NBER WORKING PAPER SERIES

INATTENTIVE ECONOMIES

George-Marios Angeletos
Karthik Sastry

Working Paper 26413
http://www.nber.org/papers/w26413

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2019, Revised December 2019

An earlier version circulated in December 2018 under the title "Welfare Theorems for Inattentive Economies." We thank Jakub Steiner for discussing our paper at the 2019 ASSA meeting in Atlanta; and Daron Acemoglu, BenjaminHébert, John Leahy, Jennifer La'O, Stephen Morris, Alessandro Pavan, and Harald Uhlig for comments and stimulating discussions. Angeletos acknowledges the support of the National Science Foundation under Grant Number SES-1757198. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Inattentive Economies
George-Marios Angeletos and Karthik Sastry
NBER Working Paper No. 26413
October 2019, Revised December 2019
JEL No. D5,D6,D8,E7

ABSTRACT

We study the efficiency of competitive markets when people are rationally inattentive. Appropriate amendments of the Welfare Theorems hold if attention costs satisfy an invariance condition, which amounts to free disposal of decision-irrelevant aspects of the state of nature. This condition is satisfied by the Shannon mutual information formulation of attention costs. More generally, inefficiency emerges and Hayek's (1945) argument about the informational optimality of prices fails. Markets are the best means of allocating scarce attention when agents gain nothing from directly contemplating prices rather than the entire state of nature.

George-Marios Angeletos
Department of Economics, E52-530
MIT
50 Memorial Drive
Cambridge, MA 02142
and NBER
angelet@mit.edu

Karthik Sastry
Department of Economics, MIT
77 Massachusetts Avenue E52-300
Cambridge, MA 02139
ksastry@mit.edu

1 Introduction

People are inattentive, forgetful, impulsive, and otherwise "cognitively constrained." They overlook some pieces of information and overreact to others. They use simplifications and heuristics.

In such circumstances, it is natural to question the efficacy of the market mechanism. Recent reviews by Maćkowiak, Matějka, and Wiederholt (2018) and Gabaix (2019), for instance, take for granted that the Welfare Theorems fail if people are inattentive.[1] This is true when comparing the outcomes obtained in the presence of inattention to those attainable in its absence. But if inattention is an unavoidable fact of life, the right question is whether welfare can be improved by means *other* than eliminating inattention.

A related question concerns Hayek's (1945) classic argument about the informational optimality of the price system. In his words:

> We must look at the price system as a mechanism for communicating information if we want to understand its real function. [...] The most significant fact about this system is the economy of knowledge with which it operates, or how little the individual participants need to know in order to be able to take the right action.

This argument presumes not only that markets are complete (Grossman, 1981) but also that prices are observed and decoded perfectly and costlessly. But if attention is a scarce resource, what exactly is the "economy of knowledge" achieved by markets? And, could welfare then be improved by regulating markets or even replacing them with other mechanisms?

We address these questions by augmenting the Arrow-Debreu framework with a generalized form of *rational inattention* and revisiting the two Fundamental Theorems of Welfare Economics.

Following Sims (1998, 2003) and the related literature, we treat "attention," "information processing," and "cognition" as interchangeable notions and model them as the choice of a signal subject to a cost. We allow great flexibility on what these signals and costs may be, as well as on how attention choices interact in general equilibrium.[2] The amended notion of Pareto optimality maps to a planner who may regulate people's choices and replace markets with other mechanisms but internalizes cognitive limitations.

The main results can be summarized as follows. If attention is modeled as the choice of a signal about the exogenous state of nature, as often done in the literature, the appropriately amended Welfare Theorems hold necessarily. If instead people pay attention directly to prices, market data, or the behavior of others, as seems plausible in practice, the amended Welfare Theorems require that attention satisfy a certain invariance condition.

This condition holds in the familiar benchmark that ties the cost of attention to Shannon mutual information. Away from this benchmark, though, a cognitive externality is possible: one agent's ease of decoding equilibrium objects can be endogenous to others' choices. Welfare may then be improved by various, context-specific policies, including both "noising up" and "stabilizing" prices—or, relatedly, by introducing new means of communication in addition to or in place of complete markets.

[1] In a similar vein, Sims (2010) writes: "If both sides of the market react to prices with rational inattention, then neither side is reacting precisely and immediately. Prices therefore cannot play their usual market-clearing role."

[2] We thus connect two disparate strands of the literature: one that focuses on decision theory and behavioral puzzles (Caplin et al., 2018; Kőszegi and Matějka, 2018; Steiner and Stewart, 2016; Woodford, 2019), and another that focuses on equilibrium interactions and policy (Colombo et al., 2014; Maćkowiak and Wiederholt, 2015; Myatt and Wallace, 2012; Tirole, 2015).

Example and main ideas (Section 3). We prove our main results in a modified Arrow-Debreu model without functional form restrictions. But we start with a closed-form example to build intuition.

There are two goods and a continuum of consumers. The demand for "coconuts" is subject to rational inattention; the consumption of "money" adjusts so as to meet the consumer's budget. The aggregate endowment of coconuts is the economy's only random fundamental. The cost of attention is (some transformation of) the Shannon mutual information between a noisy, idiosyncratic signal and a primitively specified *tracked object*. The latter is either the coconut endowment or their price.

An equilibrium of this "inattentive economy" exists, is unique, and is invariant to which of the two objects agents track. This equilibrium features various "pathologies" relative to the first best, including misallocation in coconuts and, in the case of endogenous production, excess volatility in aggregate output. But the equilibrium is constrained efficient in the sense that any attempt to manipulate the market of coconuts or the agents' attention choices can only reduce welfare.

The generality of this lesson cannot be fully addressed until the second part of our paper. Two exercises, however, provide intuition for what assumptions *can* open the door to inefficiency.

Suppose first that people struggle to discern smaller price changes. This is true, for example, under the axiomatic foundations of attention costs proposed by Pomatto et al. (2018), as adapted to our context. In this case, a policy that induces larger fluctuations in the price of coconuts is optimal. Conversely, if people struggle to track volatile or "complex" objects, welfare can be improved by stabilizing or "simplifying" prices. And, there can exist *cognitive traps*, or multiple Pareto-ranked equilibria in which suboptimal prices and suboptimal attention choices reinforce each other.

Consider next the possibility that cognitive mistakes can be correlated at zero or small enough cost. This opens the door to an inefficiency of a subtler, yet related, form. The equilibrium described earlier, with uncorrelated mistakes, continues to exist. But it is now Pareto dominated by other equilibria which economize cognitive resources by correlating mistakes and "noising up" prices.

The common thread behind these instances of inefficiency is the externality that emerges in cognition when the objects agents try to track and decode are endogenous to others' behavior. But why was efficiency preserved under *some* conditions? It must be these conditions were muting the cognitive externality. To clarify this logic, and identify general conditions for efficiency, we move on to the second part of our paper.

General framework (Section 4). Our rational-inattention extension of the Arrow-Debreu framework is extremely flexible. On the one hand, we let costs depart from the Shannon benchmark, nesting recent developments in decision theory. On the other hand, we accommodate rich equilibrium interactions and higher-oder uncertainty, connecting to the macroeconomic and game-theory literatures on rational inattention. What is crucial for our purposes is only the question of what objects people track. We say that "an agent tracks object z" if his cognitive cost is a functional C of the joint density of her signal and z. We then structure the formal arguments by considering different specifications for z.

Welfare Theorems for state-tracking economies (Section 5). We start with a case that is ubiquitous in the literature: we let z be the state of nature.

In this case, an appropriate amendment of the First Welfare Theorem is provided under arbitrary C. Any equilibrium of the inattentive economy is shown to coincide with the allocation preferred by a planner free to regulate people's choices but banned from eliminating or disregarding their cognitive costs.

The converse, or our version of the Second Welfare Theorem, also holds provided that the relevant convexity requirements are extended from the (familiar) domain of goods to the (new) domain of attention strategies. This is immediately guaranteed if C belongs in the class of posterior separable costs of information (as defined in Caplin and Dean, 2015).

What is going on? As long as people track directly the state of nature, *all* externalities are pecuniary. It is well known that the pecuniary externalities that pertain to consumption and production choices net out thanks to complete markets. Here we show that the same logic extends to the additional externalities that originate in attention choices, even though these choices are not directly priced.

This result clarifies the following points. First, attention does not give rise to the inefficiencies associated with information discovery, or innovation, because it is effectively a rival, non-transferable good. Second, the instances of inefficiency documented in Colombo, Femminis, and Pavan (2014) and Tirole (2015) derive not from inattention per se but rather from its interaction with other distortions, such as missing or non-competitive markets. And third, the failure of Welfare Theorems claimed in Gabaix (2014) depends on the use of either an *irrational* form of inattention or the "wrong" efficiency benchmark.

Importantly, this result also serves as a stepping stone for the next part of our analysis, which gets to the heart of Hayek's (1945) argument.

Welfare Theorems for price-tracking economies (Section 6). Consider now the case in which people have the option to obtain a signal *directly* about the prices they care about as opposed to the state of nature, and attention costs are appropriately adjusted to factor in the joint distribution of the signal and the prices. This maps to letting z contain the equilibrium price vector (the relevant endogenous outcome).

Because z is now endogenous to others' choices, the cost of observing or decoding it is also (generically) endogenous to others' choices. This opens the door to the cognitive externality mentioned earlier on. It also means that the price system plays a dual role. It not only clears markets but also has the potential for reducing cognitive costs, providing new formal meaning to Hayek's "economy of knowledge."

This begs the question of whether welfare could be improved by complementing or even replacing markets with other means of encoding and communicating the socially optimal course of action, in direct contradiction of Hayek's (1945) thesis. For instance, could welfare be improved by introducing a statistical agency that collects and disseminates data about the behavior of others? Or, could a "mediator" do better by directly instructing people what to do?

These questions make little sense in an Arrow-Debreu world without inattention, because a complete price system conveys *all* relevant information even if the state of nature is not directly observed (Grossman, 1981). But they become meaningful once inattention is added to the picture, *even if* markets are complete. A related issue is that, in such circumstances, taxes (on goods) may not only regulate people's attention to prices but also expand the very set of objects people can, or have to, pay attention to.

Addressing these issues requires a further amendment of the efficiency concept: the planner continues to internalize attention costs but may now send arbitrary messages in place of prices and/or the state of nature. The planner's problem thus resembles an information-design problem à la Bergemann and Morris (2013, 2019) and Kamenica and Gentzkow (2011), freed of incentive compatibility but ridden with costly information processing.[3]

[3] Recent contributions on Bayesian persuasion with an inattentive receiver have a similar flavor but are focused on substantially

The amended Welfare Theorems now hold under a certain restriction on the costs of attention, which is stated as Assumption 3 in the main text and can be informally summarized as follows:

INFORMATIONAL INVARIANCE. *Adding or subtracting irrelevant information in the tracked object z, or taking any invertible transformation of the relevant information, has no effect on attention costs.*

In other words, any excess information can be freely disposed of and the scaling or "framing" of the essential information does not matter. This restriction is necessarily satisfied if attention costs are measured by Shannon mutual information, but not more generally. It has the same basic, positive content as the axiom of "invariance under compression" defined by Caplin, Dean, and Leahy (2017) for decision problems, but plays a new, normative role in our equilibrium context: it guarantees *both* that there are no gains from manipulating the informational content of prices, or replacing them with arbitrary messages, *and* that there no losses from forcing people to track the entire state of nature.

Bottom line. Consider the following two questions:

Q1. *Are free markets the best mechanism for utilizing scarce attention in society?*

Q2. *Are there attention gains from tracking the relevant prices instead of the underlying state of nature?*

If the answer to Q2 is negative, which is the case implicitly imposed by Sims's original formulation of rational inattention, then the answer to Q1 is also negative. But if the answer to Q2 is positive, which seems more plausible in reality, then there is room for policies that aim at regulating markets and attention choices, *even if* markets are competitive and complete.

A corollary of this lesson is that, once the "economy of knowledge" is formalized in terms of rational inattention, Hayek's (1945) argument appears to contain an oxymoron: markets are the best mechanism for economizing on scarce attention only when they are no better than a direct mechanism that tells people the entire state of nature. When instead markets *strictly* economize on attention costs, there is generally room for further improvement.

Another corollary is a new perspective on the growing decision-theoretic and experimental literature exploring departures from Shannon mutual information. This literature focuses on how such departures are needed in order to capture certain choice patterns. Here we have shown how these departures are also necessary for making sense of prices' role in economizing cognitive costs.

2 Related Literature

The literature on rational inattention spurred by Sims (1998, 2003) is voluminous. Some works focus on single-agent behavior (Matějka, 2016; Matějka, Steiner, and Stewart, 2015); others study specific macroeconomic models (Maćkowiak and Wiederholt, 2009, 2015) or games (Colombo, Femminis, and Pavan, 2014; Myatt and Wallace, 2012). Our paper's contribution vis-à-vis all this literature is to adapt the analysis of rational inattention to the Arrow-Debreu framework, to develop the appropriate amendments of the Welfare Theorems, and to identify a new kind of inefficiency.

different questions. See the discussion in Section 2.

By modeling a general form of rational inattention, we encompass a growing literature that recasts various behavioral "anomalies" as the product of rational choice under noisy internal representations of the relevant objects. See, inter alia, Köszegi and Matějka (2018) and Lian (2018) on mental accounting; Woodford (2012) and Steiner and Stewart (2016) on prospect theory and winner's curse; da Silveira and Woodford (2019) and Kohlhas and Walther (2018) on overreaction; Ilut and Valchev (2017) on imperfect perception of the optimal policy rule; and Woodford (2019) on the broader agenda. While all these works focus on decision theory, we focus on equilibrium and efficiency; and while some of them depart from the Shannon benchmark in order to account for various experimental evidence, we highlight how such departures may also have important normative ramifications in equilibrium.

The condition that guarantees equilibrium efficiency in our setting has a similar flavor as the axiom of "invariance under compression" in Caplin, Dean, and Leahy (2017), as described earlier. But whereas that paper, like those cited in the previous paragraph, studies choice in a single-agent context, here we study efficiency in a general-equilibrium context.

Letting people track prices in our market environment is akin to letting players obtain information about the actions of others in a game. Denti (2016) studies the equilibria of such a game, but not its normative properties. The link between information acquisition and efficiency in games is further explored in a recent, complementary paper by Hébert and La'O (2019). They establish that an invariance condition on information costs similar to ours is necessary and sufficient for efficiency in a class of games in which payoff externalities net out and players can track the average action of others.

Closely related are also Angeletos and La'O (2018), Colombo, Femminis, and Pavan (2014) and Gul et al. (2017). These works identify conditions for efficiency in, respectively, a macroeconomic model in which firms are rationally inattentive, a game in which players choose their attention to different signals, and an endowment economy in which consumers' information is a coarsening of the true state space. These works specialize the model of inattention in various ways, most notably the exclude the possibility that people track and decode endogenous objects such as prices. We instead allow this possibility and proceed to shed light on when the ensuing cognitive externality may or may not be muted.

The same point distinguishes our paper from Myatt and Wallace (2012), Llosa and Venkateswaran (2017), Paciello and Wiederholt (2014), and Tirole (2015). In particular, our results clarify that the instances of inefficient information acquisition found in these papers derive from distortions such as non-competitive markets and nominal rigidity. By the same token, our "cognitive traps" borrow their name from, but are of a different origin than, those in Tirole (2015).

Our analysis brings to mind the literature on noisy rational-expectations equilibria (Grossman and Stiglitz, 1980; Kyle, 1985; Laffont, 1985; Vives, 2017). In that literature, agents perfectly observe, and effortlessly extract information from, prices; and inefficiency emerges only because of missing markets and/or monopoly power. In our context, instead, markets are complete and competitive; agents are optimally inattentive to prices; and inefficiency is of a different origin.

Finally, Vives and Yang (2018) share our emphasis on inattention and prices but model a different friction: they let inattention interfere with how people extract information from prices about fundamentals. In our context, learning about fundamentals from prices is allowed but is not essential for the results. Instead, the key friction is that people are inattentive to prices themselves.

3 An Inattentive Economy

We start with a tractable example that foreshadows our subsequent and more general Welfare Theorems and also sheds light on when inefficiency can be obtained.

3.1 Frictionless Benchmark

There are two goods, "coconuts" and "(real) money," and a continuum of agents, indexed by $i \in [0,1]$. Each agent has linear-quadratic preferences, represented by

$$\mathscr{U}(x_1, x_2) = x_1 - \tfrac{1}{2}x_1^2 + x_2, \tag{1}$$

where $x_1 \in \mathbb{R}$ and $x_2 \in \mathbb{R}$ denote the consumption of, respectively, coconuts and money. Each consumer receives respective endowments ξ and 1, where $\xi \sim N(\mu, \sigma^2)$. Let $\pi(\cdot)$ denote the associated prior density. For now, ξ is also the entire state of nature.

We let markets operate after ξ is realized, normalize the price of money to 1, and denote the (relative) price of coconuts by p. Momentarily, we also abstract from inattention.[4]

An equilibrium is an allocation $(x_{1i}(\xi), x_{2i}(\xi))_{i\in[0,1],\xi\in\mathbb{R}}$ and prices $(p(\xi))_{\xi\in\mathbb{R}}$ such that: each i maximizes (1) subject to her budget, $p(\xi)x_{1i} + x_{2i} \le p(\xi)\xi + 1$; and markets clear, or $\int_0^1 x_{1i}(\xi)\,\mathrm{d}i = \xi$ and $\int_0^1 x_{2i}(\xi)\,\mathrm{d}i = 1$.

Because of the symmetry in preferences and endowments, it is clear that "autarky" is the only equilibrium: $x_{1i}(\xi) = \xi$ and $x_{2i}(\xi) = 1$ for all $i \in [0,1]$. And because the agent's first-order conditions give her demand for coconuts as $x_{1i} = 1 - p$, the equilibrium price is $p = P(\xi) = 1 - \xi$.

3.2 Adding Rational Inattention

Now suppose agents cannot perfectly observe either ξ or p. Instead, each agent i conditions her demand of coconuts on a noisy signal, denoted by ω.

Each agent chooses a joint density $\phi(\cdot)$ over signal realizations $\omega \in \mathbb{R}$ and state realizations $\xi \in \mathbb{R}$. Let $\mathscr{D}^2$ denote the set of continuous probability distributions in $\mathbb{R}^2$. The agent with unrestricted signal technology can freely consider any such distribution that agrees with the prior on the second dimension, or an element from the set $\Phi \equiv \left\{\phi \in \mathscr{D}^2 : \int \phi(\omega, \xi)\,\mathrm{d}\omega = \pi(\xi), \forall \xi \in \mathbb{R}\right\}$.

For the present example, we will put more structure on the problem by restricting the signal to be

$$\omega_i = \xi + r\varepsilon_i,$$

where $\varepsilon_i \sim N(0,1)$ is noise, independently distributed across agents, and r is a coefficient under i's control. Thus the agents' choice set of distributions is a strict subset of the set Φ defined above.

Attention costs. The agent incurs a non-pecuniary, or "cognitive," cost for generating any signal. Let $C : \mathscr{D}^2 \to \mathbb{R}$ be some arbitrary cost functional that "scores" any joint distribution $\phi(\cdot) \in \mathscr{D}^2$. As a leading case, we may consider cost functionals that can be expressed as the reduction in the "perceived complexity" of ξ, measured in some information units.

[4] In standard Arrow-Debreu fashion, our general framework (Section 4) assumes that markets operate before the realization of uncertainty, allowing agents to insure. In the present example, insurance is not an issue due to the quasi-linearity of preferences.

Formally, we let $H : \mathscr{D}^1 \to \mathbb{R}$, mapping one-dimensional densities to real numbers, be one such measure of complexity, and we take $C : \mathscr{D}^2 \to \mathbb{R}$ to equal the expected reduction in H after observing ω:

$$C[\phi(\cdot)] = \mathbb{E}\big[H[\pi(\cdot)] - H[\phi(\cdot \mid \omega)]\big] \tag{2}$$

where $\pi(\cdot)$ is the prior about the underlying state (here ξ), $\phi(\cdot \mid \omega)$ is the posterior given ω (obtained via Bayes rule), and the expectation is over signal realizations.

When $H[\cdot]$ returns the Shannon entropy of a random variable, expression (2) is the Shannon mutual information between the signal and state. This is the form favored by Sims (1998, 2003) and much (but not all) of the subsequent literature.

In the present example, with Gaussian fundamentals and signals, the mutual information cost has the following, even simpler representation in terms of the signal-to-noise ratio, or equivalently the correlation δ between the signal and the state:

$$C[\phi(\cdot)] = c(\delta) \equiv -\log(1-\delta) \tag{3}$$

where $\delta \equiv \sigma^2/(\sigma^2 + r^2)$. To start with, we will contemplate attention costs that can be written as a function only of δ, and no other features of the distribution $\phi(\cdot)$.

The consumer problem. Because the noise in ω_i is idiosyncratic, the aggregate demand and the price of coconuts are functions of only ξ. Thus let $p = P(\xi)$, for some $P(\cdot)$ to be determined in equilibrium. The consumer's problem can then be expressed as follows:

$$\begin{aligned} \max_{x_{1i}(\cdot), x_{2i}(\cdot), \delta_i} \int \left(x_{1i}(\omega_i) - \frac{x_{1i}(\omega_i)^2}{2} + x_{2i}(\omega_i, \xi) \right) \phi(\omega_i, \xi)\, \mathrm{d}\omega_i \mathrm{d}\xi - c(\delta_i) \\ \text{s.t. } P(\xi) x_{1i}(\omega_i) + x_{2i}(\omega_i, \xi) \le w(\xi), \forall (\omega_i, \xi) \end{aligned} \tag{4}$$

where $w(\xi) \equiv P(\xi)\xi + 1$ represents the consumer's wealth. Note that this problem contains, not only the optimal consumption of coconuts conditional on ω_i, but also the optimal choice of the joint distribution ϕ, the joint distribution of ω_i and ξ, as parametrized by the scalar δ_i. Also note that, because of the quasi-linearity in preferences, the consumer does not care to condition his consumption of coconuts on his wealth—from his perspective, any signal about ξ is *merely* a signal about the price.

3.3 Equilibrium: Definition and Characterization

We introduce the following equilibrium concept, which is self-explanatory.

Definition. An *inattentive equilibrium* is a collection $\left\{\delta, [x_1(\omega), x_2(\omega, \xi)]_{\omega,\xi}, [P(\xi)]_{\xi}\right\}$, such that:

1. δ and $[x_1(\omega), x_2(\omega, \xi)]_{\omega,\xi}$ solve the consumer's problem;
2. all markets clear, or
$$\int x_1(\omega)\phi(\omega|\xi)\mathrm{d}\omega = \xi \qquad \text{and} \qquad \int x_2(\omega, \xi)\phi(\omega|\xi)\mathrm{d}\omega = 1 \quad \forall \xi,$$
where $\phi(\omega|\xi)$ denotes the likelihood of ω conditional on ξ, as implied by the equilibrium δ.

For a given δ (i.e., information structure), one can guess and verify the following solution for the equilibrium price and consumption plan:

$$p = P(\xi) \equiv 1 - \left(1 - \frac{1}{\delta}\right)\mu - \frac{1}{\delta}\xi, \qquad x_{1i} = \omega_i, \qquad \text{and} \qquad x_{2i} = 1 + P(\xi)\xi - P(\xi)\omega_i. \tag{5}$$

Two properties are worth noting.

First, the consumption of coconuts is no more equated across agents: it is *as if* there is uninsured idiosyncratic risk, or mis-allocation of coconuts. However, as shown below, this phenomenon ceases to be a a call for market regulation once the "right" efficiency benchmark is considered.

Second, because there are only supply shocks at the aggregate level, the inverse of the price function identifies the aggregate demand function. This is true whether inattention is present or not. However, as shown in Appendix D, the area below the identified aggregate demand properly measures consumer surplus *only* in the absence of inattention. The same applies to producer surplus (in an extension with inattentive production). In a nutshell, Harberger triangles are no longer meaningful.

These properties illustrate how rational inattention invalidates standard calculations of the compensating transfers associated with any simple market correction (e.g., tax or quota). And yet, we shall show that a key normative property of the frictionless benchmark—that such corrections, as well as other manipulations of individual choices, are undesirable—remains true.

To this goal, we must compare the equilibrium to the relevant planner's problem not only in terms of the allocation of coconuts obtained for given δ but also in terms of the value of δ itself. We characterize the equilibrium value of δ in the rest of this subsection and the planner's counterpart in the next subsection.

In equilibrium, the individual's attention choice reduces to the following problem:

$$\max_{\delta_i}\{b(\delta_i,\delta)-c(\delta_i)\}$$

where $b(\delta_i,\delta)$ is the expected utility, evaluated along the equilibrium consumption plan and the equilibrium price function seen in (5), when others choose δ and the individual chooses δ_i. Computations, detailed in the appendix, show that up to scaling constants

$$b(\delta_i,\delta)=\frac{\sigma^2\delta_i}{2\delta^2}-\frac{\sigma^2}{\delta} \tag{6}$$

Think of b as the reduced-form benefit of attention and c as its cost. The dependence of $b(\delta_i,\delta)$ on δ captures the dependence of the individual's utility on the attentiveness of *others*. Clearly, an equilibrium corresponds to any δ^* such that

$$\delta^*\in\arg\max_{\delta}\left\{b(\delta,\delta^*)-c(\delta)\right\} \tag{7}$$

From an individual's perspective, more attention is always better ($b_1>0$), because it reduces the mistakes in consumption choices. But the private returns to attention are higher when others are less attentive ($b_{12}<0$), implying that attention choices are strategic substitutes: precisely when others are inattentive, and hence market prices are very volatile, there are high gains to making accurate predictions.

This substitutability property guarantees that the equilibrium is unique, provided that it exists. Existence follows from the continuity of c and uniqueness from $c'(\cdot)$ increasing with $\lim_{\delta\to 1}c'(\delta)=\infty$:

Proposition 1 (Equilibrium)**.** *The equilibrium exists and is unique. The equilibrium level of attention,* $\delta^*\in(0,1)$, *is the unique solution to* (7)*. The equilibrium price and allocation are as in* (5)*, with* $\delta=\delta^*$.

3.4 Welfare and Efficiency

The following casual argument might suggest that there is room for inefficiency. Consider the problem of a benevolent planner who can*not* intervene in the market of coconuts, and therefore takes the function b as given, but can dictate agents' attention choices. This planner solves the following problem:

$$\delta\in\arg\max\{b(\delta,\delta)-c(\delta)\}. \tag{8}$$

On the margin, this planner equates $c'(\delta)$ with $b_1(\delta,\delta)+b_2(\delta,\delta)$, where $b_1(\delta,\delta)$ and $b_2(\delta,\delta)$ measure, respectively, the marginal private value of attention and the externality imposed on others. In equilibrium, the agents instead equate $c'(\delta)$ with $b_1(\delta,\delta)$ alone. It follows that, for generic $b(\cdot)$, the equilibrium and the planner's solution won't coincide.

This argument is meaningful if the function $b(\cdot)$ is "free" for the modeler to choose, as in the class of games studied in Tirole (2015). But $b(\cdot)$ is *not* arbitrary in our context.

Proposition 2 (Efficiency)**.** *The fixed point to* (7) *coincides with the solution to* (8)*. That is, the equilibrium level of attention coincides with the socially optimal one.*

Why is this true? As anticipated in the Introduction, the relevant externalities are (so far) purely pecuniary: one's attention enters others' welfare via prices only. But as long as utility is transferable (as in the present example) or markets are complete (as in our upcoming general analysis), pecuniary externalities do not create a wedge between the private and the social value of attention. They net out on average, guaranteeing that $b_2(\delta,\delta)$ is zero for all δ.

Appendix A illustrates this logic within the present example. Section 5 establishes its greater applicability. And while the argument made above precludes the planner from intervening in the market of coconuts, the argument applies more generally in the following sense: when the planner has the option to tax or otherwise regulate the market of coconuts, $b_2(\delta,\delta)$ may cease to be zero *because* of such an intervention, but such an intervention is not worthwhile in the first place, for it only reduces $b(\delta,\delta)$. In other words, laissez-faire both minimizes the wedge in the *marginal* value of attention and maximizes the *total* value of attention.

3.5 Correlated Noise and Non-Fundamental Volatility

So far, we have equated the state of nature with the payoff-relevant fundamental, namely the endowment of coconuts, ruling out aggregate noise or correlation devices. We now sketch how one could incorporate such variables and explain why they do not, *by themselves*, upset the efficiency of the equilibrium.

Retain that $\xi \sim N(\mu,\sigma^2)$ and introduce aggregate white noise $\nu \sim N(0,1)$. The state of nature is now given by $\theta \equiv (\xi,\nu)$. Next, let the signals take the following form:

$$\omega_i = \xi + r_i\varepsilon_i + s_i\nu,$$

where $\varepsilon_i \sim N(0,1)$ is i.i.d. and the pair (r_i, s_i) is chosen by the agent, subject to some cost.

If we make no other change in the environment, there cannot exist an equilibrium in which a non-zero mass of agents set $s_i > 0$. Such an equilibrium would have aggregate demand move with ν, which would violate market clearing given that supply is fixed at ξ. If, however, we let the supply be elastic and make appropriate assumptions about c, we can support an equilibrium in which all agents choose $s_i = s > 0$ and, as a result, non-fundamental volatility emerges in both prices and quantities.[5]

Still, the equilibrium remains efficient, for the same reason as before: all externalities, including those associated with the choice of the optimal load on the aggregate noise, are purely pecuniary. The same

[5]First, introduce a technology that allows the second good to be transformed to the first and by letting an attentive firm to operate it. And second, suppose the cost of attention is decreasing in both s_i and r_i but more steeply so in s_i than in r_i (costs can be economized by substituting idiosyncratic for aggregate noise).

logic applies if we consider the more flexible information structures proposed in Myatt and Wallace (2012), Colombo, Femminis, and Pavan (2014) and Tirole (2015) These structures allow for rich, endogenous correlation in noise, but do not alone upset the efficiency of the equilibrium. They are all nested in our subsequent, more general analysis of state-tracking economies (Section 5).

3.6 Price-tracking Economies and Cognitive Externality

So far, we have focused on an economy in which agents collect signals about the state of nature. In equilibrium, such signals serve also as signals about the price. But the cost of any given signal was specified as a function of its joint density with the underlying state. This case defines what we call "state-tracking economies." The complement, referred to as "price-tracking economies," allows the cost to depend on the joint density of the signal and the price itself, capturing the idea that the difficulty of tracking prices depends on *their* stochastic properties.

In this variant, agents collect a signal ω directly about p.[6] Their cognitive cost remains some $C : \mathscr{D}^2 \to \mathbb{R}$, but now the chosen $\phi(\cdot)$ has to agree with the prior on p, which itself is endogenous to equilibrium. Our leading specification has $C[\phi(\cdot)]$ be the mutual information between ω and p. But other specifications will also be considered.

When choosing ϕ, or equivalently the likelihood for ω given p, the agent treats the marginal for p as given. But this marginal distribution is itself determined in equilibrium by the choices of others. And because this enters $C[\cdot]$, a non-pecuniary, or "cognitive," externality emerges: by affecting the equilibrium price mapping, one's choices can affect the cognitive costs and the attention choices of others.

This has a similar flavor as the informational externality found in the literature that follows the traditions of Grossman and Stiglitz (1980), Grossman (1981) and Laffont (1985). But it differs from it in two ways. First, it does not derive from missing markets. And second, its bite depends crucially on the "units" of information as embedded in the specification of $C[\cdot]$. We illustrate this point in the rest of this section by providing first a benchmark in which this externality is muted and then two examples in which it is active.

3.7 An Efficient Price-Tracking Economy

Assume the following two restrictions on the primitives of the economy.

A1. The entire state is $\theta = \xi$, which rules out correlated noise.

A2. The cognitive cost is an increasing, convex function of the mutual information between ω and p as long as both objects are Gaussian, and infinite otherwise.

By A1, p must be a function of ξ. By A2, (ω, p) must be jointly Normal. It follows that p must be a linear function ξ. And because the mutual information between ω and any monotone function of ξ is the same as the mutual information between ω and ξ itself, it is *as if* agents are tracking ξ instead of p.

Proposition 3. *Under restrictions A1 and A2, the equilibrium of the price-tracking economy coincides with that of the corresponding state-tracking economy.*

[6] In the present example, agents care to know *only* the price of coconuts: because of the quasi-linearity in preferences, their endowment of coconuts ("their wealth") is irrelevant for their optimal consumption of coconuts. Of course this is not a generic property, and in our more general analysis we will let agents track *at least* their "own" fundamentals in addition to prices.

Introduce now a planner. Suppose, for the present purpose, that this planner cannot replace the market mechanism. But let him manipulate, via taxes or other instruments, the agents' consumption and attention choices, subject to A1: the state cannot be expanded to include variables other than ξ.

By manipulating the consumption and attention choices of all agents, the planner can induce a different mapping from ξ to p, thus also manipulating agents' prior about p. In general, this could have allowed the planner to economize cognitive costs and improve upon the equilibrium. This is, however, not the case here due to A2. First, any *non-linear* mapping from ξ to p cannot be optimal, because it induces infinite cognitive costs. Second, any (non-flat) *linear* mapping from ξ to p entails the same cognitive costs as the equilibrium one. We thus reach the following conclusion.

Proposition 4. *Under restrictions A1 and A2, the equilibrium of the price-tracking economy coincides with the solution to the planner's problem described above.*

In a nutshell, A1 and A2 make sure that the cognitive externality is muted and, hence, that efficiency is preserved (at least in the sense described above). But what if we relax these assumptions?

3.8 Inefficiency I: Externalities and "Complexity"

Let us now return to the more general form of "reduction in uncertainty or perceived complexity" captured by (2), without committing to the specific "units" imposed by Shannon entropy.

Say agents are performing physical experiments that have a fixed precision in the problem's *cardinal* units: it is easier to tell $p = 1$ from $p = 2$ than $p = 1$ from $p = 1.1$. Pomatto et al. (2018) show how to construct a large family of cost functionals that embody this intuition and obey a sensible set of other axioms for information acquisition or cognition.[7] One particularly convenient form, which is their leading case and also a common "reduced-form" choice in the literature (e.g., Wilson, 1975; Van Nieuwerburgh and Veldkamp, 2010) is cost proportional to the *precision* of the generated signal (i.e., $1/r^2$). In our notation from (2), this is tantamount to having $H[\cdot]$ return the precision (inverse variance) of the random variable represented by the given density. It follows that the cost of attention can be expressed as

$$c(\delta_z, \sigma_z^2) = \frac{1}{\sigma_z^2} \frac{\delta_z}{1-\delta_z}$$

where σ_z^2 is the variance of the tracked object and δ_z is, as before, the correlation of the signal with the tracked object (or, up to a monotone transformation, the signal-to-noise ratio).

Similarly to the benchmark with Shannon mutual information, attention costs are increasing and convex in δ_z (i.e., $c_1 > 0$ and $c_{11} > 0$). But unlike that benchmark, both the total and the marginal cost are now decreasing in variance σ_z^2 (i.e., $c_1 < 0$ and $c_{12} < 0$).

When z is the state of nature, this dependence is irrelevant. But when z is the equilibrium price, it creates both a negative externality and a source of strategic substitutability in attention choices. Higher levels of aggregate attention translate into less variable prices, which in turn translate to higher total cognitive costs and higher marginal costs of paying attention. The latter property preserves the uniqueness of the equilibrium, the former upsets its efficiency. A planner internalizing the externality would favor taxing attention and/or noising up prices.

[7]These include constant returns to scale in running additional physical or "cognitive" (thought) experiments, and appropriate notions of monotonicity and continuity.

We can generalize these ideas by allowing for cost representation $c(\cdot)$ to depend arbitrarily on (δ_z, σ_z^2). The opposite case than that studied above, or $c_1 > 0$ and $c_{12} > 0$, seems quite plausible as well: on the margin, it may be harder to pay attention when prices are more volatile or more erratic, because larger changes across states are confusing.

This case switches the sign of the cognitive externality and of the desirable policy intervention: the planner favors corrections that reduce price volatility, for instance, by introducing a tax on the consumption of coconuts that increases with ξ. This case also introduces strategic complementarity in attention choices: "when others pay less attention, prices are more confusing, and I find it harder to pay attention myself." If this force is sufficiently strong to overcome the substitutability described in equation (6), the "game" of choosing δ may admit multiple equilibria. Low δ equilibria have more price volatility, larger cognitive externalities, and lower welfare than high δ equilibria. Hence the former are "cognitive traps":

Proposition 5 (Cognitive Traps I)**.** *There exist c, with $c_2 > 0$ and $c_{12} > 0$, such that the economy admits multiple, Pareto-ranked equilibria, each corresponding to a different value for δ^*.*

These results illustrate how departures from the Shannon benchmark open the door to a particular kind of inefficiency, one that lets the "complexity" of the price system matter. The particular examples considered above tied "complexity" to the *variance* of p. But one can also think of examples that tie "complexity" to "sparsity" (Gabaix, 2014), "perceptual distance" (Hébert and Woodford, 2018), and other traits that describe how easy it is for people to track and decode decision-relevant objects such as prices.

3.9 Inefficiency II: Correlated Mistakes

Let us re-embrace the units of Shannon entropy but allow "cognitive noise" to be correlated across agents. We shall show that this, too, allows for inefficiency and cognitive traps, although of a more subtle kind. We shall also discuss why this may depend on whether the correlated noise is interpreted as internal or external to the cognitive process.

First, let the state be $\theta = (\xi, \nu)$ and express the signal as

$$\omega = p + r\varepsilon_i + s\nu + t\xi \tag{9}$$

where ν is aggregate noise, ε is idiosyncratic noise, and (r, s, t) are scalars under the control of the agent (we are henceforth suppressing the i index). Second, let the cost depend only on the mutual information between ω and p. Under the Gaussian restriction, this means that the cost can be written as $c(\delta_p)$, where δ_p is the signal-to-noise ratio between ω and p.

Conjecture now that prices depend on both elements of the state:

$$p = -g\xi + h\nu,$$

for some scalars g and h. An equilibrium is indexed by $(r, s, t, g, h, \delta_p)$. In Appendix A, we show that equilibrium imposes only five restrictions over these six scalars: there is a one-dimensional continuum of equilibria. The lessons are even sharpest if the cost takes the form of a "hard" capacity constraint on the mutual information between ω and p, which amounts to fixing δ_p exogenously.

Proposition 6 (Cognitive Traps II)**.** *Fix any $\delta_p \in (0, 1)$. There exist a continuum of equilibria, indexed by $h \in [0, \bar{h}]$, such that (i) non-fundamental volatility and welfare both increase in h and (ii) when $h = \bar{h}$, the allocation of goods is first-best.*

The Appendix works out the math. It also extends the argument to the case in which $c(\cdot)$ is a smooth increasing function of the mutual information between ω and p. This boils down to letting different equilibria be associated with different values for δ_p. Here, we sketch the main ideas treating δ_p as given. Clearly, there exists an equilibrium in which $s = t = h = 0$ and $r > 0$. This equilibrium coincides with that of the baseline, state-tracking economy studied in the beginning of this section.

Let us now show that there is another equilibrium, which attains the first best. Set $r = 0$. This gives $\omega = p + s\nu + t\xi = p_0 + (t-g)\xi + (s+h)\upsilon$. In equilibrium, the cross-sectional average of ω has to equal ξ, or else aggregate demand would not equal ξ. It follows that $g = 1$ and $s = -h$, and hence ω perfectly reveals ξ and $p = p_0 - \xi + h\upsilon$. This occurs for a unique level of non-fundamental volatility, $h = \bar{h} \equiv \sigma\sqrt{\delta_p^{-1} - 1}$, which intuitively increases with the fundamental variance.

Not surprisingly, there also exist intermediate equilibria, mixing the previous two. All of them, as well as the one that replicates the state-tracking outcomes, are inferior equilibria, or "cognitive traps."

The logic can be summarized as follows. Correlated noise is used to simultaneously *reduce* the mutual information between signals and prices and *increase* the mutual information between the signals and the underlying fundamentals. When cognitive costs come only from tracking prices, the first property economizes cognitive costs, while the second brings allocations closer to their first-best counterparts.

This begs the question of whether the correlated noise *itself* could be costly. The previous scenario, in which ν was costless, may make sense if this "common mistake" is *internal* to people's cognitive process. But if υ is an *external* impulse (e.g., a literal or metaphorical "sunspot"), it may require effort to learn about.

In Appendix A we work out an extension in which there are costs to tracking both p and ν. Whereas in the above example an individual was happy to noise up his signal of p with either ε_i or υ, she strictly prefers ε_i once υ is costly. This selects the equilibrium with $h = 0$ as the unique equilibrium.[8] But it does *not* guarantee its efficiency. In particular, if the cost of tracking υ is small enough relative to the cost of tracking p, the unique equilibrium is dominated by an allocation that features $h > 0$. Furthermore, the planner can implement this allocation by introducing a subsidy that varies with υ. That is, even though the multiplicity disappears, the argument for "noising up" prices remains.

When does this argument cease to hold? Only when the cost of tracking υ is "comparable" to the cost of tracking prices. In particular, if the total cognitive costs can be expressed as the mutual information of an individual's signal with the pair (p, υ), the efficiency of the equilibrium is restored. The same is true if the cost depends on the mutual information of the individual's signal with the pair (p, θ), where $\theta = (\xi, \upsilon)$ is the entire state of nature, inclusive of both fundamental and non-fundamental shocks. This anticipates a more general result, which we develop in the second part of the paper.

3.10 Taking Stock

The inefficiency documented in the last example does not require υ to be a "true" sunspot or "pure" noise; υ could have been a fundamental affecting some other, un-modeled market, or even the market under consideration. What opened the door to inefficiency is an asymmetry in the costs of attention: υ was

[8] This brings to mind Yang (2015) and Morris and Yang (2016). The former shows that equilibrium multiplicity survives in a global game in which players are rationally inattentive but correlated noise is effectively costless. The latter selects a unique equilibrium by making such correlation costly.

cheaper to track, or more "salient," than p or ξ.

This logic extends to our first example of inefficiency, which featured $\theta = \xi$ and $C[\cdot] = c(\delta_z, \sigma_z^2)$, for $z \in \{\theta, p\}$. As long as $c_2 = 0$ (the textbook scenario), the cost of tracking p was the same as that of tracking θ. But once $c_2 \neq 0$ (the scenario allowing for "complexity" or "perceptual distance"), this symmetry was broken, and inefficiency obtained.

This discussion hints at two insights. First, when agents are allowed to track not only the exogenous state of nature but also endogenous objects such as prices or the choices of others, efficiency requires that the costs of attention satisfy some kind of symmetry, or invariance, with respect to the various objects agents can track. Second, behavioral notions such as "complexity," "salience," and "framing" may be open the door to inefficiency *within* the rational-inattention framework if they are interpreted as violations of this kind of invariance. The precise form of this invariance, and its relation to mutual information, are made clear in Section 6.

4 General Framework

Do the preceding normative lessons extend to more general competitive economies? Does it matter what preferences are, which consumption or production choices are subject to inattention, how budgets are met, or how markets clear? What if people can track taxes or other objects in addition to prices? What if the planner could complement or even replace market prices with other means of communication?

This section lays the groundwork for clearly answering these questions by augmenting the standard Arrow-Debreu framework with a generalized form of rational inattention. The adopted formulation not only nests multiple specifications found in the literature but also helps zero in on the *only* modeling choices of essence for our purposes: whether people track the exogenous state of nature or the endogenous objects of interest, and whether doing the latter economizes costs relative to doing the former.

4.1 Frictionless Benchmark

The state of nature is represented by a random variable θ, drawn from a finite set $\Theta \subset \mathbb{R}$ according to probability distribution $\pi \in \triangle(\Theta)$. This variable is meant to contain not only payoff-relevant fundamentals, like the endowment ξ in the example of Section 3, but also aggregate noise or correlation devices, like ν in that example.[9]

There is a finite number of underlying, non-contingent commodities, indexed by $n \in \{1, ..., N\}$. In the standard Arrow-Debreu fashion, markets are complete and operate *ex ante*, before θ is revealed. We let $p(\theta) = \left(p_n(\theta)\right)_{n=1}^N \in \mathbb{R}_+^N$ be the price vector for state θ, where $p : \Theta \to \mathbb{R}_+^N$.

In the frictionless benchmark, we could easily redefine the commodity space to include all combinations of goods and states. But the separate notation for goods and states *does* matter in our formalization of rational inattention. For that reason, we use consistent notation here.

[9] At the present level of abstraction the distinction between fundamentals, noise and sunspots can be quite fuzzy. Suppose for instance that the economy is the union of two "islands" entirely disconnected from one another. Then, one island's fundamentals could serve as the other island's sunspot. This qualifies our stylized example, but does not interfere with the more general results we provide in the sequel.

Consumers. There is a unit-measure continuum of households, split into a finite number J of distinct types indexed by $j = \{1, ..., J\}$. Preferences and endowments can differ across types, but consumers of the same type are identical. The mass of type j is given by $\mu^j \in (0,1)$, with $\sum_j \mu^j = 1$.

Let $x^j(\theta) = \left(x_n^j(\theta)\right)_{n=1}^N \in \times_n X_n \equiv X \subset \mathbb{R}_+^N$ denote the consumption bundle for the typical household of type j in state θ, where $x^j : \Theta \to X \subset \mathbb{R}_+^N$. We assume that preferences are given by expected utility:

$$\sum_\theta u^j\left(x^j(\theta), \theta\right)\pi(\theta).$$

where $u^j : X \times \Theta \to \mathbb{R}$ is a type-specific, state-contingent, Bernoulli utility. We next write the budget as

$$\sum_\theta p(\theta) \cdot x^j(\theta) \leq \sum_\theta \left(p(\theta) \cdot e^j(\theta) + a^j \Pi(\theta)\right)$$

where $e^j(\theta)$ is the endowment of type j in state θ, $\Pi(\theta)$ are any state-contingent firm profits, and a^j is the profit share of household type j.

Firms. There is a unit-measure continuum of identical firms. We let $y(\theta) = \left(y_n(\theta)\right)_{n=1}^N \in \times_n Y_n \equiv Y \subset \mathbb{R}^N$ denote the production plan, or input-output vector, of the typical firm in state θ. By convention, we allow outputs to enter as positive numbers and inputs to enter as negative numbers.

The technology is given by production transformation frontier $F : \mathbb{R}^N \times \Theta \to \mathbb{R}$ such that the production plan $y(\theta)$ is feasible in state θ if and only if $F\left(y(\theta), \theta\right) \leq 0$.

4.2 Generalized Rational Inattention

We now introduce our general form of rational inattention.

Signals, tracked objects, and attention costs. Let ω be a random variable representing an individual's "cognitive state," or an internally generated signal about z. To ease notation, we suppress the indexing of ω on the identity of the individual.

For a fixed individual, we will allow *different decisions* to possibly depend on different (sub)signals. To this end, we write $\omega = (\omega_n)_{n=1}^N$, where N is the number of goods, and ω_n is the sub-component of ω upon which the demand of good n is conditioned. As discussed later on (Section 4.6), this formulation serves a dual purpose. First, it accommodates an arbitrary specification of how budgets are satisfied in the presence of inattention. And second, it helps capture "narrow thinking" and mental accounting, as in Lian (2018). For now, we assume that each sub-signal is a real number in a discrete-valued set Ω_n, for all $n \in \{1, \ldots, N\}$. Thus, $\omega \subseteq \Omega \equiv \times_n \Omega_n \subseteq \mathbb{R}^{\mathrm{N}}$.

Let z be a "tracked object," a random variable that a given agent tries to learn or reason about. In a leading case ("state-tracking economies"), z will coincide with the exogenous state of nature, or $z = \theta$. But we will also consider cases in which z contains market prices, the trades of others, or policy instruments such as taxes. In all such cases, z will ultimately be some transformation of θ, taking value in (a subspace of) $\mathbb{R}^W$ for some finite W. To accommodate all these cases, we fix a primitive, but essentially irrelevant, W and define the set of all subspaces of $\mathbb{R}^W$ as $\mathscr{Z} \equiv \{Z : Z \subseteq \mathbb{R}^W\}$.[10]

[10] Our only primitive requirement is $W \geq N + 1$. This allows us to nest the case in which $z = (\theta, p)$, namely the case in which agents are tracking the state $\theta \in \Theta \subset \mathbb{R}$ along with the price vector $p \in \mathbb{R}_+^N \subset \mathbb{R}^N$.

Costs of attention are a non-pecuniary cost of choosing a certain joint distribution between the signal ω and the tracked object z. These costs are controlled by a type-specific functional C^j that is well-defined on such distributions for *any* given identity of z and its co-domain Z:

$$C^j : \bigcup_{Z \in \mathscr{Z}} \Delta(\Omega \times Z) \to \mathbb{R} \cup \{\infty\} \tag{10}$$

The possibility of infinite costs exists to capture any possible restrictions on the space of admissible signals and/or tracked objects (e.g., distributional requirements, if desired).

State of nature. Strictly speaking, the *entire* state of nature is now given by the combination of θ with the idiosyncratic draws of ω for each and every individual. But since we shall assume that a law of large number applies within each type of agents, θ remains the only *aggregate* state variable for the economy as a whole, because all type-specific average quantities are ultimately measurable in it, as well as the only *relevant* state variable for the problem of any individual, because all prices and fundamentals are also measurable in it. With these qualifications in mind, we henceforth refer to θ simply as the state of nature.

Additional useful notation. For any given type j agent, any chosen joint distribution $\phi^j(\cdot)$ over (ω, z) implicitly defines also a joint distribution over (ω, z, θ), given by

$$f^j(\omega, z, \theta) \equiv \phi^j(\omega, z) \cdot \mathbb{I}\{z(\theta) = z\} \tag{11}$$

where $\mathbb{I}\{\cdot\}$ is the 0-1 indicator function that takes the value 1 if and only if $\{\cdot\}$ is true, and, in some abuse of notation, $z(\cdot)$ is the function that maps the value of the underlying state to the values of tracked object. Note that this embeds the restriction that an individual's cognitive state ω depends on the state of nature θ only through its relationship with the tracked object.

The implied marginal density for the pair (ω, θ) and the conditional likelihood of ω given θ are then given by, respectively,

$$g^j(\omega, \theta) \equiv \sum_{z \in Z} f^j(\omega, z, \theta) \qquad \text{and} \qquad g^j(\omega \mid \theta) \equiv \frac{g^j(\omega, \theta)}{\pi(\theta)} \tag{12}$$

4.3 New Consumer and Firm Problems

Fix a tracked object, and hence the space Z in which z takes values. Let $\pi_z \in \triangle(Z)$ be the prior distribution for this object, which of course individual agents take as given. For any such object, the consumer and firm problems can be formulated as follows.

Each agent chooses two objects: a distribution ϕ that represents the obtained signal about z; and a mapping of the realization of the signal to an action. For type any type-j consumer, this mapping is a consumption strategy $x^j : \Omega \to X$; and for firms, it is a production strategy $y : \Omega \to Y$. By construction of ω and its sub-components, these strategies can be written as "separable" across goods, or $x^j(\omega) = (x_n^j(\omega_n))_{n=1}^N$ and $y(\omega) = (y_n(\omega_n))_{n=1}^N$ for all ω and for some collection of functions $x_n^j : \Omega_n \to X_n$ and $y_n : \Omega_n \to X_n$, respectively. Let us denote the sets of all such strategies as

$$\mathscr{X} \equiv \left\{ x(\cdot) = (x_n(\cdot))_{n=1}^N \text{, for some } x_n : \Omega_n \to X_n, \forall n \right\}$$
$$\mathscr{Y} \equiv \left\{ y(\cdot) = \left(y_n(\cdot)\right)_{n=1}^N \text{, for some } y_n : \Omega_n \to Y_n, \forall n \right\}$$

Pick a consumer of type j. The optimal attention choice and the optimal consumption strategy can then be expressed as the solution to the following problem:

$$\begin{aligned} &\max_{x(\cdot),\phi(\cdot)} \sum_{\omega,\theta} u^j(x(\omega),\theta)g(\omega,\theta) - C^j[\phi(\cdot)] \\ &\text{s.t. } (x(\cdot),\phi(\cdot)) \in \mathbf{B}(p(\cdot), e^j(\cdot), a^j\Pi(\cdot)) \end{aligned} \tag{13}$$

where

$$\begin{aligned} \mathbf{B}(p(\cdot), e^j(\cdot), a^j\Pi(\cdot)) \equiv \Big\{ &\text{functions } x(\cdot) \in \mathscr{X} \text{ and } \phi(\cdot) \in \Delta(\Omega \times Z) \text{ such that :} \\ &\sum_{\omega,\theta} (p(\theta)\cdot x(\omega))g(\omega \mid \theta) \le \sum_{\theta}\left(p(\theta)\cdot e^j(\theta^j) + a^j\Pi(\theta)\right) \\ &\sum_{\omega} \phi(\omega,z) = \pi_z(z), \ \forall z \in Z \Big\} \end{aligned} \tag{14}$$

This set encodes the budget constraint, the measurability (inattention) constraints, and the natural restriction that the marginal on z agree with the prior. Also, in all the above expressions, g should be read as a transformation of ϕ: the joint density $g(\omega,\theta)$ and the conditional likelihood $g(\omega|\theta)$ vary with the agent's choice of ϕ, and is computed in the same fashion as in the objects seen in expression (12).

Each firm, on the other hand, solves the following problem:

$$\begin{aligned} &\max_{y(\cdot),\phi(\cdot)} \sum_{\omega,\theta} (p(\theta)\cdot y(\omega))g(\omega,\theta) \\ &\text{s.t. } (y(\cdot),\phi(\cdot)) \in \mathbf{F} \end{aligned} \tag{15}$$

where

$$\begin{aligned} \mathbf{F} \equiv \Big\{ &\text{functions } y(\cdot) \in \mathscr{Y} \text{ and } \phi(\cdot) \in \Delta(\Omega \times Z) \text{ such that :} \\ &F\left(y(\omega),\theta^f\right) + C^f[\phi(\cdot)] \le 0, \forall(\omega,\theta) \text{ s.t. } g^f(\omega,\theta) > 0 \\ &\sum_{\omega} \phi(\omega,z) = \pi_z(z), \ \forall z \in Z \Big\} \end{aligned} \tag{16}$$

Note that "attention" enters as an additive cost to the production possibilities frontier. And as in the consumer's problem, g should be read as the joint distribution of ω and θ implied by the firm's attention choice.

4.4 Inattentive Equilibrium

Let the previous constructions of consumer and firm programs, indexed still by a tracked object $z \in Z$, constitute an "inattentive economy" We now define general equilibrium in this context:[11]

Definition 1 (Inattentive Equilibrium)**.** An equilibrium is a profile of consumption-production strategies, $\left([x^j(\cdot)]_{j=1}^J, y(\cdot)\right)$, attention choices, $\left([\phi^j(\cdot)]_{j=1}^J, \phi^f(\cdot)\right)$, and prices, $p(\cdot)$, such that the following are true.

1. *Consumers optimize*: for each j, $(x^j(\cdot),\phi^j(\cdot))$ solves program (13), fixing prices and the stochastic process for tracked object z^j.
2. *Firms optimize*: $(y(\cdot),\phi^f(\cdot)$ solves program (15), fixing prices and the stochastic process for tracked object z^f.

[11] Throughout, we focus on equilibria in which strategies are symmetric *within* types. But this is without serious loss of generality, because we can partition types into sub-types with the opportunity to make different decisions. Same point applies to our efficiency concept in the sequel.

3. *Markets clear*: for all $\theta \in \Theta$,

$$\sum_{j=1}^{J} \mu^j \bar{x}^j(\theta) = \sum_{j=1}^{J} \mu^j e^j(\theta) + \bar{y}(\theta)$$

where

$$\bar{x}^j(\theta) \equiv \sum_{\omega} x^j(\omega) g^j(\omega \mid \theta) \qquad \text{and} \qquad \bar{y}(\theta) \equiv \sum_{\omega} y(\omega) g^f(\omega \mid \theta).$$

are, respectively, the aggregate demand of type-j consumers and the aggregate supply of firms.

This definition is self-explanatory. The only notable subtlety is that the tracked objects $(z^j)_{j=\{1,\ldots,J,f\}}$ are potentially endogenous and form part of the equilibrium fixed point. The definition of these tracked objects indexes the "type" of equilibrium being studied.

4.5 On the Tracked Object

The sense of "agents' tracking some object" relates to the identity of z and, thus, to how the cost functional C penalizes learning. In the stylized example of Section 3, z was variously the aggregate endowment of coconuts or their price. Here, we have allowed for a flexible specification of z. For our main analysis, though, we shall concentrate on the two scenarios described in the following definitions:

Definition 2. A *state-tracking economy* is an economy in which $z = \theta$ for all j, and $Z = \Theta \subseteq \mathbb{R}$.

Definition 3. A *price-tracking economy* is an economy in which $z = (\theta, p)$ for all j, and $Z = \Theta \times \mathbb{R}^N \subseteq \mathbb{R}^{N+1}$.

As mentioned in the Introduction, the first scenario captures the vast majority of the existing macroeconomic and game-theoretic applications of rational inattention.

The second scenario is more subtle. In a given equilibrium, p will always be some function of the state θ. What the second scenario accommodates is the possibility that the particular transformation $p(\theta)$ obtained in equilibrium may have a *different* cost to track than θ itself, and that this aspect of the cost endogenously changes *across* different equilibria (either multiple equilibria of the same primitive set-up, or equilibria indexed by policy interventions that manipulate attention choices and/or market prices).

The same subtlety applies if we consider cases in which z contains other equilibrium objects, such at the trades of other agents, or policy instruments, such as taxes. Although all these objects are ultimately functions of the state of nature, the costs of tracking the *could* be different than the cost of tracking the state of nature itself.

Finally, a modification of the second scenario that feels even more natural in our context is one in which each agent tracks only the sub-component of (θ, p) that directly enters her decisions (e.g., a consumer's own endowment and the prices of the particular basket of goods she consumes). As discussed at the end of Section 6, this scenario can either be treated as a separate primitive case (by redefining z accordingly), or be obtained as an equilibrium *implication* of our invariance condition within the more flexible price-tracking scenario defined above (by exploiting free disposal of any irrelevant information in z).

4.6 Clarifications and Remarks

We close this section with a couple of clarifying examples of what our formulation can accommodate and a few additional remarks.

Information acquisition. Here we provide a specialization of our setting that clarifies how it connects to the growing decision-theoretic literature on rational inattention.

Suppose that all agents (consumers and firms) have no measurability constraints on their choice of the last good, $n = N$. This is an "adjustment good," like the second good in our simple example or saving (tomorrow's consumption) in Sims (2003, 2006) Suppose further that the choice of all other goods must be jointly measurable in a single (scalar) cognitive state s drawn from some set $S \subseteq \mathbb{R}$. The cognitive state thus has the restriction $\omega_n \equiv s$ for $n < N$ and $\omega_N = (s, z)$.

In this case, it is natural to write the cognitive cost in terms of the joint density between s and z. Furthermore, we can use the budget to solve out for the consumption of the last good and to express the realized utility as a function of the remaining goods, which are all measurable in s. We can thus transform our constrained optimization problem to the kind of unconstrained optimization problems typically found in the more abstract, decision-theoretic literature on rational inattention.

We can then also allow $C[\cdot]$ to fall into a number of familiar forms for cost of information. Consider first the class of posterior-separable costs introduced by Caplin and Dean (2015) and Caplin et al. (2017), and axiomatized by Denti (2018) and Mensch (2018). Let $\phi(\cdot \mid s) \in \triangle(Z)$ denote the posterior distribution over a specific tracked object space induced by a specific signal; $\pi(\cdot) \in \triangle(Z)$ denote the prior; $\phi(s) \in \Delta(S)$ denote the marginal likelihood of a given cognitive state realization; and $G : \triangle(Z) \times \triangle(Z) \to \mathbb{R}$ be a functional strictly convex in the first argument for any value of the second.[12] We can write a posterior-separable cost functional as

$$C[\phi(\cdot)] = \sum_{s \in S} G[\phi(\cdot \mid s), \pi(\cdot)]\phi(s) - G[\pi(\cdot), \pi(\cdot)]$$

When $G[\cdot]$ has no dependence on the second argument, the form is uniformly posterior separable and also admits a "reduction in complexity" representation like (2). This posterior separable class nests both Shannon mutual information and several attractive deviations thereof (e.g., Tsallis, 1988; Hébert and Woodford, 2018; Morris and Strack, 2019).

Modeling decision-specific learning. The aforementioned decision-theoretic works allow multiple actions but constrain them to depend on the same signal. Our formulation relaxes this constraint in part because this is needed for budget constraints to be satisfied and in part because this helps capture an imperfection in how the same agent coordinates his different decisions (Lian, 2018).

To fix ideas, consider the following functional form that merges costs of information with "cross-decision restrictions." Let $C^n : \triangle(\Omega_n \times Z) \to \mathbb{R}$ be a decision-specific cost of information and let $X : \Delta(\Omega) \to \mathbb{R} \cup \infty$ be a functional that rewards or penalizes certain relationships between the cognitive states themselves. Assume the cognitive cost can be written as

$$C^j[\phi(\cdot)] = \sum_{n=1}^{N} C^{jn}[\phi^n(\cdot)] + X^j[\phi^\omega(\cdot)] \tag{17}$$

where $\phi^n(\cdot)$ is the joint density of ω_n with z, and $\phi^\omega(\cdot)$ is the marginal density over ω. The first term captures arbitrary costs for collecting each internal signal ω_n, which could even be zero (to nest the "adjustment good" example). The second flexibly rewards or penalizes various correlation structures between the signals—for instance, to reward (or punish) high correlation between the signals.

[12] It is simple, but notationally cumbersome, to extend all of these definitions to arbitrary spaces of the tracked objects.

Different specifications could encode how budgets are met in the more elaborate fashion assumed in Gabaix (2014), by returning infinite cost when signal realizations are in conflict with the budget constraint. Alternatively, they could help capture "narrow bracketing" and "mental accounting" as in (Lian, 2018). We omit precise "constructions" of $C[\cdot]$ to match these or other behavioral models. But we hope to convey that our model is quite general.

Markets and games. Because firms and consumers make their choices under imperfect observation of prices, they effectively play a game of incomplete information. See Angeletos and Pavan (2007) and Angeletos and La'O (2010) for concrete examples of how competitive economies with informational frictions can be mapped to games of incomplete information.

In such a "competitive" game, actions (demands or supplies) can be either strategic substitutes or strategic complements. This may affect attention choices in the ways indicated by Hellwig and Veldkamp (2009) and Myatt and Wallace (2012). Furthermore, because θ is arbitrary, the environment may feature rich higher-order uncertainty, as in the literature spurred by Morris and Shin (1998, 2002) and Woodford (2003) and reviewed in Angeletos and Lian (2016).

As it will become clear, none of these elements are relevant for understanding efficiency in our class of economies. At the same time, thinking of the class of economies we study as games hints at how some of our results can be extended to games. We return to this point in the concluding section.

5 State-Tracking Economies

In this section, we focus on the scenario in which agents track only the exogenous state of nature ($z=\theta$ and $Z=\Theta$). We first define an efficiency concept that imposes that all agents track the state of nature and show that, relative to such a benchmark, state-tracking economies are efficient.

5.1 Constrained Efficiency

We envision a planner who cannot alter the underlying physical environment (inclusive of the cognitive costs and the restriction that agents only track the exogenous state), but can freely control people's consumption and production choices as well as their attention strategies. This is formalized by modifying the familiar feasibility and efficiency concepts as follows.

Definition 4 (Feasibility)**.** A profile of consumption-production choices, $([x^j(\cdot)]_{j=1}^J, y(\cdot))$, and attention strategies, $([\phi^j(\cdot)]_{j=1}^J, \phi^f(\cdot))$, is feasible in a state-tracking economy if it satisfies the following restrictions:

$$\sum_{j=1}^J \mu^j \sum_\omega x^j(\omega)\phi^j(\omega \mid \theta) = \sum_{j=1}^J \mu^j e^j(\theta) + \sum_\omega y(\omega)\phi^f(\omega \mid \theta), \forall \theta \in \Theta \tag{18}$$

$$F(y(\omega),\theta) + C[\phi^f(\cdot)] \leq 0,\ \forall(\omega,\theta) \text{ s.t. } \phi^f(\omega,\theta) > 0 \tag{19}$$

$$x^j \in \mathscr{X}, \forall j \in \{1,\ldots,J\} \quad \text{and} \quad y \in \mathscr{Y} \tag{20}$$

$$\phi^j(\cdot) \in \triangle(\Omega \times \Theta),\ \forall j \in \{1,\ldots,J,f\} \tag{21}$$

$$\sum_\omega \phi^j(\omega,\theta) = \pi(\theta),\ \forall \theta \in \Theta, \forall j \in \{1,\ldots,J,f\} \tag{22}$$

Definition 5 (Efficiency)**.** A profile of consumption-production choices and attention strategies is *efficient* in a state-tracking economy if there exists no other such profile that is feasible in the sense of Definition 4, strictly preferred by a positive mass of agents, and weakly preferred by all other agents.

The first two restrictions in Definition 4 give the economy's resource constraints and production technology. The third captures the choice-specific measurability constraints. The fourth gives the domain of the available information structures. The fifth clarifies that the information structures need to agree with the prior. A final restriction, implicit in the adopted notation but of critical importance, is that each agent's decision have to be measurable in her own, noisy signal. By the same token, Definition 5 thus departs from standard Pareto optimality in two ways. First, it embeds the informational constraints through the amended notion of feasibility. And, second, it counts the cognitive costs of any informational structure in the evaluation of welfare by respecting the agents' own preferences over different information structures.[13]

Our version of the First Welfare Theorem will establish that, regardless of the cost functional C, any inattentive equilibrium in a state-tracking economy is an efficient allocation in the sense of the above definition. Our version of the Second Welfare Theorem will establish that the converse is also true under additional convexity restrictions. Efficiency can then be represented in the following planner's problem.

PLANNER'S PROBLEM. *An efficient allocation is a solution to the following problem:*

$$\max_{[x^j(\cdot),\phi^j(\cdot)]_{j=1}^J,(y(\cdot),\phi^f(\cdot))} \sum_{j=1}^{N} \chi^j \mu^j \left[\sum_{\omega,\theta} u^j(x^j(\omega),\theta)\phi^j(\omega,\theta) - C^j[\phi^j(\cdot)] \right] \tag{23}$$

$$s.t.\ (18),\ (19),\ (20),\ (21),\ and\ (22).$$

for some Pareto weights $(\chi^j)_{j=1}^J$.

Had information been exogenous (i.e., had $C^j[\cdot]$ been infinite but for a single ϕ^j for all j), the planner's problem would be similar to that studied in Angeletos and Pavan (2007). In that benchmark, the planner dictates how agents *use* their dispersed information, but has not control over the information structure itself. The key novelty here is precisely that the planner chooses a socially optimal information structure, taking into account the associated information costs.

5.2 Intuition with First-order Conditions

Our proofs of the amended Welfare Theorems do not require differentiability with respect to either the goods or the attention choices. Differentiability of C with respect to ϕ is not even well defined at the level of generality we have afforded so far. But to gain intuition we start with a simple, informal argument in terms of first-order conditions.

Consider first the planner's first-order condition for a specific good n, type j, and cognitive state ω:

$$\mathbb{E}\left[\frac{\partial u^j(x^j(\omega),\theta)}{\partial x_n} \mid \omega\right] = \mathbb{E}\left[\frac{\lambda_i(\theta)}{\chi^j} \mid \omega\right] \tag{24}$$

[13]This means that the following is true: although a first-best allocation of goods may be *feasible* in the sense of Definition 4, it does *not* have to efficient in the sense of Definition 5. Intuitively, this is true whenever a signal perfectly revealing of θ is available (i.e., has finite cost for all j) but too costly to be optimally chosen.

where $\lambda_n(\theta)$ is the Lagrange multiplier on the resource constraint for good n. Consider next the corresponding equilibrium condition of a type-j household:

$$\mathbb{E}\left[\frac{\partial u^j(x^j(\omega),\theta)}{\partial x_n} \mid \omega\right] = \mathbb{E}\left[m^j p_i(\theta) \mid \omega\right] \tag{25}$$

where m^j is the marginal value of wealth for type j (the Lagrange multiplier on type j's budget constraint).

Clearly, these two conditions coincide if $\lambda_n(\theta) = p_n(\theta)$ and $\chi^j = \frac{1}{m^j}$, meaning that the planner's shadow value coincides with equilibrium prices and that the Pareto weights equal the reciprocal of the equilibrium marginal values of wealth. Both of these requirements are satisfied here in the exact same manner as in the textbook version of the Welfare Theorems. The only novelty is the presence of the expectation operator in conditions (24) and (25). This reflects the informational, or cognitive, friction.

In the language of Angeletos and Pavan (2007), the coincidence of conditions (24) and (25), which obtains holding ϕ^j constant, means that the equilibrium *use* of information is efficient. We now show that efficiency extends to equilibrium *acquisition* of information.

Suppose that C^j is a differentiable function of each $\phi^j(\omega,\theta)$, evaluated at a pair $(\omega,\theta) \in \Omega \times \Theta$. We can then write the planner's first-order condition for the choice of attention as follows:

$$u^j(x^j(\omega),\theta) - \frac{\partial C^j}{\partial \phi^j(\omega,\theta)} = \frac{\lambda(\theta)}{\chi^j} \cdot x^j(\omega) \tag{26}$$

This again parallels the consumer's first-order condition. We can do a similar exercise for the choices of firm production and attention.

As long as first-order conditions and feasibility constraints, at equality, are sufficient for characterizing a solution to (23), we have a basic proof of the Welfare Theorems. Of course, in asserting the sufficiency of first-order conditions, we are presuming convexity with respect to both the goods and the attention strategies. But such convexity is actually needed only for the Second Welfare Theorem. Furthermore, while the above argument requires differentiability of the cost function with respect to the attention choice, our actual proofs dispense with it and thus bypass the need to even define what such differentiability means in the space of arbitrary attention choices.

5.3 The First Welfare Theorem

In a standard Arrow-Debreu economy, one proves that competitive equilibria are Pareto efficient using only local non-satiation in preferences. A sufficient extension of this condition to our case is the following:

Assumption 1. *For every* $j \in \{1,\ldots,J\}$, $x(\cdot):\Omega \to \mathbb{R}^N$, $\phi(\cdot) \in \Delta(\Omega\times\Theta)$, *and* $\varepsilon > 0$, *there exists some* $x'(\cdot) \in \mathscr{B}_\varepsilon(x(\cdot)) \equiv \{x''(\cdot) : \left\|[x''(\omega) - x(\omega)]_{\omega\in\Omega}\right\| < \varepsilon\}$ *and some* $\phi'(\cdot) \in \Delta(\Omega\times\Theta)$ *such that* j *strictly prefers* $(x'(\cdot),\phi'(\cdot))$ *to* $(x(\cdot),\phi(\cdot))$.

Under the maintained simplification that attention costs are separable from the utility of goods, this assumption is immediately satisfied if u^j itself features non-satiation. In any event, with this assumption in hand, we can extend the First Welfare Theorem to the presence of rational inattention.

Theorem 1 (First Welfare Theorem for state-tracking economy)**.** *Let Assumption 1 hold. Then, any inattentive equilibrium that has strictly positive prices is efficient (in the sense of Definition 5).*

It is obvious from our reformulation of the consumer problem that, in any inattentive equilibrium, resources are optimally allocated across different realizations of ω, within each type. The problem that remains, of allocating resources across the types j, is familiar to the analogue without rational inattention. Generating a Pareto improvement requires expanding the budget sets of all agents; combining this with the result of profit maximization generates the familiar contradiction and proves the result. More succinctly, inefficiency is ruled out because all externalities, inclusive of the new ones that pertain to attention choices, are purely pecuniary and net out thanks to complete, competitive markets.

Seen from this perspective, Theorem 1 is not terribly surprising. Indeed, our reformulations of consumer's and firm's problems equates the choice attention to, respectively, a form of "home production" and the use of an non-traded, firm-specific input. This in turn clarifies why rational inattention is quite different from information discovery or innovation, each of which is generally associated with inefficiency in otherwise competitive GE settings.

Theorem 1 also clarifies the following points about the existing literature on inattention. First, the failure of the Welfare Theorems observed in Gabaix (2014) rests on a deviation from rational attention. Second, the related point made by Maćkowiak et al. (2018) is valid only if one moves away from settings in which agents only track the exogenous state of nature (as we do in Section 6 below). Third, Sims's (2010) concern that "prices cannot play their usual market-clearing role" may be relevant for the existence of the equilibrium but not for its efficiency. And fourth, the instances of inefficiency found in Colombo, Femminis, and Pavan (2014) and Tirole (2015) derive, not from rational inattention (or imperfect cognition) per se, but rather from its interaction with other distortions, such as missing or non-competitive markets.

5.4 The Second Welfare Theorem

The standard version of the Second Welfare Theorem requires convexity of preferences and production sets. These convexity assumptions can be dispensed within our setting, because there is a continuum of agents per type and because the planner can use the noise in the agents' signals to replicate lotteries. But because different signals induce different costs, a convexity assumption is required in their domain.

Assumption 2. *The cognitive cost is (weakly) convex over the distribution of posteriors induced by any given signal ω about the physical state θ.*

Theorem 2 (Second Welfare Theorem for state-tracking economy)**.** *Impose Assumption 2. Any efficient allocation in the sense of Definition 5 can be supported as a state-tracking equilibrium.*

Note that the required convexity is naturally tied to convexity of the cost of information.[14] Moreover, assuming convex costs of information is also the most natural way to assume compatibility with the Blackwell ordering of experiments.[15] All posterior separable representations (as discussed in Subsection 4.6) have this property.

[14]This is immediate when we have a representation like (17), the information cost functionals $C^{jn}[\cdot]$ are convex, and and the cross-state restriction functional $X^j[\cdot]$ is also convex for all agents j.

[15]The reason follows from Jensen's inequality—a Blackwell garbling, in this context, is a mean-preserving spread of the distribution of posteriors, and this increases costs because of convexity.

6 Price-tracking Economies

We now extend the analysis to price-tracking economies (Definition 3). To simplify the exposition, we shut down production and focus on pure exchange economies.

6.1 Constrained Efficiency Revisited: Adding Messages

In the present context, the notion of efficiency introduced in the previous section is unappealingly restrictive. Insofar as markets allow agents to economize on cognitive costs by tracking prices rather than the underlying state of nature, it seems natural to give the planner also a larger toolkit to expand or even replace such "market communication."

We capture this idea by allowing the planner to send a *message* m^j, which arbitrarily depends on the state of nature, to each agent of type j. The collection of these messages, $m = (m^j)_{j\in\{1,\ldots,J,f\}}$, replaces the market prices, or more generally the "tracked object" z^j, in each individual's information-processing problem. In an nutshell, where the price-tracking equilibrium has each agent of type j tracking $z^j = (\theta, p)$, the planner's problem has that agent tracking $z^j = m^j$.

We assume effectively no restrictions on what these messages are. To be more concrete, we let $\mathscr{M}$, the common feasible set of message choices to send to *any* agent type, be

$$\mathscr{M} = \left\{ m : \Theta \to Z \text{ for any } Z \in \mathscr{Z} \right\} \tag{27}$$

where, to recall previous notation, $\mathscr{Z}$ indicates all possible subsets of $\mathbb{R}^W$ for some primitive and finite (but arbitrarily large) W. Thus the planner is choosing mappings from the state to real vectors, which are precisely the tracked objects for which we have defined agents' costs of attention.

Why is this a natural formulation of the problem? We suggest two reasons. The first relates to Hayek's (1945) argument about the role of the price system as a communication mechanism. If we manage to show that the planner cannot improve upon market outcomes by complementing or even replacing prices with any other conceivable combination of messages, then we will have shown the robustness of Hayek's (1945) argument to rational inattention.

The second reason is more practical. When we study the equilibrium of a competitive economy *without* policy intervention, it is natural to restrict attention to the scenario in which people track the prices they care about. But when taxes, regulations, and government spending are added, there is no immediate and natural extension to allow "automatic" tracking of these objects.

To illustrate, suppose that the policy maker may impose a rich set of state-, type- and good-specific ad valorem taxes, but cannot communicate anything but these taxes. Then, it would be natural to study a *restricted* version of our planner's problem where the messages are constrained to satisfy, for all j and θ,

$$m^j(\theta) = (\theta, p(\theta), \tau^j(\theta))$$

where $\tau^j(\cdot)$ is the tax schedule faced by type j. This would not only have restricted the space of the available messages, but also tie the instrument used for communication with the one used for incentives. By ignoring all such restrictions and letting the planner send arbitrary messages, we are solving a relaxed problem. If we manage to show that the equilibrium in the absence of policy intervention coincides with the solution to this relaxed problem, then we will have also guaranteed that no conceivable policy intervention can be welfare improving, regardless of how this policy affects incentives and/or cognition.

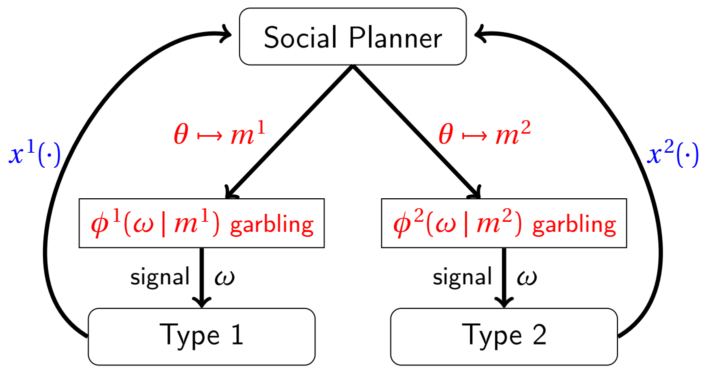

Figure 1: The planner's problem as a communication problem.

6.2 The Planner's Problem as a Communication Problem

The messages introduced above summarize how the aggregate state θ determines the agents' signals and actions. Put differently, there can be no correlation in signals ω between two distinct agents (either within or across types) conditional on the messages. For certain specification of C, this may provide a potential trade-off for the planner: too much information could confuse agents, while too little would render them unable to make good decisions.[16]

Figure 1 illustrates the planner's problem in a two-type example. The choices in red, and the "inner" part of the diagram, represents the planner's communication. But there is an additional step, in which the planner recommends allocations $(x^1(\cdot), x^2(\cdot))$. These are effectively free from any incentive-compatibility or implementability constraints, because the planner is implicitly assumed to have an arbitrarily rich toolkit of tax instruments. But allocations must be measurable in ω, because people receive noisy signals of the planner's messages. The planner can effectively choose these signals, but fully internalizes the associated cognitive costs.

All in this, this problem can be thought of as a multi-agent information-design problem (Bergemann and Morris, 2019), freed of any mis-alignment of incentives between the principal and the agents but augmented with costly communication.[17] A key question is then whether cognitive costs can be economized, and better, more informed, choices can be induced, by sending different messages than those corresponding to market prices. This is the new angle we offer to Hayek's (1945) classic argument.

These ideas are formalized by amending the earlier notions of feasibility and efficiency as follows:

Definition 6 (Feasibility with arbitrary message)**.** A combination of messages, $[m^j(\cdot)]_{j=1}^J$, attention choices, $[\phi^j(\cdot)]_{j=1}^J$, and consumption choices, $[x^j(\cdot)]_{j=1}^J$, is *feasible* if it satisfies the following restrictions:

[16] Formally, this restriction is nested in the construction of the joint densities $\phi(\cdot)$ and $g(\cdot)$ in Definition 6 below.

[17] This recalls a few recent contributions that study Bayesian persuasion with an inattentive receiver (Lipnowski et al., 2019; Wei, 2018; Bloedel and Segal, 2018). These works have a similar flavor as ours but are focused on a substantially different question: how the sender's optimal communication depends on the misalignment of incentives between him and the receiver. Such misalignment could be present even in the absence of inattention (as in Kamenica and Gentzkow, 2011). But inattention itself becomes the source of mis-alignment in these papers because the sender does *not* internalize the receiver's attention cost. In our context, there is no incentive problem to start with (thanks to complete, competitive markets) and the planner fully internalizes people's attention costs. The last point also distinguishes our paper from Farhi and Gabaix (2019).

$$g^j(\omega \mid \theta) = \frac{\sum_m \phi^j(\omega, m) \cdot \mathbb{I}\{m^j(\theta) = m\}}{\pi(\theta)}, \forall j \in \{1, \ldots, J, f\} \tag{28}$$

$$\sum_{j=1}^{J} \mu^j \left(\sum_{\omega} x^j(\omega) g^j(\omega \mid \theta) \right) = \sum_{j=1}^{J} \mu^j e^j(\theta) + \left(\sum_{\omega} y(\omega) g^f(\omega \mid \theta) \right), \forall \theta \in \Theta \tag{29}$$

$$F(y(\omega), \theta) + C^f[\phi^f(\cdot)] \leq 0, \ \forall (\omega, \theta) \text{ s.t. } \phi^f(\omega, m^f(\theta)) > 0 \tag{30}$$

$$x^j(\cdot) \in \mathscr{X}, \forall j \in \{1, \ldots, J\} \quad \text{and} \quad y(\cdot) \in \mathscr{Y} \tag{31}$$

$$m^j(\cdot) \in \mathscr{M}, \ \forall j \in \{1, \ldots, J, f\} \tag{32}$$

$$\phi^j(\cdot) \in \triangle(\Omega \times M^j), \ M^j = m^j(\Theta), \ \forall j \in \{1, \ldots, J, f\} \tag{33}$$

$$\sum_{\omega} \phi^j(\omega, m) = \sum_{\theta} \pi(\theta) \mathbb{I}\{m^j(\theta) = m\}, \ \forall m \in M^j, \forall j \in \{1, \ldots, J, f\} \tag{34}$$

Definition 7 (Efficiency with messages)**.** A profile of messages, $[m^j(\cdot)]_{j=1}^J$, attention choices, $[\phi^j(\cdot)]_{j=1}^J$, and consumption choices, $[x^j(\cdot)]_{j=1}^J$, is *efficient* if there is no other profile that is feasible in the sense of Definition 6, strictly preferred by a positive mass of agents, and weakly preferred by all other agents.

Definition 6 is the same as Definition 4 but with the added flexibility about how people's signals and hence their decisions can depend on θ through the messages that the planner sends in place of prices. The efficiency concept from Definition 5 carries over with the same amendment, yielding Definition 7.

Provided convexity, we can then represent efficiency as the solution to the following problem.

Planner's Problem. *An efficient profile is a solution to the following problem:*

$$\max_{[m^j(\cdot)]_{j=1}^J, [x^j(\cdot), \phi^j(\cdot,\cdot)]_{j=1}^J} \sum_{j=1}^{N} \chi^j \mu^j \left[\sum_{\omega, \theta} u^j(x^j(\omega), \theta) g^j(\omega, \theta) - C^j[\phi^j(\cdot)] \right] \tag{35}$$

s.t. (28), (29), (30), (31), (32), (33), *and* (34)

for some Pareto weights $(\chi^j)_{j=1}^J$.

Compared to the planner's problem we considered earlier on for state-tracking economies, the one stated above is more relaxed. The old problem amounts to restricting $m^j(\cdot)$ to be the identity function for all j. The new problem allows the planner to send different messages, including any of the following: the prices that would have obtained in equilibrium; the aggregate quantities of all types; and any other transformation, or coarsening, of the state. Whether this extra option affords a welfare improvement ultimately depends, as shown in the sequel, on the properties of the cost functional C.

Intuitively, if agents can effortlessly dispose of any decision-irrelevant information, and can readily go back back and forth between different transformations of the state that contain the same information vis-à-vis their decisions, there should be no gain from sending a message different than θ. But if that's the case, there should also be no gain in equilibrium from tracking prices rather than tracking the entire state itself. This logic suggests that the efficiency of price-tracking equilibria is tied to the coincidence of price-tracking and state-tracking equilibria. We make these ideas precise in the remainder of this section.

6.3 Burden of Tracking and Informational Invariance

Let ω be a signal of some particular $z = f(\theta)$, with joint density $\phi(\omega, z)$. Let $\tilde{z} = h(z) = h(f(\theta))$ be a different object, which has to be weakly "coarser" than z. Let us define $\tilde{\omega}$ as the "projection" of ω that gives the best signal of $\tilde{z}$. The precise construction of the density $\tilde{\phi}(\tilde{\omega}, \tilde{z})$ is

$$\tilde{\phi}(\tilde{\omega}, \tilde{z}) = \sum_z \phi(\tilde{\omega}, z) \cdot I\{\tilde{z} = h(z)\} \tag{36}$$

The new signal $\tilde{\omega}$ has the same domain as ω and induces the same posteriors about $\tilde{z}$. It also depends on z (and the state θ) *only* through its relationship with $\tilde{z}$. In this way it obeys our restrictions on attention choice for consumers and firms, conditional on tracking $\tilde{z}$. But it may not induce the same posteriors about z (or θ), because it contains strictly less information about these objects.

We can now ready to state our invariance condition:

Assumption 3 (Informational Invariance). *Let z be a random variable that can be defined as $z = f(\theta)$ for some $f : \Theta \to Z$ and $Z \in \mathscr{Z}$, and ω be some signal of z with joint density $\phi(\cdot,\cdot)$. Let $\tilde{z} \in \tilde{Z}$ be a second random variable that can be defined as $\tilde{z} = h(z)$ for some $h : Z \to \tilde{Z}$ and $\tilde{Z} \in \mathscr{Z}$. Construct $\tilde{\omega}$ and $\tilde{\phi}(\cdot)$ as in (36). Then $C[\cdot]$ satisfies informational invariance if*

(i) $C^j[\phi(\cdot)] \geq C^j[\tilde{\phi}(\cdot)]$ always;

(ii) $C^j[\phi(\cdot)] = C^j[\tilde{\phi}(\cdot)]$ if and only if $\tilde{z}$ is a sufficient statistic for z about ω.

The first part of this assumption imposes that tracking a coarser object yields weakly lower cognitive cost. This is highly plausible, but not sufficient for our purposes.

In the price-tracking example of Section 3, the key observation was that costs were *identical* when written with respect to the state or the prices. How does this reconcile with the above logic about coarsening? There must be an upper bound to the "value" of coarsening z for a fixed cognitive state ω. This bound is formalized by the second part of the assumption. Intuitively, this is exactly the condition under which the construction of $\tilde{\omega}$ in (36) looses no information.[18]

The same argument, since θ is always a sufficient statistic for any $z = f(\theta)$, implies that the agent is equally burdened by tracking $\tilde{z}$ or the entire vector θ. Any additional information in θ can be freely disposed of. Agents' choices and payoffs, by implication, are invariant to expanding the state space in ways that do not affect decision-relevant variables (and hence payoffs).

As mentioned in the Introduction, our invariance condition brings to mind the axiom of "invariance under compression" from Caplin, Dean, and Leahy (2017). But whereas that paper shows how the combination of this axiom with another one ("uniform posterior separability") provides a foundation for mutual-information costs in a single-agent, decision-theoretic context, here we shall show how our invariance condition suffices for efficiency in a general-equilibrium context.

6.4 "Free Disposal" in Action

It is simple to show that, under the invariance condition formalized above, the equilibria of price- and state-tracking economies naturally coincide.

[18] Finally, on a more technical level, the combination of the two conditions ensures that any restrictions on the signal space (i.e., existence of signals that return infinite cost) are properly "nested" as one changes the tracked object z.

Theorem 3 (Coincidence of price-tracking and state-tracking equilibria)**.** *Impose Assumption 3.*

(a) Let

$$A = \left\{x^j(\cdot), \phi^j(\cdot), p(\cdot)\right\}_{j\in\{1,\dots,J\}}$$

be the equilibrium of a price-tracking economy, where $z = (\theta, p(\theta))$ Next, let $\varphi^j(\omega, \theta)$ be the induced joint distribution between ω and θ in this equilibrium. Then, the following is an equilibrium of a state-tracking economy with the same endowments, preferences, and cognitive costs:

$$A' = \left\{x^j(\cdot), \varphi^j(\cdot), p(\cdot)\right\}_{j\in\{1,\dots,J\}}$$

(b) Conversely, let

$$B = \left\{x^j(\cdot), \phi^j(\cdot), p(\cdot)\right\}_{j\in\{1,\dots,J\}}$$

be an equilibrium of a state-tracking economy and let $\varphi^j(\omega, z)$ be the induced joint distribution with respect to $z = (\theta, p(\theta))$. Then, the following is an equilibrium of a price-tracking economy with the same endowments, preferences, and cognitive costs:

$$B' = \left\{x^j(\cdot), \varphi^j(\cdot), p(\cdot)\right\}_{j\in\{1,\dots,J\}}$$

An almost identical argument shows that the social planner would never strictly want to send a message different from $m^j \equiv \theta$ for all types j.

Corollary 1 (Restricting messages is without loss)**.** *Impose Assumption 3. Let $A = \left\{x^j(\cdot), \phi^j(\cdot), m^j(\cdot)\right\}$ be efficient in the sense of Definition 7 (i.e.,* with *messages) and let $\varphi^j(\cdot)$ be the induced joint distribution between ω and θ. Then, $A' = \left\{x^j(\cdot), \varphi^j(\cdot)\right\}$ is efficient in the sense of Definition 5 (i.e.,* without *messages).*

This suggests a simple strategy of *adapting* the previously proven Welfare Theorems to this context.

6.5 Welfare Theorems for Price-tracking Economies

We now state our main results for price-tracking economies, starting with the following version of the First Welfare Theorem.

Theorem 4 (First Welfare Theorem for price-tracking economies)**.** *Impose Assumptions 1 and 3. Any inattentive equilibrium with positive prices is efficient in the sense of Definition 7.*

Proof. Let

$$A = \left([x^j(\cdot)]_{j=1}^J, [\phi^j(\cdot)]_{j=1}^J, p(\cdot)\right)$$

be an equilibrium of the price-tracking economy and let $\varphi^i(\omega, \theta)$ be the associated density with respect to the state. By Theorem 3, the pair $\left([x^j(\cdot)]_{j=1}^J, [\varphi^j(\cdot)]_{j=1}^J\right)$ is part of an equilibrium of an equivalent state-tracking economy. From the previously proven First Welfare Theorem, this pair is a solution to problem (23), where the planner is restricted to sending the entire state as the only message. But this also solves the unrestricted problem, in which the planner can send arbitrary messages, by Corollary 1. □

Similar logic allows a version of the Second Welfare Theorem.

Theorem 5 (Second Welfare Theorem for price-tracking economies)**.** *Impose Assumptions 2 and 3. Let*

$$A = \left([x^j(\cdot)]_{j=1}^J, [\phi^j(\cdot)]_{j=1}^J, (m^j(\cdot))_{j=1}^J\right)$$

be an efficient combination of messages, attention plans, and consumption plans in the sense of Definition 8. The attention and consumption plans are implementable as a price-tracking equilibrium.

Proof. From Corollary 1, A is efficient in the sense of (23). From Theorem 2, A' can be implemented as an equilibrium with transfers in a state-tracking variant of the same economy. From Theorem 3, the allocations and prices in this state-tracking economy equilibrium are also part of an equilibrium in the equivalent price-tracking economy given the same endowments (i.e., including transfers). □

Theorems 4 and 5 together generalize the logic of Proposition 2 along two dimensions that were not available in the toy model of Section 3: for optimality relative to the planner's arbitrary choice of messages, and no constraints on the price system for implementation. The formal argument also makes clear how a generalized version of Proposition 3 (in the form of Theorem 3) directly implies a generalized version of Proposition 4 (in the form of Theorems 4 and 5). Basically, the same conditions that guarantee the coincidence of price-tracking and state-tracking equilibria also guarantee the efficiency of the former.

6.6 Optimal Messages

While the previous results do formally answer our original question, they do not provide an immediately satisfying answer about implementation. Efficiency is provided precisely because the social planner has no *strict preference* for sending a more economized message than: "this is the entire state θ, from which you can compute all relevant equilibrium quantities."

Can the planner send any (subjectively, in these authors' view) more "natural" message? As a first step, let us introduce some new terminology for what parts of the state of nature are *directly* relevant to a given agent type j. Let θ^j denote the subset of the state vector that summarizes agent type j's preferences and endowment. First, one can show as a result that a consumer with information cost satisfying the invariance properties described above would weakly prefer to learn just about this θ^j and the price p, the latter of which summarizes all useful information about other parts of θ:

Lemma 1 (Agents track prices)**.** *Impose Assumption 3. Let the bundle $(x^j(\cdot),\phi^j(\cdot))$ solve the state-tracking endowment economy consumer problem given some prices $p(\theta)$. Then the signal associated with this bundle is such that $(p(\theta),\theta^j)$ is a sufficient statistic for θ in the joint distribution $\phi^j(\omega,\theta)$.*

Proof. Let $(x(\cdot),\phi(\cdot))$ be a proposed solution of the state-tracking consumer problem that does not have the sufficient statistic property. The state θ enters the problem only via θ^j and $p(\theta)$, the former defined to capture the agent's *own* preferences and endowments. By Assumption 3, there is a strictly lower cognitive cost to another feasible bundle with the same consumption choice and new attention choice $\tilde{\phi}(\tilde{\omega},(p(\theta),\theta^j))$, defined by (36). Thus $\mathscr{U}(x(\cdot),\tilde{\phi}(\cdot)) > \mathscr{U}(x(\cdot),\phi(\cdot))$, which is a contradiction to the optimality of the first bundle. □

This is like an extension of the logic in Section 3.9, in which we showed that adding small but nonzero costs of tracking aggregate noise variables immediately induced agents to obtain purely private signals of the decision-relevant variable (which, in that case, was the price p). Idiosyncratic randomization is cheaper, cognitively, than randomization based on (payoff-irrelevant) entries of θ.

Assumption 3 guarantees there is actually no loss from sending this economized signal (θ^j,p) in the planner's implementation of one of the previously described optimal allocations:

Corollary 2 (Prices as the Planner's Signal)**.** *To implement the optimal allocation of a price-tracking economy with price function $p(\cdot)$, the social planner can send the following messages: $m^j(\theta) = (\theta^j, p(\theta))$ and $m^f(\theta) = (\theta^f, p(\theta))$.*

In this sense, the "invisible hand" both optimally allocates goods and produces an informally efficient *message* about that implementation.

Tracking only prices. We can flip this and also speculate about whether efficiency extends when agents in equilibrium are tracking "smaller" objects than the joint of the full state of nature and prices. This is necessarily true as long as agents track something complex enough to implement the planner's preferred allocation in the way described by Corollary 2.

In particular consider an economy in which all agents, in some sense, fully internalize the previous constraint and worry *only* about tracking θ^j and p. Let us introduce this scenario as a "price-only-tracking" economy, in the following sense.

Definition 8. A *price-only-tracking economy* is an economy in which $z^j = (\theta^j, p)$ for all j. and a type-j agent's signal, conditional on z^j, cannot be correlated with the signal of any other agent.

This is, of course, an implementation in the optimal set of the planner with *unrestricted* messages according to Corollary 2. So, after verifying that the restriction on messages does not change competitive equilibria, it is simple to show an extension of our theorems to this context.

Corollary 3. *Consider a price-only-tracking economy in the sense of Definition 8. Theorems 4 and 5 continue to hold.*

From the perspective of cognitive externalities, this is perhaps a more subtle result than Theorems 4 and 5. Even with the invariance condition on cognitive costs, the cost of an *arbitrary* attention strategy depends on the actions of others. For instance, if all agents coordinated on some strategy that made prices constant across states of nature, everyone would enjoy cognitive savings. But the externality is muted exactly at the *optimum* attention strategy. Lemma 1 showed that this strategy has a very specific structure, which we could exploit in the proof to show equivalence with state-tracking economies, extending Theorem 3, and ultimately extend the welfare theorems.

The informational role of policy. We can finally revisit the thought experiment of Subsection 6.1 regarding how the taxes or other policy messages could themselves play the role of "messages." So far we have shown that, under our invariance condition, the equilibria in which z is the state of nature coincides with the equilibria in which z is the equilibrium prices. What if agents can also track any taxes or other policy tools that a government may have in its disposal? If tracking such objects was easier than tracking prices, then equilibria with policy intervention could improve welfare, not by regulating the price system so as to influencing the costs of decoding it, but rather by letting policy tools themselves serve as better means of communication than market prices. Our invariance condition makes sure that this is not the case.

6.7 The Role of Mutual-Information Costs

How do the results presented here relate to those obtained in Section 3, particularly with regard to the conditions under which (Shannon) mutual-information costs ensure efficiency?

In that section, we highlighted that mutual information *alone* does not suffice for efficiency: when we specified the cognitive costs as a function of the mutual information of the decisions and *only* the price, there was room for multiple Pareto-ranked equilibria, supported by different levels of correlation in the agents mistakes. We complemented that sect with an extension in the Appendix, which let costs be

$$C[\phi(\cdot)] = (1-K)I[\omega, p] + K\,I[\omega,(p,v)],$$

for some $K \in [0,1]$ and where each mutual information functional is evaluated with respect to the "right" transformation of the joint distribution $\phi(\cdot)$. When $K=0$, this reduces to $C = I(\omega, p)$, which is the specification considered in the Section 3. When instead $K=1$, this reduces to $C = I[\omega_i,(p,v)]$ and defines the sense in which the cost of tracking v is the "same" as that of tracking p. We can then show that inefficiency survives for $K<1$ but disappears when $K=1$.

These results can be understood through the lens of our more general findings as follows. In the example under consideration, p is always a linear combination of $\theta = (\xi, v)$, implying that $I[\omega,(p,v)] = I[\omega,(p,\theta)]$ and therefore

$$C[\phi(\cdot)] = (1-K)I[\omega, p] + K\,I[\omega, z],$$

with $z = (p,\theta)$. When $K=1$, the cost satisfies our invariance condition and, as a result, efficiency is guaranteed. When instead $K=0$, or more generally $K \in [0,1)$, our invariance condition is violated and inefficiency is possible.

The next result generalizes these ideas and explains a more precise condition under which mutual-information costs suffice for efficiency:

Corollary 4. *Consider an economy in which cognitive costs for all types have form* (17) *such that, for all types* j*, (i) each* $C^{jn}[\cdot]$ *across goods* $n \in \{1,\ldots,N\}$ *is a transformation of Shannon mutual information; and (ii)* $X^j[\cdot]$ *is a transformation of the (multidimensional) Shannon mutual information of the tuple* $(\omega_1,\ldots,\omega_N)$*. Our invariance condition is guaranteed and the Welfare Theorems hold for state-tracking and price-tracking variants of the same setting.*

Theorems 4 and 5 show the path to efficiency when the cognitive cost is the mutual information of the signal with the *joint* of prices and the full state of nature. Here, the cost of tracking any any non-fundamental information is "just right." Corollary 3 allows for the possibility of mutual information with *only* prices (and one's own, directly-relevant fundamentals) when correlated randomization is infeasible. Here, the same cost is infinite.

What if we depart from Shannon mutual information altogether? Section 3 illustrated how this can open the door to inefficiency by letting "scale" or "complexity" to matter. But not every relaxation of Shannon mutual information does this. The precise characterization of the broader class of costs that satisfy our invariance condition and their axiomatic underpinnings are outside the scope of our paper.

7 Discussion

This paper's main result connected the efficiency of markets when attention is a scarce resource, and Hayek's (1945) casual argument about the informational optimality of the price system, with a sharp condition on the *units* of attention. Our position is that this condition, much like market completeness, is an unrealistic but intellectually useful benchmark. In this final section we discuss how this benchmark helps contextualize ongoing debates about policy intervention to "correct" inattention and suggests avenues for future research.

Revisiting the price system's "economy of knowledge." Consider the following thought experiment in the spirit of Hayek. A central planner wants to improve upon the laissez-faire outcome, which corresponds to a price-tracking equilibrium as studied above. To learn the state θ, the planner first surveys all agents in the economy about their own preferences, endowments, and technologies. Armed with this knowledge, the planner crafts messages $m^j(\theta)$ to send back to the agents. Such messages encode instructions about what each agent oughts to do.[19]

Agents are obedient but inattentive. An optimal mechanism takes this into account and fine tunes the messages so as to utilize the agents' scarce attention as efficiently as possible. This is true regardless of whether the invariance condition of Assumption 3 holds or or not. But whenever it holds, the following is also true: free markets is an optimal mechanism (and prices are optimal messages).

Markets "economize on knowledge" relative to many feasible but suboptimal alternatives. For instance, consider a specific contingency that is essentially irrelevant for decisions, like whether US GDP measured in whole dollars is an even or odd number. In equilibrium, agents may optimally choose to pay no attention to this contingency, and hence completely "dry up" the associated contingent claims market that would have otherwise operated. By contrast, a centrally planned allocation that *forced* agents to very precisely learn about this contingency would be suboptimal because it allocates people's scarce attention to the wrong target.

But the perhaps paradoxical additional result is that the same condition that guarantees the optimality of markets also implies that such optimality is weak, not strict. Corollary 1 showed that, if the central planner sent every agent a message equal to the entire state of nature, there would be no difference in welfare. And while we did not develop formal tools to show the generic necessity of Assumption 3, our intuition about cognitive externalities and the narrow situations in which they might cancel out suggests that this is essentially the *only* case in which free-market prices are informationally optimal.

Away from our invariance property, central planning can do better than the invisible hand, or markets can be improved upon. For example, a planner can contemplate the *removal* of unnecessarily confusing markets as strictly welfare-improving. Alternatively, the planner might internalize the value of "improving" both price and non-price data (e.g., macroeconomic aggregates) so as to contribute to better decisions. Engaging these issues substantially requires departing from either complete markets (this is known) or mutual-information costs (this is novel).

On the units of information. Herbert Simon, in his early work on information processing in organizations (Simon, 1971), highlighted the importance of "attention management" for efficient allocations:

> In an information-rich world, the wealth of information means a dearth of something else: a scarcity of whatever it is that information consumes. [...] Hence a wealth of information

[19] For simplicity, we assume that agents know for free their own preferences, endowments, and technologies (i.e., each agent of type j knows θ^j). But the argument extends even if agents must incur a cost in order to figure out such information.

> creates a poverty of attention and a need to allocate the attention efficiently among the overabundance of information sources that might consume it.

Simon recognized also the delicate relationship between this intuition and an appropriate "measurement" of information or attention:

> To formulate an allocation problem properly, ways must be found to measure the quantities of the scarce resource; and these quantities must not be expandable at will.

Our paper shares Simon's goal of understanding how the *micro problem* of measuring attention relates to the *macro problem* of social efficiency. And it offers a sharp answer for competitive economies.

Borrowing Simon's language, our results say the following. If the units of attention are such that our invariance condition is satisfied, there is no mismatch between where the social planner would direct attention and where agents naturally direct it. Furthermore, manipulating that set of learnable things, by altering the messages in central planning or refining policy instruments like taxes in a market implementation, does not help economize on the scarce resource of attention. This is a subtle, and non-generic, property of measuring cognition in particular units. Recall the general form of "reduction in perceived complexity" introduced as expression (2) and restated here:

$$C[\phi(\cdot)] = \mathbb{E}\big[H[\pi(\cdot)] - H[\phi(\cdot \mid \omega)]\big] \tag{37}$$

For such costs, the condition underlying our Welfare Theorems is that payoff-irrelevant manipulations of the state-space uniformly shift the two terms, leaving their expected difference constant. This property causes any intuition about externalities changing the "cost of paying attention" to fail. This is the precise sense in which Shannon's formulation might be a poor model of "attention scarcity" in the aggregate economy, to use Simon's language.[20]

We hope these insights contribute to the decision-theoretic (e.g., Caplin et al., 2017; Pomatto et al., 2018; Hébert and Woodford, 2018) and experimental (e.g., Dewan and Neligh, 2017; Dean and Neligh, 2017) literature on measuring attention costs that push beyond the Shannon benchmark. One application that we highlighted in the simple, closed-form example of Section 3 was the importance of attention costs with some notion of continuity, or difficulty of distinguishing numerically close states (e.g., Pomatto et al., 2018; Hébert and Woodford, 2018). But we were not able to provide a full taxonomy of how this or other deviations behave in substantially more general settings. This is a promising area for future research.

Markets for information. We considered no explicit *markets* for information, for sale by experts or news media (e.g., as in Grossman and Stiglitz, 1980; Admati and Pfleiderer, 1990; Veldkamp, 2006). This was natural in our formalism of signals as internal cognition, and indeed there was a formal link between cognition

[20] Simon arrives at a similar conclusion for the different (but compatible) reason that Shannon's measure is motivated by too narrow a context:

> Can we use the bit [the unit of Shannon entropy] as a measure of an information-processing system's capacity for attention? Unfortunately, it is not the right unit. Roughly the trouble is that the bit capacity of any device (or person) for receiving information depends entirely on how the information is encoded. Bit capacity is not an invariant, hence is an unsuitable measure of the scarcity of attention.

in our models and private consumption (and home production) in a classic Arrow-Debreu framework. We conjecture that, in a world of tradable information, one could extend our invariance condition to become also a restriction on the technology of producing this commodity sufficient for the related production externality to be muted—the information supply twin of our present conditions for information demand.

Of course, it is possible that this model of an efficient "informational economy" would be quite uninteresting. Our invariance condition for cognitive costs already encodes zero preference for simplification of messages or external "oracles" providing useful recommendations. We hope that this negative result provides guidance on how to contemplate a more sophisticated theory of information markets and the scope for policy intervention.

Incomplete markets. In many dimensions, our conditions for constrained efficiency may be "loose." Consider first the assumption of complete markets. Our main intuition was that pecuniary externalities induced by attention choice net out. This netting out generically fails in incomplete market models, but there exist appropriate conditions, identified in Geanakoplos and Polemarchakis (1986), under which it still holds. We conjecture that a version of their constrained-efficiency results (i.e., holding fixed available markets) could hold after our constraints on cognitive costs are applied.

From markets to games. Had information or attention been exogenous, the economies studied here could have been mapped to a class of games in which the equilibrium use of information is efficient in the sense of Angeletos and Pavan (2007). Indeed, the condition on payoffs that guarantees efficiency in that paper corresponds to our setting's property that pecuniary externalities net out (see Appendix E for a direct analogy to the closed-form example from Section 3).

Under this lens, our results hint to a possible link between the efficiency of the *use* of information and that of the *acquisition* of information in games. This link is further explored in a recent paper by Hébert and La'O (2019). These authors show that efficiency obtains in a large game in which players choose how much information to collect about the state of nature and the average action of others if a generalization of the condition on payoffs provided in Angeletos and Pavan (2007) is complemented by an invariance condition similar to ours. The combination of our paper and that of Hébert and La'O (2019) thus provide a unified approach to the welfare implications of rational inattention in markets and games alike.

The case for policy intervention. Our benchmark case clarifies that inattention is not *by itself* a justification for correcting markets in an economy that would otherwise be efficient (i.e., with competitive and complete markets). But, as we have already discussed, the benchmark case fails on many intuitive levels when confronted with real-life experience—from the "smell test" of whether individuals can truly dispose of additional information for free, to the experimental evidence that requires a departure from mutual-information costs, to the basic observation that people do demand simplified information. In practice, then, our results open a new door to corrective policy.

We have offered a few concrete examples what this could mean. If people economize cognitive costs by tracking objects that are less uncertain or more coarse, there can be room for stabilizing prices or "simplifying" markets. And if people can pay attention to certain objects (e.g., noisy news in the media) at a lower cost than others (e.g., big data or expert opinion), equilibria in which market outcomes are "noisy" could be superior to equilibria in which market outcomes are more tightly connected to "hard" fundamentals.

Of course, a more direct justification for policy intervention can be made if inattention is *ir*rational, as in a segment of the behavioral literature (Chetty et al., 2009; Gabaix, 2014, 2019). But our analysis also hints at how some lessons from that literature *could* be fruitfully recast and studied in a rational-inattention context. For instance, the notions of "salience" and "context dependence" (Bordalo et al., 2013; Köszegi and Szeidl, 2012) may be accommodated within the rational-inattention framework as violations of our invariance condition. Similarly, a "desire for sparsity" à la Gabaix (2014), recast in a rational-inattention manner, could open the door for regulation *without* the form of irrationality assumed in that paper.

These ideas circle back to our earlier point about the value of departing from mutual-information costs within the rational-inattention framework. Such departures offer the promise of understanding jointly choice data, the focus of the growing decision-theoretic and experimental literatures we cited earlier on, and efficiency, the focus of our paper.

References

ADMATI, A. R. AND P. PFLEIDERER (1990): "Direct and Indirect Sale of Information," *Econometrica*, 58, 901–928.

ANGELETOS, G.-M. AND J. LA'O (2010): "Noisy Business Cycles," in *NBER Macroeconomics Annual 2009, Volume 24*, University of Chicago Press, 319–378.

——— (2018): "Optimal Monetary Policy with Informational Frictions," *Journal of Political Economy*, forthcoming.

ANGELETOS, G.-M. AND C. LIAN (2016): "Incomplete Information in Macroeconomics: Accommodating Frictions in Coordination," in *Handbook of Macroeconomics*, ed. by J. B. Taylor and H. Uhlig, Elsevier, vol. 2, chap. Chapter 14, 1065 – 1240.

ANGELETOS, G.-M. AND A. PAVAN (2007): "Efficient Use of Information and Social Value of Information," *Econometrica*, 75, 1103–1142.

BERGEMANN, D. AND S. MORRIS (2013): "Robust Predictions in Games with Incomplete Information," *Econometrica*, 81, 1251–1308.

——— (2019): "Information Design: A Unified Perspective," *Journal of Economic Literature*, 57, 44–95.

BLOEDEL, A. W. AND I. R. SEGAL (2018): "Persuasion with Rational Inattention," *Available at SSRN 3164033*.

BORDALO, P., N. GENNAIOLI, AND A. SHLEIFER (2013): "Salience and Consumer Choice," *Journal of Political Economy*, 121, 803–843.

CAPLIN, A. AND M. DEAN (2015): "Revealed Preference, Rational Inattention, and Costly Information Acquisition," *American Economic Review*, 105, 2183–2203.

CAPLIN, A., M. DEAN, AND J. LEAHY (2017): "Rationally Inattentive Behavior: Characterizing and Generalizing Shannon Entropy," Working Paper 23652, National Bureau of Economic Research.

——— (2018): "Rational Inattention, Optimal Consideration Sets, and Stochastic Choice," *The Review of Economic Studies*, 86, 1061–1094.

CHETTY, R., A. LOONEY, AND K. KROFT (2009): "Salience and Taxation: Theory and Evidence," *American Economic Review*, 99, 1145–77.

COLOMBO, L., G. FEMMINIS, AND A. PAVAN (2014): "Information Acquisition and Welfare," *Review of Economic Studies*, 81, 1438–1483.

DA SILVEIRA, R. A. AND M. WOODFORD (2019): "Noisy Memory and Over-Reaction to News," *NBER Working Paper No. 25478*.

DEAN, M. AND N. L. NELIGH (2017): "Experimental Tests of Rational Inattention," .

DEBREU, G. (1954): "Valuation Equilibrium and Pareto Optimum," *Proceedings of the National Academy of Sciences of the United States of America*, 40, 588–592.

DENTI, T. (2016): "Games with Unrestricted Information Acquisition," *MIT mimeo*.

——— (2018): "Posterior-Separable Cost of Information," Tech. rep., working paper.

DEWAN, A. AND N. NELIGH (2017): "Estimating information cost functions in models of rational inattention," *unpublished, Columbia University*.

FARHI, E. AND X. GABAIX (2019): "Optimal Taxation with Behavioral Agents," *American Economic Review*.

GABAIX, X. (2014): "A Sparsity-Based Model of Bounded Rationality," *Quarterly Journal of Economics*, 129, 1661–1710.

——— (2019): "Behavioral Inattention," in *Handbook of Behavioral Economics*, ed. by B. D. Bernheim, S. DellaVigna, and D. Laibson, Elsevier, vol. 2.

GEANAKOPLOS, J. AND H. M. POLEMARCHAKIS (1986): "Existence, Regularity, and Constrained Suboptimality of Competitive Allocations When the Asset Market Is Incomplete," *Uncertainty, Information and Communication: Essays in Honor of K.J. Arrow*, 3, 65–96.

GROSSMAN, S. J. (1981): "An Introduction to the Theory of Rational Expectations Under Asymmetric Information," *The Review of Economic Studies*, 48, 541–559.

GROSSMAN, S. J. AND J. E. STIGLITZ (1980): "On the impossibility of informationally efficient markets," *The American Economic Review*, 70, 393–408.

GUL, F., W. PESENDORFER, AND T. STRZALECKI (2017): "Coarse Competitive Equilibrium and Extreme Prices," *American Economic Review*, 107, 109–37.

HAYEK, F. A. (1945): "The use of knowledge in society," *The American economic review*, 519–530.

HÉBERT, B. AND J. LA'O (2019): "Information Acquisition, Efficiency, and Non-Fundamental Volatility," *Stanford GSB and Columbia University mimeo*.

HÉBERT, B. AND M. WOODFORD (2018): "Information Costs and Sequential Information Sampling," *NBER Working Paper No. 25316*.

HELLWIG, C. AND L. VELDKAMP (2009): "Knowing what others know: Coordination motives in information acquisition," *The Review of Economic Studies*, 76, 223–251.

ILUT, C. AND R. VALCHEV (2017): "Economic Agents as Imperfect Problem Solvers," *Duke University mimeo*.

KAMENICA, E. AND M. GENTZKOW (2011): "Bayesian Persuasion," *American Economic Review*, 101, 2590–2615.

KOHLHAS, A. AND A. WALTHER (2018): "Asymmetric Attention," *IIES miméo*.

KÖSZEGI, B. AND F. MATĚJKA (2018): "An attention-based theory of mental accounting," .

KÖSZEGI, B. AND A. SZEIDL (2012): "A Model of Focusing in Economic Choice," *The Quarterly Journal of Economics*, 128, 53–104.

KYLE, A. S. (1985): "Continuous auctions and insider trading," *Econometrica*, 1315–1335.

LAFFONT, J.-J. (1985): "On the Welfare Analysis of Rational Expectations Equilibria with Asymmetric Information," *Econometrica*, 53, 1–29.

LIAN, C. (2018): "A Theory of Narrow Thinking," Working paper, Massachusetts Institute of Technology.

LIPNOWSKI, E., L. MATHEVET, AND D. WEI (2019): "Attention Management," *Available at SSRN 3161782*.

LLOSA, L.-G. AND V. VENKATESWARAN (2017): "Efficiency with Endogenous Information Choice," *NYU mimeo*.

MAĆKOWIAK, B. AND M. WIEDERHOLT (2009): "Optimal Sticky Prices under Rational Inattention," *American Economic Review*, 99, 769–803.

——— (2015): "Business Cycle Dynamics under Rational Inattention," *The Review of Economic Studies*, 82, 1502–1532.

MAĆKOWIAK, B. A., F. MATĚJKA, AND M. WIEDERHOLT (2018): "Survey: Rational Inattention, a Disciplined Behavioral Model," CEPR Discussion Papers 13243, C.E.P.R. Discussion Papers.

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

MATĚJKA, F., J. STEINER, AND C. STEWART (2015): "Rational Inattention Dynamics: Inertia and Delay in Decision-Making," *CERGE-EI mimeo*.

MENSCH, J. (2018): "Cardinal representations of information," *Available at SSRN 3148954*.

MORRIS, S. AND H. S. SHIN (1998): "Unique Equilibrium in a Model of Self-fulfilling Currency Attacks," *American Economic Review*, 587–597.

——— (2002): "Social Value of Public Information," *American Economic Review*, 92, 1521–1534.

MORRIS, S. AND P. STRACK (2019): "The Wald Problem and the Relation of Sequential Sampling and Ex-Ante Information Costs," *Available at SSRN 2991567*.

MORRIS, S. AND M. YANG (2016): "Coordination and the Relative Cost of Distinguishing Nearby States," *Princeton mimeo*.

MYATT, D. P. AND C. WALLACE (2012): "Endogenous Information Acquisition in Coordination Games," *Review of Economic Studies*, 79, 340–374.

PACIELLO, L. AND M. WIEDERHOLT (2014): "Exogenous Information, Endogenous Information, and Optimal Monetary Policy," *Review of Economic Studies*, 81, 356–388.

POMATTO, L., P. STRACK, AND O. TAMUZ (2018): "The Cost of Information," .

SIMON, H. A. (1971): "Designing Organizations for an Information Rich World," in *Computers, Communications, and the Public Interest*, ed. by M. Greenberger, Baltimore, MD: The Johns Hopkins Press.

SIMS, C. A. (1998): "Stickiness," *Carnegie-Rochester Conference Series on Public Policy*, 49, 317–356.

——— (2003): "Implications of Rational Inattention," *Journal of Monetary Economics*, 50, 665–690.

——— (2006): "Rational inattention: Beyond the linear-quadratic case," *The American Economic Review*, 96, 158–163.

——— (2010): "Rational Inattention and Monetary Economics," *Handbook of Monetary Economics*, 3, 155–181.

STEINER, J. AND C. STEWART (2016): "Perceiving Prospects Properly," *American Economic Review*, 106, 1601–31.

TIROLE, J. (2015): "Cognitive Games and Cognitive Traps," *Toulouse School of Economicss mimeo.*

TSALLIS, C. (1988): "Possible generalization of Boltzmann-Gibbs statistics," *Journal of Statistical Physics*, 52, 479–487.

VAN NIEUWERBURGH, S. AND L. VELDKAMP (2010): "Information acquisition and under-diversification," *The Review of Economic Studies*, 77, 779–805.

VELDKAMP, L. L. (2006): "Media frenzies in markets for financial information," *The American Economic Review*, 577–601.

VIVES, X. (2017): "Endogenous public information and welfare in market games," *Review of Economic Studies*, 84, 935–963.

VIVES, X. AND L. YANG (2018): "Costly Interpretation of Asset Prices," mimeo, IESE/University of Toronto mimeo.

WEI, D. (2018): "Persuasion Under Costly Learning," .

WILSON, R. (1975): "Informational Economies of Scale," *Bell Journal of Economics*, 6, 184–195.

WOODFORD, M. (2003): "Imperfect Common Knowledge and the Effects of Monetary Policy," *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps.*

——— (2012): "Prospect Theory as Efficient Perceptual Distortion," *American Economic Review*, 102.

——— (2019): "Modeling Imprecision in Perception, Valuation and Choice," *NBER Working Paper No. 26258.*

YANG, M. (2015): "Coordination with flexible information acquisition," *Journal of Economic Theory*, 158, 721–738.

Appendices

A Proofs for Section 3

Proofs of Propositions 2, 3, and 4

Solving for competitive equilibrium. Let us first specialize to agents' tracking the random variable ξ. Agents solve

$$\begin{aligned} \max_{x_{1i},x_{2i},\delta_i} \; & \mathbb{E}\big[\mathbb{E}_i\big[x_{1i} - x_{1i}^2/2 + x_{2i}\big]\big] - c(\delta_i) \\ \text{s.t. } & p x_{1i} + x_{2i} \le p\xi + 1 \end{aligned} \tag{38}$$

where δ_i is the "attention level" for agent i . The informational precision δ is now a choice variable, and $c(\cdot)$ is an increasing and differentiable cost function. The outer expectations reflects the fact the information level is chosen before the revelation of the signal.

Carry over all the previous "shortcuts" which allowed for easy computation of the (ultimately unique) equilibrium. The equilibrium price function remains affine, $p = p_0 - g\xi$, and agents choose symmetric strategies in equilibrium. Substituting in the budget constraint at equality gives the expression

$$\max_{X_{1i},X_{2i},\delta_i} \mathbb{E}\left[(1-p)x_{1i} - \frac{x_{1i}^2}{2} + p\xi + 1\right] - c(\delta_i)$$

The first order condition in terms of X_{1i} is

$$\mathbb{E}[1 - p - x_{1i}] = 0$$

and the measurability constraint, which was suppressed in the short-hand notation of the main text, is that x_{1i} is measurable in some i-specific information set (or $\mathbb{E}_i[x_{1i}] = x_{1i}$). An equivalent formulation of this first-order condition is thus

$$\mathbb{E}[\mathbb{E}_i[1-p] - x_{1i}] = 0$$

Thus $x_{1i} = \mathbb{E}_i[1-p]$.

Let δ denote the information choice of other agents. The resource constraint is $\int x_{1i}\,\mathrm{d}i = \xi$. These are the same two equations by which we earlier derived the expressions:

$$x_{1i} = \mu + \frac{\delta_i}{\delta}((\xi-\mu)+\varepsilon_i)$$
$$p = 1 - \mu - \frac{\xi-\mu}{\delta}$$

where we have been careful to differentiate the attention choice δ_i of a given agent i from the attention choice δ of all other agents.

Each term of the objective is:

$$\mathbb{E}[(1-p)x_{1i}] = \mu^2 + \frac{\delta_i}{\delta^2}\sigma^2$$
$$\mathbb{E}[-x_{1i}^2/2] = -\frac{1}{2}\left(\mu^2 + \left(\frac{\delta_i}{\delta}\right)^2(\sigma^2 + r_i^2)\right) = -\frac{1}{2}\left(\mu^2 + \frac{\delta_i}{\delta^2}\sigma^2\right)$$
$$\mathbb{E}[p\xi] = \mu(1-\mu) - \frac{\sigma^2}{\delta}$$

Collecting these terms, it is convenient to write the objective, up to a constant, as

$$b(\delta_i;\delta) - c(\delta_i)$$

with the first part defined as

$$b(\delta_i, \delta) \equiv \frac{\sigma^2 \delta_i}{2\delta^2} - \frac{\sigma^2}{\delta} \tag{39}$$

$b(\delta, \delta)$ is exactly the welfare in an economy with exogenously incomplete information and signal-to-noise ratio δ.

To get the value of δ in competitive equilibrium, we take the first order condition and subsequently substitute $\delta = \delta_i$. In math, that is

$$\frac{\partial}{\partial \delta_i} b(\delta_i, \delta)|_{\delta=\delta_i} = c'(\delta)$$

which becomes

$$\frac{\sigma^2}{2\delta^2} = c'(\delta) \tag{40}$$

This has a unique solution for $\delta \in (0,1)$ if costs satisfy the stated "Inada" conditions are satisfied.

State-tracking efficiency (Proposition 2). We first prove the result by brute force. We then illustrate how it relates to pecuniary externalities, anticipating our subsequent, more general version of the First Welfare Theorem.

Consider a planner that can dictate the agents what δ to choose, but cannot otherwise regulate markets or replace them without mechanisms. Optimality requires that

$$\frac{\mathrm{d}}{\mathrm{d}\delta} b(\delta, \delta) = c'(\delta) \tag{41}$$

with a total, not partial, derivative on the left hand side. Using our earlier characterization of B, we get

$$\frac{\mathrm{d}}{\mathrm{d}\delta} b(\delta, \delta) = \frac{\mathrm{d}}{\mathrm{d}\delta}\left[-\frac{\sigma^2}{2\delta}\right] = \frac{\sigma^2}{2\delta^2},$$

which is the same as the first partial of b evaluated in equilibrium. Thus, the equilibrium choice of δ coincides with the planner's solution.

Let us now under the broader logic behind this "coincidence." Take δ as the given action of other agents and consider the ex ante utility an agent after optimization:

$$V_i(\delta) \equiv \max_{\delta_i} \{b(\delta_i, \delta) - c(\delta_i)\},$$

or, equivalently,

$$V_i(\delta) = \max_{x_{1i}(\cdot), \delta_i} \int_\omega \int_\xi [u(x_{1i}(\omega)) + (1 + p(\xi; \delta)(\xi - x_{1i}))] \phi(\omega, \xi; \delta_i) \mathrm{d}\xi \, \mathrm{d}\omega \tag{42}$$

By the standard envelope-theorem argument, the total derivative of V_i is given by the corresponding partial derivative of the objective, or

$$\frac{\mathrm{d}V_i(\delta)}{\mathrm{d}\delta} = \int_\omega \int_\xi \frac{\mathrm{d}p(\xi, \delta_i)}{\mathrm{d}\delta} (\xi - x_{1i}(\omega)(\omega)) \phi(\omega, \xi; \delta_i) \mathrm{d}\xi \, \mathrm{d}\omega$$

Since $p(\xi, \delta_i)$ does not depend on ω,

$$\frac{\mathrm{d}V_i(\delta)}{\mathrm{d}\delta} = \int_\xi \frac{\mathrm{d}p(\xi, \delta_i)}{\mathrm{d}\delta} \left\{ \int_\omega (\xi - x_{1i}(\omega)) \phi(\omega, \xi; \delta_i) \mathrm{d}\omega \right\} \mathrm{d}\xi.$$

In equilibrium, $\delta_i = \delta$ and $x_{1i}(\omega) = x_1(\omega)$ by symmetry, and $\int_\omega (\xi - x_1(\omega)) \phi(\omega, \xi; \delta) \mathrm{d}\omega = 0$ by market clearing. It follows that

$$\frac{\mathrm{d}V_i(\delta)}{\mathrm{d}\delta} = 0,$$

which verifies that the pecuniary externalities induced by the choice of attention net out.

Price-tracking efficiency (Propositions 3 and 4). The calculations of A continue to hold when agents track p, because the signal-to-noise ratio is the same between signals and ξ or between signals and p. This continues to define the competitive equilibrium, proving 3. Similarly, for the planner, the thought experiment of manipulating the attention of others exactly corresponds to the thought experiment outlined prior to the statement of Proposition 4. Hence this Proposition is proved by the same calculation.

Proof of Proposition 5

Consider the cost-benefit calculation of Appendix section A, but with the altered cost function described above:

$$\frac{\partial}{\partial \delta_i} b(\delta_i, \delta_{-i})|_{\delta_{-i}=\delta_i} = \frac{\partial}{\partial \delta_i} c(\delta_i, \sigma_p^2)|_{\delta_{-i}=\delta_i}$$

Note that, in the class of linear equilibria, $\sigma_p^2 = \sigma^2/\delta_{-i}^2$ and is decreasing in δ_{-i}. For the sake of generating an example, let $c(\delta_i, \sigma_p^2) = k(\sigma_p^2)c(\delta_i)$ for some convex $c(\cdot)$ and increasing $k(\cdot)$. The right-hand-side argument is then

$$\frac{\partial}{\partial \delta_i} c(\delta_i, \sigma_p^2)|_{\delta_{-i}=\delta_i} = k\left(\frac{\sigma^2}{\delta_i^2}\right) c'(\delta_i)$$

which is not necessarily upward sloping. This means there could be two intersections with the marginal benefits curve, both of which define possible equilibria.

To be even more concrete, let $k(y) = k_0(1 - y^{-1/2}\sigma)$ and $c(y) = y^2/2$. Then the condition defining equilibria is

$$k_0\delta(1-\delta) = \frac{\sigma^2}{2\delta^2}$$

A sufficient condition for this to have two solutions within $[0,1]$ is $k_0 > 8\sigma^2$.

Proof of Proposition 6

This section first proves claims in an environment with exogenously fixed signal precision and mutual information costs between actions and p, and then discusses a generalization with endogenous signal precision and and/or different costs of information.

Exogenous signal precision. Conjecture that the price has the form

$$p = p_0 - g\xi + h\nu$$

for some scalars (c, d), and the signal has the form

$$\begin{aligned} \omega_i &= p + r\varepsilon_i + s\nu + t\xi \\ &= p_0 + (t-g)\xi + (s+h)\nu + r\varepsilon_i \end{aligned}$$

Note that, given a fixed "budget" of precision, it is optimal to get a signal whose residual is orthogonal to p. This means that $\mathbb{E}[(\omega_i - p)(p - p_0)] = 0$, or

$$-gt\sigma^2 + sh = 0. \tag{43}$$

The signal-price correlation is then given by

$$\delta_p = \frac{g^2\sigma^2 + h^2}{(t-g)^2\sigma^2 + (s+h)^2 + r^2} \tag{44}$$

and the capacity constraint $I[\omega_i, p] \leq M$ reduces to $\delta_p \leq \bar{\delta} \equiv 1 - \exp^{-M}$. Since expected utility is increasing in δ_p, this amounts to fixing δ_p exogenously, to the value $\delta_p = \bar{\delta}$.

Finally, in equilibrium, market clearing is

$$1 - \bar{\mathbb{E}}[p] = 1 - \delta\left[p_0 + (t-g)\xi + (s+h)\nu\right] = \xi \tag{45}$$

This implies that $p_0 = 1/\delta$, $t = -1/\delta + g$, and $s = -h$. The last implies that ω_i does not move with ν, because the noise in the signal and price cancel each other out.

It remains only to solve for the triplet (g, h, r). Equations (43) and (44) re-arrange as follows:

$$\begin{aligned} \frac{h^2}{\sigma^2} &= g\left(\frac{1}{\delta} - g\right) \\ \frac{1}{\delta} + \frac{r^2}{\sigma^2}\delta &= g^2 + \frac{h^2}{\sigma^2} \end{aligned} \tag{46}$$

Obviously the signs of r and h are indeterminate (and meaningless, since each loads on a zero-mean noise term); without loss, we let $r, h \geq 0$. Furthermore, because there are two equations and three unknowns, there are generally multiple solutions; let us index the solutions by c. which means solving the above equations for r and h as functions of g. It is then immediate that a solution exists if and only if $g \in [1, 1/\delta]$, and that the solution is then given by

$$h = \sigma\sqrt{g\left(1/\delta - g\right)} \qquad \text{and} \qquad r = \sigma/\delta\sqrt{g-1}.$$

Note that $g = 1$ yields $r = 0$ and $h = h_{fb} \equiv \sigma\sqrt{(1/\delta - 1)}$. This identifies the equilibrium that replicates the first-best allocation. Also note that, as we vary g within the admissible range, h reaches its maximum at $g = \bar{g} \equiv \max\{1, 1/2\delta\}$ and this maximum is given by $h = \bar{h} \equiv \sigma\sqrt{\bar{g}\left(1/\delta - \bar{g}\right)}$. Finally, note that $h_{fb} = \bar{h}$ for $\delta \geq 1/2$ but $h_{fb} < \bar{h}$ for $\delta < 1/2$. Either way, the result holds with $\bar{h} = h_{fb}$.

Choice of δ. To find whether the previous can be supported as a symmetric competitive equilibrium, it is sufficient to check whether a given δ is a best response to all others' having signal precision δ.

Let $g(\delta)$ denote the equilibrium slope of prices as a function of others' attention level δ. The value of information for the agent, up to scale, is

$$b(\delta_i; \delta) = \frac{\delta_i}{\delta^2} + \sigma^2 g(\delta)\frac{\delta_i}{\delta} - \frac{\delta_i^2}{2\delta^2}\left(1 + \sigma^2\right)$$

The marginal benefit of paying a little more attention (cost of paying a little less) is

$$b_1(\delta_i, \delta) = \frac{1}{\delta^2} + \frac{\sigma^2 g(\delta)}{\delta} - \frac{\delta_i}{\delta^2}\left(1 + \sigma^2\right)$$

which, evaluated at the fixed-point condition, is

$$b_1'(\delta_i, \delta)|_{\delta_i = \delta} = \delta^{-1}(\delta^{-1} - 1) + \delta \cdot \sigma^2 \delta^{-2}(g(\delta) - 1) = \delta^{-1}(\delta^{-1} - 1) + \delta r^2 \tag{47}$$

Meanwhile, the marginal cost of information continuously increases from 0 to infinity on the domain $\delta \in (0, 1)$.

Generalization to other information costs. First consider the case $h = 0$. In this case, $g(\delta) = 1/\delta$ from solving (46). The marginal benefits curve when $h = 0$ is imposed, r^2 is solved for as a function of δ, is

$$(1 + \sigma^2)\delta^{-1}(\delta^{-1} - 1)$$

which is continuously decreasing from ∞ to 0 and thus has one (unique) intersection with the marginal cost curve, defining an equilibrium with uncorrelated noise.

Consider now imposing $r^2 = 0$. From the previous subsection, assuming the derivative of the cost function is invertible, then there exists a unique equilibrium attention level which solves $C'[\hat{\delta}] = \hat{\delta}^{-1}(\hat{\delta}^{-1} - 1)$.

Note that increasing r^2 unambiguously shifts down the marginal benefits curve defined by (47). Hence the equilibrium associated with any intermediate r^2 features strictly more attention. We cannot say *exactly*, however, whether these cases feature more or less non fundamental-volatility, because of the delicate interaction with changing equilibrium δ.

Positive costs of tracking v. Consider now the case in which the agent pays to track the vector (p, v). Assume that the cognitive cost has the following representation

$$C[\phi(\cdot)] = (1-K)\cdot I[\omega_i, p] + K\cdot I[\omega_i, (p, v)],$$

for some $K \in [0,1]$ and where $I[x, y]$ denotes the mutual information between random variables x and y, as induced by the specified joint distribution between ω and the state. When $K = 0$, this reduces to $C = I[\omega_i, p]$, which is the specification considered in the main text. When instead $K = 1$, this reduces to $C = I[\omega_i, (p, v)]$ and defines the sense in which the cost of tracking v is the "same" as that of tracking p. By the same token, $K \in (0,1)$ represents a situation in which the cost of tracking v is positive but lower than that of tracking p.

Expressing the signal in the form (9), we obtain

$$C = c(r, s, t; g, h, K) = -\log\big(1 - \delta_p(r, s, t; g, h)\big) - K\log\left(\frac{r^2}{r^2 + s^2 + \sigma^2 t^2}\right)$$

The first term is the mutual information of ω_i and p. The second term captures the "marginal" cost of tracking v, or the mutual information between ω_i and v conditional on p.

Fix (g, h) and consider how an individual decides to construct her signal. The capacity constraint is $C(r, s, t; g, h) \le M$, for some constant M. This can be re-written as

$$\delta_p \le 1 - e^M \left(\frac{r^2 + s^2 + \sigma^2 t^2}{r^2}\right)^K \tag{48}$$

The agent can freely pick (δ_p, r, s, t) subject to 43 and 44. For any $K > 0$, picking $s \neq 0$ or $t \neq 0$ strictly tightens the above constraint, thus reduces the highest attainable value of δ_p. Since utility is strictly increasing in δ_p, it follows that the agent find its strictly optimal to set $s = t = 0$. By market clearing, $h = 0$, which selects the equilibrium with no correlation as the unique equilibrium.

This equilibrium, however, need not be efficient. To see this, consider a planner that dictates agents what combination of (δ_p, r, s, t) to choose, subject to (43), (44) and (48), and that internalizes the market clearing conditions. Take now the limit as $K \downarrow 0$. The marginal effect on δ_p of increasing s^2 and t^2, while decreasing r^2, is approximately zero around the point $s^2 = t^2 = 0$. Meanwhile, the planner's objective can be shown to be proportional to $\delta_p^{-1} - \frac{r^2}{2}$, so there is a first-order benefit to decreasing r. It follows that, at least around the point $s^2 = 0$, social welfare is increasing in s^2 and hence the optimal s^2 is greater than 0. For K positive but small enough, the unique equilibrium is therefore dominated by an allocation with $s < 0$ and $h = -s > 0$.

B Proofs for State-Tracking Economies

Proof of Theorem 1

Let $((x^j(\omega))_{j=1}^J, y(\omega))$ be the competitive equilibrium allocation of goods and $((\phi^j)_{j=1}^J)$ be the levels of attention. Assume (counterfactually) that there exists some $((x'^j(\omega))_{j=1}^J, y'(\omega))$ and $((\phi'^j)_{j=1}^J, \phi') \in (\Delta(\Omega \times \Theta))^{J+1}$ that is feasible and Pareto dominates the previous. This means there exists some j such that $(x'^j(\omega), \phi'^j) \succ^j (x^j(\omega), \phi^j)$. For all other $i \neq j$, $(x'^i(\omega), \phi'^i) \succsim^i (x^i(\omega), \phi'^i)$.

Because of consumer optimization, it must be that $(x'^j(\omega), \phi'^j) \notin \mathbf{B}(p(\theta), e^j(\theta), a^j\Pi(\theta))$. Recall that this set embeds both the budget constraint and the measurability constraints associated with the cognitive friction. If these constraints are violated, then the proposed allocation is not feasible and the proof is done. If instead these constraints are satisfied, it must be the case that the budget is violated, or

$$\sum_{\omega,\theta}(p(\theta)\cdot x'^j(\omega))\phi'^j(\omega,\theta) > \sum_{\theta}(p(\theta)\cdot e^j(\theta) + a^j\Pi(\theta))$$

Because of local non-satiation, it must further be the case that, for all $i \neq j$,

$$\sum_{\omega,\theta}(p(\theta)\cdot x'^i(\omega))\phi'^i(\omega,\theta) \geq \sum_{\theta}(p(\theta)\cdot e^i(\theta) + a^i\Pi(\theta))$$

Denote the aggregate demand of a given type as $x^j(\theta) \equiv \sum_\omega x^j(\omega)\phi^j(\omega \mid \theta)$ in the first allocation and $x'^j(\theta) \equiv \sum_\omega x'^j(\omega)\phi'^j(\omega \mid \theta)$ in the proposed better one. Let aggregate supply similarly be $y(\theta) \equiv \sum_\omega y(\omega)\phi(\omega \mid \theta)$. Summing these expressions, using population weights ξ^j, and substituting in the expression for profits gives

$$\sum_{i=1}^{J}\sum_{\theta}\mu^j(p(\theta)\cdot x'^j(\theta)) > \sum_{i=1}^{J}\sum_{\theta}\mu^j(p(\theta)\cdot e^j(\theta)) + p(\theta)\cdot y(\theta)$$

Since it is part of an equilibrium, $(y(\omega, \phi(\omega,\theta))$ maximizes profits among all feasible combinations of production plans and cognitions, given prices $p(\theta)$. By construction, $(y'(\omega), \phi'(\omega,\theta))$ is feasible. Hence,

$$\sum_{\omega,\theta}(p(\theta)\cdot y(\omega))\phi(\omega,\theta) \geq \sum_{\omega,\theta}(p(\theta)\cdot y'(\omega))\phi'(\omega,\theta),$$

or equivalently

$$p(\theta)\cdot y(\theta) \geq p(\theta)\cdot y'(\theta).$$

Combining the above yields

$$\sum_{j=1}^{J}\sum_{\theta}\mu^j(p(\theta)\cdot x'^j(\theta)) > \sum_{j=1}^{J}\sum_{\theta}\mu^j(p(\theta)\cdot e^j(\theta)) + p(\theta)\cdot y'(\theta)$$

Provided $p(\theta) > 0$ for all $\theta \in \Theta$, this contradicts the resource-feasibility of the proposed allocation. That is, a Pareto dominating allocation cannot exist.

Proof of Theorem 2

The main insight from our First Welfare Theorem proof—the more easily verifiable optimality of allocations *within* types—suggests it will suffice to establish the requisite convexity properties for the "outer" preferences and production sets defined over the type-specific "team" problem. Below, we formalize this idea and show how the assumed convexity condition on C suffices for a version of the Second Welfare Theorem to apply in our setting even if the primitive preferences and technology are not convex. Once this step is completed, we can prove Theorem 2 by applying the standard Second Welfare Theorem on the economy defined by the outer preferences and technologies.

Outer preferences and technologies. Let $\bar{x}(\theta)$ be a type-specific demand for goods in state θ. This mapping $\bar{x}(\theta)$, given a fixed tracked variable z (and its relationship with the state θ), is compatible with a set of possible inattentive demands $x(\omega):\Omega\to\mathbb{R}^N$ and attention distributions $\phi(\omega,z)$. Denote this set of inner choices as

$$\mathbf{G}(\bar{x}(\cdot))\equiv\left\{x(\omega),\phi(\omega,\theta):\sum_{\omega\in\Omega}x(\omega)\phi(\omega\mid\theta)=x(\theta),\forall\theta\in\Theta\right.$$
$$\left.x\in\mathscr{X}\right\}$$

Now consider defining preferences over $\bar{x}(\cdot)$ bundles that "concentrate out" the choices of $x(\cdot)$ and $\phi(\cdot)$: these is the best choice of "aggregate demand" across states conditional on optimizing the other parameters. Denote this "outer preference" ordering as $\succsim^{j,Out}$ The outer preferences are represented by the following utility function:

$$\begin{aligned}U^j(\bar{x}(\cdot))\equiv\max_{x(\cdot),\phi(\cdot)}\sum_{\omega,\theta}u^j(x(\cdot),\theta)\phi(\omega,\theta)-C^j[\phi(\cdot)]\\ \text{s.t. }(x(\cdot),\phi(\cdot))\in\mathbf{G}(\bar{x}(\cdot))\end{aligned}\tag{49}$$

Finally let $\mathbf{X}$ denote some technologically feasible set for the outer bundles $\bar{x}(\theta)$.

Similarly, for the firms, we define the aggregate production set as

$$\mathbf{Y}:=\left\{\bar{y}(\cdot):\exists\left[(\phi(\cdot),y(\cdot))\in\mathbf{F}\right]\text{ s.t. }\sum_{\omega\in\Omega}y(\omega)\phi(\omega\mid\theta)=y(\theta)\right\}\tag{50}$$

where $\mathbf{F}$ is defined in (16) as the "regular" production frontier (here, specialized to $Z=\Theta$). These are aggregate production plans that are feasible under any choice of cognition. Note that, because of the linearity of the firm's problem, we can redefine profit maximization as selecting a bundle $y(\theta)\in Y$ to maximize $\sum_\theta p(\theta)\cdot y(\theta)$.

Convexity. We are now ready to show that the convexity assumption invoked in Theorem 2 suffices for the "outer" preferences and technologies defined above to be convex. This is formalized in the following:

Proposition 7. *Impose Assumption 2, i.e., let $C[\cdot]$ be (weakly) convex over the distribution of posteriors induced by a given signal ω about the physical state θ. Then:*

1. (Convexity of outer preferences) For every $j\in\{1,\dots,J\}$ and every pair $x(\theta),x'(\theta)\in\mathbf{X}$, $x(\theta)\succ^{j,Out}x'(\theta)$ implies that $a\,x(\theta)+(1-a)x'(\theta)\succ^{j,Out}x'(\theta)$ for all $a\in(0,1)$.
2. (Convexity of outer technology) $\mathbf{Y}$ is convex.

Proof. Here we prove that that the invoked assumption on C suffices for the "outer" preferences to be convex, even if the primitive preferences, U, are not. The proof of the convexity of $\mathbf{Y}$ is omitted because it follows from a similar argument.

Let $(x^0(\cdot),\phi^0(\cdot))$ and $(x^1(\cdot),\phi^1(\cdot))$ be the maximum arguments of (49) for $\bar{x}^0(\cdot)$ and $\bar{x}^1(\cdot)$, respectively. Define a new signal which is a compound lottery over the previous two signals: agents receive $\tilde{\omega}\equiv(\omega,\xi)$, where $\xi\in\{0,1\}$ indicates which of the two previous distributions ω has. In the space of posteriors, this is a convex combination of the previous signal structures. If this signal has joint distribution $\phi(\tilde{\omega},\theta)$, then by our convexity assumption on cognitive costs, $-C[\phi(\cdot)] \geq -aC[\phi^0(\cdot)]-(1-a)C[\phi^1(\cdot)]$.

Assume the allocation is such that agents consume $x^i(\omega)$ when they receive (ω, i). This strategy's feasibility is evident from the linearity of the budget constraint. The utility, net of cognitive costs, of this strategy is strictly higher than the convex combination of utilities from options 0 and 1 and hence the utility of option 1 given $x^0(\theta) \succsim^{j,Out} x^1(\theta)$. The utility of $ax^0(\theta)+(1-a)x^1(\theta)$ must be weakly higher than that of the constructed strategy, since it involves optimization over all possible feasible strategies. □

Putting everything together. From here, it is fairly straightforward to arrive at our version of the Second Welfare Theorem by following similar steps as in Debreu (1954).[21]

Let $((x^j(\cdot))_{j=1}^J, y(\cdot))$ be a Pareto optimal allocation of goods and $((\phi^j(\cdot))_{j=1}^J, \phi^f(\cdot))$ be the associated levels of attention. Denote the aggregate demand of a given type as $\bar{x}^j(\theta) \equiv \sum_{\omega\in\Omega} x^j(\omega)\phi^j(\omega \mid \theta)$ and aggregate supply as $\bar{y}(\theta) \equiv \sum_{\omega} y(\omega)\phi(\omega \mid \theta)$. From Theorem 2 in Debreu (1954), there exists a linear functional $v(\bar{x}(\theta)) \equiv \sum_{\theta} \lambda(\theta)\cdot\bar{x}(\theta)$ such that $U^j(\bar{x}'(\theta)) \geq U^j(\bar{x}(\theta))$ implies $v(\bar{x}'(\theta)) \geq v(\bar{x}(\theta))$ for all j and, for all feasible production plans $\bar{y}'(\theta)$, $\sum_{\theta}\lambda(\theta)\cdot y(\theta) \geq \sum_{\theta}\lambda(\theta)\cdot y'(\theta)$. This implies that the "aggregate" Pareto allocations $x(\theta)$ and $\bar{y}(\theta)$ solve the "outer" consumer problem and "outer" producer problem, respectively, for prices equal to $\lambda(\theta)$. Further, the Pareto consumer allocations $(x^j(\cdot),\phi^j(\cdot))_{j=1}^J$ are (possibly non-unique) solutions to the inner problem (49), and that the proposed firm allocations $(y(\cdot),\phi(\cdot))$ are (possibly non-unique) solutions to the inner problem (50). This is a sufficient condition for individual optimality. Thus the allocation can be supported as a competitive equilibrium.

C Proofs for Price-Tracking Economies

Proof of Theorem 3

Let us start with the first direction. By consumer optimality in the price-tracking economy, all bundles $(x^j(\cdot),\phi^j(\cdot))$ solve the price-tracking consumer problem. Assumption 3 ensures that these bundles have the same information-cost-inclusive utility with respect to θ or $(\theta, p(\theta))$. This is because the former is necessarily a sufficient statistic for the latter. Finally there are no differences in feasibility in terms of goods or signals. Hence the same allocations solve each consumer's problem in a state-tracking economy evaluated at prices $p(\theta)$; markets clear; and this is a state-tracking equilibrium.

The second result (state-tracking to price-tracking) follows from almost identical logic. Consumer optimality in the state-tracking economy implies that the bundles give weakly higher utility than any other bundle when evaluated with respect to $(p(\theta),\theta)$; agents thus optimize with the same actions and equilibrium prices; markets clear; and thus this is a price-tracking equilibrium.

Proofs of Theorem 4, Theorem 5 and Lemma 1

See main text.

[21] There are two additional assumptions that are part of the primitive model set-up. First, the space of consumption possibilities is convex. Second, the goods space is finite dimensional (because, by assumption, $N|\Theta| < \infty$). Third, outer preferences are continuous, or for every $(x(\theta), x'(\theta), x''(\theta))$, the sets $\{a \in [0,1] : x(\theta) \succsim^{j,Out} ax'(\theta)+(1-a)x''(\theta)\}$ are closed. This is a trivial consequence of having a continuous Bernoulli utility function.

Proof of Corollary 3

The difficult part is writing an extended version of Theorem 3, which says that equilibria in the price-only-tracking economy are interchangeable with equilibria in a replicating state-tracking economy. Armed with such a result, it is very clear how to re-prove Theorems 4 and 5. The remainder of this proof will thus extend Theorem 3.

Let us start with the first direction, from the price-only tracking economy to the state-tracking economy. By consumer optimality in the price-tracking economy, all bundles $(x^j(\cdot), \phi^j(\cdot))$ solve the price-only-tracking consumer problem. *A fortiori*, these bundles give weakly higher utility than any feasible bundle that has the $(p(\theta), \theta^j)$ sufficient statistic property in signals. By Lemma 1, a sufficient condition for a bundle to solve the θ-tracking consumer problem is that it weakly dominates all bundles with signals that satisfy the property by the state-tracking criterion.

Assumption 3 ensures that, if we take the same bundle with a signal for which $(p(\theta), \theta^j)$ is sufficient for θ, and generate a new "direct" signal of θ, the same *numerical* utility value will be the same.

Thus, we can conclude that $(x^j(\cdot), \phi^j(\cdot))$ has weakly higher utility than any feasible bundle with the $(p(\theta), \theta^j)$-sufficient-statistic property for the cognitive state. This, plus the fact that feasibility is the same in both the θ and price-tracking problems, proves that $(x^j(\cdot), \phi^j(\cdot))$ is optimal for each consumer evaluated at prices $p(\theta)$. Since the allocation is optimal for each consumer, and markets clear (carried over from the price-tracking economy), this is a price-tracking equilibrium.

The second result (state-tracking to price-only-tracking) has almost identical logic. Consumer optimality in the state-tracking economy implies that the bundles give weakly higher utility than all bundles with the sufficient statistic property in signals, which (combined with Assumption 3) *a fortiori* implies optimality in the price-only-tracking economy. Markets clear and all agents optimize; thus this is a price-only-tracking equilibrium.

D Consumer Surplus in the Example Economy

Since there are only supply (endowment) shocks, the equilibrium price function $P(\cdot)$ can be read as the inverse of the aggregate demand function. In textbook microeconomics, the area under the demand curve up to the equilibrium quantity measures consumer surplus. The same is true here in the absence of rational inattention ($\delta = 1$), but not in its presence ($\delta < 1$).

Proposition 8 (Consumer surplus)**.** *Consider the area under the demand curve, $\int_{\xi_0}^{\xi_1} P(X)\,\mathrm{d}X$. In the presence of rational inattention, this area ceases to measure either ex post consumer surplus (the increase in experienced utility from an increase in the realized value of the endowment from ξ_0 and ξ_1) or ex ante consumer surplus (the increase in expected utility from an increase in its prior mean from ξ_0 to ξ_1).*

This is perhaps most obvious in the limit of vanishingly little attention, or $\delta \downarrow 0$ (obtained as the equilibrium outcome in a limit as C becomes arbitrarily high for all $\delta > 0$). In this limit, demand is highly inelastic (and prices highly variable). But this inelasticity, and the correspondingly vast area under the demand curve, does not imply that the good is particularly "essential."

Proof of Proposition 8

Coincidence with complete information. Let $V(\xi) \equiv \mathscr{U}(x_{1i}^*, x_{2i}^*)$ the (identical) value function for each agent i, behaving optimally. The change in realized welfare is

$$V(\xi_1) - V(\xi_0) = (\xi_1 - \xi_0) + \frac{\xi_0^2 - \xi_1^2}{2}$$

Now let ξ be a random variable with mean ξ_0 and variance $\sigma^2 < \infty$. The change in expected utility from shifting the mean to ξ_1, maintaining all other properties of the distribution (including variance), is equal to the area under the demand curve $\int_{\theta_0}^{\theta_1} P(X)\,\mathrm{d}X$. $\mathbb{E}[V(\xi)] = 1 + \mathbb{E}[\xi] - (\mathrm{Var}[\xi] + (\mathbb{E}[\xi])^2)/2$. Thus the effect of a mean shift is $(\xi_1 - \xi_0) + (\xi_0^2 - \xi_1^2)/2$.

Finally, it is straightforward to calculate that the area under the demand curve is:

$$\int_{\xi_0}^{\xi_1} 1 - x\,\mathrm{d}x = \left[x - \frac{x^2}{2}\right]_{\xi_0}^{\xi_1} = (\xi_1 - \xi_0) + \frac{\xi_0^2 - \xi_1^2}{2}$$

This equivalence is *exact* given the lack of income effects in the model (and hence the equivalence of Marshallian and Hicksian demand).[22]

Non-coincidence with incomplete information. Let us define $V(\xi; r)$ as the cross-sectional expectation of utility given a certain value of ξ an a signal noise r. Equivalently, it is a welfare functional with utilitarian Pareto weights. It is straightforward to calculate

$$V(\xi; r) = \int_i (x_{1i} - x_{1i}^2/2 + x_{2i})\,\mathrm{d}i = 1 + \xi - (\xi^2 + r^2)/2$$

Thus change and realized utility and the change in expected utility from a mean shift remain equal to $\xi_1 - \xi_0 + (\xi_0^2 - \xi_1^2)/2$ irrespective of r (hence δ).

The area under the incomplete information demand curve changes with δ:

$$\begin{aligned}\int_{\xi_0}^{\xi_1} (p_0 - p_1 X)\,\mathrm{d}X &= p_0(\xi_1 - \xi_0) + (p_1/2)(\xi_0^2 - \xi_1^2) \\ &= \left(1 + \left(\frac{1-\delta}{\delta}\right)\mu\right)(\xi_1 - \xi_0) + \frac{1}{2\delta}(\xi_0^2 - \xi_1^2)\end{aligned}$$

Only in the case of $\delta = 1$ (i.e., complete information) does this coincide with the welfare measures.

E From Market Economies to Games: An Example

Consider the two-good example of Section 3. For the present purposes, let us add a representative, fully attentive, competitive firm that can transform the second good to the first good according to the following technology: $-y_2 = \frac{1}{2\beta}(y_1)^2$, where $y_1 \geq 0$ is the firm's output of the first good ("coconuts"), $-y_2 \geq 0$ is its input of the second good, and $\beta > 0$ is a technological parameter. Because the firm is attentive, it maximizes $p y_1 - y_2$ state-by-state. It follows that her optimal supply of coconuts is given by $y_1 = \beta p$ and that the market for coconuts is given by $\xi + y_1 = \bar{x}_1$, or equivalently $\xi + \beta p = \bar{x}_1$, where $\bar{x}_1$ is the aggregate consumption of coconuts. Solving this for the price, we have

$$p = P(\xi, \bar{x}_1) = \tfrac{1}{\beta}(\xi - \bar{x}_1)$$

[22] We use the entire area under the demand curve, instead of the "Harberger triangle," because there is an implicit producer surplus above the (vertical) supply curve in the endowment economy.

Finally, using the budget constraint of the household to express his consumption of he second good ("money") $x_2 = 1 + p\xi - px_1$, replacing p from the above market-clearing condition, and dropping the "1" index from x_1, we can express the realized utility of the typical consumer as

$$u = \mathscr{V}(x, \bar{x}, \xi)$$

where $\mathscr{V}(x, \bar{x}, \xi) \equiv \mathscr{U}(x, 1 + P(\xi, \bar{x})(\xi - x)) = x - \frac{1}{2}x^2 + 1 + \frac{1}{\beta}(\xi - \bar{x})(\xi - x)$. That is, the Walrasian economy under consideration maps to a game in which the players are the consumers, a player's action is her demand for coconuts, and her payoff is given by $\mathscr{V}(x, \bar{x}, \xi)$, or a function of her action, the average action, and an exogenous fundamental.

For a moment, abstract from the *endogeneity* of information, as done in Angeletos and Pavan (2007). For *any* given information structure, the equilibrium is the fixed point to

$$\mathbb{E}[V_x(x, \bar{x}, \xi) | \omega] = 0,$$

while the planner's solution is the fixed point to

$$\mathbb{E}[V_x(x, \bar{x}, \xi) + V_{\bar{x}}(x, \bar{x}, \xi) | \omega] = 0,$$

where $V_x(\cdot)$ and $V_{\bar{x}}(\cdot)$ denote the derivatives of $V(\cdot)$ with respect to, respectively, the individual and average actions. Intuitively, the planner instructs the agents to internalize their payoff externality, to the best possible degree given their information. This is true for arbitrary V and, in this sense, for arbitrary games. But for the particular game under consideration, the equilibrium "happens" to satisfy $\mathbb{E}[V_{\bar{x}}(x, \bar{x}, \xi) | \omega] = 0$ and, therefore, the equilibrium use of information coincides with the efficient use of information, regardless of the information structure. As already explained, this property follows in our closed-form example from the quasi-linearity of utility, and in our general framework from complete markets.

This explains the sense in which the class of inattentive economies studied in this paper map to the class of games that, under the taxonomy of Angeletos and Pavan (2007), feature efficiency in the equilibrium use of information. Now take this class of games, let the players choose their information structure in the flexible manner we have described in this paper, and compare the equilibrium to the solution of planner that can control both actions and information. Provided that our invariance condition holds, the equilibrium and the planner's solution continue to coincide. In other words, our invariance condition suffices for efficiency to extend from the use of information to the acquisition of information.

The broader validity of this insight is explored in a recent paper by Hébert and La'O (2019). But whereas our invariance condition is strong enough to make sure that the equilibrium attains the same outcomes as a planner who can regulate not only people's signal choices but also the very objects they track, Hébert and La'O (2019) show that a *weaker* invariance condition is needed under the restriction that the planner can do the former but not the latter.[23]

[23] To see what we mean, consider again the game-representation of our simple example. In Hébert and La'O (2019), the planner may economize on the cognitive cost of one's learning about the average action $\bar{x}$ by manipulating the mapping from the state of nature to $\bar{x}$, but is precluded from sending *other* messages in place of $\bar{x}$. This is akin to a *restricted* version of the planner's problem we studied in Section 6, which in turn helps explain why a weaker invariance condition is needed for efficiency in the context of Hébert and La'O (2019). But note that such a restriction is not natural in our own context, because, as noted in the beginning of Section 6, other market data or even taxes could themselves play the role of messages in place of or in addition to prices (or the average action).