is equivalent to complementarity — and are subject to taste shocks. Consistent with our perspective, Simonson argues that naive diversification is due to the combination of taste uncertainty and the desire to simplify the decision.

Our model predicts a type of naive diversification also for substitute products when the agent's attention cost is so high that she does not obtain any information. In this case, her consumption of all products is fixed at the ex-ante optimal level, and therefore the mix of products is fixed as well. But the more interesting type of naive diversification above, whereby the agent does pay some attention to her decision problem and still naively diversifies, is — for the case of independent shocks — only possible for complementary products.

The observation that the agent reacts to ex-ante known changes exactly as in the full-information case (Fact 1) qualifies Proposition 5 in an interesting way. For instance, suppose that an investor distinguishes between stock and bond funds, and knows that stocks are more valuable investments for her. Then she chooses more stock funds than bond funds, or might choose only stock funds. But if she considers stock funds as ex-ante identical, then she still naively diversifies within the class of stock funds. More generally, if the agent sees a reason to invest in only a handful of funds, but treats these funds as ex-ante equally good investments, then she may naively diversify between these funds. Huberman and Jiang (2006) find some evidence of such a conditional $1/N$ rule.

For simplicity of presentation, we have treated the case of substitute products and the case of complementary products separately. But it is easy to combine the two problems into one grand decision problem. In particular, suppose that a subset of the products are substitutes as above, while the rest are complements as above, with preferences over the two subsets being separable. Because the two problems are then separable, our results apply unchanged to each subset.

[8] In one study, for instance, students chose snacks to be received at the end of three different classes. When choosing the snacks one at a time at the beginning of these classes, 9% of students chose three different snacks. But when simultaneously choosing three snacks ahead of time, 64% of students chose three different snacks. To the extent that in the sequential-choice conditions students know more about their momentary tastes, the former choices better reflect their true preferences.

5 Price Uncertainty and the Nature of Budgets

In this section, we establish a version of our budgeting result for situations characterized by price uncertainty, and identify plausible conditions under which a budget is optimally set in terms of monetary spending rather than consumption.[9] The agent has the same utility function as in Section 4.1, and for tractability we assume that $L = 2$ — there is a single category of substitute products — and the goods are symmetric. Furthermore, while the agent's tastes and consumption opportunities are deterministic, the prices of the goods, p_1 through p_N, are i.i.d. normally distributed random variables. Hence, the agent's consumption utility is

$$-\sum_m y_m^2 - \sum_{m \neq n} \theta y_m y_n + \sum_m \overline{x} y_m - \sum_m p_m y_m. \tag{11}$$

We assume that the agent can obtain information about the p_m in the same costly way as about the x_m in Section 4. As we have noted, an alternative interpretation of attention costs is reoptimization costs when the agent observes the price shocks, but must exert costly cognitive effort to figure out what these imply for optimal consumption.

We conceptualize the problem of whether the agent might want a budget for consumption or for spending by asking a more fundamental question: whether she wants to think — i.e., make plans and execute decisions — in terms of the consumption levels of the goods, or the amounts of spending on the goods.[10] Formally, in the former case she chooses consumption y_m for each good, and in the latter case she chooses spending $Y_m = p_m y_m$ on each good. While these two ways of thinking are equivalent in a classical problem with known prices, in our model — in which the agent does not fully learn prices before making decisions — they are not equivalent. For instance, deciding to buy a front-row ticket to a concert no matter how much it costs will in general not result in the same consumption as deciding to spend $100 on the concert no matter where one sits.

The case in which the agent thinks in terms of consumption levels reduces to our previous analysis by setting $x_m = -p_m$, so our budgeting results from Section 4 apply, and mean that the agent sets consumption budgets. We now compare this to thinking in terms of spending. Denoting

[9] In our main application for naive diversification, retirement investment, decisions are naturally denominated in dollars of investment into funds. This corresponds to prices that equal 1, so there is no price uncertainty.

[10] This type of question is almost never considered in the literature on rational inattention, but one notable exception is Reis (2006). Analyzing a consumption-savings problem in which a consumer does not know her wealth perfectly, Reis asks whether the consumer prefers to make decisions in terms of consumption or savings.

the means of p_m and y_m by $\overline{p}$ and $\overline{y}$, respectively, and linearly approximating spending Y_m as $Y_m = (\overline{p} + (p_m - \overline{p}))(\overline{y} + (y_m - \overline{y})) \approx \overline{p} y_m + (p_m - \overline{p})\overline{y}$, we get

$$y_m \approx \frac{Y_m}{\overline{p}} - \frac{p_m - \overline{p}}{\overline{p}} \overline{y}. \tag{12}$$

To keep our model within the quadratic framework of Section 3, we work with this approximation. The approximation retains a general property of thinking in terms of spending: that by fixing spending when she does not know the price, the agent makes the consumption level responsive to the unknown price. It is this general property, and not our use of an approximation, that drives the logic of Proposition 6 below.

To state our results, we define two measures of how the agent would optimally respond to information if it was costless. Assuming for the definition that $\lambda = 0$, let

$$\epsilon^1 = \frac{\partial E[y_m - y_n | p_m - p_n = x]/\partial x}{\overline{y}/\overline{p}} \quad \text{and} \quad \epsilon^2 = \frac{\partial E[\sum_m y_m | \sum_m p_m = x]/\partial x}{\overline{y}/\overline{p}},$$

which are the optimal elasticity of substitution between products and the optimal elasticity of total consumption with respect to the total price, respectively. We find:

Proposition 6. *For any $\lambda, \sigma_0^2, \epsilon^1, \epsilon^2$, thinking in terms of spending yields strictly higher expected utility than thinking in terms of consumption if (a) $\epsilon^1, \epsilon^2 > 1/2$ or (b) $\epsilon^1 > 1/2$ and N is sufficiently large, and the converse holds if (c) $\epsilon^1, \epsilon^2 < 1/2$.*

Proposition 6 identifies two sufficient conditions for thinking in terms of nominal spending to be optimal. Both conditions require that the products are relatively good substitutes ($\epsilon^1 > 1/2$). To understand the logic of Condition (a), suppose first that $N = 1$, i.e., there is a single product. Then, the condition says that the price elasticity of consumption of the single product must be greater than 1/2. Intuitively, fixing nominal spending generates a price elasticity of consumption of 1 (from (12), $[(y_m - \overline{y})/\overline{y}]/[(p_m - \overline{p})/\overline{p}] = 1$) while fixing consumption generates a price elasticity of consumption of 0, so the former is optimal if and only if the optimal price elasticity is closer to 1 than to 0. Extending the logic to $N > 1$ gives Condition (a): thinking in terms of spending is optimal if both relevant elasticities are greater than 1/2. And the converse gives Condition (c): thinking in terms of consumption is optimal if both relevant elasticities are less than 1/2.

If $\epsilon^1 > 1/2$ and $\epsilon^2 < 1/2$, then the above logic is not sufficient to determine whether thinking in terms of spending is optimal. Still, Condition (b) says that it is optimal if N is sufficiently large. Intuitively, this occurs because with many products, the predominant manner in which the agent wants to adjust consumption to shocks is by substituting between products — not by adjusting total consumption — so this substitution elasticity is more important in determining how she wants to think. Crucially, what matters for all of these results is the optimal *full-information* price elasticity, not the price elasticity the agent exhibits under costly attention.

Our next proposition extends the budgeting result in Proposition 3 and Corollary 1 to spending:

Proposition 7. *Suppose that the agent thinks in terms of spending, and* $\epsilon_1\epsilon_2 > 1$*. Then, there are* λ_1, λ_2 *satisfying* $0 < \lambda_2 < \lambda_1$ *such that if* $\lambda_2 \leq \lambda < \lambda_1$*, then total spending* $\sum_m Y_m$ *is constant, but the individual spending levels* Y_m *are not constant.*

The logic also parallels that before: the most valuable pieces of information to know about are price differences, so often this is all the agent pays attention to. As a result, she restricts adjustments to substitutions between products, fixing total spending.[11]

Thinking in terms of nominal spending, and having nominal budgets, is therefore optimal if the price elasticity of total consumption is sufficiently high, or it is not too low and product categories feature many closely substitutable products. These results explain the general prevalence of spending budgets, and — since lower-income individuals have higher price elasticities of consumption — the greater prevalence of spending budgets among lower-income individuals.[12] Nevertheless, our model predicts that individuals who do not care much about prices are more likely to have budgets expressed in terms of quantities. Consider, for instance, a rich person whose primary constraint in entertainment consumption is time, not money. Since she is therefore not price sensitive, she is more likely to choose a budget in entertainment quantity. Anecdotally, some people do seem to set consumption budgets, such as when deciding to go out twice a month or take two weeks of vacation per year. Relatedly, as Krishnamurthy and Prokopec (2010) note, in some self-control

[11] The intuition for the qualifier $\epsilon_1\epsilon_2 > 1$ derives from the central property of thinking in terms of spending, that it forces consumption to be sensitive to unanticipated price shocks. If the optimal (full-information) price elasticity of category consumption, ϵ_2, is low, then it is important for the agent to pay attention to the price level to reduce unanticipated price shocks. Hence, in that case paying attention to the price level is more important than paying attention to price differences, so trading off only within the category is never optimal.

[12] This tendency is reinforced to the extent that lower-income individuals also have higher costs of attention. An experiment by Mani et al. (2013), and a variety of other evidence discussed in Schilbach et al. (2016), indicates that poverty impedes cognitive performance, which means that lower-income individuals have a higher λ. A classical account, however, would suggest that lower-income individuals have a lower opportunity cost of time due to lower wages, and therefore have a lower λ.

settings people tend to have quantity budgets, for instance in the number of weekly desserts or Weight Watchers points they allow themselves. While our model does not formalize a self-control motive, this is another setting in which the primary cost of consumption is not the price, so that we predict quantity budgets rather than spending budgets.

Having a spending budget leads to an interesting pattern in how a person reacts to price shocks:

Corollary 2. *Suppose that the agent thinks in terms of spending, and* $\epsilon_1\epsilon_2 > 1, \lambda_2 \leq \lambda < \lambda_1$*. A decrease in the price of good* m *lowers spending on good* m *and increases spending on other goods.*

With full information, a decrease in the price of a good would lead to an increase in the consumption of that good and a decrease in the consumption of substitutes. In direct contrast, Corollary 2 says that the agent increases the consumption of substitutes as well. Experimental results by Heath and Soll (1996) and Heilman et al. (2002) are evocative (though not precise confirmations) of this prediction.[13]

Note that our model assumes linear disutility of money. Since thinking in terms of spending rather than quantities reduces risk in one's total spending, a budget constraint over nominal spending, or more generally a concave utility function over nominal savings provides an additional reason to think in terms of spending and therefore to have spending budgets.[14] Once again, this is especially likely to apply to low-income individuals, who typically face tighter constraints.

6 Unit Demand

In our main model, the consumer chooses consumption levels from a continuum. In a number of prototypical consumer decisions, however, a person is better described as having unit demand, choosing the one item she needs from a selection. For instance, in the medium run the car a person

[13] Roughly consistent with our prediction, shoppers in the experiment of Heilman et al. (2002) who were given $1 off an item increased their purchases of products related to the discounted item. But unlike in our model, the discount applied only to one item and hence was not a price decrease, and the discount also increased purchases of unrelated "treats." Similarly, Heath and Soll (1996) find in hypothetical choices that MBA students reduce their entertainment consumption more if they had spent $20 on a sports ticket than if they had received the same ticket as a gift. But again, a gift is not identical to a price shock.

[14] The simplest formal way to make this point is to assume mean-variance preferences over spending. Start with our model above, in which the agent does not care about the variance of spending. Suppose that the agent wants to set budgets, and is indifferent between thinking in terms of spending and thinking in terms of consumption. Now suppose that she also derives disutility from the variance of her spending. Then, her achievable level of utility is strictly lower when thinking in terms of quantities, but the same if she thinks in terms of spending. This means that she strictly prefers to think in terms of spending.

uses, and how much she uses her car, are fixed, so that she needs to buy a fixed amount of gasoline. When a consumer's computer breaks, she needs to buy exactly one new computer to replace it. And when shopping for a new bedroom, a homeowner may be looking for exactly one mattress and one comforter. We now analyze the implications of our framework for such purchases.

Suppose that there are N categories of products. The consumer is looking to buy exactly one product in each category (with her utility being $-\infty$ if there is a category in which she does not purchase) from a continuum of options with different quality levels. In category m, product $y_m \in \mathbb{R}$ has utility y_m, and a random price p_m, with total utility $\sum_m (y_m - p_m)$. The shape of prices is determined by the differentiable, strictly increasing and strictly convex function $p(\cdot)$ that has full range and satisfies $\lim_{y\to-\infty} p'(y) < 1$ and $\lim_{y\to\infty} p'(y) > 1$. But the price of a specific product is subject to shocks: $p_m = p(y_m) + x_m$ with probability s, and $p_m = p(y_m + x_m)$ with probability $1 - s$, where x_m is a random variable with mean zero. This specification incorporates both vertical and horizontal shifts in prices.[15]

We consider an agent who has sufficiently costly attention (a sufficiently high λ) such that she does not want to think about price shocks, and therefore makes a plan that is independent of price realizations. An alternative interpretation is that the price uncertainty is the residual uncertainty after the agent has thought about the problem. Similarly to the previous section, we ask whether the agent wants to fix the level of quality or the amount of spending for each category. For computers, for instance, she could decide on a specific computer brand and configuration no matter how much it costs, or she could ask for the best $2,000 computer no matter what specific machine that is. And for gasoline, she could buy the same grade each time, or she could decide how much she is willing to spend on gas, and choose the grade that is closest to that amount. These choice variables seem equally easy to implement in practice: in the former case the agent needs to remember the version she wants to buy in each category, and in the latter case she needs to remember the price she is aiming for in each category.

Proposition 8. *For any $p(\cdot)$ and any shock distribution, there is an $S \in (0, 1)$ such that fixing the quality level is optimal for $s < S$, and fixing spending is optimal for $s > S$.*

[15] In terms of consumption utility, this model is an extremely simplified variant of our basic model with price uncertainty, in which there is one product in each category, there is no quadratic term in utility, and the utility from different products is separable. Prices, however, are more complicated here. In our basic model, where y_m represents the consumption quantity of a single divisible product, prices are linear. In the current model, where y_m represents quality, prices are not linear.

If all shifts in prices are vertical ($s = 1$), then fixing quality is optimal. In this case, the marginal price of increasing y_m is constant, so choosing a fixed y_m is optimal. In contrast, if all shifts in prices are horizontal ($s = 0$), then fixing spending is optimal. In this case, a decrease in the price level also decreases the marginal price of increasing quality, so fixing spending better aligns marginal value with marginal price. Extending this logic, thinking in terms of spending is optimal if the price and marginal price of quality are sufficiently positively correlated.

Again, our model assumes linear disutility of money. If the agent has a budget constraint over nominal spending, or more generally her utility over nominal savings is concave, then she is more prone to think in terms of spending rather than consumption to reduce risk in her total spending.

For a consumer who thinks in terms of spending ($s > S$), the implications of our unit-demand model contrast in an interesting way with those of our continuous-demand model above. When the agent has a budget in the continuous model, an equal increase in prices for a category leaves the agent's spending levels unchanged in expectation for all products. In the unit-demand model, in contrast, any increase in prices means that the agent must substitute to a lower-quality product to keep her spending constant.

This last prediction provides a potential explanation for the findings of Hastings and Shapiro (2013), although more research seems necessary to determine how compelling the explanation is. Hastings and Shapiro document that when gasoline prices rise, there is a shift in demand from premium to regular gasoline — i.e., a cheaper product in the same category. Importantly, they document such behavior for price shifts for which the price and marginal price of quality are approximately uncorrelated (a setting with a low s), while our model predicts the behavior only if the price and marginal price of quality are sufficiently positively correlated for the agent to think in terms of spending ($s > S$). Nevertheless, the correlation between the price and marginal price of quality is plausibly positive for many products consumers have experience with. (Even for gasoline, Hastings and Shapiro focus on the short run, and the correlation may be positive in the longer run.) In as much as this is the case, and a consumer does not think about the correlation separately for gasoline, our model accounts naturally for the evidence, especially for budget-constrained consumers.

7 Related Literature

In this section, we discuss theoretical work most closely related to our paper.

In applications of prospect theory, the term "mental account" is used to refer to the set of monetary outcomes that are evaluated jointly in the context of a single decision (e.g., Kahneman and Tversky, 1984, Thaler, 1985, Henderson and Peterson, 1992). For instance, a person is more willing to drive 20 minutes for a $5 saving if it comes off of a $15 purchase than if it comes off of a $125 purchase (Tversky and Kahneman, 1981), presumably because she evaluates the saving together with the purchase to which it is applied. Our paper is instead about mental accounts that serve as a decision-making aid when there are multiple competing uses for money.[16]

The main existing explanation for mental accounts as a decision-making aid is self-control problems — attempting to use budgets or accounts to mitigate overconsumption in the future. Our theory provides a different, complementary, reason for mental accounts, with a number of distinct features. Most importantly, in theories on mental accounting and self-control, money is exogenously assumed to be non-fungible in the sense that spending from different accounts is subject to different constraints or preferences. In our model, mental accounts emerge despite money being fully fungible. Due to the different foundation, research on mental accounting and self-control also does not generate many of our other predictions, such as the connection we find between mental budgeting and naive diversification.

In a classic paper, Shefrin and Thaler (1988) develop a life-cycle consumption-savings model in which the individual's "planner" self would like to control the "doer" self's tendency to consume too much. Shefrin and Thaler assume that the individual can separate money into different mental accounts, current spendable income, current assets, and future income. They exogenously assume that the marginal propensity to consume out of these accounts is different.

In the context of goal setting under self-control problems, Koch and Nafziger (2016) assume that an individual can decide between broad and narrow goals, and that falling short of one's chosen goal(s) leads to sensations of loss. The motive to avoid such losses creates an incentive that

[16] Prelec and Loewenstein (1998) propose a model of mental accounting in which paying for a good or experience is painful but buffered by thoughts of future consumption, and the pleasure of consumption is lowered by thoughts of future payments. This model predicts a strong aversion to paying for consumption with debt, and has a number of other implications for how an individual might want to time payments relative to consumption. Once again, these results pertain primarily to individual transactions, and capture a completely different aspect of mental accounting than do we.

mitigates self-control problems. A broad goal diversifies the risk of failure due to shocks, but it also lowers incentives because underperformance in one task can be offset by good performance in another task. As a result, narrow bracketing can be optimal, especially when uncertainty is low. Hsiaw (2018) qualifies this insight for multi-stage projects when uncertainty is resolved over time, showing that an increase in late uncertainty makes incremental goals more appealing.[17]

Galperti (forthcoming) compares good-specific and total-expenditure budgets for a person who is subject to self-control problems as well as both intratemporal and intertemporal taste shocks. Good-specific budgets can be useful for an agent with a mild self-control problem, as they help curb overconsumption of all goods. But for an agent with severe self-control problems, effective good-specific budgets would distort intratemporal consumption too much, so a total-expenditure budget is superior.

Gorman (1959) identifies circumstances under which it is optimal for a standard utility maximizer to make consumption decisions using a two-step procedure similar to that in Sections 4 and 5, whereby she first allocates fixed budgets to different consumption categories, and then optimizes within each category given the allocated budget. Unlike in our model, the budgeting in the first stage requires the agent to know with certainty all the relevant price indices for the categories, and there is no taste uncertainty. Even so, the conditions under which two-stage budgeting is optimal are extremely strict.

In predicting that the agent may completely ignore some aspects of her decision environment, our model is similar in spirit to the sparsity-based model of bounded rationality by Gabaix (2014). In Gabaix's setting, the variables that the agent may choose to look at are exogenously given, whereas in ours the agent can choose any combination of variables. We also apply the model to different questions than does Gabaix.

Since our theory predicts unambiguous budgets based on economic preferences and fundamentals, it fails to capture some subtle context dependence in how individuals categorize outlays. For instance, Cheema and Soman (2006) find that individuals categorize a restaurant dinner flexibly as either food or entertainment depending on which budget has more money left over in it. The authors interpret such malleability in mental accounting as an attempt to justify spending.

[17] See also Pagel (2017) for other implications of loss aversion for consumption-savings behavior.

8 Conclusion

While our models explain a number of findings, there are phenomena that are usually interpreted in terms of mental accounting that we have not covered. The most important of these is the consumption effect of transfers that can only be used on a subset of products. The rational consumer model with full information implies that if such a transfer is inframarginal — i.e., if the consumer would have spent more than the transfer on the products in question — then it is equivalent to cash. Yet experimental work by Abeler and Marklein (2016) and empirical work by Hastings and Shapiro (2017) document that inframarginal transfers have larger effects on the consumption of targeted products than cash transfers. Even when a transfer is not inframarginal, it can have a surprisingly large effect: for instance, incentives for health-improving behaviors that are minute relative to the health benefits can significantly influence behavior (Volpp et al., 2008, Dupas, 2014).[18] While not predicted by our current framework, there is a plausible attention-based account for these phenomena. Namely, there are many things that a person could consider doing, but that she deems not worthwhile to think about due to costly attention, and that she therefore does not do. Receiving a transfer or subsidy can induce the person to think about the potential benefits, increasing the effect of the transfer. In ongoing work, we formalize this mechanism, and also consider what it implies for the optimal design of transfers.

Of course, we do not believe that mental accounting is solely about costly attention. As we have mentioned, a likely motive for creating mental accounts is self-control problems. It would be interesting to combine the attention-based and self-control-based explanations of mental accounting to identify interactions. For example, a person may use the costly nature of her attention to improve self-control by creating plans that she is unwilling to reconsider later. And when it comes to implementing a mental-accounting-based consumption plan, researchers understand that if the budget for an account becomes a reference point, then loss aversion helps stick with the plan. Our theory provides one possible foundation for which outcomes are evaluated jointly in a reference-dependent model. Once again, it would seem fruitful to combine the attention-based view with loss aversion.

[18] A related finding in political economy is the flypaper effect (Hines and Thaler, 1995): when a local government receives a grant earmarked for a specific purpose, it tends to increase spending on that purpose by the amount of the grant.

References

Abeler, Johannes and Felix Marklein, "Fungibility, Labels, and Consumption," *Journal of the European Economic Association*, 2016, *15* (1), 99–127.

Ameriks, John, Andrew Caplin, and John Leahy, "Wealth Accumulation and the Propensity to Plan," *Quarterly Journal of Economics*, 2003, *118* (3), 1007–1047.

Antonides, Gerrit, I. Manon de Groot, and W. Fred van Raaij, "Mental Budgeting and the Management of Household Finance," *Journal of Economic Psychology*, 2011, *32* (4), 546–555.

Benartzi, Shlomo and Richard H. Thaler, "Naive Diversification Strategies in Defined Contribution Saving Plans," *American Economic Review*, 2001, *91* (1), 79–98.

_ and _ , "Heuristics and Biases in Retirement Savings Behavior," *Journal of Economic Perspectives*, 2007, *21* (3), 81–104.

Caplin, Andrew and Mark Dean, "Revealed Preference, Rational Inattention, and Costly Information Acquisition," *American Economic Review*, 2015, *105* (7), 2183–2203.

Cheema, Amar and Dilip Soman, "Malleable Mental Accounting: The Effect of Flexibility on the Justification of Attractive Spending and Consumption Decisions," *Journal of Consumer Psychology*, 2006, *16* (1), 33–44.

Cover, Thomas M. and Joy A. Thomas, *Elements of Information Theory*, Hoboken, NJ: Wiley, 2006.

Dean, Mark and Nate Leigh Neligh, "Experimental Tests of Rational Inattention," 2017. Working Paper.

Dupas, Pascaline, "Getting Essential Health Products to their End Users: Subsidize, but How Much?," *Science*, 2014, *345* (6202), 1279–1281.

Fosgerau, Mogens, Emerson Melo, and Matthew Shum, "Discrete Choice and Rational Inattention: A General Equivalence Result," 2017. Working Paper.

Gabaix, Xavier, "A Sparsity-Based Model of Bounded Rationality," *Quarterly Journal of Economics*, 2014, *129* (4), 1661–1710.

_ and David Laibson, "Myopia and Discounting," 2017. Working Paper.

Galperti, Simone, "A Theory of Personal Budgeting," *Theoretical Economics*, forthcoming.

Gorman, W. M., "Separable Utility and Aggregation," *Econometrica*, 1959, *27* (3), 469–481.

Hastings, Justine S. and Jesse M. Shapiro, "Fungibility and Consumer Choice: Evidence from Commodity Price Shocks," *Quarterly Journal of Economics*, 2013, *128* (4), 1449–1498.

_ and _ , "How Are SNAP Benefits Spent? Evidence from a Retail Panel," 2017. Working Paper.

Heath, Chip and Jack B. Soll, "Mental Budgeting and Consumer Decisions," *Journal of Consumer Research*, 1996, *23* (1), 40–52.

Heilman, Carrie M., Kent Nakamoto, and Ambar G. Rao, "Pleasant Surprises: Consumer Response to Unexpected In-Store Coupons," *Journal of Marketing Research*, 2002, *39* (2), 242–252.

Henderson, Pamela W. and Robert A. Peterson, "Mental Accounting and Categorization," *Organizational Behavior and Human Decision Processes*, 1992, *51* (1), 92–117.

Hines, James and Richard Thaler, "The Flypaper Effect," *Journal of Economic Perspectives*, 1995, *9* (4), 217–226.

Hsiaw, Alice, "Goal Bracketing and Self-Control," *Games and Economic Behavior*, 2018, *111*, 100–121.

Huberman, Gur and Wei Jiang, "Offering versus Choice in 401(k) Plans: Equity Exposure and Number of Funds," *Journal of Finance*, 2006, *61* (2), 763–801.

Kahneman, Daniel and Amos Tversky, "Choices, Values, and Frames," *American Psychologist*, 1984, *39* (4), 341–350.

Koch, Alexander K. and Julia Nafziger, "Goals and Bracketing under Mental Accounting," *Journal of Economic Theory*, 2016, *162* (C), 305–351.

Krishnamurthy, Parthasarathy and Sonja Prokopec, "Resisting That Triple-Chocolate Cake: Mental Budgets and Self-Control," *Journal of Consumer Research*, 2010, *37* (1), 68–79.

Lave, Jean, *Cognition in Practice: Mind, Mathematics, and Culture in Everyday Life*, Cambridge University Press, 1995.

Luo, Yulei and Eric R. Young, "Signal Extraction and Rational Inattention," *Economic Inquiry*, 2014, *52* (2), 811–829.

Mackowiak, Bartosz; and Mirko Wiederholt, "Optimal Sticky Prices under Rational Inattention," *The American Economic Review*, June 2009, *99.*

Mackowiak, Bartosz, Filip Matějka, and Mirko Wiederholt, "Rational Inattention: A Disciplined Behavioral Model," 2018. Working Paper.

Mani, Anandi, Sendhil Mullainathan, Eldar Shafir, and Jiaying Zhao, "Poverty Impedes Cognitive Function," *Science*, 2013, *341* (6149), 976–980.

Matějka, Filip, "Rationally Inattentive Seller: Sales and Discrete Pricing," *Review of Economic Studies*, 2016, *83* (3), 1125–1155.

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

Miao, Jianjun, Jieran Wu, and Eric Young, "Multivariate Rational Inattention," 2019. Working Paper.

Morris, Stephen and Ming Yang, "Coordination and Continuous Stochastic Choice," 2016. Working Paper.

— **and Philipp Strack**, "The Wald Problem and the Equivalence of Sequential Sampling and Static Information Costs," 2017. Working Paper.

Pagel, Michaela, "Expectations-Based Reference-Dependent Life-Cycle Consumption," *Review of Economic Studies*, 2017, *84* (2), 885–934.

Pomatto, Luciano, Philipp Strack, and Omer Tamuz, "The Cost of Information," 2019. Working Paper.

Prelec, Drazen and George Loewenstein, "The Red and the Black: Mental Accounting of Savings and Debt," *Marketing Science*, 1998, *17* (1), 4–28.

Rainwater, Lee, Richard P. Coleman, Gerald Handel, and W. Lloyd Warner, *Workingman's Wife: Her Personality, World and Life Style*, Oceana Publications, 1959.

Rajagopal, Priyali and Jong-Youn Rha, "The Mental Accounting of Time," *Journal of Economic Psychology*, 2009, *30* (5), 772–781.

Reis, Ricardo, "Inattentive Consumers," *Journal of Monetary Economics*, 2006, *53* (8), 1761–1800.

Schilbach, Frank, Heather Schofield, and Sendhil Mullainathan, "The Psychological Lives of the Poor," *American Economic Review*, 2016, *106* (5), 435–440.

Shefrin, Hersh M. and Richard H. Thaler, "The Behavioral Life-Cycle Hypothesis," *Economic Inquiry*, 1988, *26* (4), 609–643.

Simonson, Itamar, "The Effect of Purchase Quantity and Timing on Variety-Seeking Behavior," *Journal of Marketing Research*, 1990, *27* (2), 150–162.

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

Telatar, Emre, "Capacity of Multi-antenna Gaussian Channels," *European transactions on telecommunications*, 1999, *10* (6), 585–595.

Thaler, Richard H., "Mental Accounting and Consumer Choice," *Marketing Science*, 1985, *4* (3), 199–214.

— , "Mental Accounting Matters," *Journal of Behavioral Decision Making*, 1999, *12* (3), 183–206.

Tversky, Amos and Daniel Kahneman, "The Framing of Decisions and the Psychology of Choice," *Science*, 1981, *211*, 453–458.

Veldkamp, Laura L., "Information Markets and the Comovement of Asset Prices," *Review of Economic Studies*, 2006, *73* (3), 823–845.

Verstyuk, Sergiy, "Log, Stock and Two Simple Lotteries," 2019. Working Paper.

Volpp, Kevin G., Leslie K. John, Andrea B. Troxel, Laurie Norton, Jennifer Fassbender, and George Loewenstein, "Financial Incentive-Based Approaches for Weight Loss," *Journal of the American Medical Association*, 2008, *300* (22), 2631–2637.

Woodford, Michael, "Information-Constrained State-Dependent Pricing," *Journal of Monetary Economics*, 2009, *56* (Supplement 1).

_ , "Inattentive Valuation and Reference-Dependent Choice," 2012. Working Paper.

A LQ Multivariate Setup

Proof of Proposition 1. The quadratic utility function can be rewritten as

$$U(\mathbf{y},\mathbf{x}) = -\Big(\mathbf{y} - \frac{C^{-1}B'}{2}\mathbf{x}\Big)' C\Big(\mathbf{y} - \frac{C^{-1}B'}{2}\mathbf{x}\Big) + \frac{\mathbf{x}'BC^{-1}B'\mathbf{x}}{4}. \tag{13}$$

If the posterior mean is $\tilde{\mathbf{x}}$, then the agent chooses an action (maximizing expected utility):

$$\mathbf{y} = \frac{C^{-1}B'}{2}\tilde{\mathbf{x}}. \tag{14}$$

This is because certainty equivalence applies in a quadratic setup. Plugging (14) into (13), the realized utility $\tilde{U}$ for a state $\mathbf{x}$, but a posterior mean $\tilde{\mathbf{x}}$ is:

$$\tilde{U}(\tilde{\mathbf{x}},\mathbf{x}) = -(\tilde{\mathbf{x}}-\mathbf{x})'\Omega(\tilde{\mathbf{x}}-\mathbf{x}) + \mathbf{x}'\Omega\mathbf{x}, \tag{15}$$

where $\Omega = BC^{-1}B'/4$. The first term is the loss from imperfect posterior beliefs, $(\tilde{\mathbf{x}}-\mathbf{x})$ is the misperception. Given the variance-covariance matrix Σ for the distribution of $(\tilde{\mathbf{x}}-\mathbf{x})$, the expectation of the first term equals the trace of $\Omega\Sigma$. Since the second term in (15) depends on the realized state $\mathbf{x}$ only, i.e., it is independent of the agent's strategy, then the original problem takes the form:

$$\max_{\psi\succeq\Sigma} -Tr(\Omega\Sigma) + \frac{\lambda}{2}\log|\Sigma|. \tag{16}$$

The second term in (16) is the cost of information, it is a log of the determinant of Σ.[19] The larger the posterior uncertainty is, the lower the cost. The cost term here includes entropy of the posterior only, since entropy of a fixed prior amounts to an additive constant only. The condition $\psi \succeq \Sigma$ requires that $(\psi - \Sigma)$ is positive semi-definite, which means that acquisition of Gaussian signals cannot make beliefs less precise, i.e., signals must have non-negative precision.

[19] Entropy of a multivariate $N(\mu,\Sigma)$ of dimension n is $\frac{n}{2}(\log(2\pi)+1) + \frac{1}{2}\log|\Sigma|$.

To explore what signals the agent collects, let us decompose the loss matrix Ω, which is symmetric and thus has an orthonormal basis of eigenvectors. Let $\Omega = U\Lambda U'$, where U is a unitary matrix (the columns of which are eigenvectors of Ω), and Λ is a diagonal matrix with its elements Λ_{ii} equal to the eigenvalues Λ_i of Ω.

$$\begin{aligned} -Tr(\Omega\Sigma) + \frac{\lambda}{2}\log|\Sigma| &= -Tr(U\Lambda U'\Sigma) + \frac{\lambda}{2}\log|\Sigma| = \\ &= -Tr(\Lambda U'\Sigma U) + \frac{\lambda}{2}\log|U'\Sigma U| = \\ &= -Tr(\Lambda S) + \frac{\lambda}{2}\log|S|, \end{aligned} \tag{17}$$

where $S = U'\Sigma U$ is the posterior variance-covariance matrix in the basis of eigenvectors of Ω. The condition $(\psi \succeq \Sigma)$ takes the form of $(U'\psi U \succeq S)$; note that $\psi = \sigma_0^2 I$.

Now we show by contradiction that S is diagonal. Let the optimal S were not diagonal, and let S^D be the matrix constructed from its diagonal, i.e., $S^D_{ii} = S_{ii}$ for all i and $S^D_{ij} = 0$ for all $i \neq j$.

First, since $\sigma_0^2 I - S$ is positive semi-definite, then $\sigma_0^2 I - S^D$ is also positive semi-definite. This is because for a diagonal S^D it suffices to check that $S^D_{ii} \leq \sigma_0^2$, which is implied by the fact that $\sigma_0^2 I - S$ is also positive semi-definite. Second, Hadamard's inequality implies:

$$\frac{\lambda}{2}\log|S| \leq \frac{\lambda}{2}\sum_i \log S_{ii} = \frac{\lambda}{2}\log|S_D|, \tag{18}$$

where the equality holds if and only if S is diagonal. Third, $Tr(\Lambda S) = Tr(\Lambda S^D)$, since Λ is diagonal. Therefore, putting this together implies that S not cannot be the optimum, since S^D delivers a higher objective due to the lower information cost, (18), and is feasible.

Therefore, S is diagonal. Using (17), the original problem takes the form:

$$\max_{S_{ii} \leq \sigma_0^2} -\sum_{i=1}^{N}\Big(S_{ii}\Lambda_i + \frac{\lambda\log(S_{ii})}{2}\Big). \tag{19}$$

The first order condition with respect to S_{ii} implies:

$$-\Lambda_i + \frac{\lambda}{2S_{ii}} = 0,$$

and the solution is

$$S_{ii} = \min\Big(\sigma_0^2, \frac{\lambda}{2\Lambda_i}\Big).$$

$\square$.

<u>Proof of Proposition 2.</u> Part 1: Proposition 1 implies that the space of posterior means $\tilde{\mathbf{x}}$ is spanned by all eigenvectors v^i for which $\lambda/(2\Lambda_i) < \sigma_0^2$, the statement is then a trivial implication.

Part 2: Let $\xi_i = 1 - \frac{S_{ii}}{\sigma_0^2}$ be the relative reduction of uncertainty about the component $v^i \cdot \mathbf{x}$. ξ is also the linear weight on a signal (as opposed to on the prior) in Bayesian updating with Gaussian signals. This means that in 1D Bayesian updating, if the random variable $v^i \cdot \mathbf{x}$ moves by Δx, then the posterior mean about this variable moves in expectation by $\xi_i \Delta x$.

Since the agent chooses independent signals on $v^i \cdot \mathbf{x}$, Bayesian updating does in fact take the 1D form. Responsiveness then is:

$$\varepsilon_i^\lambda = \frac{|H\xi_i v^i|}{|v_i|} = \frac{|\xi_i H v^i|}{|v_i|} = \xi_i \varepsilon_i^0.$$

This equation together with Proposition 1 implies the expression (7).

Part 3: Differentiating $\varepsilon_i^\lambda / \varepsilon_j^\lambda$ with respect to λ then implies the statement. $\square$.

B Consumption and Spending Budgets

Let Θ have the structure described in the main text, i.e., given by symmetrically nested categories, and let $R^{k,l}$ denote a category number k on level l, size of which is r_l.

Lemma 1. *Θ has a base of eigenvectors $\{v^{k,l,r'}\}_{l\in\{2..L\},k\in\{1..N/r_l\},r'\in\{1..(r_l-1)\}}$, and $(1,..,1)$. $v^{k,l,r'}$ is associated with a category k of a level l, and it has the following properties*

$$v_m^{k,l,r'} = 0 \qquad \forall m \notin R^{k,l} \tag{20}$$

$$\sum_{m\in R^{k,l}} v_m^{k,l,r'} = 0 \tag{21}$$

$$v_m^{k,l,r'} = v_n^{k,l,r'} \qquad \forall m,n; \exists k' : m,n \in R^{k',l-1}. \tag{22}$$

Moreover, eigenvalues μ^l of $v^{k,l,r'}$ are given by:

$$\mu^l = \sum_{n \in R_m^{l-1}} \left(\Theta_{m,n} - \gamma^l\right) = \mu^{l-1} + (\gamma^{l-1} - \gamma^l) r_{l-1},$$

and $\mu^2 = \gamma^1 - \gamma^2$.

Proof of Lemma 1. Let us fix m, and apply Θ to an eigenvector associated with $R^{k,l}$; we drop the index m of the vector. Let R_m^{l-1} be the category on level $(l-1)$ that the good m belongs to.

$$\begin{aligned}
\sum_n \Theta_{m,n} v_n^{k,l} &= \left(\sum_{n \notin R^{k,l}} \Theta_{m,n} v_n^{k,l}\right) + \left(\sum_{n \in R^{k,l}/R_m^{l-1}} \Theta_{m,n} v_n^{k,l}\right) + \left(\sum_{n \in R_m^{l-1}} \Theta_{m,n} v_n^{k,l}\right) = \\
&= 0 + \gamma^l \left(\sum_{n \in R^{k,l}/R_m^{l-1}} v_n^{k,l}\right) + \left(\sum_{n \in R_m^{l-1}} \Theta_{m,n} v_n^{k,l}\right) = \\
&= \gamma^l \left(\sum_{n \in R^{k,l}} v_n^{k,l} - \sum_{n \in R_m^{l-1}} v_n^{k,l}\right) + \left(\sum_{n \in R_m^{l-1}} \Theta_{m,n} v_n^{k,l}\right) = \\
&= \gamma^l \left(0 - \sum_{n \in R_m^{l-1}} v_m^{k,l}\right) + \left(\sum_{n \in R_m^{l-1}} \Theta_{m,n} v_m^{k,l}\right) = \\
&= v_m^{k,l} \sum_{n \in R_m^{l-1}} \left(\Theta_{m,n} - \gamma^l\right)
\end{aligned}$$

The first equality is a simple decomposition into terms with elements within different categories. In the second, we used (20). Third is based on a decomposition of elements of $R^{k,l}$ into a sub-category with m and the other elements. The fourth equality uses (21) for the first term, and (22) is applied for the other two terms to substitute elements $v_j^{k,l}$ indexed by j by a constant $v_i^{k,l}$, since $v_j^{k,l}$ is constant in R_i^{l-1}.

Eigenvalue μ^l is therefore $\sum_{n \in R_m^{l-1}} (\Theta_{m,n} - \gamma^l)$. For $l = 2$ the only sub-category including m is m itself, $\mu^2 = \gamma^1 - \gamma^2$. And for $l > 2$:

$$\begin{aligned}
\mu^l &= \mu^{l-1} + \gamma^{l-1} r_{l-2} - \gamma^l r_{l-1} - (r_{l-2} - r_{l-1}) \gamma^{l-1} = \\
&= \mu^{l-1} + (\gamma^{l-1} - \gamma^l) r_{l-1}.
\end{aligned} \tag{23}$$

Therefore, each is $v^{k,l,r'}$ is an eigenvector, and they form a basis. This is because they are all mutually orthogonal. Vectors associated with distinct categories due to (20), and vectors associated with a nested categories due to (22) and (21). And for vectors of the same category, the dimensionality is due to (22) equal the number of sub-categories minus one lower dimensionality due to (21), $r_l/r_{l-1} - 1$. The total number of vectors associated with level $l > 1$ is $N/r_{l-1} - N/r_l$, and the total number of these orthogonal eigenvectors on all levels is $N - 1$, which together with the eigenvector $(1, .., 1)$ delivers N orthogonal eigenvectors, and thus a basis. □

Proof of Proposition 3. This proposition follows from the propositions 1 and 2, and Lemma 1.

First, since in this case $\Omega = \Theta^{-1}/4$, then Ω has the same eigenvectors as Θ, and the corresponding eigenvalues Λ_i are proportional to the inverse of the eigenvalue of Θ that is associated with the same vector. Specifically, for v^i associated with a level l, the eigenvalue is:

$$\Lambda_i = \frac{1}{4}(\mu^l)^{-1} \tag{24}$$

The expression below is derived from (24) and (7) with $\sigma_0^2 = \lambda/2\Lambda_i$ at which ξ_i hits zero.

$$\lambda_{l-1} = 2\sigma_0^2\Lambda_i = \frac{\sigma_0^2}{2\sum_{n\in R_m^{l-1}}(\Theta_{m,n} - \gamma^l)}, \tag{25}$$

where m is a good in $R^{k,l-1}$. We denote this threshold cost for attention to vectors associated with level l by λ_{l-1} rather than by λ_l, because for λ lower than this quantity the total consumption in each category on levels $(l-1)$ and lower is constant.

The eigenvectors satisfy (21), which means that the agent pays attention to vectors that keep "budgets" of the random elements of $\mathbf{x}$ across the corresponding category fixed, but not across categories on lower levels. Moreover, since the action matrix H equals $\Theta^{-1}/2$, the eigenvectors of Ω are its eigenvectors as well. Thus, the fixed budgets of $\mathbf{x}$ translate to the fixed budgets of actions $\mathbf{y}$ across the same categories.

Notice that for nested substitutes, since $\gamma^l > \gamma^{l-1}$, eigenvalues μ^l of Θ are increasing in l, and thus eigenvalues of Ω are decreasing in l. □

Proof of Proposition 4.

The variance-covariance matrix of posterior means (describing correlations of beliefs about x_i and x_j) is $P = (\psi - \Sigma)$. This matrix is diagonal in the basis of eigenvectors v^k, i.e, $P = UQU^{-1}$, where the columns of U are v^i. The diagonal elements of $Q_{kk} \equiv Q_k$ equal $\sigma_0^2 - \sigma_k^2$, which is the reduction of uncertainty about $v^k \cdot \mathbf{x}$. The reduction $Q_k = max(0, \sigma_0^2 - \lambda/(2\Lambda_k))$ is weakly increasing in Λ_k and weakly decreasing in λ, see Proposition (1).

The resulting variance-covariance matrix of actions is $A = HPH^{-1}$, where $P_{ij} = \sum_k Q_k v_i^k v_j^k$, and v^k are eigenvectors of H, too, $2\Lambda_k$ is the eigenvalue. The matrix A thus is:[20]

$$A_{ij} = \sum_k Q_k (2\Lambda_k)^2 v_i^k v_j^k. \tag{26}$$

And finally, the correlation of interest, that of y_m and $Y_{-m} = \sum_{n \in R^{k,l} \setminus m} y_n$, is given by

$$\frac{Cov(y_m, Y_{-m})}{\sqrt{Var(y_m)Var(Y_{-m})}} = \frac{(n-1)A_{i,j}}{\sqrt{A_{ii}\Big((n-1)(n-2)A_{i,j} + (n-1)A_{ii}\Big)}} = \frac{\sqrt{n-1}\rho_{ij}}{\sqrt{(n-2)\rho_{ij}+1}},$$

where $\rho_{ij} = A_{ij}/\sqrt{A_{ii}A_{jj}} = A_{ij}/A_{ii}$ is the correlation between y_i and y_j such that $i \neq j$. The correlation of y_m and Y_{-m} is thus an increasing function of ρ_{ij}. To prove the statement of Proposition 4, it now suffices to show that ρ_{ij} is decreasing in λ.

Next, using (26) we express derivative of the correlation:

$$\frac{\partial \rho_{ij}}{\partial \lambda} = \frac{\Big(-\sum_k 2\Lambda_k v_i^k v_j^k\Big)\Big(\sum_k Q_k(2\Lambda_k)^2 (v_i^k)^2\Big) - \Big(\sum_k Q_k (2\Lambda_k)^2 v_i^k v_j^k\Big)\Big(-\sum_k 2\Lambda_k (v_i^k)^2\Big)}{\sum_k Q_k (2\Lambda_k)^2 (v_i^k)^2},$$

where the sums are over all k such that $\lambda < 2\Lambda_k$. Due to Lemma 1, the eigenvectors can be selected such that $v_i^k v_j^k = -1$ for some vectors v^k that are associated with the smallest level on which goods i and j are in the same category, let the level be l^* and the number of such vectors be ψ_{l^*}. Similarly, $v_i^k v_j^k = 1$ for some v^k that are associated with levels higher than l^*, and let ψ_s be the number of such vectors on the level s. For all other vectors v^k: $v_i^k v_j^k = 0$.

Let $\hat{L} \geq (l+1)$ be the largest s such that $\lambda < 2\Lambda_s$. The numerator of the RHS of equation

[20] Equation (26) implies that actions y_i and y_j are more positively correlated if more uncertainty is reduced in a direction of v^k, for which the signs of entries v_i^k and v_j^k are the same.

above then equals eight times the following quantity:

$$
\begin{aligned}
\Big(\psi_{l^*}\Lambda_{l^*} - \sum_{s=l^*+1}^{\hat{L}} \psi_s\Lambda_s \quad \Big) \quad & \Big(\sum_{s=l^*}^{\hat{L}} \psi_s Q_s(\Lambda_s)^2\Big) - \\
- \qquad & \Big(-\psi_{l^*}Q_{l^*}(\Lambda_{l^*})^2 + \sum_{s=l^*+1}^{\hat{L}} \psi_s Q_s(\Lambda_s)^2\Big)\Big(-\sum_{s=l^*}^{\hat{L}} \psi_s\Lambda_s\Big) = \\
= \;\; & \psi_{l^*}\Lambda_{l^*}\Big(\sum_{s=l^*+1}^{L} \psi_s Q_s(\Lambda_s)^2\Big) - \psi_{l^*}Q_{l^*}(\Lambda_{l^*})^2\Big(\sum_{s=l^*+1}^{\hat{L}} \psi_s\Lambda_s\Big) = \\
= \;\; & \psi_{l^*}\Lambda_{l^*}\sum_{s=l^*+1}^{\hat{L}}\Big(\psi_s Q_s(\Lambda_s)^2 - \psi_s Q_l \Lambda_s \Lambda_l\Big) < 0.
\end{aligned}
$$

In the last step we used the fact that both $Q_s\Lambda_s$ are decreasing in the level s, see the proof of Proposition 3. This together with the positivity of the denominator of the RHS of (27) concludes the proof. □

Proof of Proposition 5. Analogous to the proof of Proposition 3. Eigenvectors take the same form, but ranking of magnitudes of eigenvalues, given by (25) is the opposite because for complements $\gamma^2 < \cdots < \gamma^L < 0$. □

Lemma 2. Losses from uncertainty of a fixed form: *for utility function* $-\boldsymbol{y}\Theta\boldsymbol{y} + \Big(\bar{\boldsymbol{x}} - \boldsymbol{p}\Big)' \cdot \boldsymbol{y}$, *losses from imperfect information of a fixed form about* $\boldsymbol{p}\cdot v^i$ *are lower when thinking in terms of spending than when thinking in terms of consumption if and only if*

$$\epsilon^i > 1/2.$$

Proof of Lemma 2.

We first express the analog of expected losses from imperfect information, (15), for the spending choice variables using transformation (6). Let $\mathbf{p'} = \mathbf{p} - \bar{\mathbf{x}}$.

$$U(\mathbf{u}, \mathbf{p}) = -\Big(\mathbf{y} + \frac{\Theta^{-1}}{2}\mathbf{p'}\Big)'\Theta\Big(\mathbf{y} + \frac{\Theta^{-1}}{2}\mathbf{p'}\Big) + \frac{\mathbf{p'}'\Theta^{-1}\mathbf{p'}}{4} =$$
$$= -\frac{1}{\bar{\mathbf{p}}^2}\Big(\mathbf{Y} - \bar{\mathbf{y}}(\mathbf{p} - \bar{\mathbf{p}}) + \frac{\bar{\mathbf{p}}\Theta^{-1}}{2}\mathbf{p'}\Big)'\Theta\Big(\mathbf{Y} - \bar{\mathbf{y}}(\mathbf{p} - \bar{\mathbf{p}}) + \frac{\bar{\mathbf{p}}\Theta^{-1}}{2}\mathbf{p'}\Big) + \frac{\mathbf{p'}'\Theta^{-1}\mathbf{p'}}{4}. \tag{27}$$

Therefore, the optimal spending $\mathbf{Y}$ conditional on posterior beliefs (with a mean $\tilde{\mathbf{p'}}$) is:

$$\mathbf{Y} = \bar{\mathbf{y}}(\mathbf{p} - \bar{\mathbf{p}}) - \frac{\bar{\mathbf{p}}\Theta^{-1}}{2}(\mathbf{p} - \bar{\mathbf{x}}).$$

The utility loss from imperfect beliefs thus equals to:

$$(\tilde{\mathbf{p}} - \mathbf{p})'\Omega^N(\tilde{\mathbf{p}} - \mathbf{p}), \tag{28}$$

where

$$\Omega^N = \frac{1}{\bar{p}^2}\Big(\bar{p}\Theta^{-1}/2 - \bar{y}\Big)'\Theta\Big(\bar{p}\Theta^{-1}/2 - \bar{y}\Big),$$

which can be rearranged to:

$$\Omega^N = \Omega - \Big(\frac{\bar{y}}{\bar{p}}\Big)I + \Big(\frac{\bar{y}}{\bar{p}}\Big)^2\Theta. \tag{29}$$

The matrix $\Omega = \Theta^{-1}/4$, as defined right under (5). The loss matrices Ω and Ω^N have the same eigenvectors since Θ and Θ^{-1} have the same eigenvectors. However, their eigenvalues Λ_i and Λ_i^N, which also drive the extent of losses, can differ:

$$\Lambda_i^N = \Lambda_i - \Big(\frac{\bar{y}}{\bar{p}}\Big) + \Big(\frac{\bar{y}}{\bar{p}}\Big)^2\frac{1}{4\Lambda_i} = \Lambda_i - \Big(\frac{\bar{y}}{\bar{p}}\Big)\Big(1 - \frac{\bar{y}}{\bar{p}}\frac{1}{4\Lambda_i}\Big). \tag{30}$$

Spending choice variables thus imply a lower eigenvalue associated with v^i, and according to (6) lower losses in this direction if and only if

$$\frac{4\bar{p}\Lambda_i}{\bar{y}} > 1. \tag{31}$$

To provide interpretation to this expression, let us introduce elasticity of consumption with respect to the price eigenvector v^i, i.e., the ratio of relative changes of consumption with respect to relative changes of prices along v^i(note that $Hv^i = \Theta^{-1}/2 = 2\Lambda_i v^i$),

$$\epsilon^i = \frac{\partial(\mathbf{y} \cdot v^i | \mathbf{p} \cdot v^i = x)/\partial x}{\bar{y}/\bar{p}} = \frac{2\bar{p}}{\bar{y}}\Lambda_i. \tag{32}$$

Condition (31) then takes the form of:

$$\epsilon^i > 1/2. \tag{33}$$

□

<u>Proof of Proposition 6.</u> Part (a) is an immediate implication of Lemma 2. This is because if both $\epsilon^1, \epsilon^2 > 1/2$, then losses are lower when thinking in terms of spending for any given form of information. Therefore, whatever information strategy the agent chooses when thinking in terms of consumption, then the agent can generate a higher objective when thinking in terms of spending by replicating the same information stately.

Part (b) is more involved. Consider the decomposition into 1D problems as in (6). The objective is then

$$\sum_i \max\Big(-\Lambda_i\sigma_0^2, -\frac{\lambda}{2} - \frac{\lambda}{2}\log\frac{2\Lambda_i\sigma_0^2}{\lambda}\Big). \tag{34}$$

The first element in the bracket is the objective if no information is processed, while the second is the utility from imperfect posterior beliefs less the cost of information.

If $\lambda > \lambda_1$, i.e., no information is processed, then the difference between the objective under spending and under consumption is $\sum_i(\Lambda_i^N - \Lambda_i)\sigma_0^2$, which equals

$$\sigma_0^2\Lambda_1\frac{2}{\epsilon^1}\Big(1 - \frac{1}{2\epsilon^1}\Big)(N-1) + \sigma_0^2\Lambda_2\frac{2}{\epsilon^2}\Big(1 - \frac{1}{2\epsilon^2}\Big). \tag{35}$$

where we used (38) to express Λ_i^N in terms of ϵ^i.

If $\lambda_2 < \lambda < \lambda_1$ then under consumption the agent processes information about $\mathbf{x} \cdot v^1$, but does not process information about $\mathbf{x} \cdot v^2$. We now express the difference between the objective under spending and under consumption when in both cases the information acquisition is optimal for the problem with consumption. The difference is:

$$\frac{\lambda}{2}\frac{2}{\epsilon^1}\Big(1 - \frac{1}{2\epsilon^1}\Big)(N-1) + \sigma_0^2\Lambda_2\frac{2}{\epsilon^2}\Big(1 - \frac{1}{2\epsilon^2}\Big). \tag{36}$$

The difference between objectives when information is chosen optimally under spending, too, is thus at least as high as this quantity. The second term in (36) is the same as in (35) since no information is processed about $\mathbf{x} \cdot v^2$ in either case. However, the first term is $-\frac{\lambda}{2} - \frac{\lambda}{2} \log \frac{2\Lambda_1 \sigma_0^2}{\lambda}$ for consumption and $-\frac{\lambda}{2}\frac{\Lambda_1^N}{\Lambda_1} - \frac{\lambda}{2} \log \frac{2\Lambda_i \sigma_0^2}{\lambda}$ for spending. The cost of information is the same in both cases, and drops out, and the losses from the same posterior beliefs are scaled by the corresponding eigenvalues.

Finally, if $\lambda < \lambda_2$ then the difference between the objectives under spending and consumption is higher than

$$\frac{\lambda}{2}\frac{2}{\epsilon^1}(1 - \frac{1}{2\epsilon^1})(N-1) + \frac{\lambda}{2}\frac{2}{\epsilon^2}(1 - \frac{1}{2\epsilon^2}), \tag{37}$$

which is again the difference between the objectives for information under spending being held at the optimal information under consumption.

All three differences between the two objectives (35)-(37) are for $\epsilon^1 > 1/2$ positive for sufficiently large N. In each of the expressions, the second term is independent of N, while the first terms are positive for $\epsilon^1 > 1/2$ and increasing linearly with N.

□

Proof of Proposition 7. We replicate the proof of Proposition 3 as long as the ordering of eigenvalues is the same regardless of whether thinking in terms of spending or consumption.

Plugging (32) into (30) we get

$$\Lambda_i^N = \Lambda_i \Big(1 - \frac{2}{\epsilon^i}(1 - \frac{1}{2\epsilon^i})\Big). \tag{38}$$

Using (30) we can express differences between eigenvalues for nominal variables:

$$\Lambda_i^N - \Lambda_j^N = (\Lambda_i - \Lambda_j)\Big(1 - \Big(\frac{\bar{y}}{2\bar{p}}\Big)^2 / (\Lambda_i \Lambda_j)\Big).$$

The right-hand side has the same sign as $(\Lambda_i - \Lambda_j)$, i.e., the ordering is the same for both decision variables, if and only if

$$\epsilon^i \epsilon^j > 1.$$

If this condition holds, then the analog of Proposition 3 applies. □

Proof of Corollary 2. We normalize $\overline{p} = 1$. Substituting our approximation $y_m = Y_m - (p_m - 1)\overline{y}$ in the agent's utility function, dropping terms the agent cannot influence, and rearranging gives the objective function

$$-\sum_m Y_m^2 - \sum_{m \neq n} \theta Y_m Y_n + \sum_m \underbrace{\left(\overline{x} - 1 + 2(p_m - 1)\overline{y} + \sum_{n \neq m} 2(p_n - 1)\overline{y}\theta\right)}_{\equiv X_m} Y_m. \tag{39}$$

Denote by $\tilde{X}_m$ the agent's posterior mean of X_m, and let $\mathbf{Y} = (Y_1, \ldots, Y_N)', \tilde{\mathbf{X}} = (\tilde{X}_1, \ldots, \tilde{X}_N)'$. We know that $\mathbf{Y} = \Theta^{-1}\tilde{\mathbf{X}}/2$, so $E[\mathbf{Y}] = \Theta^{-1}E[\tilde{\mathbf{X}}]/2$. Notice that $X_m - X_n = 2(1-\theta)(p_m - p_n)$. Since $\lambda_2 \leq \lambda < \lambda_1$, the agent acquires information about $p_m - p_n$, which is equivalent to acquiring information about $X_m - X_n$ but not about the sum of the X_m. Hence, a decrease in p_m lowers $E[\tilde{X}_m]$ and increases $E[\tilde{X}_n]$ for all $n \neq m$, leaving the sum unchanged. This lowers $E[Y_m]$ and raises $E[Y_n]$ for all $n \neq m$. □

Proof of Proposition 8.

The expected utility from choosing consumption y_m is

$$y_m - E[s(p(y_m) + x_m) - (1-s)p(y_m + x_m)] = y_m - sp(y_m) - (1-s)E[p(y_m + x_m)]. \tag{40}$$

Since $p(\cdot)$ is strictly convex, $E[p(y_m + x_m)] > p(y_m)$. Hence, the expected utility from choosing y_m is strictly increasing in s. As a result, the maximum of the above expression is also strictly increasing in s.

Let $q(\cdot)$ be the inverse of $p(\cdot)$. Note that $q(\cdot)$ is strictly concave. The expected utility from choosing spending Y_m is

$$E[sq(Y_m - x_m) + (1-s)(q(Y_m) - x_m)] - Y_m = sE[q(Y_m - x_m)] + (1-s)q(Y_m) - Y_m. \tag{41}$$

Since $q(\cdot)$ is strictly concave, $E[q(Y_m - x_m)] < q(Y_m)$. Hence, the expected utility from choosing Y_m is strictly decreasing in s. As a result, the maximum of the above expression is also strictly decreasing in s.

To complete the proof, we show that for $s = 0$ choosing spending is optimal, and for $s = 1$ choosing version is optimal. If $s = 0$, then Expression (40) is strictly less than $y_m - p(y_m)$, which is exactly Expression (41) with $Y_m = p(y_m)$, so fixing spending dominates fixing the version. If $s = 1$, then Expression (41) is strictly less than $q(Y_m) - Y_m$, which is exactly Expression (40) for $y_m = q(Y_m)$, so fixing the version dominates fixing spending. $\square$

C Asymmetries

Proposition 9. *Let $L = 2$ and $N = 4$ and let $\sigma^2_{0,i}$ denote the prior variance of x_i. If $\sigma^2_{0,1} \neq \sigma^2_{0,2} = \sigma^2_{0,3} = \sigma^2_{0,4}$, then there exist α, $\lambda_1 > 0$ and $\lambda_2 > \lambda_1$ such that*

$$\alpha y_1 + y_2 + \cdots + y_N = \textit{constant}$$

for all $\lambda > \lambda_1$, and no other non-trivial independent linear combination of the y_i's is constant for $\lambda < \lambda_2$.

Proof of Proposition 9. WLOG $\sigma^2_{0,2} = 1$. We transform the state-space such that in the new coordinates, $x'_1 = x_1/\sqrt{\sigma^2_{0,1}}$ and $x'_i = x_i$ for all $i > 1$, the prior variance-covariance matrix $\Psi = \sigma^2_{0,2} I$. The only other change to the original choice problem is that B_{11} of the matrix interacting actions and states is equal to $a = \sqrt{\sigma^2_{0,1}}$.

Now, we can compute the loss matrix Ω and its eigenvectors and eigenvalues. The eigenvalue Λ_s of the smallest absolute value is

$$\Lambda_s = \frac{\Big(-2a^2\theta - a^2 - \sqrt{4a^4\theta^2 + 4a^4\theta + a^4 + 12a^2\theta^2 - 4a^2\theta - 2a^2 + 1} - 1\Big)}{4(\theta - 1)(3\theta + 1)},$$

with the corresponding eigenvector:

$$v^* = \Big(\frac{\frac{1}{2} + a^2(-\frac{1}{2} - \theta) + \frac{1}{2}\sqrt{1 + 4a^4(\frac{1}{2} + \theta)^2 + a^2(-2. - 4\theta + 12\theta^2)}}{a\theta}, 1, 1, 1\Big).$$

This is the dimension of the state-space, to which the agent pays the least attention and for which the threshold cost of information above which no attention is paid to this dimension is the lowest.

Applying the matrix H to this dimension thus yields the direction in the action space that is relatively the most fixed, we get:

$$Hv^* = \Big(\alpha, 1, 1, 1\Big),$$

where

$$\alpha = \frac{-\frac{3}{2}a(1-\theta)^2\theta - \Big(\frac{1}{2} + \theta^2(-\frac{3}{2}+\theta)(-\frac{1}{2} + a^2(\frac{1}{2}+\theta) - \frac{SQ}{2})\Big)/\theta}{\frac{1}{2} - (1-\theta)^2\theta - \frac{3}{2}\theta^2 + \theta^3 + \Big(\frac{1}{2}(1+(-2+\theta)\theta)(-\frac{1}{2} + a^2(\frac{1}{2}+1\theta) - \frac{SQ}{2})\Big)/a},$$

where

$$SQ = \sqrt{1 + 4a^4(\frac{1}{2}+\theta)^2 + a^2(-2+\theta(-4+12\theta))}.$$

□

One natural question is whether the agent still engages in soft budgeting in the sense of the text, that an increase in the consumption of a good is associated with a decrease in the consumption of other goods much more than with full information. To measure this, we analyze the relative volatility of the budget, i.e., ratio of the variance of $\sum_i y_i$ and the sum of variances of the y_i. Intuitively, this answers how much the agent changes her budget relative to how much she changes the consumption levels of the individual goods. For $\lambda = 0.3$, for instance, this ratio is zero for $a = 1$ and increasing fairly slowly. For $\theta = \frac{1}{4}$ we find that the volatility of x_1 needs to be quadrupled, i.e., $\sigma_{0,1}^2 = 4\sigma_{0,2}^2$, to make the relative volatility of the budget a half of the relative volatility under perfect information.