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

We thank Jakub Steiner for discussing our paper at the 2019 ASSA meeting; and Benjamin Hébert, John Leahy, Jennifer La'O, Stephen Morris, and various seminar participants for useful comments. Angeletos acknowledges the support of the National Science Foundation under Grant Number SES-1757198. We have no other financial interests or conflicts to disclose. 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
JEL No. D5,D6,D8,E7

ABSTRACT

We study the efficiency of competitive markets in the presence of a general form of rational inattention. The appropriate amendments of the Fundamental Welfare Theorems are shown to hold if rational inattention is modeled as an arbitrary cost for obtaining signals about the exogenous state of nature. If instead rational inattention is modeled as a cost for observing prices or other endogenous outcomes, inefficiency can arise because of a cognitive externality: people do not internalize how their choices affect the complexity of the price system and thereby others' cost of tracking or decoding it. This externality is muted in an important special case, when cognitive costs are given by the mutual information of agents' decisions with the joint of the price system and the entire state of nature. For more general costs, however, there is room for policies aimed at simplifying or otherwise regulating markets.

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, and otherwise "cognitively constrained." They appear to overlook some pieces of information and over-react to others. They use simplifications and heuristics. And they can be manipulated by nudges or attention-grabbing interventions.

In such circumstances, it is tempting to question the efficacy of the market mechanism. After all, both Hayek's (1945) classic argument about the superiority of markets over central planning and the textbook versions of the Welfare Theorems abstract from the aforementioned impediments.[1] But is there room for improving upon markets barring the elimination of the cognitive constraints themselves?

We explore this question by augmenting the Arrow-Debreu framework with a generalized form of *rational inattention* (Sims, 2003) and revisiting the two Fundamental Theorems of Welfare Economics.

We model attention, or cognition, as the choice of a noisy signal subject to a non-pecuniary cost.[2] But we allow great flexibility on what these signals and costs may be. We thus encompass two growing literatures: one that shows how various departures from the textbook version of "homo economicus" can be recast as rational choice with costly observation of prices or other decision-relevant objects;[3] and another that studies the axiomatic and experimental foundations of costs other than mutual information.[4] But whereas these literatures study single-agent decision theory, we study efficiency in equilibrium.

Our main results can be summarized as follows. If rational inattention is modeled as a cost for tracking, or understanding, the underlying state of nature, the (appropriately amended) Welfare Theorems hold *necessarily*. If instead inattention is modeled as a cost for tracking, or understanding, equilibrium objects such as prices and income, the Welfare Theorems may or may not hold: efficiency hinges on whether cognitive costs satisfy a certain invariance condition.

The sole origin of any inefficiency is the cognitive externality produced when the entropy, sparsity, or other cognitively relevant traits of the objects that agents try to track or understand are endogenous to the actions of others. Our invariance condition makes sure that this externality is effectively muted.

This condition is necessarily satisfied if cognitive costs are measured by the Shannon mutual information between an agents' decisions and *every* other object in the economy. Away from this benchmark, the cognitive externality can be active and policy intervention can be warranted. In some instances, welfare is improved by stabilizing or "simplifying" prices; in others, by "noising up" markets.

Example and main ideas (Section 3). Before diving into the full Arrow-Debreu formalism, we start with a simple example that transparently delivers some of the main insights.

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. Both goods are in fixed supply. The aggregate endowment of coconuts is the economy's only stochastic fundamental.

[1] Similarly, Sims (2010) writes: "If both sides of the market react to market prices with rational inattention, then neither side is reacting precisely and immediately. Prices therefore cannot play their usual market-clearing role." And Mackowiak, Matejka, and Wiederholt (2018) write: "Agents make mistakes. Welfare theorems do not apply."

[2] For our purposes, "attention," "cognition," and "information-processing" are interchangeable notions.

[3] See, inter alia, Koszegi and Matejka (2018) and Lian (2018) on mental accounting; Caplin, Dean, and Leahy (2018) on categorization; Woodford (2009) on bounded recall; Kohlhas and Walther (2018) on differential attention; Ilut and Valchev (2017) on imperfect contemplation of the optimal policy rule; and Gossner, Steiner, and Stewart (2019) on manipulation.

[4] Such as Caplin, Dean, and Leahy (2018), Hébert and Woodford (2018), and Dean and Neligh (2017).

The cost of attention is an increasing, convex function of the (Shannon) mutual information between the signal received and the object tracked. In one scenario, this object is the aggregate endowment of coconuts. In another, it is their price. Finally, cognitive mistakes (noises) are uncorrelated across agents.

Under these assumptions, an "inattentive" equilibrium exists, is unique, and is the same *regardless* of whether agents track the endowment of coconuts or their price. The first best is unsurprisingly not attained. Relative to it, various "pathologies" emerge, including misallocation in the consumption of coconuts and (in the case of endogenous production) excessive volatility in aggregate output. Nevertheless, 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.

Does this conclusion extend to more general economies? Could a planner perhaps raise welfare by replacing the market of coconuts with another mechanism? Such questions cannot be fully addressed until the second part of our paper. Two exercises, however, provide intuition for what assumptions about the form of rational inattention *can* open the door to inefficiency.

Suppose first that people have hard time tracking volatile or "complex" objects (as measured, for instance, by their entropy). This makes no difference when agents track the state of nature, because the latter's properties are fixed. But it can render the equilibrium inefficient when agents track prices. Cognitive costs can then be economized by stabilizing or "simplifying" prices. And attention choices can be strategic complements, opening the door to multiple, Pareto-ranked equilibria, or "cognitive traps."

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

The cognitive externality and our invariance property. The common thread behind the above instances of inefficiency is that cognitive resources can be economized by manipulating the price system. This is possible in general because markets fail to internalize the cognitive externality mentioned earlier on. But why was efficiency preserved under *some* conditions on the cognitive costs? It must be these conditions were somehow muting the cognitive externality. To clarify this logic, and state precise conditions that allow 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 very flexible. We can restrict the available signals to be Gaussian (Morris and Shin, 2002; Mackowiak and Wiederholt, 2009), sparse (Gabaix, 2014), or coarse (Gul et al., 2017); or we can leave them entirely unrestricted (Matejka and Sims, 2011). We can accommodate public signals (Morris and Shin, 2002; Nimark et al., 2019), endogenous attention to such signals (Colombo, Femminis, and Pavan, 2014; Myatt and Wallace, 2012), and rich higher-order uncertainty (Angeletos and Lian, 2018; Tirole, 2015). We do not need to take a stand on which decisions adjust to meet budgets. And we can allow for decisions to be based on non-nested information sets, capturing bounded recall and mental accounting (Lian, 2018).

What is crucial for our purposes is the modeling choice of which objects people "track" and try to learn, or reason, about. We say that "an agent tracks object z" if his cognitive cost can be expressed as a functional C of the joint density of z and her signal (equivalently, her decisions). We then structure the formal arguments by considering two specifications for z.

Welfare Theorems for state-tracking economies (Section 5). We start with the case in which z equals the exogenous state of nature. This case is ubiquitous in the literature.[5] Depending on one's preferred specification of C, agents may be able to learn about only a subset of the entire state, may be restricted to receiving independent signals about different pieces of it, or may face no such restrictions.

In this context, an appropriate amendment of the First Welfare Theorem is shown to hold under arbitrary C: any inattentive equilibrium coincides with the solution to the problem of a planner who is free to regulate people's attention to z, as well as their consumption-production decisions, but is banned from eliminating, sidestepping, or ignoring people's cognitive costs.[6] 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 and the (new) domain of attention strategies.

What is going on? By affecting equilibrium prices, one's attention choice influences the payoff and the attention choices of others. The equilibrium attention choices can be represented as the Nash equilibrium of a game like those studied in Myatt and Wallace (2012), Colombo, Femminis, and Pavan (2014) and Tirole (2015). But whereas abstract games allow for arbitrary externalities and inefficiencies, the type of Walrasian economies studied here are more "disciplined."

As long as people only track the state of nature, *all* externalities are pecuniary. It is well known that the "traditional" pecuniary externalities, which pertain to consumption and production choices, net out thanks to complete markets. As shown here, the same is true for the additional pecuniary externalities that originate in attention choices. It follows that, in state-tracking economies, the efficiency of the equilibrium is guaranteed regardless of the form, severity, and empirical footprint of inattention.

This result clarifies that the failure of the Welfare Theorems claimed in Gabaix (2014) rests on a deviation from *rational* inattentionSimilarly, the concerns voiced by Sims (2010) and Mackowiak et al. (2018) about the workings of the market mechanism (see footnote 1) can be relevant only if one specifies the cognitive cost as a function of *endogenous* objects, as we discuss next.[7]

Welfare Theorems for price-tracking economies (Section 6). Consider now the case in which people can track prices. This opens the door to the cognitive externality mentioned earlier on. It also means the price system plays a dual role: it not only clears markets but also has the potential of economizing cognitive costs relative to a world in which people are forced to track the state of nature.

This begs the question of whether welfare could be improved by replacing markets with other means of encoding and communicating information, in direct contradiction of Hayek's (1945) thesis. Addressing this question requires a further amendment: the planner is allowed to send arbitrary messages in place of market prices. The updated efficiency concept resembles an information-design problem à la Bergemann and Morris (2013, 2018), freed of incentive compatibility but ridden with imperfect communication: rational inattention amounts to people receiving the planner's messages with endogenous noise.

[5]It nests, inter alia, Colombo et al. (2014), Mackowiak and Wiederholt (2009, 2015), Myatt and Wallace (2012), Paciello and Wiederholt (2014), and Tirole (2015). A notable exemption is Denti (2016), which allows the players of a game to collect signals about the endogenous actions of others but does not study the normative questions we are concerned with.

[6]This distinguishes our approach to normative questions from that of Farhi and Gabaix (2019), which does not incorporate cognitive costs in the welfare criterion.

[7]Our result for state-tracking economies also generalizes Angeletos and La'O (2018) and Gul, Pesendorfer, and Strzalecki (2017), and qualifies Colombo, Femminis, and Pavan (2014) and Tirole (2015). See the discussion in Section 2.

In this context, the amended Welfare Theorems hold under the following restriction on people's cognition, or information processing:

Informational Invariance. *C is such that the cost of tracking z is the same as tracking any $\tilde{z}$ that is a sufficient statistic of z with respect to the agent's signals (or decisions).*

This result is linked to the first part of our paper as follows. The above condition (and hence efficiency) necessarily holds if (i) C is a transformation of Shannon mutual information and (ii) z includes the entire state of nature. Otherwise, our invariance condition may fail and inefficiency may emerge.

More succinctly, the above condition means that it makes no difference if market signals are replaced with other means of communication that add or subtract decision-irrelevant information, or merely with a complete description of the underlying state of nature. This coincidence let us extend our Welfare Theorems from state-tracking to price-tracking economies. But it also means that, in these circumstances, people make no cognitive gains from tracking the relevant prices as opposed to the entire state of nature.

Bottom line. Our results therefore tie together the answers to the following questions:

Q1. *Does rational inattention alone invite policy intervention?*

Q2. *Are there welfare gains from tracking market signals instead of the underlying state of nature?*

If the answer to Q2 is negative, then the answer to Q1 (the question raised in the beginning) is also negative. This is the scenario captured by our invariance condition and the one implicitly imposed in the existing macroeconomic applications of rational inattention.

But if the answer to Q2 is positive, which seems quite plausible, then there is generally room for policies that aim at manipulating attention choices and market signals—*even if* all the familiar assumptions of the standard Welfare Theorems, such as complete and competitive markets, are satisfied.

2 Related Literature

The literature on rational inattention spurred by Sims (2003, 2006) is voluminous. Some works focus on decision theory (Caplin and Dean, 2015; Matejka and McKay, 2015; Matejka, Steiner, and Stewart, 2015); others study specific macroeconomic models (Mackowiak and Wiederholt, 2009, 2015) or games (Colombo, Femminis, and Pavan, 2014; Myatt and Wallace, 2012). Our paper's main contributions 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.

Our invariance condition has a similar flavor as the axiom of "invariance under compression" in Caplin, Dean, and Leahy (2017): both properties embed free disposal of decision-irrelevant information. But whereas that paper, like the broader decision-theoretic literature on rational inattention, studies the implications of this axiom in a single-agent context, here we show how a related (but not identical) invariance property suffices for efficiency in a general-equilibrium context.

The link between information-processing costs and equilibrium efficiency is further explored in a recent, complementary paper by Hébert and La'O (2019). They establish that a weaker invariance condition is necessary and sufficient for efficiency in a large game where players track the state and the average action of others.

Closely related are also Angeletos and La'O (2018) and Gul et al. (2017), which establish efficiency in, respectively, a macroeconomic model in which firms are rationally inattentive and an endowment economy in which consumers' information is a coarsening of the true state space. Relative to these papers, we make two key contributions. First, we accommodate a more general form of rational inattention, which nests multiple specifications in the literature. Second, and more crucially, we accommodate the possibility that people track endogenous objects, show how this opens the door to a cognitive externality, and identify conditions under which this externality may or may not be muted.

The same points distinguish our paper from Colombo, Femminis, and Pavan (2014), Myatt and Wallace (2012), Paciello and Wiederholt (2014), and Tirole (2015). In particular, our results clarify that the instances of inefficiency found in these works derive not from rational inattention per se but rather from its combination with other distortions, such as missing or non-competitive markets. This also explains why our "cognitive traps" borrow their name from, but are of a different origin than, those in Tirole (2015).

Our analysis also recalls an older literature that studied efficiency in "island economies" (Prescott and Townsend, 1984; Prescott and Rios-Rull, 1992). In that literature, the informational friction is the byproduct of the geographical segmentation of markets. But agents always observe, perfectly and without any cost, both their own fundamentals and all the prices in the markets they participate—and they do not make any information choice. The endogeneity of information and the friction in observing prices are the core elements that distinguish our work from that literature.

These elements bring to mind the literature on noisy rational-expectations equilibria (Grossman and Stiglitz, 1980; Laffont, 1985; Amador and Weill, 2010). In that literature, agents can perfectly observe, and extract information from, prices. As a result, information can be imperfect, and inefficiency can emerge, only if certain markets are missing (e.g., if there are no futures markets). In our context, markets are complete; agents are optimally inattentive to prices; and inefficiency is of a different origin.

3 An Inattentive Economy

We start with a tractable example that foreshadows our subsequent and more general Welfare Theorems and that 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

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

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 $\xi \sim N(\mu, \sigma^2)$. For now, ξ is also the entire state of nature.

Momentarily, abstract from inattention. Suppose further that markets operate after ξ is realized.[8] Normalize the price of money to 1 and let p denote the (relative) price of coconuts.

[8] 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.

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 agent maximizes utility, $U(x_{1i}, x_{2i})$, subject to her budget, $p(\xi)x_{1i} + x_{2i} \leq p(\xi)\xi + 1$; 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 FOCs 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 for coconuts only on a private signal of ξ, contaminated with idiosyncratic noise. Let this signal be

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

where $\varepsilon_i \sim N(0,1)$ is independently distributed across agents and r_i is a non-random variable, under the control of agent i. In equilibrium, ω_i will also serve as a signal of p. But agents are not allowed, for the time being, to observe a *direct* signal of p.

Let $\delta_i \equiv \sigma^2/(\sigma^2 + r_i^2)$ be the correlation between the signal and fundamental. The agent can pick any $r_i \in \mathbb{R}_+ \cup \infty$, or equivalently any $\delta_i \in [0,1]$. A higher δ_i helps reduce mistakes. But it also has a cognitive cost equal to $C(\delta_i)$, where $C(\cdot) : [0,1] \to \mathbb{R} \cup \{\infty\}$ is increasing and convex. Given the Gaussian specification, this is akin to the cost being an increasing and convex function of the mutual information between ω_i and ξ.[9] By the same token, δ_i measures the level of attention in mutual-information units.

Because the noise in ω_i has been assumed to be idiosyncratic, the aggregate demand and the price of coconuts are only functions of 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:

$$\begin{aligned} \max_{x_{1i}(\cdot),x_{2i}(\cdot),\delta} \quad & \int_{s_i,\xi} \left(x_{1i}(\omega_i) - \frac{x_{1i}(\omega_i)^2}{2} + x_{2i}(\omega_i,\xi) \right) \phi(\omega_i,\xi)\,\mathrm{d}\omega_i \mathrm{d}\xi - C(\delta_i) \\ \text{s.t.} \quad & p x_{1i}(\omega_i) + x_{2i}(\omega_i,\xi) \leq p(\xi)\xi + 1, \forall(\omega_i,\xi) \end{aligned} \tag{2}$$

where $\phi(\omega_i,\xi)$ is the joint density between ω_i and ξ implied by δ. Note that this problem contains, not only the optimal consumption of coconuts conditional on ω_i, but also the optimal choice of ϕ, or δ.

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.

[9] Just let $C(\delta) = K(I(\omega_i,\xi))$, where $I(\omega_i,\xi) \propto -\log(1-\delta)$ is the mutual information of the agent's signal and the underlying state and K is an increasing and convex function.

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

$$p = P(\xi) \equiv 1 - \left(1 - \frac{1}{\delta}\right)\mu - \frac{1}{\delta}\xi, \qquad x_{1i} = \omega_i, \qquad \text{and} \qquad x_{2i} = 1 + P(\xi)\xi - P(\xi)\omega_i. \tag{3}$$

Note that the consumption of coconuts is no more equated across agents; it is *as if* there is uninsured idiosyncratic risk. Furthermore, the price of coconuts is more volatile than in the frictionless benchmark. Had production been endogenous, this could have translated to more volatility in their aggregate quantity ("amplified business cycles").

We can next express the optimal choice of δ_i as the solution to the following outer problem:

$$\delta_i \in \arg\max_{\delta_i} \left\{ B(\delta_i, \delta) - C(\delta_i) \right\},$$

where $B(\delta_i, \delta)$ is the expected utility evaluated along the consumption plan and the price function seen in (3). Think of B as the equilibrium "benefit" of attention and C as its cost. The dependence of $B(\delta_i, \delta)$ on δ captures the dependence of the agent's utility on the attentiveness of *others*.

Clearly, an equilibrium corresponds to any δ^* such that

$$\delta^* \in \operatorname{argmax}_{\delta} \left\{ B(\delta, \delta^*) - C(\delta) \right\}. \tag{4}$$

Computations, detailed in the appendix, show that

$$B(\delta_i, \delta) = \frac{\sigma^2 \delta_i}{2\delta^2} - \frac{\sigma^2}{\delta} \tag{5}$$

More attention is always better ($B_1 > 0$) because it reduces the mistakes in consumption choices. But the 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. To guarantee an interior solution, $\delta^* \in (0, 1)$, we henceforth impose $C(0) = 0$ and $\lim_{\delta \uparrow 1} C(\delta) = \infty$. We thus reach the following result.

Proposition 1 (Equilibrium). *The equilibrium exists and is unique. The equilibrium level of attention is given by $\delta^* \in (0, 1)$, where δ^* is the unique solution to* (4)*. The equilibrium price and allocation are given by* (3)*, with $\delta = \delta^*$.*

3.4 Welfare and Efficiency

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 2 (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."

This underscores how rational inattention can complicate welfare inferences out of observables, or the evaluation of policy counterfactuals.[10] But does it interfere with the welfare questions we are interested in, namely the efficiency of the equilibrium attention strategies?

The following casual argument might suggest that there is "generically" room for inefficiency. Consider the problem of a benevolent planner who can not manipulate the competitive equilibrium for given δ, but can directly control the latter. This planner solves the following problem:

$$\delta \in \arg\max \left\{ B(\delta,\delta) - C(\delta) \right\}. \tag{6}$$

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 except for the knife-edge case in which $B_2 = 0$, the equilibrium and the planner's solution won't coincide.

This argument would have been correct if B was "free" for the modeler to choose, as in the works of Colombo et al. (2014), Myatt and Wallace (2012), and Tirole (2015). But B is *not* arbitrary in our context.

Proposition 3 (Efficiency)**.** *The fixed point to 4 coincides with the solution to 6. 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 only via prices. But as long as utility is transferable (as in the present example) or markets are complete (as in our general analysis), pecuniary externalities do not create a wedge between private and social valuations. They net out on average, guaranteeing that B_2 is indeed zero in equilibrium (even though it can be non-zero away from equilibrium). Appendix A illustrates this logic within the present example. Section 5 establishes is greater applicability.

3.5 Correlated Noise

So far, we have equated the state of nature with the payoff-relevant fundamental (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 the 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:

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

where $\varepsilon_i \sim N(0,1)$ is i.i.d. and where the pair (r_i, b_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 $b_i > 0$. Such an equilibrium would have aggregate demand move with v, 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

[10] For instance, it is an open question how the presence of rational inattention affects optimal tax formulas which, in its absence, depend on the observed demand or supply elasticities.

$b_i = b > 0$ and, as a result, non-fundamental volatility emerges in both prices and quantities.[11]

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

3.6 State- vs Price-tracking Economies and Cognitive Externalities

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 per se. 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.

The precise, and more general, concepts will be developed in Section 4. Here, we sketch the main ideas by considering a simple variant of the example studied so far.

In this variant, agents collect a signal ω directly about p, and their cognitive cost is a functional C of ϕ, where ϕ denotes the joint density of ω and p. Our leading specification has $C[\phi(\cdot)]$ be the mutual information between ω and p. But other specifications will also be considered.[12]

Clearly, ϕ can be expressed as the product of the marginal for p and the conditional of ω given p. When choosing ϕ, or equivalently the conditional for ω given p, the agent treats the marginal for p as given: this is her prior about prices. But this prior is itself determined in equilibrium by the choices of others. And because this prior enters C, a non-pecuniary, or "cognitive," externality emerges: by affecting the equilibrium price system, one's choices can affect the information-processing costs of others.

In the sequel, we first explain why this externality is muted in a special case that is commonplace in the literature. We then explore two specific departures from it.

3.7 An Efficient Price-tracking Economy

Suppose 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.

[11] 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, let C be decreasing in both r_i and b_i but more steeply so in b_i than in r_i (costs can be economized by substituting idiosyncratic for aggregate noise).

[12] 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. But in the present context the case of tracking p *alone* is a plausible thought experiment.

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

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 5. *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, the assumptions made thus far have made sure that the cognitive externality is muted and, hence, that efficiency is preserved. But what if we relax these assumptions?

3.8 Inefficiency I: Allowing "Complexity" to Matter

Imagine that the cost of information depended also on the "complexity" of the tracked object, as measured by its entropy. To fix ideas, let

$$C(\phi) = K\left(I[\omega, z], H[z]\right),$$

where I stands for mutual information, H for entropy, K for some increasing and convex function, and z is either θ or p, depending on whether agents track the state or the price. When ω and z are Gaussian, this can be re-expressed (up to a redefinition of the function K) as

$$C(\phi) = K(\delta_z, \sigma_z)$$

where δ_z is the signal-to-noise ration between ω and z and σ_z is the standard deviation of z.

The dependence of C on σ_z helps capture the idea that it may be harder for agents to track objects that are more uncertain or more complex. Here, complexity is equated to variance, or scale. But similar arguments can be made in a non-Gaussian context were complexity is equated to coarseness or sparsity.

When $z = \theta$, σ_z is exogenously fixed by nature and the dependence of C on it is irrelevant. When instead $z = p$, σ_z is endogenously determined in equilibrium and the dependence of C on it becomes crucial. Intuitively, if agents are "confused" or their bandwidths are overwhelmed by high volatility in prices, there is room for policy interventions aimed at stabilizing prices.

Consider first the case in which $K(\cdot, \cdot)$ increases in both arguments but is separable in them: $K_1 > 0$, $K_2 > 0$, and $K_{12} = 0$. This preserves the uniqueness of the equilibrium but breaks its efficiency in particular direction: a planner would like to subsidize attention or otherwise stabilize prices.

Now let $K_{12} > 0$. This case captures the idea that, on the margin, it is harder to pay attention when the price system is "more confusing" (high variance). Crucially, this case 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 (5), the "game" of choosing δ may admit multiple equilibria. And because equilibria with low δ display more volatility in prices and, thereby, a larger cognitive externality, these equilibria attain low welfare than equilibria with higher δ. In other words, equilibria with low attention represent "cognitive traps."

Proposition 6 (Cognitive Traps I)**.** *There exist K, with $K_2 > 0$ and $K_{12} > 0$, such that the economy admits multiple, Pareto-ranked equilibria, each corresponding to a different value for δ^*.*

The lessons obtained so far can be summarized as follows. In the previous subsection, the cognitive externality was muted because the cost of tracking prices was "scale-free," in the sense of depending only on the mutual information of ω and p. Here, instead, the cognitive externality is active because the cost also depends on the scale, variance, or entropy of the tracked object. Beyond the Gaussian realm, such a dependence could invite for prices that are, not only less volatile, but also more sparse—or for markets that are "less sophisticated."

Of course, one could also imagine the opposite situation than that captured by the preceding analysis, namely a situation in which $K_2 < 0$ and $K_{12} < 0$, or higher entropy making it *easier* to learn. Hébert and Woodford (2018) and Pomatto et al. (2018) propose cost functionals that seem to have this flavor and argue they help explain certain evidence from experimental psychology. So the broader lesson from the example considered here is not to argue for a particular policy intervention but rather to illustrate how plausible relaxations of mutual-information costs may have, not only interesting decision-theory implications (the focus of the existing literature), but also crucial welfare implications.

3.9 Inefficiency II: Correlated Mistakes

Let us re-embrace the "scale-free" property discussed above but allow noise to be correlated across agents. We shall show that this, too, allows for inefficiency and cognitive traps, although of a more subtle kind than those obtained above.

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

$$\omega = p + r\varepsilon_i + a\xi + b\nu \tag{7}$$

where v is aggregate noise, ε is idiosyncratic noise, and (r, a, b) 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 θ :

$$p = -c\xi + dv,$$

for some scalars c and d. An equilibrium can be indexed by the tuple $(r, a, b, c, d, \delta_p)$. 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 in the following case, in which δ_p is exogenously fixed, or derived in equilibrium from an information cost function that puts zero cost to $\delta \leq \delta_p$ and an arbitrarily large one otherwise:

Proposition 7 (Cognitive Traps II)**.** *Fix an attention level δ_p. There exist a continuum of equilibria, indexed by $d \in [0, \bar{d}]$, such that (i) non-fundamental volatility and welfare both increase in d and (ii) when $d = \bar{d}$, the allocation of goods is first-best.*

The Appendix works out all the math, in a slightly more general case that allows for δ_p to change in equilibrium as attention costs are smooth. Here we sketch the extreme cases. First, there clearly exists an equilibrium in which $a = b = d = 0$ 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 exists another equilibrium, which indeed attains the first best. Set $r = 0$. This gives $\omega = p + a\xi + b\nu = p_0 + (a - c)\xi + (b + d)v$. In equilibrium, the cross-sectional average of ω has to equal ξ, or else aggregate demand would not equal ξ. It follows that $c = 1$ and $b = -d$, and hence ω perfectly reveals ξ and $p = p_0 - \xi + dv$. This occurs for a uniquely pinned down level of non-fundamental volatility, or for $d = \bar{d} \equiv \sigma\sqrt{\delta_p^{-1} - 1}$, which intuitively increases in fundamental variance and decreases in the total level of attention.

Finally, there also exist intermediate equilibria which mix the two previous cases. All of these feature strictly more consumption inequality than the case with $r = 0$ and the same (zero) cognitive cost — thus they are inferior "cognitive trap" equilibria.

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. More succinctly, the above analysis assumes that tracking p was costly, but tracking v was not. This may make sense if, similarly to the idiosyncratic mistake ε_i, the common mistake v is internal to people's cognitive process. But if v is an exogenous sunspot, tracking it may be costly.

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 v, now she strictly prefers ε_i because v is no longer costless. This selects the equilibrium with $d = 0$ as the unique equilibrium. But it does *not* guarantee its efficiency.

In particular, if the cost tracking v is small enough relative to the cost of tracking p (or, relatedly, the cost of tracking the underlying endowment), then the unique equilibrium is dominated by an allocation that features $d > 0$. Furthermore, the planner can implement this allocation by introducing a subsidy that varies with v, thus adding non-fundamental volatility in prices (and quantities too, if supply is elastic). That is, even though the multiplicity disappears, the argument for "noising up" prices remains.

When does this argument ceases to hold? Only when the cost of tracking v is sufficiently high relative to the cost of tracking prices. In particular, if the total cognitive costs can be expressed as the mutual information of an individual's consumption with the pair (p, v), the efficiency of the equilibrium is restored. This actually follows from a more general result, which we develop in the second part of the paper.

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 the planner can send other messages in place of prices?

To address these questions as clearly and flexibly as possible, in this section we augment the standard Arrow-Debreu framework with a generalized form of rational inattention, which nests multiple specifications found in the literature. In the subsequent sections, we then revisit the two Fundamental Theorems of Welfare Economics.

4.1 Frictionless Benchmark

Let θ, the state of nature, be a random variable drawn from a finite set Θ.[13] This 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. Its has probability distribution is denote by $\pi(\theta)$.

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 $p(\theta) = (p_n(\theta))_{n=1}^N \in \mathbb{R}_+^N$ 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 *will* matter in our formalization of cognitive frictions. For that reason, we use consistent notation here.

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 $x^j(\theta) = \left(x_n^j(\theta)\right)_{n=1}^N \in X \subset \mathbb{R}_+^N$ denote the consumption bundle (across N goods) for the representative 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:

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

where $u^j : X \times \Theta \to \mathbb{R}$ is a type-specific, state-contingent, Bernoulli utility. We finally 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 $y(\theta) = (y_n(\theta))_{n=1}^N$ 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 $F\left(y(\theta), \theta\right) \leq 0$.

[13] Finite states allow for simpler proofs, though conceptually they seem unimportant for the main results.

4.2 Generalized Rational Inattention

We now introduce our generalization of rational inattention. This involves a definition of the tracked object, the measurability constraints on decision making, and the information choice.

Tracked object. As explained in the Introduction, we say that "agents tracks an object" if their cognitive costs are a function C of the joint distribution of this object and of their signals (or "cognitive states"). In the stylized example of Section 3, this object was variously the aggregate endowment of coconuts or their price. Here, we allow agents to track a possibly type-specific object, denoted by z^j. For our main analysis, we then consider two possible specifications,

Definition 1. A *state-tracking economy* is an economy in which $z^j = \theta$ for all j.

Definition 2. A *price-tracking economy* is an economy in which $z^j = (\theta, p)$ for all j.

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 like taxes).

The underlying state. It is worth reiterating that the state θ is a complete collection of all aggregate or type-specific random variables, whether fundamental or non-fundamental. In the language of our earlier example, if there were a fundamental shifter ξ and a noise term ν, the state would be $\theta = (\xi, \nu)$. At the present level of abstraction , however, the distinction between fundamentals, noise and sunspots can be quite fussy. 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.

Attention, or cognition. Agents *choose* how much effort they put in understanding what's going on around them and in figuring out their optimal response. Following the tradition of Sims (2003) and the subsequent literature on rational inattention, we formalize these ideas as the choice of a signal upon which decisions have to measurable.

Pick any consumer of any type j. Firm, we let ω be a random variable representing the agent's cognitive state, or the signal upon which her decisions have to be measurable. Second, we denote by ϕ^j the joint density of ω and z^j. Third, we let the consumer *choose* this density out of some set, Φ^j, subject to some non-pecuniary cost, represented by a functional $C^j : \Phi^j \to \mathbb{R}$. Finally, we let this cost be separable from, and additive to, her expected utility from consuming goods. The "extended" preferences over the spaces of goods and attention choices are thus represented by

$$\mathbb{E}\left[u^j(x^j, \theta)\right] - C^j\left[\phi^j\right],$$

where u^j is the original utility function (as in the frictionless benchmark).[14]

[14] Such additive separability is common place in the literature but not strictly need for our results.

We make similar assumptions for the firms. In particular, we think a firm's cognitive cost as forgone output and express its "extended" production frontier as

$$F(y, \theta) + C^f[\phi] \leq 0,$$

where F is the original technology and $C^f : \Phi^f \to \mathbb{R}$ is the firm's cognitive cost.

Naturally, for every $j \in \{1, ..., J\} \cup \{f\}$, Φ^j embeds the constraint that, for any $\phi^j \in \Phi^j$, the corresponding marginal equals the unconditional distribution of z, which the agent takes as given. Φ^j may be given by the set of *all* probability densities satisfying this natural property, or by any subset obeying any additional constraint the modeler may deem appropriate (e.g., sparsity, conditional normality, etc).

One can always maps the choice of an element in Φ^j to a choice of a posterior about z^j. This is commonplace in the rational-inattention literature. The only subtlety here is that the agent's prior about z^j, and hence the cost associated with any chosen posterior, becomes endogenous to the choices of others as we move from state-tracking to price-tracking economies. Whether this contribution to inefficiency or not is the subject of our subsequent analysis.

Signals, or cognitive states. We assume that the realizations of ω are identically and independently distributed *within* any given type, but allow for arbitrary heterogeneity and correlation *across* types through the state θ. Since we can arbitrarily refine the partition of the population to types, this is only a technical assumption that lets us invoke a law of large numbers to the continuum within each type.

We let the received signal have separate components, one for each good: $\omega = (\omega_n)_{n=1}^N$. These components can have different (finite) supports, given by Ω_i^j, so the entire vector ω has support $\Omega^j \equiv \prod_{n=1}^N \Omega_n^j$. The consumption (or, in the firm's case, the production) of good i is required to be measurable in ω_n.

A flexible form. As anticipated in the Introduction, the above formalism allows an arbitrary specification of which decisions are subject to rational inattention and which ones adjust to meet the budget. For instance, it could be that the demand for the "last" good is attentive (measurable in θ), as in the example of Section 3 and various applications in the literature; or that budgets are met in the more elaborate fashion assumed in Gabaix (2014); or that rational inattention takes the form of an arbitrary tremble along the consumer's budget.

Furthermore, we can restrict the available signals to belong in a particular class, or let them completely unrestricted. We can accommodate the leading case of Shannon mutual-information costs, as well as the variant cases that, according to a growing decision-theoretic literature, are needed in order to explain certain choice data. And last but not least, we can encompass another growing literature that shows how various "behavioral" phenomena can be cast within the rational-inattention framework by making appropriate assumptions about the available signals and/or the associated costs (see the references in 3).

Additional notation. By construction, the joint distribution of (ω, z, θ), for a given type j, is given by

$$f^j(\omega, z, \theta) \equiv \phi^j(\omega, z) \cdot \mathbb{I}\{z^j(\theta) = z\} \tag{8}$$

where $\mathbb{I}\{\cdot\}$ is the 0-1 indicator function that takes the value 1 only if $\{\cdot\}$ is true. The implied marginal density for the pair (ω, θ) and the conditional likelihood of ω given θ are then given by, respectively,

$$g^j(\omega, \theta) \equiv \sum_z f^j(\omega, z, \theta) \qquad \text{and} \qquad g^j(\omega \mid \theta) \equiv \frac{g^j(\omega, \theta)}{\pi(\theta)} \tag{9}$$

Complete markets and budgets. Complete markets are, for many applications, unrealistic. They are, however, central to proving the Welfare Theorems in our paper, just as they are in the standard Arrow-Debreu framework. The only subtlety relative to the standard framework is that, in our context, complete markets means that transfers can be contingent not only on θ but also on ω. We can thus write the ex-ante budget constraint of a type-j consumer as follows:

$$\sum_{\omega,\theta}(p(\theta)\cdot x^j(\omega))g^j(\omega \mid \theta) \leq \sum_{\theta}\left(p(\theta)\cdot e^j(\theta) + a^j\Pi(\theta)\right)$$

where, as already explained, g^j denotes the likelihood of the signal conditional on the state, as implied by ϕ^j, the consumer's attention choice.

The complete-markets assumption does *not* eliminate the bite of rational inattention on choice and equilibrium. It also does not mean that risks can be shared as perfectly as in a first-best world: recall the inequality in the consumption of coconuts induce by inattention in the toy model of Section 3. It only let us isolate rational inattention as the sole origin of *any* departure from the first best.

The equivalent assumption in the toy model of Section 3 was quasi-linear (or "transferable") utility. In macroeconomic applications, the same objective is often achieved by allowing for "big families" in the tradition of Lucas (1975). Understanding the interaction of rational inattention with missing markets is of course important, but it is not the subject of this paper.

4.3 New Consumer and Firm Problems

We express the consumption bundle for a type j consumer as $x^j(\omega) = (x^j_n(\omega_n))_{n=1}^N$ and the production bundle for a firm as $y(\omega) = (y_n(\omega_n))_{n=1}^N$. We let

$$\mathcal{X}^j \equiv \left\{\text{functions } x : \Omega^j \to \mathbb{R}^N_+\right\} \qquad \text{and} \qquad \mathcal{Y} \equiv \left\{\text{functions } y : \Omega \to \mathbb{R}^N\right\}$$

denote the relevant choice sets for these consumption and production bundles.

The goods and attention choice of a type-j consumer can then be written in the following form:

$$\begin{aligned} &\max_{x(\cdot),\phi(\cdot)} \sum_{\omega,\theta} u^j(x(\omega),\theta)g(\omega,\theta) - C^j[\phi(\cdot)] \\ &\text{s.t. } (x(\cdot),\phi(\cdot)) \in \mathbf{B}(p(\cdot), e^j(\cdot), a^j\Pi(\cdot)) \end{aligned} \tag{10}$$

where

$$\begin{aligned} \mathbf{B}(p(\cdot), e^j(\cdot), a^j\Pi(\cdot)) \equiv \Bigg\{ &\text{functions } x(\cdot) \in \mathcal{X}^j \text{ and } \phi(\cdot) \in \Phi^j \text{ such that :} \\ &\sum_{\omega,\theta}(p(\theta)\cdot x(\omega))g(\omega \mid \theta) \leq \sum_{\theta}\left(p(\theta)\cdot e^j(\theta^j) + a^j\Pi(\theta)\right); \text{ and} \\ &x_n(\omega) = x_n(\omega') \text{ whenever } \omega_n = \omega'_n, \forall n \Bigg\} \end{aligned} \tag{11}$$

This set encodes both the budget and the measurability (inattention) constraints. 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 (9).

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

$$\begin{aligned} \max_{y(\cdot),\phi(\cdot)} \ & \sum_{\omega,\theta} (p(\theta)\cdot y(\omega))g(\omega,\theta) \\ \text{s.t. } & (y(\cdot),\phi(\cdot)) \in \mathbf{F} \end{aligned} \tag{12}$$

where

$$\begin{aligned} \mathbf{F} \equiv \Big\{ & \text{functions } y \in \mathcal{Y} \text{ and } \phi \in \Phi^f \text{ such that :} \\ & F\left(y(\omega),\theta^f\right) + C^f[\phi(\cdot)] \leq 0, \forall(\omega,\theta) \text{ s.t. } g^f(\omega,\theta) > 0; \text{ and} \\ & y_n(\omega) = y_n(\omega') \text{ whenever } \omega_n = \omega'_n, \forall n \Big\} \end{aligned} \tag{13}$$

This set encodes the joint technology over production and attention. 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, including the cognitive aspects of each, constitute an "inattentive economy." We now define general equilibrium in this context:[15]

Definition 3 (Inattentive Equilibrium)**.** An equilibrium is a combination of consumption-production choices, $([x^j(\cdot)]_{j=1}^J, y(\cdot))$, attention strategies, $([\phi^j(\cdot)]_{j=1}^J, \phi^f(\cdot))$, and prices, $p(\cdot)$, such that the following are true.

1. *Consumers optimize*: for each j, $(x^j(\cdot), \phi^j(\cdot))$ solves program (10), fixing prices and the stochastic process for tracked object z^j.
2. *Firms optimize*: $(y(\omega), \phi^f(\omega, z^f)$ solves program (12), fixing prices and the stochastic process for tracked object z^f.
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

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

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 z^j is endogenous in price-tracking economies. Whenever added clarity is need, we qualify the equilibrium as "state-tracking" or "price-tracking," depending on whether $z^j = \theta$ or $z^j = (\theta, p)$.

5 State-Tracking Economies

In this section, we focus on the scenario in which agents only track the exogenous state of nature ($z^j = \theta$ for all j). 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.

[15] 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.

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-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 combination 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:

$$\sum_{j=1}^J \mu^j \sum_\omega x^j(\omega)\phi^j(\omega \mid \theta) = \sum_{j=1}^J \mu^j e^j(\theta) + \sum_\omega y(\omega)\phi^f(\omega \mid \theta), \forall\theta \in \Theta \tag{14}$$

$$F(y(\omega), \theta) + C[\phi^f(\omega, \theta)] \leq 0, \ \forall(\omega, \theta) \text{ s.t. } \phi^f(\omega, \theta) > 0 \tag{15}$$

$$x_n^j(\omega) = x_n^j(\omega') \text{ if } \omega_n = \omega_n', \ \forall j, \quad \text{and} \quad y_n(\omega) = y_n(\omega') \text{ if } \omega_n = \omega_n' \tag{16}$$

$$\phi^j(\cdot) \in \Phi^j, \ \forall j \in \{1, \dots, J, f\} \tag{17}$$

Definition 5 (Efficiency)**.** A combination of consumption-production choices and attention strategies is *efficient* in a state-tracking economy if there exists no other such combination 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. A fifth 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, first-best, 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).[16]

Our version of the First Welfare Theorem will establish that, regardless of the domains $\left[\Phi^j\right]_{j=1}^J$ of the available signal structures and 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:*

$$\max_{[x^j(\cdot), \phi^j(\cdot,\cdot)]_{j=1}^J, (y(\cdot), \phi^f(\cdot))} \sum_{j=1}^N \chi^j \mu^j \left[\sum_{\omega,\theta} u^j(x^j(\omega), \theta)\phi^j(\omega, \theta) - C^j[\phi^j(\omega, \theta)]\right] \tag{18}$$

s.t. (14)*,* (15)*,* (16) *and* (17)*,*

for some Pareto weights $\left[\chi^j\right]_{j=1}^J$.

[16]This means, inter alia, 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., contained in Φ^j for all j) but too costly to be optimally chosen.

Had information been exogenous (i.e., had Φ^j been a singleton 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. Compare this also to Bergemann and Morris (2013), who let a planner choose the information structure in game but abstract from 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. To gain intuition, however, 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:

$$\mathbb{E}\left[\frac{\partial u^j(x^j(\omega),\theta)}{\partial x_n} \mid \omega\right] = \mathbb{E}\left[\frac{\lambda_i(\theta)}{\chi^j} \mid \omega\right] \tag{19}$$

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:

$$\mathbb{E}\left[\frac{\partial u^j(x^j(\omega),\theta)}{\partial x_n} \mid \omega\right] = \mathbb{E}\left[m^j p_i(\theta) \mid \omega\right] \tag{20}$$

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 (19) and (20). This reflects the informational, or cognitive, friction.

In the language of Angeletos and Pavan (2007), the coincidence of conditions (19) and (20), 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:

$$u^j(x^j(\omega),\theta) - \frac{\partial C^j}{\partial \phi^j(\omega,\theta)} = \frac{\lambda(\theta)}{\chi^j} \cdot x^j(\omega) \tag{21}$$

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 (18), 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 $j \in \{1, \ldots, J\}$, $x(\cdot) : \Omega \to \mathbb{R}^N$, $\phi(\cdot) \in \Phi^j$, and $\varepsilon > 0$, there exists some $x'(\cdot) \in \mathscr{B}_\varepsilon(x(\cdot)) \equiv \{x''(\cdot) : \|[x''(\omega) - x(\omega)]_{\omega\in\Omega}\| < \varepsilon\}$ and some $\phi'(\cdot) \in \Phi^j$ 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-rival, non-traded input. But this kind of understanding is part of our contribution. Furthermore, as anticipated in the Introduction, the result helps clarify the following points.

First, the failure of the Welfare Theorems claimed in Gabaix (2014) rests on a deviation from rational attention. Second, the related point made by Mackowiak 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.*

The requisite convexity is satisfied for any member of the class of "posterior separable" cost functionals (Caplin and Dean, 2015; Caplin et al., 2017). This class includes the familiar Shannon mutual information, a generalization known as Tsallis mutual information, and other plausible specifications of C as a convex transformation of some "distance" between priors and posteriors. Away from this class, the Second Welfare Theorem may fail. But the First Welfare Theorem remains valid: in state-tracking economies, the efficiency of any equilibrium is guaranteed for entirely arbitrary cost functionals.

6 Price-tracking Economies

We now extend the analysis to price-tracking economies (Definition 2). To simplify the exposition, we shut down production and focus on pure exchange economies.

6.1 Constrained Efficiency, with 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 replace such "market communication."

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

$$\mathcal{M} = \{\text{any function from } \Theta \text{ to } \mathbb{R}^{|\Theta|+N}\} \tag{22}$$

where $|\Theta|$ is the (finite) number of unique primitive states and the dimensionality of the target space just ensures, in a simple way, that the planner *could* send something as rich as (θ, p).

The messages 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. This provides a potential trade-off for the planner: too much information could confuse agents, while too little would render them unable to make good decisions.[17]

Figure 1 illustrates this scenario in a two-type example. The choices in red, and the "inner" part of the diagram, resembles a pure information-design problem, as in Bergemann and Morris (2018). 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, but must be measurable in ω.

Formally, the amended version of feasibility given an arbitrary message is the following:

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

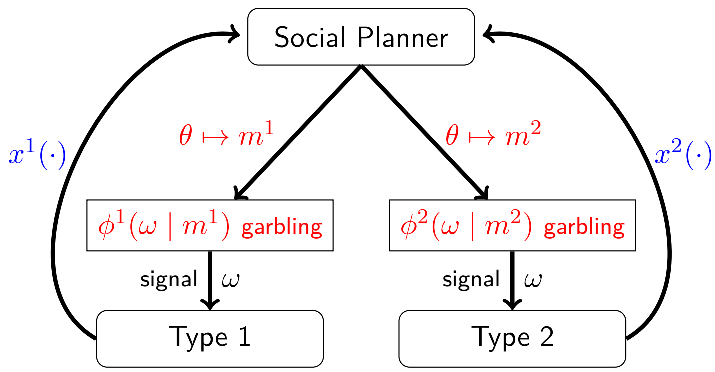

Figure 1: The planner's problem as a "communication problem."

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:

$$g^j(\omega \mid \theta) = \frac{\sum_m \phi^j(\omega, m) \cdot \mathbb{I}\{m^j(\theta) = m\}}{\pi(\theta)}, \forall j \in \{1, \ldots, J, f\} \tag{23}$$

$$\sum_{j=1}^J \mu^j \left(\sum_\omega x^j(\omega) g^j(\omega \mid \theta) \right) = \sum_{j=1}^J \mu^j e^j(\theta) + \left(\sum_\omega y(\omega) g^f(\omega \mid \theta) \right), \forall \theta \in \Theta \tag{24}$$

$$F(y(\omega), \theta) + C^f[\phi^f(\cdot)] \leq 0, \ \forall (\omega, \theta) \text{ s.t. } \phi^f(\omega, m^f(\theta)) > 0 \tag{25}$$

$$x_n^j(\omega) = x_n^j(\omega') \text{ if } \omega_n = \omega_n', \ \forall j, \quad \text{and} \quad y_n(\omega) = y_n(\omega') \text{ if } \omega_n = \omega_n' \tag{26}$$

$$\phi^j(\cdot) \in \Phi^j, \ \forall j \in \{1, \ldots, J, f\} \tag{27}$$

This is the same as Definition 4 but with the added flexibility about how the garbling/signals and hence the decisions can depend on θ through the sent messages.

The efficiency concept from Definition 5 carries over with the same amendment.

Definition 7 (Efficiency with arbitrary messages)**.** 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 *efficient* if there is not other such combination that is feasible in the sense of Definition 6, strictly preferred by a positive mass of agents, and weakly preferred by all other agents.

Provided convexity, we can then express the associated planner's problem as follows.

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

$$\max_{[m^j(\cdot)]_{j=1}^J, [x^j(\cdot), \phi^j(\cdot,\cdot)]_{j=1}^J} \sum_{j=1}^N \chi^j \mu^j \left[\sum_{\omega,\theta} u^j(x^j(\omega), \theta) g^j(\omega, \theta) - C^j[\phi^j(\cdot)] \right] \tag{28}$$

$$\text{s.t. (23), (24), (25), (26), and (27)}$$

for some Pareto weights $\{\chi^j\}$.

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 relative to either the aforementioned restriction or the equilibrium of the price-tracking economy ultimately depends 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-a-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.

6.2 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 (i.e., each value of $\tilde{z}$ corresponds to at least one value of 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

$$\tilde{\phi}(\tilde{\omega}, \tilde{z}) = \sum_{z} \phi(\tilde{\omega}, z) \cdot I\{\tilde{z} = h(z)\} \tag{29}$$

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 $z = f(\theta)$ be some random variable, for an arbitrary function f, and ω be some signal of z with joint density $\phi(\omega, z)$. Next, let $\tilde{z} = h(z)$ be some coarsening of z, for an arbitrary function h, and construct $\tilde{\omega}$ and $\tilde{\phi}(\tilde{\omega}, \tilde{z})$ as in (29). Then:*
(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 (29) looses no information.

The same argument, from a different perspective, 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.

6.3 "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.

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

(a) Let

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

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:

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

(b) Conversely, let

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

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:

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

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 $A = \{x^j(\cdot), \phi^j(\cdot), m^j(\cdot)\}$ 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, $A' = \{x^j(\cdot), \varphi^j(\cdot)\}$ 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.

6.4 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 (18), 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 (18). 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., transfers). □

Theorems 4 and 5 together generalize the logic of Proposition 3 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 4 (here, Theorem 3) directly implies a generalized version of Proposition 5 (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.

6.5 Prices as 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 $(\hat{x}(\cdot), \phi(\cdot))$ 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 is defined to summarize the role of θ in the Bernoulli utility function). By Assumption 3, there is a strictly lower cognitive cost to another (feasible) bundle with the same allocation and new attention choice $\tilde{\phi}(\tilde{\omega}, (p(\theta), \theta^j))$, defined by (29). Thus $\mathcal{U}(x(\cdot), \tilde{\phi}(\cdot)) > \mathcal{U}(x(\cdot), \phi(\cdot))$. This is a contradiction to the optimality of the first bundle. $\square$

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.

6.6 Two Variants

We can flip the logic in the previous subsection 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.

Consider first 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.*

A second case which is mathematically equivalent, but has a different interpretation, is constraining agents in equilibrium to track the net trades of all types along with the own fundamental.

Definition 9. A *net-trades-tracking economy* is an economy in which $z^j = (\theta^j, (\bar{x}^i(\theta) - e^i(\theta))_{i=1}^J)$ for all j, and a type-j agent's signal, conditional on z^j, cannot be correlated with the signal of any other agent.

Inspection of the market clearing condition reveals that the former is sufficient for calculating all equilibrium prices. Hence, the logic of the previous case extends to this case:

Corollary 4. *Consider a net-trades-tracking economy in the sense of Definition 9. Theorems 4 and 5 continue to hold.*

This case also hints at how our methods could be adapted to games: letting agents track the net trades of others in a market is akin to letting players track the actions of others in game.

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

$$C = (1 - K)I(\omega, p) + KI(\omega, (p, v)),$$

for some $K \in [0, 1]$. 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 under the lenses 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

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

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 the precise condition under which mutual-information costs suffice for efficiency.

Corollary 5. *If cognitive costs are a transformation of the Shannon mutual information of ω and the* joint *of p and θ, our invariance condition is guaranteed, and so do our welfare theorems.*

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

We conclude with a few take-home lessons and avenues for future research.

The case for policy intervention. Our results have tied the efficiency of competitive markets to the question of whether people can effortlessly go back and forth between tracking the equilibrium objects of interest (prices) and tracking the underlying state of nature. But such an approach also means that people suffer no cognitive costs from being *forced* to track the entire state of nature as opposed to tracking merely the market prices or other equilibrium objects they care about. In a nutshell, when our invariance condition holds, markets do not *strictly* economize cognitive costs.

This is both what it takes for efficiency to be guaranteed, and what it is *implicitly* assumed in much of the applied literature (e,g., Angeletos and La'O, 2018; Colombo, Femminis, and Pavan, 2014; Mackowiak and Wiederholt, 2015). But this need not be what is relevant in practice. If one expects people to be greatly confused in a counterfactual world that replaces the available market signals with a complete, explicit description of the underlying state of nature, then one is lead by our results to conclude that there is likely room for manipulating people's choices and the available market signals.

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 random variables (e.g., sunspots, noisy news in the media) at a lower cost than others, equilibria in which market outcomes vary with such variables could be superior to equilibria in which market outcomes are tied to "hard" fundamentals.

While these examples are illuminating, they are not exhaustive. For instance, we can imagine examples where welfare is improved by regulating trade volume or even shutting down certain markets. A further investigation of these possibilities and of their empirical relevance is beyond the scope of our paper. But we hope to have open the door to such investigations by showing what it takes, in terms of a departure from "conventional" modeling choices, for inefficiency to obtain.

From axiomatic/experimental foundations to welfare implications, and back. Our results provide a new context for 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) literatures on rational inattention. These literatures offer axiomatic and empirical foundations for mutual-information costs or plausible departures thereof. But they are focused on the positive question of which costs best capture individual choice data. Here, we have instead shifted the focus to normative questions regarding the equilibrium interaction of multiple inattentive agents.

Our paper is thus the first to build a bridge from the aforementioned literatures to the welfare implications of rational inattention. We thus highlighted that a departure from the standard, mutual-information specifications is *necessary* in order to open the door to inefficiency, and illustrated by example what this inefficiency may look like. But we did not provide a "taxonomy" of the welfare implications of the different cost functionals found in the aforementioned literatures, nor have we identified an "ideal" experimental test for the policy questions we have been concerned with. These seem important directions for future research. A recent step in the first direction is made by Hébert and La'O (2019).

Incomplete or thin markets. Like the standard welfare theorems, ours relied on the assumption that markets are complete. In Section 3, though, we illustrated how rational inattention can *itself* be the source of inequality in consumption, or what looks like incomplete risk-sharing. A direct extension of this logic is that cognitive constraints can explain why certain markets may be thin or even inexistent.

To given an example, suppose that a fraction λ of the consumers either cannot, or optimally choose not to, comprehend certain contingencies, in the sense that it is sufficiently costly for them to receive any signal about them (or to "reason about them"). Then, their consumption choices have to be entirely independent from these contingencies, which in turn means that these consumers will not trade on such contingencies even if markets for them exist. And when $\lambda \to 1$, these markets disappear.

Perhaps more interestingly, our results also shed light on when such endogenously thin or absent markets can be the symptom of inefficiency: only when the model of rational inattention violates the invariance condition identified here. This circles back to our earlier point about the value of pushing the research frontier past mutual-information costs.

A complementary direction for future work is, of course, to extend the analysis to a setting with *exogenously* incomplete markets. We suspect that constrained efficiency may be preserved if the conditions on cognition identified in our paper are combined with appropriate homotheticity and spanning assumptions, of the kind identified in Geanakoplos and Polemarchakis (1986).

From markets to games. Had information (attention) been exogenous, the economies studied here can be mapped to a class of games in which the equilibrium use of information is efficient in the sense of Angeletos and Pavan (2007): the condition on payoffs that guarantees efficiency in that paper corresponds to our setting's property that pecuniary externalities net out. 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). Efficiency is shown to obtain in a large game where players track average actions if a generalization of the condition on payoffs provided in Angeletos and Pavan (2007) is combined with an invariance condition on information costs similar to ours.[18] 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.

Rational vs irrational inattention. 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 (Gabaix, 2014; Chetty et al., 2009). But our analysis also hints at how some lessons from that literature *could* be fruitfully recast and studied in a rational-inattention context.

In particular, the notion of "salience" that is prevalent in this literature may be recast as a violation of our invariance condition. Similarly, we have indicated how, in economies where people economize cognitive costs by tracking prices instead of the underlying, possibly incomprehensible, state of nature, a "desire for sparsity" à la Gabaix (2014) could open the door for policy intervention *without* the form of irrationality assumed in that paper.

[18] There is a subtle difference in our set-ups. In Hébert and La'O (2019), average actions are a sufficient statistic for the GE interaction and the social planner is not allowed to replace this *informationally* with another message. This helps explain why their invariance condition is less stringent than ours.

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 (the focus of the existing decision-theoretic and experimental literatures) and efficiency (the focus of our paper).

References

Amador, M. and P.-O. Weill (2010): "Learning from Prices: Public Communication and Welfare," *Journal of Political Economy*, 118, 866–907.

Angeletos, G.-M. and J. La'O (2018): "Optimal Monetary Policy with Informational Frictions," *Journal of Political Economy*, forthcoming.

Angeletos, G.-M. and C. Lian (2018): "Forward Guidance without Common Knowledge," *American Economic Review*, 108, 2477–2512.

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.

——— (2018): "Information Design: A Unified Perspective," *Cowles Foundation Discussion Paper No. 2075R3*.

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.

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*.

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.

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.

Gossner, O., J. Steiner, and C. Stewart (2019): "Attention Please!" *mimeo*.

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*.

Ilut, C. and R. Valchev (2017): "Economic Agents as Imperfect Problem Solvers," *Duke University mimeo*.

Kohlhas, A. and A. Walther (2018): "Asymmetric Attention," *IIES miméo*.

Koszegi, B. and F. Matejka (2018): "An attention-based theory of mental accounting," .

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 Local Thinking," Working paper, Massachusetts Institute of Technology.

Lucas, R. E. J. (1975): "An Equilibrium Model of the Business Cycle," *Journal of Political Economy*, 83, 1113–44.

Mackowiak, 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.

Mackowiak, B. A., F. Matejka, and M. Wiederholt (2018): "Survey: Rational Inattention, a Disciplined Behavioral Model," CEPR Discussion Papers 13243, C.E.P.R. Discussion Papers.

Matejka, F. and A. McKay (2015): "Rational Inattention to Discrete Choices: A New Foundation for the Multinomial Logit Model," *American Economic Review*, 105, 272–98.

Matejka, F. and C. A. Sims (2011): "Discrete Actions in Information-Constrained Tracking Problems," *CERGE-EI mimeo*.

Matejka, F., J. Steiner, and C. Stewart (2015): "Rational Inattention Dynamics: Inertia and Delay in Decision-Making," *CERGE-EI mimeo*.

Morris, S. and H. S. Shin (2002): "Social Value of Public Information," *American Economic Review*, 92, 1521–1534.

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

Nimark, K., R. Chahrour, and S. Pitschner (2019): "Sectoral Media Focus and Aggregate Fluctuations," *mimeo*.

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," .

Prescott, E. C. and J.-V. Rios-Rull (1992): "Classical competitive analysis of economies with Islands," *Journal of Economic Theory*, 57, 73 – 98.

Prescott, E. C. and R. M. Townsend (1984): "General Competitive Analysis in an Economy with Private Information," *International Economic Review*, 25, 1–20.

Sims, C. A. (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.

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

Woodford, M. (2009): "Information-Constrained State-Dependent Pricing," *Journal of Monetary Economics*, 56, 100–124.

Appendix

A Proofs for Section 3

Proof of Proposition 2

Coincidence with complete information. Let $V(\xi) := U(X^*_{1i}, X^*_{2i})$ 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 $\int_{\theta_0}^{\theta_1} P(X)\,\mathrm{d}X$. $\mathbb{E}[V(\xi)] = 1 + \mathbb{E}[\xi] - (\text{Var}[\xi] + (\mathbb{E}[\xi])^2)/2$. 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).[19]

Non-coincidence with incomplete information. Let us define $V(\xi; r)$ as the cross-sectional expectation of utility given a certain value of ξ. 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:

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

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

Proofs of Propositions 3, 4, and 5

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

$$\begin{aligned}\max_{x_{1i}, x_{2i}, \delta_i} \quad & \mathbb{E}\left[\mathbb{E}_i\left[x_{1i} - x_{1i}^2/2 + x_{2i}\right]\right] - C(\delta_i) \\ \text{s.t. } & px_{1i} + x_{2i} \le p\xi + 1\end{aligned} \tag{30}$$

[19] 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.

where δ_i is the "attention level" for agent i . The informational precision δ is now a choice variable, and $C(\delta)$ 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, $p = a - b\xi$, 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

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

One solution to this problem is to set $x_{1i} = \mathbb{E}_i[1-p]$.

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 choice of a given agent i with the choice of all other agents $-i$.

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 those terms, it is convenient to write the objective as

$$B(\delta_i;\delta) - C(\delta_i)$$

with the first part defined as

$$B(\delta_i, \delta) := \frac{\sigma^2\delta_i}{2\delta^2} - \frac{\sigma^2}{\delta} \tag{31}$$

$B(\delta, \delta)$ is exactly the utilitarian 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

$$\frac{\sigma^2}{2\delta^2} = C'(\delta) \tag{32}$$

This has a unique solution for $\delta \in [0, 1]$ if costs are strictly convex and infinite for full revelation of the state (like an "Inada condition" to ensure an interior solution).

State-traking efficiency (Proposition 3). 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

$$\frac{\mathrm{d}}{\mathrm{d}\delta} B(\delta, \delta) = C'(\delta) \tag{33}$$

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 δcoincides 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:

$$V_i(\delta) := \max_{\delta_i} \left\{B(\delta_i, \delta) - C(\delta_i)\right\},$$

or, equivalently,

$$V_i(\delta) = \max_{x_{1i}(\cdot),\delta_i} \int_\omega \int_\xi \left[u(x_{1i}(\omega)) + (1 + p(\xi; \delta)(\xi - x_{1i}))\right] \phi(\omega, \xi; \delta_i)\, \mathrm{d}\xi\, \mathrm{d}\omega \tag{34}$$

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

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

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

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

In equilibrium, $\delta_i = \delta$ and $x_{1i}(\omega) = x_1(\omega)$ by symmetry, and $\int_\omega \left(\xi - x_1(\omega)\right) \phi(\omega, \xi; \delta)\, \mathrm{d}\omega = 0$ 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 4 and 5). 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 4. 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 5. Hence this Proposition is proved by the same calculation.

Proof of Proposition 6

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

$$\frac{\partial}{\partial \delta_i} B(\delta_i, \delta_{-i})|_{\delta_{-i}=\delta_i} = \frac{\partial}{\partial \delta_i} K(\delta_i, \sigma_p)|_{\delta_{-i}=\delta_i}$$

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

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

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 7

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

$$p = p_0 - c\xi + d\nu$$

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

$$\begin{aligned} \omega_i &= p + r\varepsilon_i + a\xi + b\nu \\ &= p_0 + (a-c)\xi + (b+d)\nu + r\varepsilon_i \end{aligned}$$

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

$$-ac\sigma^2 + db = 0. \tag{35}$$

The signal-to-price correlation is then given by

$$\delta_p = \frac{c^2\sigma^2 + d^2}{(a-c)^2\sigma^2 + (b+d)^2 + r^2} \tag{36}$$

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

$$1 - \bar{\mathbb{E}}[p] = 1 - \delta\left[p_0 + (a-c)\xi + (b+d)\nu\right] = \xi \tag{37}$$

This implies that $p_0 = 1/\delta$, $a = -1/\delta + c$, and $b = -d$. 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 c and d and r. Equations (35) and (36) re-arrange to

$$\begin{aligned} d^2 &= \sigma^2 c(1/\delta - c) \\ \frac{\sigma^2}{\delta} + r^2\delta &= \sigma^2 c^2 + d^2 \end{aligned} \tag{38}$$

Obviously the signs of δ and r are indeterminate, since each loads on a symmetric noise term. But this generically has a continuum of solutions that can be indexed by values of $d > 0$ that solve:

$$r^2 = \frac{\sigma^2}{\delta^2}(c-1)$$
$$c = \frac{1}{2\delta} + \frac{1}{2}\sqrt{\frac{1}{\delta^2} - 4\frac{d^2}{\sigma^2}} \tag{39}$$

Some comments are in order. First, if d^2/σ^2 is too large, this equilibrium will not exist because there will not exist a $c > 1$ that solves the second equation. The precise condition for existence is

$$\begin{cases} d^2/\sigma^2 \leq 1/(4\delta^2) & \text{if } 1/(2\delta) \geq 1 \\ d^2/\sigma^2 \leq 1/\delta - 1 & \text{if } 1/(2\delta) < 1 \end{cases}$$

Second, if we impose $r = 0$, it is simple to show that the equilibria indexed by d features $c = 1$and $\delta = (d^2/\sigma^2 + 1)^{-1}$.

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 $c(\delta)$ 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

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

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

$$B'(\delta_i) = \frac{1}{\delta^2} + \frac{\sigma^2 c(\delta)}{\delta} - \frac{\delta_i}{\delta^2}\left(1 + \sigma^2\right)$$

which, evaluated at the fixed-point condition, is

$$B'(\delta_i)|_{\delta_i=\delta} = \delta^{-1}(\delta^{-1} - 1) + \delta \cdot \sigma^2\delta^{-2}(c(\delta) - 1) = \delta^{-1}(\delta^{-1} - 1) + \delta r^2 \tag{40}$$

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 $d = 0$. In this case, $c(\delta) = 1/\delta$ from solving (39). The marginal benefits curve when $d = 0$ 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 (40). 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 ν. Consider now the case in which the agent pays to track the vector (p, ν). Assume that the cognitive cost has the following representation

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

for some $K \in [0, 1]$. 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 (7), we obtain

$$C = C(a, b, r; c, d, K) = -\log\left(1 - \delta_p(a, b, r; c, d)\right) - K \log\left(\frac{r^2}{r^2 + \sigma^2 a^2 + b^2}\right)$$

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

Fix (c, d) and consider how an individual decides to construct her signal. The capacity constraint is $C(a, b, r; c, d) \leq M$, for some constant M. This can be re-written as

$$\delta_p \leq 1 - e^M \left(\frac{r^2 + a^2\sigma^2 + b^2}{r^2}\right)^K \tag{41}$$

The agent can freely pick (δ_p, r, a, b) subject to 35 and 36. For any $K > 0$, picking $a \neq 0$ or $b \neq 0$ 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 $a = b = 0$. By market clearing, $d = 0$, 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, a, b) to choose, subject to 35, 36 and (41), and that internalizes the market clearing conditions. Take now the limit as $K \downarrow 0$. The marginal effect on δ_p of increasing b^2 and a^2, while decreasing r^2, is approximately zero around the point $a^2 = b^2 = 0$. 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 $b = 0$, social welfare is increasing in b^2 and hence the optimal b^2 is greater than 0. For K positive but small enough, the unique equilibrium is therefore dominated by an allocation with $b < 0$ and $d = -b > 0$.

B Proofs for State-Tracking Economies

Proof of Theorem 1

Proof. 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 $((\phi'^j)_{j=1}^J, \phi') \in \Phi^{J+1}$ 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 the case that $(x'^j(\omega), \phi'^j) \notin \mathbf{B}(p(\theta), e^j(\theta), a^j\Pi(\theta))$. Recall that the latter set embeds both the budget constraint of household j 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, and the aforementioned finite mass deviation property, 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 $x^j(\theta) := \sum_\omega x^j(\omega)\phi^j(\omega \mid \theta)$ in the first allocation and $x'^j(\theta) := \sum_\omega x'^j(\omega)\phi'^j(\omega \mid \theta)$ in the proposed better one. Let aggregate supply similarly be $y(\theta) := \sum_\omega y(\omega)\phi(\omega \mid \theta)$. 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

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

$$x_n(\omega) = x_n(\omega') \text{ if } \omega_n = \omega'_n \Bigg\}$$

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:

$$\begin{gathered} U^j(\bar{x}(\cdot)) \equiv \max_{x(\cdot),\phi(\cdot)} \sum_{\omega,\theta} u^j(x(\cdot),\theta)\phi(\omega,\theta) - C^j[\phi(\cdot)] \\ \text{s.t. } (x(\cdot),\phi(\cdot)) \in \mathbf{G}(\bar{x}(\cdot)) \end{gathered} \tag{42}$$

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

$$\mathbf{Y} := \left\{ \bar{y}(\cdot) : \exists\ [(\phi(\cdot), y(\cdot)) \in \mathbf{F}] \text{ s.t. } \sum_{\omega\in\Omega} y(\omega)\phi(\omega \mid \theta) = y(\theta) \right\} \tag{43}$$

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 on C invoked in Theorem 2 suffices for the "outer" preferences and technologies defined above to be convex. This is formalized in the following:

Proposition 8. *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 $ax(\theta) + (1-a)x'(\theta) \succ^{j,Out} x'(\theta)$ 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 (42) 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, $-C[\phi(\tilde{\omega},\theta)] \geq -aC[\phi^0(\omega,\theta)] - (1-a)C[\phi^1(\omega,\theta)]$.

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. $\square$

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).[20]

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 $\bar{x}^j(\theta) := \sum_{\omega\in\Omega} x^j(\omega)\phi^j(\omega \mid \theta)$ and aggregate supply as $\bar{y}(\theta) := \sum_\omega y(\omega)\phi(\omega \mid \theta)$. From Theorem 2 in Debreu (1954), there exists a linear functional $v(\bar{x}(\theta)) := \sum_\theta \lambda(\theta)\cdot\bar{x}(\theta)$ 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 (42), and that the proposed firm allocations $(y(\cdot), \phi(\cdot))$ are (possibly non-unique) solutions to the inner problem (43). 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

Proof. 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 **??** 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. $\square$

Proofs of Theorems 4 and 5

See main text.

[20] 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 Lemma 1

See main text.

Proof of Corollary 3

Proof. 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 a have the $(p(\theta), \theta^j)$ sufficient statistic property in signals (a subset of the feasible set) by the price-tracking criterion. By Lemma 1, a sufficient condition for a bundle to solve the θ-consumer problem is that it weakly dominates all bundles with signals that satisfy the property by the state-tracking criterion.

Generically, we cannot compare bundles across the two cases with two different utility functions (inclusive of costs). But Assumption **??** ensures that all bundles with signals for which $(p(\theta), \theta^j)$ is sufficient for θ. have the same *numerical* utility value when evaluated by either criterion.

Thus, we can conclude that $(x^j(\cdot), \phi^j(\cdot))$ has weakly higher utility than any feasible bundle with a $(p(\theta), \theta^j)$-coarse signal by the state-tracking criterion. 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 **??**) *a fortiori* implies optimality in the price-only-tracking economy. Markets clear and all agents optimize; thus this is a price-only-tracking equilibrium. □