Acquisition, (Mis)use and Dissemination of Information: The Blessing of Cursedness and Transparency[*]

Franz Ostrizek[†] Elia Sartori[‡]

Wednesday 18th October, 2023

Abstract

Adapting cursed equilibrium to a beauty contest game, we study the impact of information policies in settings where agents underinfer from equilibrium statistics. To discipline information acquisition with mislearning, we propose a *subjective envelope condition* which allows for a tractable analysis while maintaining behaviorally plausible assumptions: agents correctly anticipate their actions but incorrectly deem them optimal. We show that this condition characterizes the rest points of a simple learning process.

Cursed agents use and acquire more private information, creating a positive externality. Welfare increases for low degrees of cursedness, as these gains exceed the losses from incorrect use. Transparency crowds out private information but always increases welfare. Policies targeting fundamental information may backfire as they distract cursed agents from a source of information they already underuse. Finally, we investigate the behavior and welfare of an atomistic rational agent in a cursed economy.

[*]We thank Roland Bénabou, Mira Frick, Mathijs Janssen, Stephen Morris, Marco Pagnozzi, Antonio Rosato, Sebastian Schweighofer-Kodritsch, Federico Vaccari and audiences at University of Copenhagen, LUISS, Humboldt University Berlin, AMES 2021, the NOeG Annual Meeting 2021, the 8th Young Researchers Workshop of the CRC TR 224, EEA Congress 2021, GRASS XV, the VfS Annual Conference 2021, and the SEA Annual Meeting 2022 for helpful comments and discussions. We also thank Nicolas Fajardo and Ruichao Pan for excellent research assistance. Ostrizek: Funding by the Deutsche Forschungsgemeinschaft (DFG) through CRC TR 224 (Project B02) is gratefully acknowledged.

[†]Sciences Po, Department of Economics; franz.ostrizek@gmail.com

[‡]CSEF; elia.sartori@unina.it

1 Introduction

Many economic decisions are taken in environments with interdependent payoffs and uncertainty both about fundamental states and the actions of others. To guide such decisions, players rely on information that includes direct sources about the fundamental as well as signals about aggregate statistics of the actions of others. Such statistics ("aggregative signals") arise naturally in many economic settings, ranging from the transaction price in a financial market (Grossman and Stiglitz, 1980; Diamond and Verrecchia, 1981; Kyle, 1985) and the level of activity and number of infections during a pandemic, to inflation statistics (Lucas, 1972; Morris and Shin, 2005).

Aggregative signals do not just contain information about the actions of others but also disseminate information about the fundamental. The amount of fundamental information conveyed by such statistics depends both on their precision around the true aggregate moment —for which we adopt the moniker of transparency— and on the informativeness of the moment itself. This informativeness, in turn, depends on how much private information the players use and disseminate through their strategies and ultimately on how much information they possess to begin with.[1] These intricate interactions are at the center of lively debates both in policy circles and in the academic literature, especially surrounding policy instruments that directly target the very availability and transmission of information.

One key though hitherto neglected aspect of these interactions is that extracting fundamental information from an equilibrium statistic is no easy feat. In particular, it requires understanding how the actions of others reflect their private information. Ample evidence (e.g. on the winner's curse and underinference in social learning) indicates that agents often fail to take this inferential step.[2] A natural implication is that agents will also underuse aggregative statistics and place an excessive weight on private information. This observation is empirically supported in several contexts. Inflation expectations strongly depend on noisy private signals (individual supermarket prices) even when precise aggregative information is provided (Cavallo et al., 2017); traders in an experimental financial market fail to incorporate information from prices (Ngangoué and Weizsäcker, 2021). This evidence suggests that a complete analysis of such environments should take into account both the endogenous dissemination of

[1] Deviations from full transparency may arise, e.g., from measurement error, intentional coarsening of information, or delays in reporting. For example, a fully transparent financial market would be one where a trader knows the transaction price before submitting his order. A lower level of transparency would correspond to a market where traders have only noisy information about the transaction price, say they observe the current price at another similar market place or in the past.

[2] In the winner's curse, agents fail to appreciate that they are more likely to win a common value auction when their private information leads them to overestimate the value of the prize. Implicitly, they dismiss the information contained in the fact that they won the auction. See Kagel and Levin (2002) for a review of the experimental evidence. For social learning, see Weizsäcker (2010).

information as well agents' limited grasp of the information contained in aggregative signals.

Towards this end we study a beauty contest game with information acquisition and adapt cursed equilibrium (Eyster and Rabin, 2005) as a tractable model of incomplete inference from aggregative signals. The beauty contest and its generalizations are a stylized yet rich laboratory used to study dispersed information across a range of applications (Hellwig and Veldkamp, 2009; Vives, 2017; Angeletos and Lian, 2018). Cursed equilibrium provides a parsimonious solution concept capturing the range from rational to fully cursed agents who fail to take into account that the actions of other players are a result of their private information and hence consider the aggregate outcomes to be uninformative about fundamentals. It has been used successfully to account for overbidding in common value auctions and has also found applications to financial markets (Eyster et al., 2019). We introduce cursed equilibrium in a simple beauty contest game in which agents target a combination of the state and the average action. Simplifying the strategic environment (relative to, e.g., Angeletos and Pavan (2007) and many specific applications) allows us to focus on the novel interaction between cursedness and transparency.

To evaluate the effects of transparency and other information policies we need a model which includes all the relevant adjustment margins. Those margins encompass both information use and information acquisition: Increased transparency, for instance, may backfire because it crowds out private information acquisition and thereby undermines the very source of information in the aggregate moment. To capture such adjustments, our beauty contest game features an information acquisition stage wherein agents choose the precision of their private information at a cost (as in Colombo et al., 2014). To even represent this decision, one has to describe how agents assess the value of information ex-ante. While this is conceptually straightforward for rational agents, it becomes tricky when dealing with cursed agents. Specifically, how do agents perceive the impact of altering the same information environment that they misperceive when acting within it? To what extent are they aware of their misuse, and how does such awareness affect information acquisition? Our modeling challenge is to address these issues with behaviorally plausible assumptions that yield a tractable analysis of the role of transparency in markets with cursed agents.

We propose a notion — *cursed expectations equilibrium with information acquisition* — based on the following behavioral desiderata: Agents correctly anticipate the equilibrium relationships as well as their expected welfare and play, but they do not consider their future information use to be erroneous. They are, therefore, not meta-rational and do not use information acquisition to fix their wrong use. To operationalize these desiderata, we assume that the amount of private information chosen by agents satisfies a *subjective envelope* condition. We provide a further foundation for the subjective envelope condition by grounding it in a natural learning process whose rest point corresponds to its solution(Theorem 3). Throughout the process, agents choose a target level of information acquisition, which is implemented with trembles. Agents do not adjust their actions to the tremble, but record its realization together with the realization of welfare; this data set indicates a direction for improvement for all targets that fail the subjective envelope condition. The learning process also clarifies that cursedness is the only cognitive limitation of our agents, but that it manifests itself in both stages of the game: at the action stage, through the misuse implied by the cursed updating rule; at the acquisition stage, through the conviction that the cursed updating rule is the correct updating rule. The subjective envelope condition is a representation of this dual nature of cursedness.

Having developed this theoretical tool, therefore, we can address our original question of the effects of transparency in markets where agents under-appreciate the information content of aggregative statistics. We now preview these results. The equilibrium is characterized by a vector of loadings on the different sources of information. As agents become more cursed, they substitute away from the aggregative signal and increase the use of private information. This is because cursedness makes them perceive the aggregative signal as less informative, so they need to rely more on their remaining information sources. The subjective envelope condition implies that the use of private information is a sufficient statistic for its acquisition; in particular, the comparative statics of information acquisition and information use coincide. Moreover, the acquisition channel creates a feedback loop, as the precision of private information is a function of its perceived value: using and acquiring more information, cursed agents disseminate it more effectively.

Cursedness and transparency have opposed impacts on the equilibrium loadings: For any degree of cursedness, an increase in transparency makes agents substitute from private information towards the aggregate signal but crowds out private information use and acquisition. The crowding out effect, however, never overturns the direct positive effect and the aggregative signal becomes more informative about the state as transparency increases. This monotonicity does not necessarily hold for other measures of informational efficiency, such as the total precision of information available to agents or the realized covariance between the aggregate action and the state. Along the latter metric, cursedness increases the inflow of private information into the aggregative signal but hinders its extraction, thereby reducing the efficiency of dissemination. With information acquisition, these forces balance exactly. Contrary to the intuition that inferential naivety hampers information aggregation and to the result with exogenous private information in Eyster et al. (2019), the covariance between the aggregate action and the state is independent of cursedness and transparency. While cursed agents extract less from the aggregative signal, they also inject more private information into it.

The tractability of our framework allows an exhaustive analysis of the welfare consequences of cursedness and of information policies. In the rational equilibrium, the use and acquisition of private information is inefficiently low because of an information dissemination externality. If agents are cursed enough, however, they may use (and acquire) at or even above the efficient level, though they simultaneously misuse their signals. Indeed, welfare is nonmonotonic in the degree of cursedness. Local to rationality, cursedness is bliss: An increase in cursedness causes a (first-order) welfare gain from improved dissemination that dominates the (second-order) welfare loss from privately suboptimal use. As cursedness grows further, agents are increasingly unable to reap the gain from increased dissemination and significantly misuse their information; eventually, cursedness reduces equilibrium welfare.

Lower information acquisition costs and more public information have an ambiguous effect. While both increase welfare in the rational benchmark for our payoff specification (Bayona, 2018; Colombo et al., 2014), they can reduce welfare with cursed agents: They cause agents to substitute away from the aggregative signal and towards fundamental information (private or public, resp.), exacerbating suboptimal information use when agents are cursed. We show that this effect can dominate the direct effect of cheaper and more precise information, causing the paradoxical comparative statics. More transparency, by contrast, unambiguously improves the dissemination of information and does not exacerbate cursed agents' misuse. Hence, transparency is the only policy instrument that increases welfare across the whole parameter space – our second blessing. Somewhat paradoxically, therefore, only increasing the precision of the very source of information that cursed agents underestimate makes them systematically better off.

Although their welfare increases with transparency, their inability to correctly process all available information implies that cursed agents fail to reap its full benefits. A natural question that arises is how an agent who is able to extract all the information from the environment—such as proverbial smart money in a financial market— would behave in an environment with cursed agents. Does he benefit from interacting with less rational agents? How is he impacted by information policies? We address these questions by studying the behavior of an atomistic rational agent facing equilibrium play in a economy of cursed agents. Such a shrewd agent benefits from the large amount of information disseminated by the cursed crowd, sometimes even abstaining from acquiring private information. However, the shrewd agent is also harmed by their misuse of information, as strategic complementarity forces him to follow the crowd and distort his actions away from the fundamental. Compared to the rational environment, low levels of cursedness are always beneficial for the shrewd agent (whose welfare can even exceed first best). At high levels of cursedness the imitation effect dominates in games with sufficiently strong strategic complements, making excessive cursedness harmful even for the shrewd agent. The trade-off between information free riding and miscoordination creates nontrivial comparative statics in the policy parameters. A shrewd agent always profits from transparency but can be hurt by more public information and lower information acquisition costs, even when they are beneficial for the cursed crowd.

We conclude the introductory section by discussing the related literature. In Section 2 we present the model and establish existence and uniqueness of a cursed equilibrium, taking the precision of private information as given. We introduce information acquisition in Section 3: We define the notion of cursed expectations equilibrium with information acquisition (Subsection 3.1), discuss the behavioral principles behind the notion, provide a learning foundation and compare it to alternative notions (Subsection 3.2). We analyze the positive comparative statics of the model in Section 4. We turn to welfare analysis in Section 5. We analyze the behavior and welfare of an atomistic rational agent in Section 6. Section 7 concludes.

1.1 Related Literature

The impact of information policies, be they about the fundamentals of the economy or about endogenous statistics, is at the heart of several research areas of applied economics. Conducting our analysis of these questions within the workhorse class of linear quadratic models, we connect to a rich theoretical and applied literature ranging from business cycles (e.g. Hellwig and Veldkamp, 2009; Angeletos and La'O, 2010; Angeletos and Lian, 2018) and demand function competition (e.g. Vives, 1988, 2017), to political economy (e.g. Shadmehr et al., 2022). Angeletos and Pavan (2007) characterize the inefficiencies of information use in a general linear-quadratic Gaussian game. Ui and Yoshizawa (2015) classify such games according to the welfare properties of additional public and private information. Colombo et al. (2014) study how private information acquisition affects the value of information. This literature generally considers only signals of the fundamental. We analyze the value of information in the presence of a signal of the average action, providing information of endogenous precision about the state. Bayona (2018) considers an information structure with such a signal in a setting akin to Angeletos and Pavan (2007), establishing that this can lead to a dissemination inefficiency in the use of private information.[3] We focus on the role of agents' limited understanding of aggregative information for the social value

[3]Our rational benchmark nests a payoff restriction of Colombo et al. (2014), Bayona (2018), and their so far unexplored meet and establishes that the dissemination inefficiency goes hand in hand with inefficiently low information acquisition. Amador and Weill (2012) show that with such a dissemination externality more public information can cause a decrease in welfare, even without interdependent payoffs. Hebert and La'O (forthcoming) study a model with flexible information acquisition, studying which cost functions lead to efficiency and nonfundamental volatility, and show that such an externality arises

of information and transparency. Our restriction of the payoff structure to the simple beauty contest game allows us to isolate the novel sources of inefficiency in our setting.

The results of Morris and Shin (2002), who show that more precise public information can reduce welfare, have spurred extensive debate in the literature about the desirability of public information in particular in the context of central bank announcements.[4] This discussion has often been couched in the terminology of "anti-transparency" vs. "pro-transparency". This label does not correspond to our usage, as we reserve the word transparency for the precision of the public signal about the aggregate action.[5] Although we would argue that much of the information provided by central banks is aggregative in nature and explore the impact of such transparency at length, we also contribute to the original debate by demonstrating a novel channel based on cursed inference which can render public *fundamental* information undesirable: It distracts behavioral agents from other information sources whose information content they underestimate. The issue of endogenous information dissemination has been studied in the context of business cycles by Wong (2008) who show that increased transparency can be self-defeating as it reduces the information available to the central bank itself to learn about the state of the economy, a mechanism that has also been studied in Morris and Shin (2005).

Inference from a signal that aggregates information contained in individual best responses is also at the center of the literature on information aggregation in financial markets.[6] Grossman and Stiglitz (1980) show that the equilibrium informativeness of the price system is unresponsive to changes in transparency: an increase in noise leads to more information acquisition which exactly offsets the direct effect. We establish a similar invariance along a metric of informational efficiency in our setting, but show that transparency has an impact on the total precision available to (rational) agents. Updating biases have received considerable attention in the literature on behavioral

whenever the ease of acquiring information through the aggregate action depends on its responsiveness to the state, as is the case in our setting with additive noise.

[4] Svensson (2006), e.g., argues that the ratio of private to public precision required for the paradoxical welfare result is unreasonably high and Woodford (2005) calls into question the assumptions on strategic complementarity and welfare. The role of these assumptions is clarified and general conditions for such welfare results are given in Angeletos and Pavan (2007) and Ui and Yoshizawa (2015). Kool et al. (2011) show that public information can reduce information acquisition by market participants and thereby increase financial market volatility. Amador and Weill (2010) show that through a signal jamming channel public information can be welfare decreasing, as it reduces the informativeness of the price system thereby increasing uncertainty about the monetary shock.

[5] In the financial economics literature, enhanced transparency is sometimes conceptualized as the sharing of private signals between asymmetrically informed traders, e.g. Glosten and Milgrom (1985); Chowdhry and Nanda (1991). Pagano and Röell (1996) define transparency as the extent to which market makers can observe the size and direction of the current order flow, a notion that is much closer to that we use in this paper. They find that greater transparency generates lower trading costs for uninformed traders on average, although not necessarily for every size of trade.

[6] Angeletos and Sastry (2023) study flexible information acquisition, in particular about prices, in an Arrow-Debreu economy and characterize when relying on the price system for information transmission is efficient.

finance, see Barberis (2018) for recent survey. While this literature focuses on the time series properties of prices and returns, we focus on the (social) value of information.

Underinference from the actions of others has been documented across many settings in the lab and in the field, see Eyster (2019) for a survey of related evidence. Cursed equilibrium was proposed by Eyster and Rabin (2005) as a model such underinference. Eyster et al. (2019) apply a cursed analogue to rational expectations equilibrium in a trading game and show that cursed behavior can explain excessive trade volume. We adapt cursed equilibrium to a beauty contest game with endogenous aggregative and private information and devise a novel notion of information acquisition to study the impact of underinference on welfare and the social value of information. Cohen and Li (2023) and Fong et al. (2023) extend cursed equilibrium to extensive form games. A version of our game can be analyzed along the lines of Cohen and Li (2023), which however imposes independence among the conjectured actions of others which is a strong restriction in our continuum setting (see footnote 18 for a detailed discussion). An alternative equilibrium concept modeling a failure to account for the information content of others' action was proposed by Esponda (2008) in the spirit of self-confirming equilibrium. While this approach bears similarity to our learning foundation for the subjective envelope condition, the nature of learning differs: In Esponda (2008) (and more generally in Berk-Nash equilibrium (Esponda and Pouzo, 2016)) agents learn their structural beliefs to which they best-respond; our learning foundation instead directly adjusts information acquisition in the spirit of gradient ascent. We return to the implications of behavioral equilibrium in the beauty contest game as well as the differences in the learning approach in Section 3.2.

Our approach contrasts with the literature on misspecified models, which focuses on the asymptotic properties of posterior beliefs of misspecified Bayesian agents (see e.g. Heidhues et al. (2018); Bohren and Hauser (2021); Frick et al. (2023)). We focus on a specific updating bias and study active learning by non-Bayesian agents in a single period. Bohren and Hauser (2023) study the relationship between misspecified models and updating biases such as cursedness. We propose a notion of the value of information for non-Bayesian agents which does not rely on such a quasi-Bayesian representation but is instead founded on behavioral principles that are operationalized by a learning process.

2 The Model with Exogenous Private Information

The game has two stages: First, agents choose how much private information to acquire. Second, agents play a beauty contest game. We begin our description of the setting by presenting the game with *exogenous* precision of private information in this section. We then introduce information acquisition in Section 3.

2.1 Actions and Payoff

There is a unit interval of agents $i \in [0,1]$, playing a simple beauty contest game. Their payoff is given by

$$u(a_i, \bar{a}, \theta) = -(1-r)(a_i - \theta)^2 - r(a_i - \bar{a})^2, \tag{1}$$

where $a_i \in \mathbb{R}$ is the action of player i, $\bar{a} = \int_i a_i \, di$ is the average action[7] and $\theta \in \mathbb{R}$ is the state (or fundamental). We allow for both strategic complementarity ($r > 0$), and substitutability ($r < 0$), and assume that complementarity is not too strong ($r < 1$) to ensure the existence of a unique interior linear equilibrium and a planner solution.

The restriction to a simple beauty contest allows us to isolate the inefficiencies generated by the features specific to our information environment: the dissemination externality of aggregative information, and cursed updating from that source. Indeed, in our simple beauty contest game both information use and acquisition are efficient for the rational benchmark without aggregative information (Angeletos and Pavan, 2007; Colombo et al., 2014).[8]

2.2 Information, Best Response and Inference

The following information structure is common knowledge. The state θ is drawn from the prior distribution $\mathcal{N}(0, \tau_\theta^{-1})$.[9] Agents receive three signals: a *private fundamental* signal $s_i = \theta + z_{s_i} \sim \mathcal{N}(\theta, \tau_s^{-1})$, i.i.d. across agents with a precision τ_s that we will endogenize in Section 3; a *public fundamental* signal $y = \theta + z_y \sim \mathcal{N}(\theta, \tau_y^{-1})$ about the state; and a *public aggregative* signal $p = \bar{a} + z_p \sim \mathcal{N}(\bar{a}, \tau_p^{-1})$, where τ_p denotes the precision of the aggregative signal as a signal of $\bar{a}$, is our transparency parameter.[10]

Note that the timing of information implies that our game is not a Bayesian game, as agents actions are allowed to react to the realization of p which in turn is a function of their actions. We instead think of our setup in the spirit of rational expectations equilibrium. There are two foundations for this approach. First, we can think of the

[7] As is customary, we adopt a law of large numbers for the private signals as a convention, see Vives (2008, 10.3.1) for a discussion. One formal operationalization of this is to view the integrals in the sense of Pettis, see Uhlig (1996).

[8] This contrasts with the specification of the beauty contest in Morris and Shin (2002) who consider the utility function $u(a_i, \bar{a}, \theta) = -(1-r)(a_i - \theta)^2 - r(a_i - a_j)^2$, which results in a dependence of individual utility on the variance of others' actions.

[9] A prior mean of zero is merely a convenient normalization. We insist on a proper prior as we analyze the comparative statics of ex-ante welfare.

[10] The situation we have in mind is a central authority having exclusive access to the actions chosen by each player inside a market. With those data, it can perform statistical analysis (which is noisy because of missing data, imperfect reporting, etc.) and produce a report which will then be observed without further noise by everyone. Interpreted as the accuracy of the process turning actions into a report, transparency becomes a natural parameter for positive comparative statics as well as policy evaluation. More generally, τ_p represents the access agents have to information about others when they make their decisions and the degree to which they can condition their actions on an aggregate outcome, e.g. the price when submitting orders in a financial market.

model as a reduced form of a dynamic game in which the state is fixed or evolves only slowly. Second and more formally, we can consider agents who submit action schedules $a_i : p \mapsto a \in \mathbb{R}$ (see Vives, 2014). This formulation is equivalent to our equilibrium notion for rational and (partially) cursed agents and it depends on the application which seems more natural. In the context of (financial) markets, it is common to consider models of demand/supply function competition, while acting based on a realized signal seems more natural for individual consumers or workers reacting to the inflation rate.

The optimal action is given by

$$a_i(s_i, y, p) = \arg\max_{a_i} \mathbb{E}_i\left[u(a_i, \bar{a}, \theta)\right] \tag{2}$$

where $\mathbb{E}_i$ is the expectation operator with respect to agent i's information, including his updating biases. As u is quadratic, (2) takes the linear best response form

$$a_i = (1-r)\,\mathbb{E}_i(\theta) + r\mathbb{E}_i(\bar{a}) \tag{3}$$

Throughout, we focus on linear equilibria. That is, following the structure of the best response and posterior beliefs, we conjecture that the optimal action rule is a linear combination of the signals

$$a_i = \alpha_0 + \alpha_1 s_i + \alpha_2 y + \alpha_3 p \tag{4}$$

for some vector of loadings $\boldsymbol{\alpha}$. Then, we can write the true aggregate action as

$$\bar{a} = \int_0^1 a_i \,\mathrm{d}i = \delta_0 + \delta_1\theta + \delta_2 y + \delta_3 p \tag{5}$$

with aggregate weights $\boldsymbol{\delta}$. Inspection of equation (5) makes clear that the aggregative signal $p = \bar{a} + z_p$ provides information of *endogenous precision* about θ. Indeed, under the assumption that $\delta_1 \neq 0$, $\delta_3 \neq 1$ (and conditionally on y), p is informationally equivalent to

$$\hat{p} = \frac{1-\delta_3}{\delta_1}\left[p - \frac{\delta_2}{1-\delta_3}y\right] - \frac{\delta_0}{\delta_1} = \theta + \frac{1}{\delta_1}z_p \sim \mathcal{N}\left(\theta, \frac{1}{\delta_1^2\tau_p}\right) \tag{6}$$

The Bayesian posterior on θ can be written based on the three conditionally independent sources $(s, y, \hat{p})$ which determines the posterior on $\bar{a}$ through (5). The precision of the aggregative signal about the state, $\delta_1^2\tau_p$, depends both on transparency τ_p and on the equilibrium loading δ_1.

2.3 Cursed Equilibrium: Definition

As a model of the failure to update from observing the action of others, we adapt cursed equilibrium (Eyster and Rabin, 2005). In this solution concept, agents are characterized by a parameter χ, the degree of cursedness, that ranges from $\chi = 0$ for rational benchmark to $\chi = 1$ denoting fully cursed behavior. A fully cursed agent fails to perceive any correlation between other agents' actions and their private information. Instead, he thinks that others play according to the marginal distribution of their actions conditional on his private information. Consequently, according to the beliefs of a *fully cursed* agent i with information I_i, the action of agent j is

$$a_j = \mathbb{E}[a_j|\mathrm{I}_i] + \alpha_1\left(s_j - \mathbb{E}[\theta|\mathrm{I}_i]\right) = \alpha_0 + \alpha_1\mathbb{E}[\theta|\mathrm{I}_i] + \alpha_2 y + \alpha_3 p + \alpha_1\left(s_j - \mathbb{E}[\theta|\mathrm{I}_i]\right) \tag{7}$$

where the α_k are the weights used in the linear strategy of player j. The fully cursed agent treats the prediction error $s_j - \mathbb{E}[\theta|\mathrm{I}_i]$ as *independent* of the state. Therefore, in a linear symmetric equilibrium, fully cursed agents perceive the aggregate action as

$$\bar{a} = \delta_0 + \delta_1\left(\mathbb{E}[\theta|\mathrm{I}_i] + \tilde{z}\right) + \delta_2 y + \delta_3 p \tag{8}$$

where $\tilde{z}$ is a subjective noise term independent of the true state. Consequently the aggregate action is independent of θ conditional on his information. A fortiori, p do not provide additional information about the state.[11]

Partially cursed agents are characterized by an interior level of cursedness $\chi \in (0,1)$. They form expectations as a convex combination of rational and fully cursed ones, namely

$$\mathbb{E}_\chi[\theta|\mathrm{I}_i] = \chi\frac{\tau_y y + \tau_s s_i}{\tau_\theta + \tau_y + \tau_s} + (1-\chi)\frac{\tau_y y + \tau_s s_i + \delta_1^2\tau_p\hat{p}}{\tau_\theta + \tau_y + \tau_s + \delta_1^2\tau_p} \tag{9}$$

$$\mathbb{E}_\chi[\bar{a}|s_i, y, p] = \delta_0 + \delta_1\left(\chi\frac{\tau_y y + \tau_s s_i}{\tau_\theta + \tau_y + \tau_s} + (1-\chi)\frac{\tau_y y + \tau_s s_i + \delta_1^2\tau_p\hat{p}}{\tau_\theta + \tau_y + \tau_s + \delta_1^2\tau_p}\right) + \delta_2 y + \delta_3 p \tag{10}$$

Note that even cursed agents do all the updating about the state θ and then turn it into a belief about $\bar{a}$ through the equilibrium condition (5): they have "equilibrium awareness". Where they go wrong is in the under-appreciation of the correlation between their private information and others' actions.

When we interpret our game as one of submitting action schedules as a function of the aggregative signal, cursedness also captures agents inability to engage in conditional or hypothetical thinking (Esponda and Vespa, 2014; Ngangoué and Weizsäcker, 2021).

[11] In contrast to other updating biases – e.g. overconfidence or dismissiveness –, cursed agents correctly perceive the relative precision of p as a signal about the aggregate action. They fail, however, to relate it to the private information of others and to extract information about the state.

In light of this evidence, we interpret the degree of cursedness not as an individual characteristic but as codetermined by the market structure.

Cursed equilibrium is defined as a solution concept for Bayesian games. Recall that due to the presence of the aggregative signal, however, the model described so far is not a Bayesian game, strictly speaking. We therefore adapt cursed equilibrium in a fashion similar to a linear rational expectations equilibrium:[12]

Definition 1. A vector of loadings $(\boldsymbol{\alpha}, \boldsymbol{\delta})$ constitutes a χ-*cursed expectations equilibrium* if $\boldsymbol{\alpha}$ satisfies the best response condition (3)-(4) with expectations formed according to (9)-(10) given $\boldsymbol{\delta}$; and the aggregate action is consistent with individual actions, $\boldsymbol{\delta} = \boldsymbol{\alpha}$.

As is customary in the literature, we restrict attention to symmetric linear equilibria. Consequently, our claims to existence and uniqueness of equilibria refer to this class throughout.[13]

2.4 Cursed Equilibrium: Characterization

This section studies the equilibrium for fixed τ_s. An equilibrium is computed by matching coefficients in the best-response function (3).

Theorem 1. *There exists a unique χ-cursed equilibrium for any τ_s. The intercept is $\delta_0 = 0$, the loading on private information $\delta_1 \in [0,1)$ is the unique real solution to*

$$\delta_1 = [1 - r + r\delta_1] \frac{\tau_s}{\tau_\theta + \tau_y + \tau_s + \delta_1^2 \tau_p} \left(1 + \chi \frac{\delta_1^2 \tau_p}{\tau_\theta + \tau_y + \tau_s}\right), \tag{11}$$

and the loadings on the public sources of information are given by

$$\delta_2 = \frac{\delta_1^2 \tau_y}{(1-r)\tau_s - \delta_1 \left(\tau_\theta + \tau_y\right)}, \delta_3 = 1 - \frac{\delta_1 (1-r)\tau_s}{(1-r)\tau_s - \delta_1 \left(\tau_\theta + \tau_y\right)} \tag{12}$$

In equation (11) the RHS has a natural interpretation as it denotes the optimal loading on private information given aggregate δ_1. First, the private signal is valuable for predicting the state, with best-response weight $1 - r$, as well as the aggregate action to the degree that it reflects the state (conditional on public signals), with best-response weight $r\delta_1$. Second, the relative precision of the private signal is the usual Bayesian weight $\frac{\tau_s}{\tau_\theta + \tau_y + \tau_s + \delta_1^2 \tau_p}$. Hence, the private signal is ignored ($\delta_1 = 0$) only if it is pure noise ($\tau_s = 0$). This term also contains the information spillover effect: the more other agents use their private information (higher δ_1), the more can be learned from the

[12] See also Eyster et al. (2019) for a similar approach in a trading game with finitely many agents.

[13] Morris and Shin (2002) show that the only equilibrium is linear in a setting without aggregative information.

aggregative signal, which reduces the weight on the private signal. Third, the final term adjusts this weight as cursed agents fail to understand that the aggregative signal is informative about the state and therefore perceives the private signal to be *relatively* more informative. In the extreme case of $\chi = 1$, the agent ignores the aggregative signal and the two final factors simplify to $\frac{\tau_s}{\tau_\theta+\tau_y+\tau_s}$, the relative precision of the private signal as if there were no aggregative information. Therefore, transparency is without effect in the fully cursed equilibrium while, symmetrically, cursedness has an impact on the equilibrium only if there is an informative aggregative signal ($\tau_p > 0$). Cursedness manifests itself as a pure updating bias and only distorts inference from the aggregative signal. Absent such signal, cursed agents act just like a rational agent as they correctly interpret all fundamental sources of information.

At first sight, that cursedness matters only in the presence of an aggregative signal may be surprising when compared with the implications of cursed equilibrium in a common value auction. In the auction, there is no aggregative information available to the agent before he chooses his action but still cursedness impacts his choice. This, however, is a natural consequence of the payoff structure: In an auction, the agent considers his payoff conditional on winning the auction, which is exactly such an aggregative conditioning event. In our model, the payoffs themselves weigh all states equally ex-ante and there is no such "implicit conditioning" embedded in them.

The fully rational case is easily obtained from Theorem 1 but doesn't lead to a simple and immediately interpretable representation. It was analyzed in depth in Bayona (2018). The fully cursed case, by contrast, results in a considerable simplification.

Corollary (**Fully Cursed Equilibrium** δ^{FC})**.** *The equilibrium with* $\chi = 1$ *has*

$$\delta_1^{\mathrm{FC}} = \frac{(1-r)\tau_s}{\tau_\theta + \tau_y + (1-r)\tau_s}, \quad \delta_2^{\mathrm{FC}} = \frac{\tau_y}{\tau_\theta + \tau_y + (1-r)\tau_s}, \quad \delta_3^{\mathrm{FC}} = 0 \tag{13}$$

The role of strategic substitutability and complementarity is directly apparent in the fully cursed equilibrium. If there are no such strategic interactions, cursed agents weigh the two signals at their (mental) disposal according to their precision. Strategic complementarity shifts weight away from the private signal s_i and towards the public signal, y, while substitutability has the opposite effect.

The fact that the fully cursed equilibrium puts no weight on the aggregative signal deserves a clarification. This does not follow from cursedness alone. Indeed, even for fully cursed agents, the aggregative signal, p, remains a valid source of the public fundamental signal, y, and of public noise, z_p. As those are relevant for coordination purposes, agents want to incorporate p into their best response as long as others do so. However, as they can directly condition on the public signal instead which provides information about the state in addition to correlated noise, they always want to put a lower weight on p than others. Hence, we only obtain $\delta_3^{FC} = 0$ as the consequence of an unraveling argument set in motion by the interplay between equilibrium and cursedness.[14]

3 Information Acquisition

In the first stage, agents simultaneously choose the precision of their private signal, τ_s, at cost $c\tau_s$. To study this decision, it is necessary to derive a representation for the agents' perceived ex-ante welfare as a function of τ_s. If our agents were Bayesian the natural modeling choice for this value would be to take the subjective ex-ante expectation of welfare according to the prior distribution. For cursed agents, however, deriving ex-ante welfare presents a modeling challenge because interim beliefs, despite being well-definite, are not derived from a prior joint distribution.

In this section we propose a notion, *cursed expectations equilibrium with information acquisition*, to overcome this challenge. In this notion, agents evaluate the impact of private information on their welfare according to the true joint distribution of state and actions, where their cursed use of information is held fixed in the spirit of the envelope theorem. We first give a formal definition and show that such an equilibrium always exists in Subsection 3.1. In Subsection 3.2 we then discuss the notion, grounding it in three behavioral principles, provide a learning foundation based on those principles, and relate its implicit assumptions to those made by alternative models of information acquisition with incorrect use, including a quasi-Bayesian approach.

3.1 χ-Cursed Expectations Equilibrium with Information Acquisition

The true ex-ante welfare of an agent who acquires precision τ_s, plays according to $\boldsymbol{\alpha}$, and faces an equilibrium $\boldsymbol{\delta}$ is given by

$$W(\boldsymbol{\alpha},\boldsymbol{\delta},\tau_s) = \mathbb{E}\left[-(1-r)(a_i-\theta)^2 - r(a_i-\bar{a})^2\right] - c\tau_s \tag{14}$$

$$= -\frac{\alpha_1^2}{\tau_s} - (1-r)\left(\left[\alpha_2 + \delta_2\alpha_3\frac{1}{1-\delta_3}\right]^2\frac{1}{\tau_y} + \left[\alpha_3\frac{1}{1-\delta_3}\right]^2\frac{1}{\tau_p} + \left(\alpha_1+\alpha_2+\alpha_3\frac{1}{1-\delta_3}\{\delta_1+\delta_2\}-1\right)^2\frac{1}{\tau_\theta}\right) - c\tau_s \tag{15}$$

[14] Vives (2017) considers limited inference, equivalent to fully cursed behavior, in a Linear-Quadratic-Normal model of competition in supply schedules with unknown costs. Fully cursed traders in his setting do not ignore the noisy signal of fundamentals, the price, as it is directly payoff relevant.

The optimal τ_s taking as given and holding fixed both the equilibrium loadings as well continuation play therefore solves

$$\frac{\partial}{\partial \tau_s} W(\boldsymbol{\alpha}, \boldsymbol{\delta}, \tau_s) = 0. \tag{16}$$

Solving this equation, we arrive at the first-order *subjective envelope condition*

$$\frac{\alpha_1^2}{\tau_s^2} = c \tag{SE}$$

According to this condition, the weight on private information in the best response, α_1, is a sufficient statistic for the marginal value of private information, independently of the level of cursedness. Cursedness only affects the calculus through $\alpha_1(\delta_1)$ and the equilibrium.

For a rational agent, this condition simply follows from the envelope theorem.[15] In the cursed case, $\boldsymbol{\alpha}$ does not maximize objective ex-ante welfare and the condition doesn't follow from an envelope theorem. Instead, we include the condition as part of our equilibrium notion to operationalize the behavioral assumption that agents (erroneously) deem their information use to be optimal, which we detail in the following subsection.

Definition 2. A tuple $(\boldsymbol{\alpha}, \boldsymbol{\delta}, \tau_s)$ constitutes a χ*-cursed expectations equilibrium with information acquisition* if $(\boldsymbol{\alpha}, \boldsymbol{\delta})$ constitute a χ-cursed expectations equilibrium given τ_s and (α_1, τ_s) satisfy the subjective envelope condition (SE).

The subjective envelope condition and equilibrium consistency give a linear dependence between τ_s and δ_1

$$\tau_s = \frac{\delta_1}{\sqrt{c}} \tag{17}$$

Taking account of endogenous information acquisition in the equilibrium condition (11), we arrive at

$$\delta_1 = [1 - r + r\delta_1] \frac{\delta_1}{\delta_1 + \sqrt{c}\left(\tau_\theta + \tau_y + \delta_1^2 \tau_p\right)} \left(1 + \chi \frac{\sqrt{c}\delta_1^2 \tau_p}{\delta_1 + \sqrt{c}\left(\tau_\theta + \tau_y\right)}\right) \tag{18}$$

Contrary to the game with exogenous τ_s which always has an interior solution, the existence of an interior equilibrium with information acquisition requires a parametric restriction. This is because we need to ensure not only that agents want to use their

[15] Indeed, the privately optimal level of private information solves a condition of this form both in our setting and in the case without an aggregative signal but a more general payoff structure studied in Colombo et al. (2014).

private information, but that they are also willing to acquire it. More precisely, we require that the best-response weight on private information (RHS) exceeds δ_1 local to $\delta_1 = 0$. This is the case if

$$\sqrt{c} < \frac{1-r}{\tau_\theta + \tau_y} \tag{19}$$

or, equivalently, if the costs of acquiring information are sufficiently small compared to the benefits of private information in the trivial candidate equilibrium. These benefits depend on the precision of prior and public information $\tau_\theta + \tau_y$ and the relative value of public versus private information, as summarized by $1-r$.[16] If this condition is not met, we are stuck in a corner solution with zero information acquisition (and therefore use). Note that (18) always has a trivial solution $\delta_1 = 0$, which describes the equilibrium in that case. We now summarize this discussion.

Theorem 2. *There exists a unique χ-cursed equilibrium with information acquisition. The acquisition of private information is proportional to its use according to (17).*

If $\sqrt{c} < \frac{(1-r)}{\tau_y+\tau_\theta}$, the loading on private information $\delta_1 \in (0,1)$ is the unique interior real solution to (18). Otherwise, we have a corner equilibrium with $\delta_1 = \tau_s = 0$ and only public fundamental information is used ($\delta_2 = \frac{\tau_y}{\tau_\theta+\tau_y}$).

3.2 Foundations of Cursed Information Acquisition

Before delving into the analysis of cursed equilibrium with information acquisition, we discuss in more detail the behavioral assumptions that lead to this notion, ground it in a learning foundation, and contrast it with possible alternatives. Recall that to model information acquisition in a setting with incorrect use, one first has to address how agents perceive their information environment and their actions from an ex-ante perspective. Our approach, which leads to Definition 2, is based on three principles. First, cursedness is the result of a systematic tendency and not of an unexpected mistake: agents correctly anticipate their information use, but they do not consider it erroneous. Therefore, agents should believe that the cursed use of information is (individually) optimal and should not try to fix their bias via information acquisition. Second, when evaluating the returns of private information, an agent should hold his coplayers' acquisition and use fixed at their equilibrium values: there can be no "magical thinking". Third, we require that—at the information acquisition stage—agents conceptualize their true ex-ante welfare. This is in line with a positive interpretation of cursed equilibrium as a representation of behavior rather than as a normatively valid cognitive model. In this interpretation, the role of cursed interim beliefs is only as a foundation of the use of information. Agents do not attempt to forecast (misspecified) interim

[16] If we instead assume convex costs with an Inada-type condition at zero, the analogue of (19) is always satisfied and we have an interior solution.

expectations to guide information acquisition, which are instead more naturally driven by experienced welfare.[17]

Definition 2 satisfies all the principles as agents take the precision of public information as well as the equilibrium loadings as given (second principle), have correct beliefs about their realized equilibrium welfare (third principle), however they do not attempt to use information acquisition to fix their bias (first principle). In χ-cursed expectations equilibrium with information acquisition, agents depart from rationality because they follow the cursed updating rule *and* erroneously deem it optimal. The two mistakes go hand in hand because they originate from the same source — cursedness — and only information acquisition gives the second mistake a stage to manifest itself.

Note that our agents are neither sophisticated nor naive in the traditional sense. They are not sophisticated as they do not anticipate making mistakes or distort their information acquisition to align with the rational updating rule. Conversely, they are not naive as they are fully aware of updating in a cursed fashion. Both these notion are unsuitable for analyzing cursed agents in this setting as cursedness affects them at all stages of the game. It is not realistic to assume that they are surprised by their actions nor that they disavow their future updating from an ex-ante perspective.

Learning Foundation

We now provide a learning foundation that operationalizes the above-described principles and show how they translate directly into the subjective envelope condition (SE). For concreteness, we detail one specific process (whose formal construction is relegated to Appendix A) that allows the agents to estimate the gains from information and update their level of acquisition accordingly. Clearly, many of its specific features are not essential for our results and are assumed for brevity and simplicity.

Consider an agent who has to choose his level of information acquisition and action for an infinite number of periods. Every period, the state and all signals are drawn anew but the parameters of the game and equilibrium δ are fixed. The agent updates his beliefs and acts in a cursed manner throughout, but has to learn the value of information. He does so based on the insights gained from implementation mistakes: in each period, the agent chooses a target level of precision $\bar{\tau}_s$, but the realized precision is $\tau_s = \bar{\tau}_s + \sigma\epsilon$, where $\sigma > 0$ scales a well-behaved zero-mean (and serially independent) implementation error ϵ. We will be particularly interested in the case of small σ.

[17] Indeed, a cursed agent forecasting his distribution over interim beliefs is problematic conceptually, as they would predict that they update their beliefs over the actions of others based on information about the state but understand that they will refuse to update in the opposite way. Furthermore, using interim welfare to guide information acquisition implies magical thinking contra the second desideratum, see later in this section and Appendix C for further discussion.

Each period, the agent uses his signal based on the *action rule associated with the target precision* $\bar{\tau}_s$, not the realized precision τ_s. This can be justified either by the timing of the decision or by costs of reoptimizing this rule: since the agent deems it optimal given the target $\bar{\tau}_s$, reoptimization gains are subjectively second order. At the end of each period, the agent records his *realized welfare* and the implementation error. With this data set at hands, the agent can infer whether positive implementation errors are systematically associated with higher or lower realized welfare and thereby estimate the slope of welfare at a target information acquisition level $\bar{\tau}_s$. If a sufficiently large sample indicates a clear direction for improvement, the agent chooses another target level; otherwise, he sticks with his choice.

This learning process, formally represented by equation (45), has two key features. First, the agent bases his decision on realized welfare. Second, by not adjusting his action to the implementation error, he gets to estimate the partial derivative of the welfare function (14). For a rational agent, who uses his information correctly, fixing the action rule is innocuous by the envelope theorem. In the cursed case, instead, it is the manifestation of an error that we deem intrinsic to cursedness, namely that at the acquisition stage individuals operate under the (erroneous) assumption that their use of information will be correct. These two features are a direct manifestation of the first and third behavioral principle discussed above (the second is satisfied automatically as the agent individually learns in fixed environment); they also constitute the essential properties of the learning process that lead the agent to the subjective envelope condition.

Consider now the long-run behavior of the learning process. Call a target level of precision a *rest point* if, conditional on reaching this level in some period, the agent remains there forever with positive probability. If a level precision satisfies the subjective envelope condition (SE), it is always a rest point of our learning process (45). Conversely, any level of precision that violates the subjective envelope condition will not be a rest point if the implementation error is small enough. Formally,

Theorem 3. *The learning process (45) satisfies*

$$\{\tau : (\text{SE})\ \textit{holds}\} = \bigcap_{\sigma>0} \{\tau : \tau\ \textit{is a rest point of}\ (45)\ \textit{for}\ \sigma\}$$

Theorem 3 establishes our learning foundation, guiding our intuition for the types of settings wherein the subjective envelope condition is most natural. In light of Theorem 3, we can interpret condition (SE) as representing the behavior of agents who actively search for information while using their natural faculties to process it. This approach appears particularly compelling for modeling, e.g., household decisions or retail investors (in financial markets) who instinctively react to information but consciously decide how much attention to pay to the relevant news.

Theorem 3 also clarifies that following the subjective envelope condition does not require high levels of sophistication. Instead, it results from a simple learning process that requires minimal cognitive resources: Agents does not compute subjective expected values and are also reluctant to reoptimize their action rule.

Alternative Models of Information Acquisition

We now discuss alternatives to the subjective envelope condition for endogenizing τ_s.

Quasi-Bayesian (Interim): One approach to modeling information acquisition by agents suffering from an updating bias is to consider them as misspecified Bayesians, the bias being the result of an incorrect prior belief about the joint distribution of states and signals (Bohren and Hauser, 2023). To arrive at this representation, we assume that cursed individuals hold correct beliefs over the distribution of signals and the marginal distribution of the state. A χ-cursed agent then corresponds to a Bayesian who believes that with probability $1-\chi$ the signals will be drawn from the true distribution while with probability χ the signal will be drawn from a distribution that renders p uninformative about the state conditional on s and y.[18] Note that this ex-ante quasi-Bayesian notion coincides with taking the expectation of interim welfare with respect to the correct prior distribution over signals.

We do not follow this approach for three reasons. First, at a practical level, the resulting mixture leads the analysis outside the family of conjugate (normal) priors selected throughout the literature for the analysis of the model, which becomes significantly less tractable as a result.[19] Secondly, at a conceptual level, we think of the partially cursed posterior as a useful tool to model the behavior of agents who

[18] This conditioning is essential. Cursedness only posits that the opponents' actions are independent of the state conditional on the agent's information. To arrive at such an interim belief with the quasi-Bayesian approach, this needs to be reflected in the cursed prior, according to which the signal p is indeed informative about the state, just such in a way that conditioning on s_i, y renders it uninformative. Furthermore, the actions of others keep their correlation. This is a key difference compared to sequential cursed equilibrium (Cohen and Li, 2023) which posits that agents do not expect that their opponents actions are correlated with each other and with signals the agent will receive in the future. In our setting, an analogue to SCE would imply that fully cursed agents view the aggregate action as deterministic: ex-ante deterministically zero, ex-interim deterministically at its expectation. This seems implausible in our setting.

[19] One way to address this issue is to instead consider a misspecified Bayesian with a normally distributed prior whose posterior replicates the behavior of the (partially) cursed agent. By assuming signals are conditionally independent and preserving the (implicit) weight on the prior mean, we arrive at the perceived precisions

$$\left(\widehat{\tau}_\theta, \widehat{\tau}_s, \widehat{\tau}_y, \widehat{\tau}_p\right) = \left(\tau_\theta, \tau_s, \tau_y, (1-\chi)\,\tau_p \frac{\tau_\theta + \tau_y + \tau_s}{\tau_\theta + \tau_y + \tau_s + \chi\delta_1^2\tau_p}\right).$$

only partially account for the information contained in aggregate statistics, but not as an accurate representation of the mental model and subjective welfare of agents. Indeed, the above mental model does not appear very plausible. Third and perhaps most problematic, in order to obtain a signal that is independent of the state conditionally on their private and public fundamental signal, the perceived correlation of the public signal p with the state θ at the ex-ante stage depends on an agent's *individually chosen* τ_s. In other words, the agent behaves as if his personal information acquisition affects the precision of public information obtained by all agents, which is a sort of magical thinking we don't commonly associate with cursedness.

In Appendix C, we formally develop this quasi-Bayesian approach and provide a numerical solution to demonstrate that the qualitative properties of the χ-cursed expectation equilibrium with information acquisition we present in the following sections remain valid when agents choose information acquisition based on this notions.

Sophisticated: Alternatively, one could consider sophisticated agents who choose τ_s to maximize true ex-ante welfare (14), that is, choose $\max_{\tau_s} W(\alpha(\tau_s, \delta), \delta, \tau_s)$ letting the action rule $\alpha(\tau_s, \delta)$ change in response to information acquisition. This corresponds to a rational agent who correctly predicts his cursed actions but desires to correct them by distorting the precision of private information available to his future biased self. This level of sophistication, together with an inability to use the aggregative signal correctly ex-interim, seems implausible to us. Our notion instead does not presume this high level of meta-rationality while preserving a subjective world view that is consistent with the atomistic position of the agent within the game.

Behavioral Equilibrium: A further alternative would be to adapt behavioral equilibrium (Esponda, 2008).[20] In this equilibrium concept, interim beliefs are required to be consistent with the marginal distribution of observables, which have to be specified by the analyst. Assuming that agents only observe what is necessary to implement their conditional strategy, namely p, the fully cursed (and many other, including highly unreasonable) beliefs and action profiles are a naive behavioral equilibrium. In the learning foundation, we instead assume that agents can observe their realized payoffs. Assuming that agents observe the ingredients of this payoff, namely θ and $\bar{a}$, no naive behavioral equilibrium exists. This is because the marginal distributions need to

Aside from being quite ad hoc, this approach suffers from conceptual drawbacks. First, the perceived precision of the exogenous signal about the aggregate action now depends on the endogenous equilibrium weight δ_1. Second, and more importantly, the perceived precision of the public signal p depends on an agent's *individually chosen* τ_s, which violates our second principle.

[20] Note that behavioral equilibrium is defined for static Bayesian games and not directly applicable in our setting. We consider an version in which agents observe s_i and y and submit action schedules conditioning on p. Still, we interpret $\bar{a}$ as the action of others and impose the independence condition of naive behavioral equilibrium with respect to this statistic.

coincide with the truth, which—under the independence condition—would imply inaccurate beliefs about the distribution of utility, which are however ruled out. The sophisticated behavioral equilibrium, instead, coincides with the rational case. The intermediate case in which agents only observe their realized utility requires finding subjective marginal distributions over $\theta, \bar{a}$ which induce the true distribution over realized welfare. There is no immediate way to ascertain whether such distribution exist.

Beyond these hurdles to applying behavioral equilibrium in our beauty contest setting, the spirit of its learning motivation differs markedly from that of our learning foundation. In a learning interpretation of behavioral equilibrium (and in Berk-Nash equilibrium more generally (Esponda and Pouzo, 2016)), individuals learn an interim belief over the structural parameters of their optimization problem (the distribution of states and the actions of others) and calculate their best-response. Our approach distinguishes between the use and acquisition of information. For former, we do not follow an approach motivated by learning interim beliefs (a very high dimensional object) but instead embrace cursed equilibrium as a model of instinctive information use of biased individuals. For the latter we do propose a notion which can be grounded in a learning process. This learning process, however, operates differently from the learning motivation behind behavioral equilibrium: Our agents disregard the structure of the problem and instead directly ascertain whether their current level of precision is optimal by estimating the derivative of their welfare, an approach akin to gradient ascent.

4 Positive Comparative Statics

In this Section, we study the comparative statics of the model. We begin by considering the impact of cursedness and transparency on the equilibrium loadings and information acquisition.

Proposition 1. *The equilibrium loadings and information acquisition respond to cursedness χ as follows*

$$\frac{\mathrm{d}\delta_1}{\mathrm{d}\chi} \geq 0, \qquad \frac{\mathrm{d}\delta_2}{\mathrm{d}\chi} \geq 0, \qquad \frac{\mathrm{d}\delta_3}{\mathrm{d}\chi} \leq 0, \qquad \frac{\mathrm{d}\tau_s}{\mathrm{d}\chi} \geq 0 \tag{20}$$

Moreover, $\frac{\mathrm{d}}{\mathrm{d}\chi}\frac{\delta_2}{\delta_1} = \frac{\mathrm{d}}{\mathrm{d}\chi}\frac{1-\delta_3}{\delta_1} = 0.$

The comparative statics w.r.t. transparency τ_p have the opposite sign, namely

$$\frac{\mathrm{d}\delta_1}{\mathrm{d}\tau_p} \leq 0, \qquad \frac{\mathrm{d}\delta_2}{\mathrm{d}\tau_p} \leq 0, \qquad \frac{\mathrm{d}\delta_3}{\mathrm{d}\tau_p} \geq 0, \qquad \frac{\mathrm{d}\tau_s}{\mathrm{d}\tau_p} \leq 0 \tag{21}$$

Moreover, the precision of the aggregative signal increases in transparency, i.e. $\frac{\mathrm{d}}{\mathrm{d}\tau_p}\delta_1^2\tau_p \geq 0$.

The above inequalities are strict if (19) holds and $\tau_p \neq 0$ *or* $\chi \neq 1$*, respectively.*

If agents are more cursed, they load more on the fundamental sources of information, either private or public, and less on the aggregative signal. Because cursedness makes agents underestimate the information content of the aggregative signal, cursed agents substitute away from this source of information and towards fundamental information. Notice that this reasoning is based only on how individuals use the signals they observe, and indeed all the comparative statics (20) hold if the level of private information τ_s were exogenous (see Appendix E.1 for this and other derivations in the fixed-τ_s environment). By the subjective envelope condition, τ_s grows proportionately to its use δ_1 and amplifies this effect: adjusting upward the precision of private information makes agents want to load even more on this source of information, reinforcing the comparative statics. The stability of the *relative* loadings $\frac{\delta_2}{\delta_1}, \frac{1-\delta_3}{\delta_1}$, is instead driven by information acquisition; if τ_s were exogenous, then cursedness would shift the relative loadings on fundamental information in favor of the public signal ($\frac{\mathrm{d}}{\mathrm{d}\chi}\frac{\delta_1}{\delta_2} \leq 0$). This is because the public signal is a closer substitute to the aggregative one as both have a public noise component. Increased information acquisition in response to higher cursedness counteracts this effect by making the private signal (relatively) more appealing.

The mirror structure of the comparative statics of χ and τ_p in Proposition 1 formalizes the intuition that cursedness and transparency are complementary antagonists: reducing the amount of information contained in the source whose interpretation is directly affected by the processing bias has (qualitatively) the same impact as increasing the processing bias itself. In particular, higher transparency decreases the loading on private information: As the private information of others is disseminated more effectively, each agent relies less on his own.[21] Nevertheless, this crowding-out effect never dominates and the precision of the aggregative signal about the fundamental is always increasing in transparency (see Figure 1). We now study how the use and acquisition of private information (δ_1 and τ_s) vary with the information and strategic environment.

Proposition 2. *The loading on private information responds as follows*

$$\frac{\mathrm{d}\delta_1}{\mathrm{d}c} \leq 0, \qquad \frac{\mathrm{d}\delta_1}{\mathrm{d}\tau_y} \leq 0, \qquad \frac{\mathrm{d}\delta_1}{\mathrm{d}r} \leq 0 \tag{22}$$

[21] Again, this result holds irrespective of whether τ_s is allowed to adjust in response of higher transparency.

The comparative statics of τ_s have the same sign as the comparative statics of δ_1, namely

$$\frac{d\tau_s}{dc} \leq 0, \qquad \frac{d\tau_s}{d\tau_y} \leq 0, \qquad \frac{d\tau_s}{dr} \leq 0 \tag{23}$$

The above inequalities are strict if (19) holds.

As costs increase, agents will acquire less precise private information, which pushes towards a decrease in its use. As everybody does so, however, the aggregative signal becomes less informative, leading to an increased reliance on private information as well as a boost in its value. This counteracting force never dominates and reliance on (as well as acquisition of) private information always decreases with cost. That δ_1 decreases in τ_y follows because when the public fundamental signal becomes more precise agents shift away from the private signal. Finally, as complementarities become stronger the public signals become more attractive relative to the private signal as they allow for better coordination with the aggregate action. Consequently, an increase in r decreases the weight on the private signal.[22]

The other equilibrium loadings, δ_2 and δ_3, have an immediate interpretation as the weight given by the agent to the public fundamental and aggregative signal, respectively. However, in terms of information, they do not fully reflect how much agents load on the different sources. For example, agents load on the public signal both directly and indirectly, through the aggregative signal. For this reason, the comparative statics of δ_2 and δ_3 in τ_s, τ_y, r (which are all ambiguous) can be misleading. We get a clear understanding of how agents use public sources of information only by analyzing the fundamental representation

$$a_i = \underbrace{\frac{\delta_1 + \delta_2}{1 - \delta_3}}_{\beta}\theta + \delta_1 z_{s_i} + \underbrace{\frac{\delta_2}{1 - \delta_3}}_{\gamma_2} z_y + \underbrace{\frac{\delta_3}{1 - \delta_3}}_{\gamma_3} z_p \tag{24}$$

The term β represents the regression coefficient of the individual action (and, a fortiori, the aggregate action) on the state and hence we interpret it as an informational efficiency metric. Notice that β is determined by the interaction of use, acquisition, and equilibrium dissemination of private information. The weights γ_2, γ_3 on the public shocks differ from the direct loadings on the signals by a factor of $\frac{1}{1-\delta_3}$: as we argued, the aggregative signal contains and amplifies both public shocks. Taking account of this amplification, we obtain the following comparative statics

[22] By the subjective envelope condition (17) δ_1 is a sufficient statistic for the gains from acquiring information, therefore the comparative statics on τ_s for all parameters but c follow directly from those of δ_1.

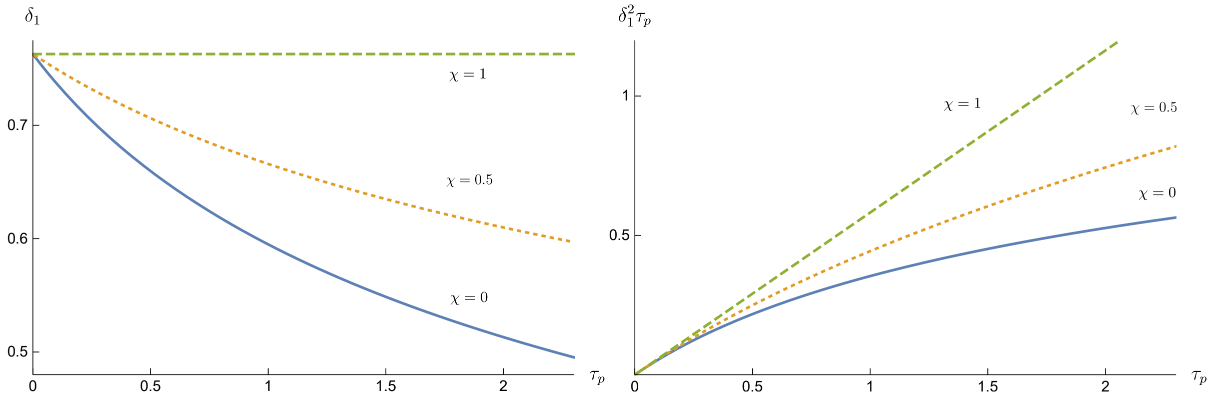

Figure 1: Crowd-out vs Reliance on Private Information: The effect of τ_p on δ_1 and $\delta_1^2\tau_P$ for different levels of cursedness.

Proposition 3. *The loadings in the fundamental representation (24) are*

$$\beta = 1 - \frac{\sqrt{c}\tau_\theta}{1-r}, \quad \gamma_2 = \frac{\tau_y\sqrt{c}}{1-r}, \quad \gamma_3 = \frac{1}{\delta_1}\left(1 - \frac{\tau_\theta + \tau_y}{1-r}\sqrt{c}\right) - 1. \tag{25}$$

The loading on the aggregative signal, γ_3, is decreasing in cursedness and increasing in transparency. It has ambiguous comparative statics in c, r and τ_y.

Let's focus first on γ_3, the fundamental loading on the aggregative signal p. Cursedness and transparency have an unambiguous effect as they impact γ_3 exclusively through δ_1: when agents perceive the aggregative signal to be less informative, either because they are more cursed or because the environment is less transparent, they will use it less. The comparative statics of γ_3 in τ_y, r, c are instead ambiguous due to a tension between the direct effect and the effect that goes through δ_1 (see Proposition 2). Consider, for example, how γ_3 responds to an increase in c. On the one hand, it reduces information acquisition and makes agents substitute towards the aggregative signal. On the other hand, the reduction in information acquisition and reliance on the private signal removes the very basis of information in p, making it less attractive. Either effect can dominate depending on the rest of the information environment. A similar intuition is the basis for the ambiguous comparative statics in the other parameters: the precision of the aggregative signal changes both in level and relative to the precision of other signals.

Notice from (25) that the state-action regression coefficient β is decreasing in c, τ_y and r, but is independent of either cursedness or transparency. That β is invariant in χ and τ_p has three consequences of economic relevance: First, we cannot identify the degree of cursedness in a market by just looking at the responsiveness of individual actions to fundamentals. Second, transparency is an ineffective tool at increasing the informational efficiency along the β metric as its effect is fully offset by lower acquisition and use of private information. This result closely resembles the invariance property with respect to the dispersion of net supply from noise traders in Grossman and Stiglitz (1980). Third, as opposed to the setting without information acquisition and the findings in Eyster et al. (2019), cursedness does not reduce the responsiveness of the aggregate action with respect to the fundamental. Even though cursed agents reduce the efficiency of information dissemination by failing to amplify the information content of p, they inject more private information into the system.

Analytically, the invariance of β results from the combination of information acquisition and use. If private information were exogenous (Appendix E.1 presents the results in this case), β would decrease in cursedness (and increase in transparency); with exogenous τ_s, the higher use of private information of cursed agents is more than offset by the fact that those same agents fail to learn from the aggregative signal and hence hamper dissemination. Once we allow agents to adjust τ_s in response to changes in the perceived value of private information caused by either a more transparent environment or a decrease in cursedness, those effects are neutralized.

The fundamental loading γ_2 on public fundamental information has intuitive comparative statics: agents load more on y if it is more precise, if private acquisition is more costly, or if the coordination motive is stronger. The degree of cursedness does not affect γ_2, not even indirectly. Even though cursed agents fail to process all information disseminated through the aggregative signal, when they can adjust τ_s, their increased demand for and use of private information exactly offsets the less efficient inference.

While cursedness does not affect the total weight on information obtained through private signals, $\beta - \gamma_2$, it changes its composition: agents substitute away from indirect inference of disseminated private information, δ_3, towards information agents have acquired themselves, $1 - \delta_3$. This decomposition (depicted in Figure 2) is apparent in the following rewriting of the fundamental representation (24)

$$a_i = (\beta - \gamma_2)\left[(1 - \delta_3)(\theta + z_s) + \delta_3 \theta\right] + \gamma_2 y + \gamma_3 z_p. \tag{26}$$

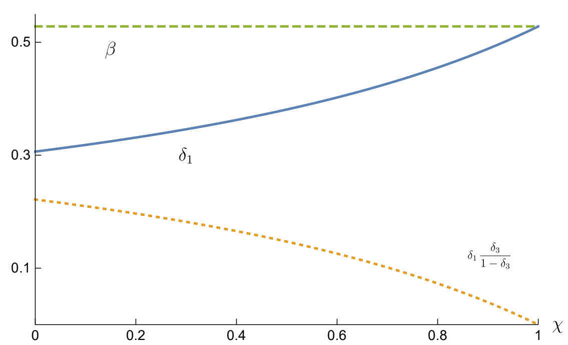

Figure 2: Equilibrium weight on information from private signals: direct and indirect.

We analyze a different metric of informational efficiency, namely the total precision available to the agents, in Appendix E.2 and show that cursedness determines whether transparency improves efficiency along this metric.

5 Welfare

In this section, we first characterize the first-best benchmark for the use and acquisition of information. Then we identify and characterize the inefficiencies of the χ–cursed equilibrium with information acquisition. Finally, we perform welfare comparative statics.

5.1 The Planner Problem

As a welfare benchmark, we consider the problem of a planner who controls both the use and acquisition of information, but cannot share information across agents.[23] To this end we impose the consistency condition $\boldsymbol{\alpha} = \boldsymbol{\delta}$ in the welfare expression (14) and, with slight abuse of notation, let $W(\boldsymbol{\delta}, \tau_s) := W(\boldsymbol{\delta}, \boldsymbol{\delta}, \tau_s)$ denote the objective function – and $W^\star$ the value – of a planner choosing

$$\left(\boldsymbol{\delta}^\star, \tau_s^\star\right) = \arg\max_{(\boldsymbol{\delta}, \tau_s)} W(\boldsymbol{\delta}, \tau_s) \tag{27}$$

that can be expressed as

$$W(\boldsymbol{\delta}, \tau_s) = -\frac{(1-r)}{(1-\delta_3)^2}\left\{\frac{\delta_2^2}{\tau_y} + \frac{\delta_3^2}{\tau_p} + \frac{(1-\delta_1-\delta_2-\delta_3)^2}{\tau_\theta}\right\} - \frac{\delta_1^2}{\tau_s} - c\tau_s. \tag{28}$$

We proceed by characterizing the solution of (27) and the comparative statics of first-best welfare.

Theorem 4. *The loading on private information in the efficient linear action rule is the unique solution of*

$$\delta_1 = (1 - r + r\delta_1)\,\tau_s^\star \frac{1}{\underbrace{\left(\frac{\tau_\theta + \tau_y + \tau_p\delta_1^2}{\tau_\theta + \tau_y + \tau_p\delta_1}\right)}_{\textit{efficiency wedge}}\left(\tau_\theta + \tau_y + \tau_p\delta_1^2\right) + \tau_s^\star}. \tag{29}$$

[23] This is the benchmark customarily adopted in the literature (Angeletos and Pavan, 2007). It avoids the unfair comparison with an economy in which agents can also share information: as there are uncountably many, this would coincide with playing a game of complete information with a trivial solution and trivial welfare properties.

The efficient precision of private information satisfies $\tau_s^\star = \frac{\delta_1^\star}{\sqrt{c}}$. *Moreover, first-best welfare has the following comparative statics*

$$\frac{dW^\star}{d\tau_\theta} > 0, \quad \frac{dW^\star}{d\tau_y} > 0, \quad \frac{dW^\star}{d\tau_p} > 0, \quad \frac{dW^\star}{dc} < 0. \tag{30}$$

Condition (29) corresponds to the rational equilibrium condition (11), modified by an efficiency wedge. The wedge accounts for the fact that using public information as the basis of action dilutes the dissemination of private information. The planner internalizes this effect and therefore downweighs public information by the adjustment term $\frac{\tau_\theta+\tau_y+\tau_p\delta_1^2}{\tau_\theta+\tau_y+\tau_p\delta_1} < 1$. This wedge is equal to one only if $\delta_1^\star = 1$ and therefore $\delta_2^\star = \delta_3^\star = 0$, i.e., if the aggregative signal is not polluted by public signals to begin with, which is impossible in equilibrium. Two implications of economic relevance follow. First, δ_1 is inefficiently low, the efficient solution features a higher weight on private information compared to the rational equilibrium.[24] Therefore, the equilibrium with $\chi = 0$ is not efficient. For a positive level of cursedness, the inefficiency of the equilibrium is an immediate consequence of the processing bias. Therefore, we have the second implication, the equilibrium is always inefficient.

The optimality condition for τ_s is our familiar envelope condition (17): The use of private information is a sufficient statistic for the gains from acquiring it, even for the planner. Both efficient and equilibrium information acquisition are fully determined by the respective use of private information. Therefore, there is under-(over)acquisition of private information in equilibrium if and only if there is under- (over)use of private information in equilibrium, i.e.

$$\operatorname{sgn}\{\tau_s - \tau_s^\star\} = \operatorname{sgn}\{\delta_1 - \delta_1^\star\}. \tag{31}$$

In particular, information acquisition in equilibrium is efficient if an only if information use in equilibrium is efficient.

Finally, the comparative statics of first-best welfare (30) are natural: as information is used efficiently by the planner, increasing precision – whatever the source – or lowering acquisition costs is always beneficial.

[24] Note that this distortion only arises because δ_1 determines the signal-to-noise ratio when inferring the state from the signal of the aggregative action. In a setting with fully flexible information acquisition, (Hebert and La'O, forthcoming) characterize the information cost functions under which this scaling does not matter and that therefore result in an efficient equilibrium.

In demand function competition, Vives (2017) shows that there is both an information dissemination externality as well as a pecuniary externality, the latter causing a potentially excessive weight on private information.

5.2 The Inefficiencies of Equilibrium

In equilibrium, by contrast, information is generally used inefficiently: agents do not internalize the dissemination externality and they are subject to a processing bias that makes them misuse the available information.

Proposition 4. *The rational equilibrium always has inefficiently low information acquisition. Sufficiently cursed agents acquire more private information than the efficient benchmark if* $\tau_p > \overline{\tau}_p$*, where* $\overline{\tau}_p = \left(\tau_\theta + \tau_y\right)\left(1 - 2\delta_1^{FC}\right)\left(\delta_1^{FC}\right)^{-3}$.

Since rational agents do not internalize the dissemination externality, the equilibrium with $\chi = 0$ features underacquisition. As τ_s is increasing in χ by Proposition 1, cursedness alleviates this inefficiency. This effect can be strong enough to lead to overacquisition relative to the efficient benchmark if the aggregative signal is sufficiently precise. Intuitively, if transparency exceeds the lower bound $\overline{\tau}_p$, then dissemination is so effective that even $\tau_s^\star$ (which is independent of χ) is low compared to the precision of information acquired by the agent in the fully cursed equilibrium (which is by construction independent of τ_p). Then, there exists an interior χ such that the equilibrium use and acquisition of private information coincide with the efficient quantity (see the left panel of Figure 3).[25] Even in this case, however, agents misperceive the information environment and hence misuse their information. To analyze this source of inefficiency, consider the gradient of welfare at equilibrium as we vary the cursedness parameter (displayed in the right panel of Figure 3).

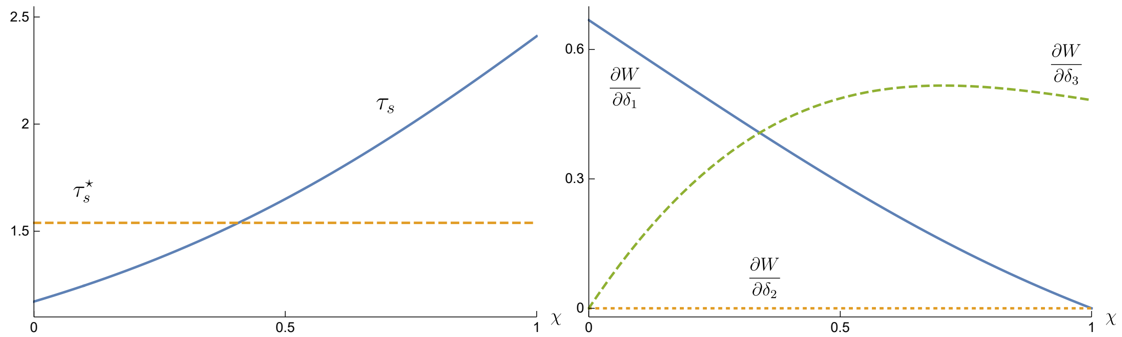

Figure 3: Equilibrium vs. efficient information acquisition (Proposition 4, left) and the gradient of W in equilibrium (Proposition 5, right) as a function of cursedness.

Proposition 5. *In any equilibrium,* δ_2 *is conditionally efficient:* $\frac{\partial W}{\partial \delta_2}(\delta, \tau_s) = 0$ *for all* χ.
In a rational equilibrium, δ_3 *is conditionally efficient:* $\frac{\partial W}{\partial \delta_3}(\delta, \tau_s) = 0$ *for* $\chi = 0$.
In a fully cursed equilibrium, δ_1 *is conditionally efficient:* $\frac{\partial W}{\partial \delta_1}(\delta, \tau_s) = 0$ *for* $\chi = 1$.

[25] The condition is always met ($\overline{\tau}_p < 0$) if incentives for private information acquisition are sufficiently high, namely $\sqrt{c} \leq \frac{1}{2}\frac{1-r}{\tau_\theta+\tau_y}$.

In a fully rational equilibrium, the only externality is the dissemination of private information. Fixing the use of private information and thereby its dissemination, the other loadings of the equilibrium are conditionally efficient. In a fully cursed equilibrium, agents ignore the aggregative signal altogether, so there is no dissemination externality and private information is used efficiently.[26] Independently of the degree of cursedness, there is no externality or misunderstanding in the use of the public fundamental signal.

5.3 The Comparative Statics of Equilibrium Welfare

The sources of equilibrium inefficiency identified in Propositions 4 and 5 provide the bedrock for analyzing the impact of cursedness and changes in the information environment on equilibrium welfare. Let $W^{EQ} := W\left(\delta^{\chi}, \tau_s^{\chi}\right)$ denote equilibrium welfare, where we introduce $\delta^{\chi}, \tau_s^{\chi}$ as shorthand for the equilibrium with χ-cursed agents.

Cursedness is Bliss

Consider a marginal increase in cursedness starting from the rational equilibrium. It has two impacts on welfare. First, agents now use their information suboptimally as they underestimate the information contained in p. The associated welfare reduction is second order, however, as δ is privately optimal in the rational equilibrium. Second, cursed agents acquire and disseminate more information. Since the rational equilibrium features underacquisition, this impact on the dissemination externality has a first order effect on welfare. Thus, local to rationality and up to first order, cursedness only has beneficial effects on welfare. When χ is already large, however, marginal increments in cursedness have a first order negative effects from additional misuse, while the underacquisition gap is narrower, if existent at all. The effect of inefficient use dominates close to $\chi = 1$ as fully cursed agents do not use the information contained in the aggregative signal.

Proposition 6 (**Cursedness is Bliss.**).

$$\left.\frac{dW^{EQ}}{d\chi}\right|_{\chi=0} > 0. \tag{32}$$

Furthermore,

$$\left.\frac{dW^{EQ}}{d\chi}\right|_{\chi=1} < 0, \tag{33}$$

[26]Again, recall that fully cursed equilibrium coincides at the action stage with fully rational equilibrium in which τ_p is set to zero. Without an aggregative signal our model is a special case of Angeletos and Pavan (2007) where payoffs satisfy the conditions for efficient use of information.

so any level of cursedness maximizing equilibrium welfare is interior.

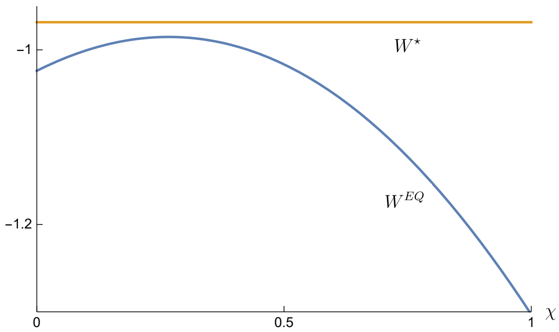

Figure 4: Cursedness is Bliss

The shape of welfare as a function of χ and the comparison to efficient welfare is shown in Figure 4. One might wonder if the comparison in the plot holds in general or whether a fully cursed economy can ever outperform full rationality. This cannot happen. Indeed, it is easy to show that in the two extreme cases $\chi \in \{0, 1\}$, welfare takes the simple form

$$\mathrm{W}^{\mathrm{EQ}} = -\sqrt{c}\,(1 + \delta_1). \tag{34}$$

and it follows from the comparative statics of δ_1 that the fully cursed equilibrium has lower welfare than the rational case: Even though acquisition and dissemination of private information are higher, cursed agents are unable to make any use of their aggregative information. The inefficiently imprecise aggregative information provided in the rational equilibrium is preferable to complete ignorance of – albeit plentiful – aggregative information.

Transparency and Other Information Policies

In contrast to the efficient solution (Proposition 4), more information and lower costs do not always increase welfare in equilibrium.

Proposition 7. *Equilibrium welfare* W^{EQ} *is always increasing in* τ_p.

W^{EQ} *is increasing in* τ_y, τ_θ *and decreasing in* c *if* χ *is sufficiently close to either* 0 *or* 1 *or* τ_p *is sufficiently small. If, however,* τ_p *is sufficiently large, there exist an interior region of* χ *such that equilibrium welfare is*

(i) decreasing in τ_y, τ_θ *if strategic complementarities are sufficiently strong* ($r > \frac{1}{2}$)*,*

(ii) and increasing in c *in a game with strategic substitutes* ($r < 0$)*.*

The proposition identifies sufficient conditions for counterintuitive comparative statics. Let us first focus on the comparative static with respect to the precision of the public fundamental signal.[27] An increase in τ_y has two equilibrium effects. First, a direct effect as the precision of the agents' information increases mechanically. Second, a substitution effect as agents reduce their use and acquisition of private information and also substitute towards y and away from p. When strategic complementarities are sufficiently strong, the second effect is particularly important and causes the paradoxical comparative static: For partially cursed agents, the use of p is already suboptimally low and a further decrease entails a welfare loss. This effect dominates the welfare calculus for interior χ. For (close to) fully cursed agents, however, the substitution effect is negligible as they disregard p and the direct effect is important as they rely heavily on y.[28]

Likewise, an increase in acquisition costs potentially benefits partially cursed agents: It causes them to rely less on private information and to substitute towards p, whose informativeness they effectively underestimate. This effect can dominate with sufficiently strong strategic substitutes, when the value of information is relatively low because agents want to anti-coordinate. As we can see in Figure 5, the effect is present for sufficiently small costs, since then the substitution is towards a relatively informative p, whereas if costs are too high, the aggregative signal itself becomes too noisy.

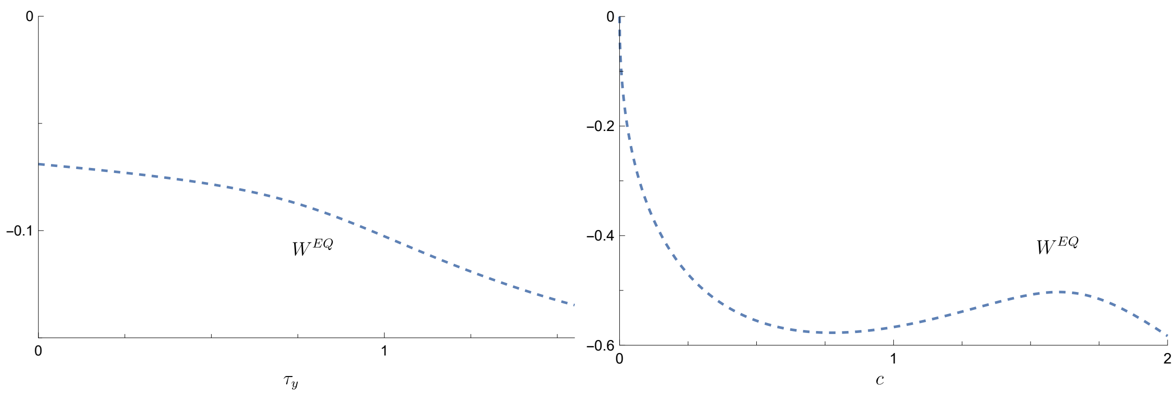

Figure 5: Counterintuitive comparative statics of welfare in τ_y (left) and c (right).

It might be surprising that transparency always increases welfare: After all, it has an ambiguous effect on total precision τ_Σ and other parameters may have perverse

[27]A similar result is also obtained in Morris and Shin (2002), but for different reasons. There, all signals are fundamental, but the increased use of public information entails a payoff externality. In our setting, payoffs are such that – absent dissemination externality and cursedness – information use is efficient (Angeletos and Pavan, 2007; Colombo et al., 2014) and hence more precise public information is always welfare improving. Natural comparative statics continue to hold even if $\tau_p > 0$ and $\chi = 0$, a case not subsumed by the literature; therefore both the dissemination externality and cursedness interact to yield the paradoxical effects listed in Proposition 7.

[28]In the proof of this result, we consider the limit economy as $\tau_p \to \infty$. The limit welfare (71) is decreasing in τ_y (or, τ_θ) if and only if $1 - 2r + r\chi < 0$, which provides a lower bound, $\chi > 2 - \frac{1}{r}$ that can be satisfied only if $r > \frac{1}{2}$. Incidentally, this is the same threshold on the degree of complementarities that Morris and Shin (2002) obtain for public information τ_y to be reduce welfare.

effects on welfare. In addition, inference from aggregative information is biased by cursedness while the fundamental sources, which generate these counterintuitive effects, are interpreted correctly. The key observation to understand this difference is that transparency renders the aggregative signal more informative relative to the fundamental sources of information, while the opposite is the case for the other parameters. Consequently, partially cursed agents substitute towards the aggregative signal, which ameliorates their bias. At worst, in the fully cursed case, the aggregative signal is not understood at all and hence irrelevant. However, as the fully cursed case makes apparent, there remain unreaped benefits from increased transparency in such economies.

6 Shrewd Agent: Behavior and Policy

Although their welfare increases with transparency, cursed agents fail to reap its full benefits. How does an agent who is able to extract all the information from the environment – such as proverbial smart money in financial markets – interact with a cursed crowd? Could it be beneficial to act in an environment with less rational agents? We address these questions by studying the behavior and welfare of a *shrewd agent*: a fully rational, atomistic agent in the model who understands its structure and is aware that all other agents (the *cursed crowd*) are χ-cursed. We discuss the results in the text, relegating much of their formal development to Appendix D in the interest of brevity.

6.1 Best Response and Information Acquisition

We continue to denote the precision of information acquired by the cursed crowd as τ_s and denote the precision acquired by the shrewd agent as τ_s^R. The shrewd agent takes the equilibrium loadings (and information acquisition) of the cursed crowd as given, and chooses both how much private information to acquire as well as the coefficients in his linear action rule $a_i^R = \alpha_1^R s_i + \alpha_2^R y + \alpha_3^R p$. Formally, he solves

$$\max_{\boldsymbol{\alpha}^R, \tau_s^R} W\left(\boldsymbol{\alpha}^R, \delta, \tau_s^R\right) \tag{35}$$

Best responding to the equilibrium in the cursed crowd, the individual loadings of the shrewd agent $\boldsymbol{\alpha}^R$ (characterized in Proposition 10 of the Appendix), will differ from δ whenever $\chi > 0$. There is again a tight connection between the use of private information, α_1^R, and its acquisition through the envelope condition

$$\tau_s^R = \frac{\alpha_1^R}{\sqrt{c}}. \tag{36}$$

Denote the total precision available to the cursed agents by $\tau_\Sigma := \tau_\theta + \tau_y + \tau_s + \delta_1^2\tau_p$ and that available to the rational agent by $\tau_\Sigma^R := \tau_\theta + \tau_y + \tau_s^R + \delta_1^2\tau_p$. We obtain an equation linking the information acquired by the shrewd agent and the cursed crowd

$$\frac{\tau_\Sigma}{\tau_\Sigma^R} = 1 + \frac{\chi\delta_1^2\tau_P}{\tau_\theta + \tau_y + \tau_s}. \tag{37}$$

The shrewd agent acquires less information. Compared to the cursed crowd, he can substitute it with a better comprehension of aggregative information. If the crowd is fully cursed, then (37) simplifies to

$$\tau_s^R = \tau_s - \delta_1^2\tau_p \tag{38}$$

that is, the shrewd agent exactly offsets the information he can glean from the aggregative signal and his precision is equal to the precision *perceived* by the crowd.

Clearly, equation (38) holds only if it delivers a positive τ_s^R. Otherwise, the shrewd agent will choose $\tau_s^R = 0$ as he is already satiated with the information he can infer from the aggregative signal. With a fully cursed crowd, this always happens with sufficiently large transparency since both τ_s and δ_1 are unresponsive to τ_p. In that case, transparency only serves as a cost-saving device for the shrewd agent. The shrewd agent continues to free-ride on the crowd's use of private information even at interior levels of cursedness.

Proposition 8. *The information acquired by the shrewd agent, τ_s^R, satisfies:*

1. *It is bounded by the precision acquired by the cursed crowd $\tau_s^R \leq \tau_s$. It grows without bounds as costs vanish but can be zero even when $\tau_s > 0$.*
2. *If τ_p is sufficiently large, then it is*
   - *(i) nonmonotonic in prior and public precision, τ_θ and τ_y, possibly with an interior* activity *region; and*
   - *(ii) nonmonotonic in costs c, possibly with an interior* inactivity *region.*
3. *It has ambiguous comparative statics w.r.t. χ,*

$$\frac{d\tau_s^R}{d\chi} \propto r - 2\sqrt{c}\delta_1\tau_p.$$

The shrewd agent acquires a strictly positive amount of information if and only if $\tau_\theta + \tau_y \in \left(\frac{1-r}{\sqrt{c}}\left(1 - \frac{1}{\tau_p\sqrt{c}}\right), \frac{1-r}{\sqrt{c}}\right)$. Therefore, there is an inactivity region whenever $\tau_p > \frac{1}{\sqrt{c}}$. In that case, the shrewd agent acquires private information only if public information is sufficiently precise. The rationale is as follows: If public fundamental information is noisy, the cursed crowd will acquire and use a lot of private information; since he can be parasitic on this information, the rational agent has no incentive to acquire information himself. As public information becomes more abundant, however, there is less information acquisition and use by the crowd. The aggregative source dries up and the shrewd agent needs to supplement it with private information acquisition. Finally, the upper bound on $\tau_\theta + \tau_y$ for the existence of a nontrivial equilibrium is the same for both classes of agents. In the trivial equilibrium $\delta_1 = \tau_s = 0$, the shrewd agent cannot utilize his advantage in understanding the aggregative source since it is uninformative: he behaves identically to the crowd.

An immediate consequence of this inactivity region is that τ_s^R is nonmonotonic in τ_y. This contrasts with the unambiguously signed comparative statics for τ_s (Proposition 1). Similarly, the effect of information acquisition costs on τ_s^R is nonmonotonic and we can have an inactivity region (see Figure 6). Again, a change in parameters affects both the availability of aggregative information provided by the cursed crowd and the shrewd agent's demand for information overall.

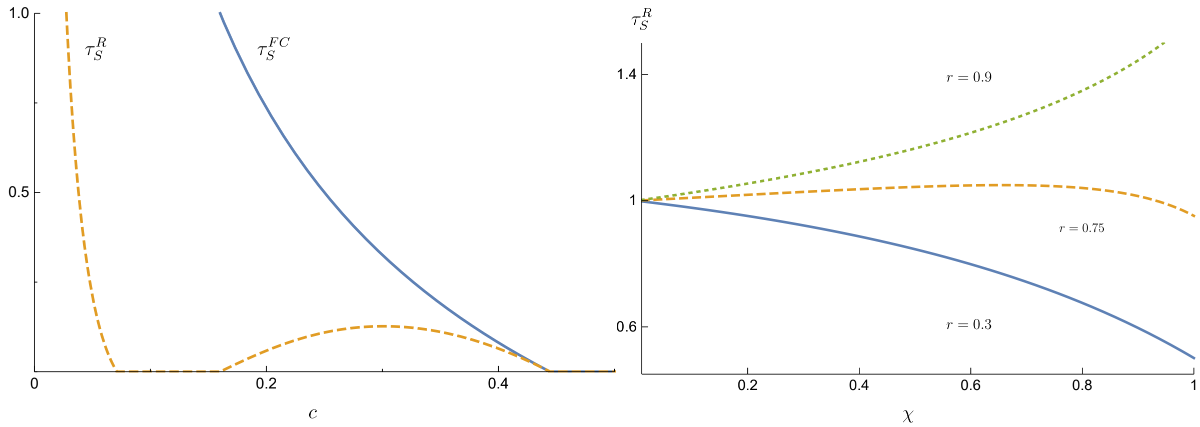

Figure 6: τ_s and τ_s^R as a function of c (left). τ_s^R (normalized to 1 at $\chi = 0$) as a function of χ for $r \in \{0.3, 0.75, 0.9\}$ (right).

The impact of cursedness on the precision of information acquired by the shrewd agent depends on the nature of strategic interactions (see Figure 6). Take as a benchmark the case of $r = 0$, i.e., all agents simply try to guess the true state, and all strategic interaction comes from the precision of aggregative information. In this case, τ_s^R is decreasing in χ as more cursed agents acquire and disseminate more information. Strategic substitutes increase this effect: As the crowd becomes more cursed, δ_1 increases, which reduces the desire to match the state and therefore the value of private information. With complements, the opposite is the case: A higher δ_1 increases the desire to match θ and therefore – if this motive is sufficiently strong – information acquisition.

6.2 Welfare

Let W_χ^R denote the welfare of the shrewd agent facing an equilibrium δ^χ. Then,

$$W_\chi^{EQ} = W\left(\delta^\chi, \delta^\chi, \tau_s^\chi\right) \leq \max_{\alpha, \tau_s^R} W\left(\alpha, \delta^\chi, \tau_s^R\right) = W_\chi^R \tag{39}$$

As he comprehends his information environment, he always obtains a higher welfare than the cursed crowd. The inequality is strict if $\chi > 0$.

We now ask whether the shrewd agent benefits from an increase in the cursedness of the crowd. This is the case for the first modicum of cursedness since for small positive ϵ

$$W_\epsilon^R > W_\epsilon^{EQ} > W_0^{EQ} = W_0^R. \tag{40}$$

The central inequality follows since in this region "cursedness is bliss" (Proposition 6).[29] In a highly cursed environment, however, the impact of cursedness depends on nature of the strategic interaction.

Proposition 9. *Suppose parameters are such that $\tau_s^R > 0$. If $r \leq 0$, then $\left.\frac{dW_\chi^R}{d\chi}\right|_{\chi=1} \geq 0$. However, for r sufficiently large, $\left.\frac{dW_\chi^R}{d\chi}\right|_{\chi=1} \leq 0$.*

If there are strategic substitutes, the shrewd agent always benefits from increasing the cursedness of the crowd: not only does he free-ride on aggregative information, but the crowd's over-reliance on the private signal helps him anti-coordinate. In the presence of complementarities, however, informational free-riding and the lack of coordination implied by cursed information misuse have opposing effects. While the shrewd agent can learn the state more precisely, his action has to follow the behavior of the less informed crowd. The latter effect can be overwhelming close to $\chi = 1$, so he would prefer an interior level of cursedness.

We conclude this section by studying the impact of precision and cost parameters on W_χ^R. If the crowd is close to rational, then policies have an impact similar to that in the rational equilibrium. We therefore focus our analysis on the other extreme case and evaluate the welfare of the shrewd agent facing a fully cursed crowd. By continuity, the results extend to a sufficiently cursed environment.

We now discuss the effect of information policies on the welfare of a shrewd agent that plays against a fully cursed crowd (formalized in Proposition 11 of the Appendix). First, notice that W_1^{EQ} is independent of τ_p as fully cursed agents do not respond

[29] Indeed, the information spillover can be strong enough to make the shrewd agent in the cursed world better off than first-best welfare (as can be checked for $r = 0$, $\tau_\theta = \tau_y = 0.1$, $\tau_p = 0.19$, $c = 0.03$, where we have $W_1^R > W^\star$). By continuity, this holds for an open set of parameters.

to transparency. Therefore, higher transparency only affects the shrewd agent by providing a more precise aggregative signal, which is clearly beneficial. Recall that more public information τ_y and lower cost c are always beneficial for the cursed crowd (Proposition 7). The shrewd agent, however, suffers from the crowding out effect, as higher τ_y decreases information acquisition and dissemination by the cursed crowd. This is especially harmful if he is largely relying on this source of information, leading to the negative welfare impact when τ_s^R is small.

Higher c can be beneficial for the shrewd agent if τ_s^R is zero, i.e. if there is no direct effect of higher costs. Suppose that this is the case, that aggregate information is relatively abundant, and that $r < 0$ (strategic substitutes). As c increases, cursed agents rely more on y which makes it easier for the shrewd agent to anti-coordinate. This action externality can dominate the harm from reduced information dissemination. When τ_s^R is positive, by contrast, the direct effect always dominates, yielding the natural comparative static.

The comparison between Proposition 7 and the results just sketched highlights qualitative differences in the impact of policy on the welfare of cursed and shrewd agents. Transparency leaves the welfare of the cursed crowd unaffected but has a strictly positive (and large) impact for the shrewd. If both types affected the aggregate outcome, this could easily turn into a redistribution result, suggesting that transparency can function as an elitist policy that gives an advantage to sophisticated agents who are able to understand and utilize aggregative information. For public information and lower cost this trade-off is already apparent in the present results.[30] A natural avenue for studying these questions further would be to extend our model to include true cognitive heterogeneity, featuring several non-atomistic groups with different levels of cursedness, all affecting the aggregate outcomes. Although the linear structure of the model makes action aggregation straightforward, the correlation between information use and acquisition affects the aggregate outcome and introduces nonlinearity. The analysis of such cognitive heterogeneity is therefore beyond the scope of this paper.

7 Conclusion

This paper studies the effect of aggregative information focusing on the interplay of two key aspects: First, that the precision of such aggregate statistics as signals of the fundamental depends on the amount of private information present in individual actions; and second, agents' well-documented difficulty in making inference based on such signals as it requires inferring others' information from their actions.

[30] This conflict of interest between experts and unsophisticated actors casts doubt on the role of expert lobbying as a source of information on the impact of such policies.

We conduct our analysis in a beauty contest game with information acquisition, adapting a notion of cursed equilibrium to model agents limited understanding of aggregative information. Though parsimonious, the model is sufficiently rich to relate to existing literature and offer alternative explanation of well-established theoretical predictions such as the detrimental effect of public information and the irrelevance of transparency for informational efficiency. Since cursedness significantly alters the positive and normative results in our setting, it would be interesting to extend the analysis to more general payoff specifications as e.g. in Angeletos and Pavan (2007) and more deeply microfounded models yielding reduced forms similar to this class, as e.g. the business cycle model considered in Colombo et al. (2014) and demand function competition in Vives (2017).

We show that there is inefficiently low acquisition and use of private information in the rational benchmark due to an information dissemination externality. Cursed agents rely more heavily on their private information, which can push information acquisition towards (or even above) its efficient level. While cursedness creates inefficiencies in information use, the increased dissemination of private information initially dominates: a bit of individual cursedness is a collective blessing. Transparency crowds out the acquisition and use of private information but always increases the endogenous precision of the aggregative signal. This is the main driving force making it the only policy instrument with an unambiguously positive effect on welfare, despite being ineffective at increasing informational efficiency (Section 4) and causing potential redistributive concerns (Section 6). Combining our results then, a policymaker considering information policies while uncertain about the specifics of the environment must weigh the risk of doing harm against the risk of having only modest effects which primarily benefit sophisticated market participants. An analysis focusing on this cognitive heterogeneity is an interesting avenue for future research.

Incorporating information acquisition into a model of incorrect information use, such as cursed equilibrium, is the conceptual novelty of this paper. Doing so requires making an assumption on how such agents assess the value of information. Our notion, *cursed expectations equilibrium with information acquisition*, is based on the principle that agents correctly anticipate both the equilibrium strategies and how they will use their information, but mistakenly consider this use to be optimal. We operationalize this principle through a subjective envelope condition, which implies that information acquisition is proportional to its use, resulting in a tractable analysis. Additionally, we demonstrate that the subjective envelope condition characterizes the rest points of a simple learning process in which the misperception at the acquisition stage is the dual error of cursedness. While alternative notions do not conform to this principle or are not tractable in our setting, the properties and predictive power of such notions across applications of cursed equilibrium (and other behavioral equilibrium notions that do not easily allow a quasi-Bayesian analysis) remain an important question for future research.

A Learning Foundation

Consider the following learning process; the heuristics of it, as well as potential microfoundations, are given in the text. In each period t, agents submit a target level of information acquisition $\bar{\tau}_t$,[31] which is implemented with error. The realized level of information acquisition is given by $\tau_t = \bar{\tau}_t + \sigma\epsilon_t$, where ϵ_t is a zero mean random variable symmetrically distributed on support $[-1,1]$ according to distribution F with continuous density f. Moreover, implementation errors are serially independent ($\epsilon_{t_1} \perp \epsilon_{t_2}$ if $t_1 \neq t_2$). To ensure that the realized precision is nonnegative, we restrict the agent to $\bar{\tau}_t > \delta$ for a small $\delta > 0$ and assume throughout that σ is small enough that $\delta - \sigma > 0$. Since (19) holds, we can choose δ such that the correct τ_s can be learned.[32]

For each τ_t, the agent uses the realized signals according to the loadings $\boldsymbol{\alpha}_t$ that are the best response for precision $\bar{\tau}_t$ of the private signal — the other information parameters, as well as the equilibrium play $\boldsymbol{\delta}$ are fixed throughout the process. Denote the realized welfare by W_t. At the end of each period the agent records $x_t = (\bar{\tau}_t, \tau_t, W_t)$. At any point in time, the data-set of the agent consists of the whole history $x^t = (x_s)_{s\le t}$.

The agent proceeds in two steps. He evaluates whether to reoptimize the target precision $\bar{\tau}_t$ based on all previous observations with the same target precision. Concretely, consider the estimates at time t of the average welfare associated to a negative (resp, positive) implementation error,

$$W^-(\bar{\tau}, t) = \sum_{s\in T^-(\bar{\tau},t)} \frac{1}{|T^-(\bar{\tau}, t)|} W_s \tag{41}$$

$$W^+(\bar{\tau}, t) = \sum_{s\in T^+(\bar{\tau},t)} \frac{1}{|T^+(\bar{\tau}, t)|} W_s \tag{42}$$

where $T^+(\bar{\tau}, t) = \{s \le t | \bar{\tau}_s = \bar{\tau}, \tau_s > \bar{\tau}_s\}$ and $T^-(\bar{\tau}, t) = \{s \le t | \bar{\tau}_s = \bar{\tau}, \tau_s < \bar{\tau}_s\}$ denote the sets of periods with positive (resp. negative) deviations from target precision $\bar{\tau}$ before time t. Denote the expected value of these processes by $\overline{W^-}(\bar{\tau})$ and $\overline{W^+}(\bar{\tau})$, respectively. Let also

$$\Delta(\bar{\tau}, t) = \frac{1}{\sigma\mathbb{E}[|\epsilon|]} \left(W^+(\bar{\tau}, t) - W^-(\bar{\tau}, t)\right) \tag{43}$$

an estimate of the derivative with expected value $\overline{\Delta}(\bar{\tau})$.

[31] Since only the precision of the private signal τ_s is chosen, we avoid double subscript and write $\bar{\tau}_t$ (and τ_t for its realized value) in lieu of $\bar{\tau}_{s,t}, \tau_{s,t}$.

[32] Notice that our procedure fails to estimate the partial derivative at $\tau = 0$ as no negative precision can be implemented. Generally, the partial derivative of W does not exist at $\tau = 0$ and the subjective envelope condition (SE) is only well-defined as the limit of $\frac{\alpha_1^2(\tau)}{\tau^2}$, which is convergent. Therefore, any learning process that could estimate the value of information at $\tau = 0$ would need to make use of this relationship, which would be more complex for the agent as it requires knowledge of the structure of the problem.

Lemma 1. *For all $\bar{\tau} > \delta$, along all histories x^∞ where $\bar{\tau}$ is chosen infinitely often,*

$$\operatorname*{plim}_{t\to\infty} \Delta(\bar{\tau}, t) = \overline{\Delta}(\bar{\tau}) = \frac{\partial}{\partial \tau} W\Big|_{\alpha(\delta,\bar{\tau}),\delta,\bar{\tau}} + \sigma \frac{\mathbb{E}[\epsilon^2]}{2\mathbb{E}[|\epsilon|]} + o\left(\sigma^2\right). \tag{44}$$

The proofs of this and the subsequent Lemmata of this section are relegated to Online Appendix E.3.

The learning process proceeds as follows. The agent keeps the same target level of precision unless he has observed at least $K(\sigma)$ draws with negative and positive implementation error. This minimum sample size $K(\sigma)$ will be pinned down in the proof of Lemma 3. If the sample is sufficient he stays with his target if the estimated derivative is smaller than a tolerance level $B(\sigma) > 0$ which shrinks to zero with the implementation tremble. This tolerance level will be pinned down in the proof of Theorem 3 to ensure that the solution of the subjective envelope condition is a rest point. If instead a sufficient sample leads to an estimate outside the tolerance bound at time t, the target is rejected and the agent draws a new target level y_{t+1} from an absolutely continuous distribution (possibly) dependent on the full history x^t.[33] Accordingly, after arbitrary initialization $\bar{\tau}_0$, the process $(\tau_t)_{t\in\mathbb{N}}$ follows

$$\bar{\tau}_{t+1} = \begin{cases} \bar{\tau}_t & \text{if } |\Delta(\bar{\tau}_t, t)| < B(\sigma) \text{ or } \min\{|T^+(\bar{\tau}, t)|, |T^-(\bar{\tau}, t)|\} < K(\sigma) \\ y_{t+1} & \text{else} \end{cases} \tag{45}$$

Definition. We say that τ is a *rest point* of $\bar{\tau}_t$ if $\mathbb{P}(\{\bar{\tau}_s = \tau, \forall s \geq t\} | \bar{\tau}_t = \tau) > 0$.

Lemma 2. *If $|\Delta(\tau)| > B(\sigma)$, then τ cannot be a rest point of $\bar{\tau}_t$.*

Lemma 3. *If $|\Delta(\tau)| < \frac{1}{2}B(\sigma)$, then τ is a rest point of $\bar{\tau}_t$.*

Proof of Theorem 3: Let τ be the solution to the subjective envelope condition. Then, set $B(\sigma) = 2\sup_{s\leq\sigma} \left|\overline{\Delta}(\tau)\right|$. Note that $B(\sigma)$ is decreasing in σ and that $B(\sigma) > 0$ since the partial derivatives in the Taylor expansion in the proof of Lemma 1 are bounded away from zero on the domain of evaluation. Hence, $B(\sigma)$ satisfies is a valid tolerance level.

Then, by Lemma 3, τ is a rest point of the learning process.

Conversely, consider a point τ that does not solve the subjective envelope condition, i.e. $|\frac{\partial W}{\partial \tau_s}(\tau)| > 0$. Then, by Lemma 1, there exists an $S > 0$ such that for any $\sigma < S$ we have $|\overline{\Delta}(\tau)| > \frac{1}{2}|\frac{\partial W}{\partial \tau_s}(\tau)|$. Furthermore, there exists an $S' > 0$ such that $B(\sigma) < \frac{1}{2}|\frac{\partial W}{\partial \tau_s}(\tau)|$ for $\sigma < S'$. Then, by Lemma 2, τ cannot be a rest point of the learning process for any $\sigma < \min\{S, S'\}$. □

[33] Note that we allow for a quite general (stochastic) reoptimization process, since the structure of this process is not essential for our results. All that is required given our definition of a rest-point below is that the reoptimization does not stubbornly return to points that have been rejected arbitrarily often, which is assured by our assumption of absolute continuity.

B Proofs

We relegate lengthy computational proofs of all Lemmas and of Proposition from the text which are not by themselves main results to Online Appendix E.4 and E.5, resp.

Proof of Theorem 1: Recall from the text that $\delta_i = \alpha_i$ and the best response

$$a_i = (1-r)\left(\chi\frac{\tau_s s_i + \tau_y y}{\tau_\theta + \tau_y + \tau_s} + (1-\chi)\frac{\tau_s s_i + \tau_y y + \delta_1^2\tau_p\frac{1-\delta_3}{\delta_1}\left[p - \frac{\delta_2}{1-\delta_3}y\right]}{\tau_\theta + \tau_y + \tau_s + \delta_1^2\tau_p}\right)$$
$$+ r\left(\alpha_0 + \alpha_1\left(\chi\frac{\tau_s s_i + \tau_y y}{\tau_\theta + \tau_y + \tau_s} + (1-\chi)\,\frac{\tau_s s_i + \tau_y y + \delta_1^2\tau_p\frac{1-\delta_3}{\delta_1}\left[p - \frac{\delta_2}{1-\delta_3}y\right]}{\tau_\theta + \tau_y + \tau_s + \delta_1^2\tau_p}\right) + \alpha_2 y + \alpha_3 p\right)$$

Matching coefficients with (4) and simplifying, we arrive at the desired expressions. From (11) we arrive at the equilibrium condition

$$0 = f(\delta_1) := \delta_1\left(\tau_\theta + \tau_y + \delta_1^2\tau_p\right) - \chi\frac{\tau_s}{\tau_\theta + \tau_y + \tau_s}\left[1 - r + r\delta_1\right]\delta_1^2\tau_p - (1-r)(1-\delta_1)\tau_s \tag{46}$$

To show that the solution is unique, note that $\frac{f(\delta_1)}{\tau_\theta+\tau_y+\tau_s+\delta_1^2\tau_p}$ evaluated at $\delta_1 = 0$ is equal to $-\frac{1-r}{\tau_\theta+\tau_y+\tau_s} < 0$ and at $\delta_1 = 1$ it is greater than $1 - \frac{\tau_s}{\tau_\theta+\tau_y+\tau_s} > 0$. Hence, there is at least one root in $(0,1)$. Furthermore, the expression is increasing at a root, as

$$\frac{\partial}{\partial\delta_1}\frac{f(\delta_1)}{\tau_\theta + \tau_y + \tau_s + \delta_1^2\tau_p} = 1 - r\tau_s\left[\frac{\chi}{\tau_\theta + \tau_y + \tau_s} + \frac{1-\chi}{\tau_\theta + \tau_y + \tau_s + \delta_1^2\tau_p}\right] + 2(1-\chi)\delta_1\tau_p\tau_s\frac{1 - r + r\delta_1}{\left(\tau_\theta + \tau_y + \tau_s + \delta_1^2\tau_p\right)^2}$$

where the first two terms are in sum positive and, at a root, we have $\operatorname{sgn}\{\delta_1\left[(1-r) + r\delta_1\right]\} = 1$ whence the final term is also positive. □

Corollary 1. *In equilibrium, we have*

$$(1-r) + r\delta_1 = \frac{\delta_1}{\tau_s\frac{\tau_\theta+\tau_y+\tau_s+\chi\delta_1^2\tau_p}{(\tau_\theta+\tau_y+\tau_s)(\tau_\theta+\tau_y+\tau_s+\delta_1^2\tau_p)}} > 0. \tag{47}$$

Proof of Theorem 2: Equation (17) is derived in the text assuming that $\delta_1 \geq 0$. There cannot be an equilibrium with $\delta_1 < 0$.

Lemma 4. *There is no equilibrium with information acquisition and $\delta_1 < 0$.*

Our system is defined by

$$f(\delta_1, \tau_s) = \delta_1\left(\tau_\theta + \tau_y + \delta_1^2\tau_p\right) - \chi\frac{\tau_s\left[(1-r) + r\delta_1\right]\delta_1^2\tau_p}{\tau_\theta + \tau_y + \tau_s} - (1-r)(1-\delta_1)\tau_s = 0 \tag{48}$$

$$g(\delta_1, \tau_s) = \delta_1^2 - c\cdot\tau_s^2 = 0 \tag{49}$$

substituting g into f we obtain that loading δ_1 must solve

$$0 = \tilde{f}(\delta_1) := \delta_1 + \sqrt{c}\left(\tau_\theta + \tau_y + \delta_1^2\tau_p\right) - [(1-r) + r\delta_1]\left(1 + \chi\frac{\sqrt{c}\delta_1^2\tau_p}{\delta_1 + \sqrt{c}\left(\tau_\theta + \tau_y\right)}\right) \tag{50}$$

Lemma 5. *We have $\tilde{f}_{\delta_1} > 0$ for all $\delta_1 \in (0,1)$.*

Moreover, $\tilde{f}(0) = \sqrt{c}\left(\tau_\theta + \tau_y\right) - (1-r)$, which is negative if and only if condition (19) holds. In that case, therefore, a solution to (50) exists. Else, we are in a corner case $\delta_1 = \tau_s = 0$. We obtain $\delta_2 = \frac{\tau_y}{\tau_\theta + \tau_y}$ and $\delta_3 = 0$ by plugging $\delta_1 = \tau_s = 0$ into the original matching coefficients equations. □

Proof of Proposition 1: Implicitly differentiating the equilibrium system (48)-(49) we get, for a generic parameter ν

$$\frac{\mathrm{d}\delta_1}{\mathrm{d}\nu} = \frac{g_{\tau_s}f_\nu - f_{\tau_s}g_\nu}{g_\delta f_{\tau_s} - g_{\tau_s}f_\delta}, \quad \frac{\mathrm{d}\tau_s}{\mathrm{d}\nu} = \frac{f_\delta g_\nu - g_\delta f_\nu}{g_\delta f_{\tau_s} - g_{\tau_s}f_\delta}$$

Lemma 6. *In equilibrium, we have $g_\delta f_{\tau_s} - g_{\tau_s}f_\delta > 0$, and hence $\frac{\mathrm{d}\delta_1}{\mathrm{d}\nu} \propto g_{\tau_s}f_\nu - f_{\tau_s}g_\nu$.*

Hence, we have

$$\frac{\mathrm{d}\delta_1}{\mathrm{d}\chi} \propto g_{\tau_s}f_\chi - f_{\tau_s}g_\chi = g_{\tau_s}f_\chi = -2c\tau_s\left(-\frac{\tau_s}{\tau_\theta + \tau_y + \tau_s}[1 + r(\delta_1 - 1)]\delta_1^2\tau_p\right) > 0.$$

$$\frac{\mathrm{d}\delta_1}{\mathrm{d}\tau_p} \propto g_{\tau_s}f_{\tau_p} - f_{\tau_s}g_{\tau_p} = g_{\tau_s}f_{\tau_p} < 0.$$

since $f_{\tau_p} > 0$. Plugging (17) into (12), we obtain

$$\delta_2 = \frac{\sqrt{c}\delta_1\tau_y}{(1-r) - \sqrt{c}\left(\tau_\theta + \tau_y\right)}, \quad \delta_3 = 1 - \frac{\delta_1(1-r)}{(1-r) - \sqrt{c}\left(\tau_\theta + \tau_y\right)}, \tag{51}$$

from which it follows immediately that $\frac{\mathrm{d}\delta_2}{\mathrm{d}\chi} \propto \frac{\mathrm{d}\delta_1}{\mathrm{d}\chi} > 0$, $\frac{\mathrm{d}\delta_3}{\mathrm{d}\chi} \propto -\frac{\mathrm{d}\delta_1}{\mathrm{d}\chi} < 0$, $\frac{\mathrm{d}\delta_2}{\mathrm{d}\tau_p} \propto \frac{\mathrm{d}\delta_1}{\mathrm{d}\tau_p} > 0$, $\frac{\mathrm{d}\delta_3}{\mathrm{d}\tau_p} \propto -\frac{\mathrm{d}\delta_1}{\mathrm{d}\tau_p} < 0$, and $\frac{\mathrm{d}}{\mathrm{d}\chi}\frac{\delta_2}{\delta_1} = \frac{\mathrm{d}}{\mathrm{d}\chi}\frac{1-\delta_3}{\delta_1} = 0$. Finally, we are left to prove that $\frac{\partial}{\partial\tau_p}\delta_1^2\tau_p > 0$. Notice

$$\frac{\mathrm{d}}{\mathrm{d}\tau_p}\tau_p\delta_1^2 = \frac{2\delta_1^4}{g_\delta f_{\tau_s} - g_{\tau_s}f_\delta}\left\{1 - r + \frac{\sqrt{c}\delta_1^2\tau_p(1-\chi r)}{\delta_1 + \sqrt{c}\left(\tau_\theta + \tau_y\right)} - \frac{\sqrt{c}\delta_1^2\tau_p(1-\chi)}{\delta_1 + \sqrt{c}\left(\tau_\theta + \tau_y + \chi\delta_1^2\tau_p\right)}\right\}$$

which follows from lengthy computation involving (47) and (17). Clearly, $1 - r > 0$ so it remains to show that the last two terms sum to a positive expression. This, however is immediate since $(1 - \chi r) > (1 - \chi) < \chi \iff r < 1$, which is assumed. □

Proof of Proposition 2: Let us begin with δ_1. Note that for all precision parameters, we have $g_\tau \equiv 0$, and for costs, we have $f_c \equiv 0$. For the cost parameter we have $\frac{d\delta_1}{dc} \propto f_{\tau_s}(\tau_s)^2 < 0$ as $f_{\tau_s} < 0$. Moreover, $\frac{d\delta_1}{d\tau_y} \propto g_{\tau_s} f_{\tau_y} < 0$ since $f_{\tau_y} > 0$, $\frac{d\delta_1}{d\tau_\theta} = \frac{\partial\delta_1}{\partial\tau_y} < 0$, and finally $\frac{d\delta_1}{dr} \propto g_{\tau_s} f_r < 0$ since $f_r \geq 0$. Using (17), we have $\frac{d\tau_s}{dv} \propto \frac{d\delta_1}{dv}$ for all $v \neq c$. For c, we get $\frac{d\tau_s}{dc} = \frac{1}{c}\frac{d\delta_1}{dc} - \tau_s c < 0$. □

Proof of Proposition 3: By plugging (17) into (67) we obtain the expressions for the fundamental loadings. For β and γ_2 the comparative statics are immediate. In particular, γ_2 is increasing in τ_y, c and r. Lengthy calculations, available upon request, show that instead δ_2 has ambiguous comparative statics in all τ_y, c and r. Notice that $\gamma_3 = \frac{\delta_3}{1-\delta_3}$ is an increasing transformation of the direct loading δ_3, so it is increasing in τ_p and carries all the following ambiguous comparative statics. For τ_y, consider

$$\lim_{\chi\to 1} \frac{\frac{\partial\delta_3}{\partial\tau_y}}{\delta_3} \propto \frac{\left(1-\delta_1(1-r)-r-\sqrt{c}\left(\tau_\theta+\tau_y\right)\right)}{1-\frac{\delta_1(1-r)}{1-r-\sqrt{c}\left(\tau_\theta+\tau_y\right)}} = 1-r-\sqrt{c}\left(\tau_\theta+\tau_y\right) > 0$$

proving $\frac{\partial\delta_3}{\partial\tau_y}$ converges to 0 from above as $\chi \to 1$. Therefore, δ_3 is increasing in τ_y for large χ. However, in the limit as $r \to 1-\sqrt{c}\left(\tau_\theta+\tau_y\right)$, then $\frac{\partial\delta_3}{\partial\tau_y} \to -\frac{1-r}{\tau_\theta+\tau_y}\delta_1 < 0$. To see that $\frac{\partial\delta_3}{\partial c}$ is of ambiguous sign, consider the limit as $\sqrt{c} \to \frac{1-r}{\tau_\theta+\tau_y}$. Then, we have $\operatorname{sgn}\left\{\frac{\partial\delta_3}{\partial c}\right\} \to \operatorname{sgn}\left\{-(1-r)^3\right\} < 0$. Furthermore, as $c \to 0$, we have

$$\frac{\partial\delta_3}{\partial c} \propto \delta_1^2(1-r)\left((1-\delta_1)\left(\tau_\theta+\tau_y\right)+(1-\chi)\delta_1^2\tau_p\right) > 0$$

Following similar arguments, $\frac{\partial\delta_3}{\partial r}$ is ambiguous: In the limit as $\sqrt{c} \to \frac{1-r}{\tau_\theta+\tau_y}$, we have $\operatorname{sgn}\left\{\frac{\partial\delta_3}{\partial r}\right\} \to \operatorname{sgn}\left\{-(1-r)^3\right\} < 0$. As $r \to -\infty$, we get

$$\frac{\partial\delta_3}{\partial r} \propto -r(1-\delta_1)\left(\delta_1+\sqrt{c}\left(\tau_\theta+\tau_y+\chi\delta_1^2\tau_p\right)\right)^2 > 0.$$ □

Proof of Proposition 15: See Online Appendix E.5. □

Proof of Theorem 4: Taking FOC in (27), we obtain

$$\begin{aligned}
W_{\delta_1} &= 2\frac{(1-r)}{(1-\delta_3)^2}\frac{(1-\delta_1-\delta_2-\delta_3)}{\tau_\theta} - 2\frac{\delta_1}{\tau_s} = 0 \\
W_{\delta_2} &= -\frac{(1-r)}{(1-\delta_3)^2}\left\{2\frac{\delta_2}{\tau_y} - 2\frac{(1-\delta_1-\delta_2-\delta_3)}{\tau_\theta}\right\} = 0 \\
W_{\delta_3} &= -\frac{2(1-r)}{(1-\delta_3)^3}\left\{\frac{\delta_2^2}{\tau_y} + \frac{\delta_3}{\tau_p} - \frac{(1-\delta_1-\delta_2-\delta_3)(\delta_1+\delta_2)}{\tau_\theta}\right\} = 0
\end{aligned} \tag{52}$$

Note that the last two equations simplify to a linear system in δ_2, δ_3, which we solve to obtain

$$\delta_2 = \frac{\tau_y(1-\delta_1)}{\tau_\theta + \tau_y + \tau_p \delta_1}, \qquad \delta_3 = \frac{\delta_1(1-\delta_1)\tau_p}{\tau_\theta + \tau_y + \tau_p \delta_1} \tag{53}$$

Simplifying and using the envelope condition we arrive at the defining equation for the efficient outcome

$$f^\star(\delta_1) = \left(\tau_\theta + \tau_y + \delta_1^2 \tau_p\right)\left(\frac{\tau_\theta + \tau_y + \tau_p \delta_1^2}{\tau_\theta + \tau_y + \tau_p \delta_1}\right) - (1-r)\frac{1}{\sqrt{c}}(1-\delta_1) = 0$$

Lemma 7. *We have $f^\star_{\delta_1} > 0$ for all $\delta_1 \in (0,1)$.*

By Lemma 7, there is a unique interior solution, as $f^\star(0) = \left(\tau_\theta + \tau_y\right) - (1-r)\frac{1}{\sqrt{c}} < 0$ by (19) and $f^\star(1) = \tau_\theta + \tau_y + \tau_p > 0$. Reformulating $f^\star(\delta_1) = 0$, we get the desired representation (29). Finally, plugging the efficiency conditions (53) and $\tau_s^\star = \frac{\delta_1^\star}{\sqrt{c}}$ into the welfare expression (28) we get

$$W^\star := W\left(\delta^\star, \tau_s^\star\right) = \max_{\delta_1} -2\sqrt{c}\delta_1 - \frac{(1-r)(1-\delta_1)^2}{\tau_\theta + \tau_y + \delta_1^2 \tau_p}. \tag{54}$$

Applying the envelope theorem, all the comparative statics (30) follow. □

Proof of Proposition 4: Since $f^\star_{\delta_1} > 0$ (Lemma 7), we know that $f^\star(\delta_1) < 0$ implies that there is underacquisition and $f^\star(\delta_1) > 0$ implies that there is overacquisition. Plugging the equilibrium δ_1 and using the fact that $f(\delta_1) = 0$

$$f^\star(\delta_1) = \left(\tau_\theta + \tau_y + \delta_1^2 \tau_p\right)\left(\frac{\tau_\theta + \tau_y + \delta_1^2 \tau_p}{\tau_\theta + \tau_y + \delta_1 \tau_p} - 1\right) + \chi \frac{1}{\sqrt{c}} \frac{1}{\tau_\theta + \frac{\delta_1}{\sqrt{c}} + \tau_y}\left[1 + r(\delta_1 - 1)\right]\delta_1^2 \tau_p$$

Note that at $\chi = 0$, this expression is negative and hence, δ_1 is inefficiently low. As δ_1 is increasing in χ, we are below the efficient initially, but may exceed it for χ sufficiently large. There exists a χ with $\delta_1^\chi = \delta_1^\star$ iff $f^\star(\delta_1^{FC}) > 0$ (by $f^\star_\delta > 0$). We get

$$f^\star(\delta_1^{FC}) = \frac{\delta_1 \tau_p}{\tau_\theta + \tau_y + \delta_1 \tau_p}\left\{2\left(\tau_\theta + \tau_y\right)\delta_1 + \delta_1^3 \tau_p - \left(\tau_\theta + \tau_y\right)\right\}$$

This is larger than zero iff $\tau_p \geq \frac{\left(\tau_\theta + \tau_y\right) - 2\left(\tau_\theta + \tau_y\right)\delta_1}{\delta_1^3}$. Since the cutoff is decreasing in δ_1 (and δ_1 is decreasing in χ), we obtain the sufficient bound $\bar{\tau}_p$.[34] □

Proof of Proposition 5. See Online Appendix (E.5). □

[34]In particular, if $\delta_1^{FC} \geq \frac{1}{2}$, i.e. $\frac{\sqrt{c}\left(\tau_\theta + \tau_y\right)}{1-r} \leq \frac{1}{2}$, the fully cursed agents always overacquires.

Proof of Proposition 6: To determine the impact of cursedness, we compute

$$\frac{\mathrm{d}W^{EQ}}{\mathrm{d}\chi} = \frac{\partial W^{EQ}}{\partial \delta_1}\frac{\mathrm{d}\delta_1}{\mathrm{d}\chi} + \frac{\partial W^{EQ}}{\partial \delta_2}\frac{\mathrm{d}\delta_2}{\mathrm{d}\chi} + \frac{\partial W^{EQ}}{\partial \delta_3}\frac{\mathrm{d}\delta_3}{\mathrm{d}\chi}$$

local to $\chi = 0$, by Proposition 5, $\frac{\mathrm{d}W^{EQ}}{\mathrm{d}\chi} = \frac{\partial W^{EQ}}{\partial \delta_1} \cdot \frac{\mathrm{d}\delta_1}{\mathrm{d}\chi}$. Therefore,

$$\frac{\mathrm{d}W^{EQ}}{\mathrm{d}\chi}|_{\chi=0} = \left\{ \frac{(1-r)}{(1-\delta_3)^2} \left\{ 2\frac{(1-\delta_1-\delta_2-\delta_3)}{\tau_\theta} \right\} - 2\sqrt{c} \right\} \frac{\mathrm{d}\delta_1}{\mathrm{d}\chi}|_{\chi=0}$$

and applying the rational δ_2, δ_3, $f(\delta_1) = 0$ and the envelope condition gives

$$\begin{aligned} \frac{\mathrm{d}W^{EQ}}{\mathrm{d}\chi}|_{\chi=0} =& \left\{ \frac{(1-r)}{(1-\delta_3)^2} \left\{ 2\frac{(1-\delta_1-\delta_2-\delta_3)}{\tau_\theta} \right\} - 2\sqrt{c} \right\} \frac{\mathrm{d}\delta_1}{\mathrm{d}\chi}|_{\chi=0} \\ =& 2\left\{ \frac{\left((1-r)\tau_s + \delta_1^2\tau_p\right)}{(1-r)\,\tau_s^2}\delta_1 - \sqrt{c} \right\} \frac{\mathrm{d}\delta_1}{\mathrm{d}\chi}|_{\chi=0} = 2\left\{ \frac{\delta_1\tau_p c}{(1-r)} \right\} \frac{\mathrm{d}\delta_1}{\mathrm{d}\chi}|_{\chi=0} > 0. \end{aligned}$$

In the fully cursed case $\chi = 1$, using Proposition 5 we know that $\frac{\mathrm{d}W^{EQ}}{\mathrm{d}\chi} = \frac{\partial W^{EQ}}{\partial \delta_3}\frac{\mathrm{d}\delta_3}{\mathrm{d}\chi}$. Plugging in the fully cursed weights δ_1^{FC}, δ_2^{FC}, and δ_3^{FC} into (52) yields $W_{\delta_3} = 2\sqrt{c}\delta_1^{FC}$. Therefore, we have $\frac{\mathrm{d}W^{EQ}}{\mathrm{d}\chi}|_{\chi=1} = \sqrt{c}\delta_1^{FC}\frac{d\delta_3}{d\chi}|_{\chi=1} < 0$, where $\frac{d\delta_3}{d\chi}|_{\chi=1} < 0$ follows from

$$\frac{\mathrm{d}\delta_3}{\mathrm{d}\chi}|_{\chi=1} = \left(\frac{\partial\delta_3}{\partial\delta_1}\frac{\partial\delta_1}{\partial\chi} \right)|_{\chi=1} + \frac{\partial\delta_3}{\partial\chi}|_{\chi=1} = 0 - \frac{\delta_1^2\tau_p\left(\tau_\theta+\tau_y+\tau_s\right)}{(1-r)\,\tau_s\left(\tau_\theta+\tau_y+\tau_s+\delta_1^2\tau_p\right)} < 0. \qquad \square$$

Proof of Proposition 7: At $\chi = 0$, $\chi = 1$ we have $W = -\sqrt{c}(1+\delta_1)$. Hence, $\frac{dW}{d\tau} \propto -\frac{d\delta_1}{d\tau}$ and the comparative statics w.r.t. τ_s follow immediately from Proposition 1. For costs, note that in the rational case, direct computation yields

$$\frac{\partial W_0^{EQ}}{\partial c} = -\frac{1-r-\sqrt{c}\left(\tau_\theta+\tau_y\right)+\sqrt{c}\delta_1\tau_p}{\sqrt{c}(1-r)+2c\delta_1\tau_p} < 0$$

which is negative by the parameter condition (19). In the fully cursed case, note that

$$\frac{\partial W_1^{EQ}}{\partial c} = -\frac{1}{\sqrt{c}}\left(\frac{(1-r)-\sqrt{c}\left(\tau_\theta+\tau_y\right)}{1-r} \right) = -\frac{\delta_1}{\sqrt{c}} < 0$$

For $\tau_p = 0$, the rational and (partially) cursed equilibria coincide, hence the above comparative statics prevail, and by continuity, this extends to small but interior τ_p. The paradoxical welfare results emerge instead for large τ_p. To make them apparent we consider the transparent limit case $\tau_p \to \infty$. By continuity, the following comparative statics hold for sufficiently large τ_p.

Lemma 8. *The limit welfare (71) is*

- *Decreasing in τ_θ, τ_y if and only if $r > 0$ and $\chi \leq \frac{2r-1}{r}$.*
- *Decreasing in costs, unless $r < 0$, when higher costs increase welfare for intermediate χ.*

It remains to be shown that welfare is always increasing in τ_p. Note that

$$\frac{\mathrm{d}W}{\mathrm{d}\tau_p} = \frac{\partial W}{\partial \tau_p} + \frac{\partial W}{\partial \delta_1}\frac{\mathrm{d}\delta_1}{\mathrm{d}\tau_p} + \frac{\partial W}{\partial \delta_2}\frac{\mathrm{d}\delta_2}{\mathrm{d}\tau_p} + \frac{\partial W}{\partial \delta_3}\frac{\mathrm{d}\delta_3}{\mathrm{d}\tau_p} + \frac{\partial W}{\partial \tau_s}\frac{\mathrm{d}\tau_s}{\mathrm{d}\tau_p} = \frac{\partial W}{\partial \tau_p} + \frac{\partial W}{\partial \delta_1}\frac{d\delta_1}{d\tau_p} + \frac{\partial W}{\partial \delta_3}\frac{d\delta_3}{d\tau_p}$$

Since we proved that $\frac{\partial W}{\partial \delta_2} = \frac{\partial W}{\partial \tau_s} = 0$ for every χ. Simplifying this expression and plugging in for $\frac{d\delta_1}{d\tau_p}$, we obtain an expression that, after removing clearly signed factors, is proportional to a sum of positive addenda plus

$$c\chi\delta_1^4\tau_p^2\left(1+\chi+\chi^2 r - 3\chi r\right),$$

but $h(\chi, r) := 1+\chi+\chi^2 r - 3\chi r \geq 0$ whenever $r \leq 1$ and $\chi \in [0,1]$,[35] as desired. □

References

Amador, Manuel and Pierre-Olivier Weill, "Learning from Prices: Public Communication and Welfare," *Journal of Political Economy*, 2010, *118* (5), 866–907.

_ **and** _ , "Learning from Private and Public Observations of Others' Actions," *Journal of Economic Theory*, May 2012, *147* (3), 910–940.

Angeletos, George-Marios and Alessandro Pavan, "Efficient Use of Information and Social Value of Information," *Econometrica*, 2007, *75* (4), 1103–1142.

_ **and Chen Lian**, "Forward Guidance without Common Knowledge," *American Economic Review*, September 2018, *108* (9), 2477–2512.

_ **and Karthik Sastry**, "Inattentive Economies," 2023.

Angeletos, Marios and Jennifer La'O, "Noisy Business Cycles," *NBER Macroeconomics Annual 2009*, April 2010, *24*, 319–378.

Barberis, Nicholas, "Psychology-Based Models of Asset Prices and Trading Volume," in "Handbook of Behavioral Economics: Applications and Foundations 1," Vol. 1, Elsevier, 2018, pp. 79–175.

Bayona, Anna, "The Social Value of Information with an Endogenous Public Signal," *Economic Theory*, December 2018, *66* (4), 1059–1087.

Bohren, Aislinn and Daniel Hauser, "The Behavioral Foundations of Model Misspecification," 2023.

[35]If $r < 0$, then $h(0,r) = 1 > 0$ and $\frac{\partial}{\partial \chi}h(\chi,r) = 1 + 2\chi r - 3r = 1 + r(2\chi - 3) > 0$. If $r \in (0,1)$, then $1+\chi+\chi^2 r - 3\chi r > r\left(1+\chi+\chi^2\right) - 3\chi r = r\left(1-2\chi+\chi^2\right) = r(1-\chi)^2 > 0$.

Bohren, J. Aislinn and Daniel N. Hauser, "Learning With Heterogeneous Misspecified Models: Characterization and Robustness," *Econometrica*, 2021, *89* (6), 3025–3077.

Cavallo, Alberto, Guillermo Cruces, and Ricardo Perez-Truglia, "Inflation Expectations, Learning, and Supermarket Prices: Evidence from Survey Experiments," *American Economic Journal: Macroeconomics*, July 2017, *9* (3), 1–35.

Chowdhry, Bhagwan and Vikram Nanda, "Multimarket Trading and Market Liquidity," *The Review of Financial Studies*, July 1991, *4* (3), 483–511.

Cohen, Shani and Shengwu Li, "Sequential Cursed Equilibrium," February 2023.

Colombo, L., G. Femminis, and A. Pavan, "Information Acquisition and Welfare," *The Review of Economic Studies*, October 2014, *81* (4), 1438–1483.

Diamond, Douglas W. and Robert E. Verrecchia, "Information Aggregation in a Noisy Rational Expectations Economy," *Journal of Financial Economics*, September 1981, *9* (3), 221–235.

Esponda, Ignacio, "Behavioral Equilibrium in Economies with Adverse Selection," *American Economic Review*, August 2008, *98* (4), 1269–1291.

_ **and Demian Pouzo**, "Berk-Nash Equilibrium: A Framework for Modeling Agents With Misspecified Models," *Econometrica*, 2016, *84* (3), 1093–1130.

_ **and Emanuel Vespa**, "Hypothetical Thinking and Information Extraction in the Laboratory," *American Economic Journal: Microeconomics*, 2014, *6* (4), 180–202.

Eyster, Erik, "Errors in Strategic Reasoning," in "Handbook of Behavioral Economics: Applications and Foundations 1," Vol. 2, Elsevier, 2019, pp. 187–259.

_ **and Matthew Rabin**, "Cursed Equilibrium," *Econometrica*, September 2005, *73* (5), 1623–1672.

_ **,** _ **, and Dimitri Vayanos**, "Financial Markets Where Traders Neglect the Informational Content of Prices," *The Journal of Finance*, 2019, *74* (1), 371–399.

Fong, Meng-Jhang, Po-Hsuan Lin, and Thomas R. Palfrey, "Cursed Sequential Equilibrium," April 2023.

Frick, Mira, Ryota Iijima, and Yuhta Ishii, "Belief Convergence under Misspecified Learning: A Martingale Approach," *The Review of Economic Studies*, March 2023, *90* (2), 781–814.

Glosten, Lawrence R. and Paul R. Milgrom, "Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders," *Journal of Financial Economics*, March 1985, *14* (1), 71–100.

Grossman, Sanford J. and Joseph E. Stiglitz, "On the Impossibility of Informationally Efficient Markets," *The American Economic Review*, 1980, *70* (3), 393–408.

Hebert, Benjamin and Jennifer La'O, "Information Acquisition, Efficiency, and Non-Fundamental Volatility," *Journal of Political Economy*, forthcoming.

Heidhues, Paul, Botond Kőszegi, and Philipp Strack, "Unrealistic Expectations and Misguided Learning," *Econometrica*, 2018, *86* (4), 1159–1214.

Hellwig, Christian and Laura Veldkamp, "Knowing What Others Know: Coordination Motives in Information Acquisition," *The Review of Economic Studies*, January 2009, *76* (1), 223–251.

Kagel, John H. and Dan Levin, *Common Value Auctions and the Winner's Curse*, Princeton, N.J: Princeton University Press, 2002.

Kool, Clemens, Menno Middeldorp, and Stephanie Rosenkranz, "Central Bank Transparency and the Crowding Out of Private Information in Financial Markets," *Journal of Money, Credit and Banking*, June 2011, *43* (4), 765–774.

Kyle, Albert S., "Continuous Auctions and Insider Trading," *Econometrica*, November 1985, *53* (6), 1315.

Lucas, Robert E., "Expectations and the Neutrality of Money," *Journal of Economic Theory*, April 1972, *4* (2), 103–124.

Morris, Stephen and Hyun Song Shin, "Social Value of Public Information," *American Economic Review*, November 2002, *92* (5), 1521–1534.

_ **and** _ , "Central Bank Transparency and the Signal Value of Prices," *Brookings Papers on Economic Activity*, 2005, *36* (2), 1–66.

Ngangoué, M. Kathleen and Georg Weizsäcker, "Learning from Unrealized versus Realized Prices," *AEJ: Microeconomics*, May 2021, *13* (2), 174–201.

Pagano, Marco and Ailsa Röell, "Transparency and Liquidity: A Comparison of Auction and Dealer Markets with Informed Trading," *The Journal of Finance*, 1996, *51* (2), 579–611.

Shadmehr, Mehdi, Sepehr Shahshahani, and Charles Cameron, "Coordination and Innovation in Judiciaries: Correct Law versus Consistent Law," *Quarterly Journal of Political Science*, January 2022, *17* (1), 61–89.

Svensson, Lars E. O., "Social Value of Public Information: Comment: Morris and Shin (2002) Is Actually Pro-Transparency, Not Con," *American Economic Review*, March 2006, *96* (1), 448–452.

Uhlig, Harald, "A Law of Large Numbers for Large Economies," *Economic Theory*, February 1996, *8* (1), 41–50.

Ui, Takashi and Yasunori Yoshizawa, "Characterizing Social Value of Information," *Journal of Economic Theory*, July 2015, *158*, 507–535.

Vives, Xavier, "Aggregation of Information in Large Cournot Markets," *Econometrica*, 1988, *56* (4), 851–876.

_ , *Information and Learning in Markets: The Impact of Market Microstructure*, Princeton: Princeton University Press, 2008.

_ , "On the Possibility of Informationally Efficient Markets," *Journal of the European Economic Association*, October 2014, *12* (5), 1200–1239.

_ , "Endogenous Public Information and Welfare in Market Games," *The Review of Economic Studies*, April 2017, *84* (2), 935–963.

Weizsäcker, Georg, "Do We Follow Others When We Should? A Simple Test of Rational Expectations," *American Economic Review*, December 2010, *100* (5), 2340–2360.

Wong, Jacob, "Information Acquisition, Dissemination, and Transparency of Monetary Policy," *The Canadian Journal of Economics*, 2008, *41* (1), 46–79.

Woodford, Michael, "Central Bank Communication and Policy Effectiveness," *Proceedings - Economic Policy Symposium - Jackson Hole*, 2005, (Aug), 399–474.

Online Appendix

Acquisition, (Mis)use and Dissemination of Information: The Blessing of Cursedness and Transparency

C The Quasi-Bayesian Approach(For Online Publication)

In this appendix, we study an alternative models of information acquisition. Instead of using the subjective envelope condition, the agent maximizes the expectation of his interim cursed-expected welfare. Formally, the interim χ-cursed posterior (Eyster and Rabin, 2005) is given by the mixture distribution $\theta|s_i,y,p \sim \mu_\chi := \chi\mathcal{N}\left(\frac{\tau_y y+\tau_s s_i}{\tau_\theta+\tau_y+\tau_s}, \frac{1}{\tau_\theta+\tau_y+\tau_s}\right)+ (1-\chi)\mathcal{N}\left(\frac{\tau_y y+\tau_s s_i+\delta_1^2\tau_p\hat{p}}{\tau_\theta+\tau_y+\tau_s+\delta_1^2\tau_p}, \frac{1}{\tau_\theta+\tau_y+\tau_s+\delta_1^2\tau_p}\right)$. Interim welfare is then given by

$$\begin{aligned} w(s_i,y,p) &= \mathbb{E}_{\mu_\chi}\left[-(1-r)(a_i-\theta)^2 - r(a_i-\bar{a})^2\right] \\ &= a_i^2 - (1-r)\left(\mathbb{V}_{\mu_\chi}[\theta] + \mathbb{E}^2_{\mu_\chi}[\theta]\right) - r\left(\mathbb{V}_{\mu_\chi}[\bar{a}] + \mathbb{E}^2_{\mu_\chi}[\bar{a}]\right) \end{aligned}$$

where a_i denotes the χ-cursed optimal action (4) which is measurable w.r.t. (s_i,y,p). To estimate the value of private information ex-ante, the agent requires a joint prior over the signals (that also takes into account their correlation through θ). Note that these beliefs over signal realizations do not depend on cursedness. Cursedness only affects the inference from the aggregative signal to the state, not the perceived distribution over signals having integrated out the state.[36] The prior distribution is therefore given by

$$\begin{pmatrix} s_i \\ y \\ p \end{pmatrix} \sim \mu(\tau_s) := \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{1}{\tau_\theta}+\frac{1}{\tau_s} & \frac{1}{\tau_\theta} & \left(\frac{\delta_1+\delta_2}{1-\delta_3}\right)\frac{1}{\tau_\theta} \\ \frac{1}{\tau_\theta} & \frac{1}{\tau_\theta}+\frac{1}{\tau_y} & \left(\frac{\delta_1+\delta_2}{1-\delta_3}\right)\frac{1}{\tau_\theta}+\left(\frac{\delta_2}{1-\delta_3}\right)\frac{1}{\tau_y} \\ \left(\frac{\delta_1+\delta_2}{1-\delta_3}\right)\frac{1}{\tau_\theta} & \left(\frac{\delta_1+\delta_2}{1-\delta_3}\right)\frac{1}{\tau_\theta}+\left(\frac{\delta_2}{1-\delta_3}\right)\frac{1}{\tau_y} & \left(\frac{1}{1-\delta_3}\right)^2\left(\frac{(\delta_1+\delta_2)^2}{\tau_\theta}+\frac{\delta_2^2}{\tau_y}+\frac{1}{\tau_p}\right) \end{pmatrix}\right) \tag{55}$$

The **quasi-Bayesian** level of information acquisition solves

$$\max_{\tau_s} \mathbb{E}_{\mu(\tau_s)}[w(s_i,y,p)] - c\tau_s,$$

where the expectation is taken under the measure (55).

[36]To see this, note that a fully cursed agent perceives the rescaled aggregative signal to be $\hat{p} = \frac{\tau_s}{\tau_\theta+\tau_y+\tau_s}s_i + \frac{\tau_y}{\tau_\theta+\tau_y+\tau_s}y + \text{noise} + \frac{1}{\delta_1}z_p$, where the noise is uncorrelated with the other signals and the state. This is because information about θ affects the agents beliefs about p even though he doesn't update vice versa. It is easy to check from this representation that the covariance of p with the other signals is unaffected by cursedness.

This notion is similar to the subjective envelope condition in that agents under both criteria correctly predict their (cursed) action rule, do not use information acquisition to fix their bias, and correctly predict the equilibrium outcome as well as understand that they cannot change it. That is, both notions satisfy the bulk of the behavioral desiderata outlined in section 3. The main difference is that agents under the subjective envelope condition consider their true welfare, while they follow the cursed mixture distribution under the objective interim condition. In terms of our learning foundation (Theorem 3), this requires that the agent records his *subjective expected* welfare—instead of the true realized welfare—and correlates it with the sign of the information acquisition implementation tremble.

This approach has three significant shortcomings. First, at a conceptual level, we see the mixture distribution as an auxiliary device to represent behavior, less as a normatively valid notion, even subjectively. Second, in order to obtain a signal that is independent of the state conditionally on their private and public fundamental signal, the perceived correlation of the public signal p with the state θ at the ex-ante stage depends on an agent's *individually chosen* τ_s. In other words, the agent behaves as if his personal information acquisition affects the precision of public information obtained by all agents, which is a sort of magical thinking we don't commonly associate with cursedness. Third, from a modeling standpoint, the mixture distribution leads posteriors outside of the family of conjugate priors selected for the analysis of the model. In our case, a mixture of normal distributions is not itself a normal distribution. In many settings of applied theory, however, the tractability of the model crucially depends on these distributional assumptions.[37] Indeed, a characterization of the comparative statics of the model under this notion has proven elusive.

To compare this notion with our results from the subjective envelope condition, we therefore resort to numerical computations. The key comparative statics results continue to hold. Higher levels of cursedness still correspond to an increased use and acquisition of private information, even though τ_s and δ_1 are no longer proportional (Figure 7). Starting with the rational case, an increase in cursedness increases welfare (Figure 8). Moreover, for intermediate cursedness and high levels of transparency, more and cheaper fundamental information has a paradoxical effect on welfare (Figure 8). Finally, notice that the two notions coincide not only in the rational case but also for fully cursed agents. This is the case since fully cursed agents act just like rational agents would in a world without the aggregative signal. The true and the subjective interim

[37]Notice that the issue is not computing the cursed optimal action, as the first moment is highly tractable. Indeed $\mathbb{E}_{\mu_\chi} = \chi\mathbb{E}_{\mu_1} + (1-\chi)\,\mathbb{E}_{\mu_0}$, while by the rule of total variance

$$\mathbb{V}_{\mu_\chi}[\theta] = \chi\mathbb{V}_{\mu_1} + (1-\chi)\,\mathbb{V}_{\mu_0} + \chi(1-\chi)\left(\mathbb{E}_{\mu_1} - \mathbb{E}_{\mu_0}\right)^2$$

It is the ex-ante expectation of the final term that limits tractability in our case.

welfare of such agents coincide, as the precision of an unused signal is immaterial for welfare.

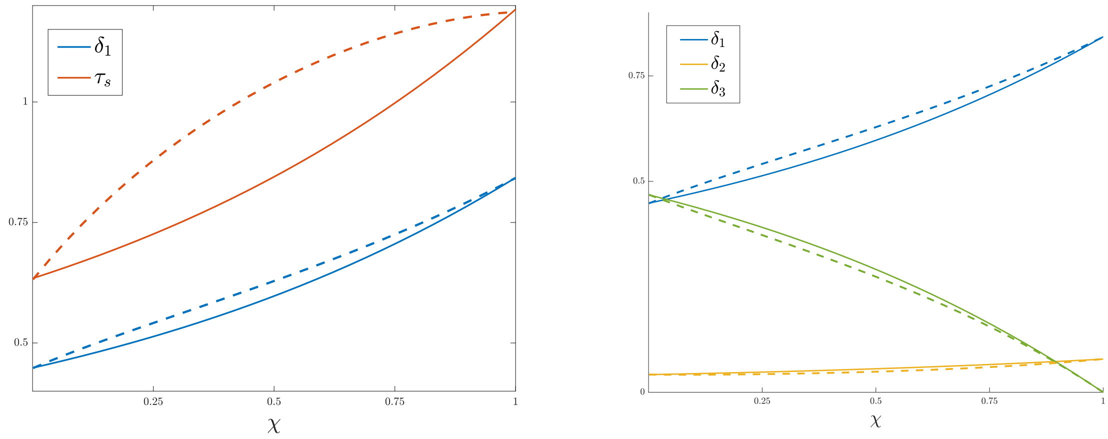

Figure 7: δ_1 and τ_s as a function of χ under different information acquisition notions (left). Equilibrium loadings as a function of χ under different information acquisition notions (right).

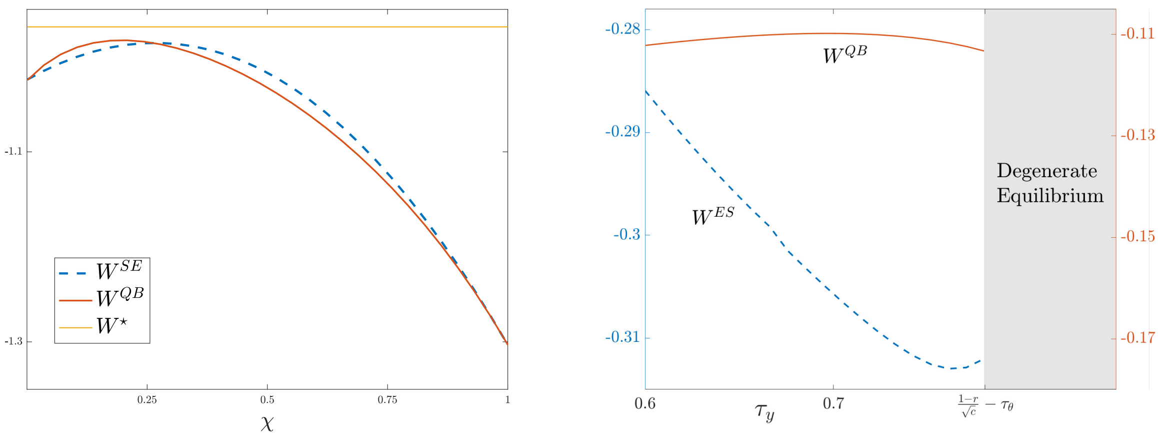

Figure 8: Welfare as a function of χ under different information acquisition notions (left). Welfare under objective interim information acquisition, as a function of τ_y. (right).

D Formal Results for Section 6 (Shrewd Agent) (For Online Publication)

Proposition 10. *The action rule of the shrewd agent is*

$$\alpha_1^R = \frac{(1-(1-\delta_1)r)\tau_s^R}{\tau_\theta + \tau_y + \tau_s^R + \delta_1^2\tau_p} \tag{56}$$

$$\alpha_2^R = \frac{(1-(1-\delta_1)r)\tau_y + \delta_2\left(r\left(\tau_\theta + \tau_y + \tau_s^R\right) - (1-r)\delta_1\tau_p\right)}{\tau_\theta + \tau_y + \tau_s^R + \delta_1^2\tau_p} \tag{57}$$

$$\alpha_3^R = \frac{\delta_1 (1-(1-\delta_1) r) \tau_p + \delta_3 \left(r \left(\tau_\theta + \tau_y + \tau_s^R \right) - (1-r) \delta_1 \tau_p \right)}{\tau_\theta + \tau_y + \tau_s^R + \delta_1^2 \tau_p} \tag{58}$$

where τ_s^R *solves*

$$\tau_s^R = \frac{\alpha_1^R}{\sqrt{c}} \tag{59}$$

with complementary slackness ensuring $\tau_s^R \geq 0$ *when required.*

Proof. This follows immediately from shrewd observation of the derivation of the matching coefficients and g equations. More directly, the welfare of an agent playing loadings $\boldsymbol{\alpha}$ and private precision τ_s^R while the rest of (the average of) others play $\boldsymbol{\delta}$ is derived in (72). By setting

$$\nabla_{\boldsymbol{\alpha}} W(\boldsymbol{\alpha}, \boldsymbol{\delta}) = 0$$

we find the best response coefficients as a function of others' loadings. We get

$$\begin{aligned}
0 =& (1-\alpha_2-\alpha_3(\delta_1+\delta_2)-\alpha_1(1-\delta_3)-\delta_3-r(1-\delta_1-\delta_2)+\delta_3(\alpha_2+r))\tau_s-\alpha_1(1-\delta_3)\tau_\theta \\
0 =& (1-\alpha_2-\alpha_3(\delta_1+\delta_2)-\alpha_1(1-\delta_3)-\delta_3-r(1-\delta_1-\delta_2)+\delta_3(\alpha_2+r))\tau_y-(\alpha_2(1-\delta_3)+\delta_2(\alpha_3-r))\tau_\theta \\
0 =& \tau_y\left[-(\delta_1+\delta_2)(1-\alpha_1-\alpha_2-\alpha_3(\delta_1+\delta_2)-(1-\alpha_1-\alpha_2)\delta_3-r(1-\delta_1-\delta_2-\delta_3))\tau_p+(\alpha_3-\delta_3 r)\tau_\theta\right] \\
&+2\tau_p\tau_\theta(\delta_2(\alpha_2(1-\delta_3)+\delta_2(\alpha_3-r)))
\end{aligned}$$

The solution to this linear system is the $\boldsymbol{\alpha}$ in the proposition, which we can plug back in welfare to obtain the expression

$$\overline{W}(\boldsymbol{\delta}) = \frac{(1-r) r}{(1-\delta_3)^2}\left[\frac{\delta_2^2}{\tau_y}+\frac{\delta_3^2}{\tau_p}+\frac{(1-\delta_1-\delta_2-\delta_3)^2}{\tau_\theta}\right]-\frac{(1-(1-\delta_1) r)^2}{\tau_\theta+\tau_y+\tau_s^R+\delta_1^2\tau_p}-c\tau_s^R$$

Differentiating this equation with respect to τ_s^R, we obtain the final condition. □

Proof of Proposition 8: We first derive equation (37) in the text. Note that

$$\begin{aligned}
\frac{(1-(1-\delta_1) r)^2}{\left(\tau_\theta+\tau_y+\tau_s^R+\delta_1^2\tau_p\right)^2}-c=0& \\
(1-(1-\delta_1) r)^2 = c\left[\tau_\theta+\tau_y+\tau_s^R+\delta_1^2\tau_p\right]^2&
\end{aligned}$$

using (47) on the LHS we get after rearranging

$$\left[\frac{\left(\tau_\theta+\tau_y+\tau_s\right)\left(\tau_\theta+\tau_y+\tau_s+\delta_1^2\tau_p\right)}{\tau_\theta+\tau_y+\tau_s+\chi\delta_1^2\tau_p}\right]=\left[\tau_\theta+\tau_y+\tau_s^R+\delta_1^2\tau_p\right]$$

From there, $\tau_s \geq \tau_s^{\mathrm{R}}$ is immediate, since

$$\frac{\tau_\theta + \tau_y + \tau_s + \delta_1^2 \tau_p}{\tau_\theta + \tau_y + \tau_s^{\mathrm{R}} + \delta_1^2 \tau_p} = 1 + \frac{\chi \delta_1^2 \tau_p}{\tau_\theta + \tau_y + \tau_s} \geq 1$$

For the fully cursed equilibrium, the above implies that $\tau_\theta + \tau_y + \tau_s = \tau_\theta + \tau_y + \tau_s^{\mathrm{R}} + \delta_1^2 \tau_p$ and hence we have

$$\tau_s^{\mathrm{R}} = \tau_s - \delta_1^2 \tau_p = \frac{1}{\sqrt{c}} \frac{1 - r - \sqrt{c}\left(\tau_\theta + \tau_y\right)}{1 - r} - \tau_p \left(\frac{1 - r - \sqrt{c}\left(\tau_\theta + \tau_y\right)}{1 - r}\right)^2$$

Solving for $\tau_s^{\mathrm{R}} \geq 0$, we get positive information acquisition iff

$$\tau_\theta + \tau_y \in \left(\frac{(1 - r)\left(1 - \frac{1}{\tau_p \sqrt{c}}\right)}{\sqrt{c}}, \frac{1 - r}{\sqrt{c}}\right) \tag{60}$$

.

Now, for the Proposition, we have established $\tau_s^{\mathrm{R}} \leq \tau_s$. For the limit result, note that

$$\tau_s^{\mathrm{R}} = \frac{\left(\tau_\theta + \tau_y + \tau_s\right)\left(\tau_\theta + \tau_y + \tau_s + \delta_1^2 \tau_p\right)}{\tau_\theta + \tau_y + \tau_s + \chi \delta_1^2 \tau_p} - \left(\tau_\theta + \tau_y + \delta_1^2 \tau_p\right) = \frac{\left(\tau_\theta + \tau_y + \tau_s\right)\tau_s - \chi \delta_1^2 \tau_p \left(\tau_\theta + \tau_y + \delta_1^2 \tau_p\right)}{\tau_\theta + \tau_y + \tau_s + \chi \delta_1^2 \tau_p}$$

and that $\tau_s \to \infty$ as $c \to 0$. From there, we have

$$\tau_s^{\mathrm{R}} = \frac{\left(\tau_\theta + \tau_y + \tau_s\right)\tau_s - \chi \delta_1^2 \tau_p \left(\tau_\theta + \tau_y + \delta_1^2 \tau_p\right)}{\tau_\theta + \tau_y + \tau_s + \chi \delta_1^2 \tau_p} \to \frac{\tau_s^2}{\tau_s} \to \infty$$

We have $\tau_s^{\mathrm{R}} = 0$ for $\tau_\theta + \tau_y \leq \frac{(1-r)\left(1 - \frac{1}{\tau_p \sqrt{c}}\right)}{\sqrt{c}}$ in the fully cursed case, where $\tau_s > 0$.

To see nonmonotonicity in $\tau_\theta + \tau_y$ in the general case, note that as we approach the limit (19), we have

$$\frac{d\tau_s^{\mathrm{R}}}{d\tau_y}\Big|_{\sqrt{c} = \frac{1-r}{\tau_\theta + \tau_y}} = \frac{\tau_\theta + \tau_y}{\sqrt{c}\left(2\tau_\theta + \tau_y\right)} \frac{d\delta_1}{d\tau_y} \propto \frac{d\delta_1}{d\tau_y} < 0$$

and hence, local to this value, we always get a positive τ_s^{R}. However, for $\tau_\theta + \tau_y \leq \chi \frac{1-r}{\sqrt{c}}$ interior, we get that for $\tau_p \to \infty$

$$\tau_s^{\mathrm{R}} = \frac{\left(\tau_\theta + \tau_y + \tau_s\right)\tau_s - \chi \delta_1^2 \tau_p \left(\tau_\theta + \tau_y + \delta_1^2 \tau_p\right)}{\tau_\theta + \tau_y + \tau_s + \chi \delta_1^2 \tau_p} \to -\delta_1^2 \tau_p < 0$$

which establishes the result. Interior nonmonotonicity follows by continuity.

To see nonmonotonicity in c, note that

1. $\frac{d\tau_s^R}{dc} < 0$ in the limit (19) ($\sqrt{c} = \frac{1-r}{\tau_\theta+\tau_y}$). To see this, take the representation above, plug 17, take the derivative, set $\sqrt{c} = \frac{1-r}{\tau_\theta+\tau_y}$, $\delta_1 = 0$, then we get

$$\frac{d\tau_s^R}{dc}\Big|_{\sqrt{c}=\frac{1-r}{\tau_\theta+\tau_y}} \to \frac{2(1-r)^4 \frac{\partial\delta_1}{\partial c}}{\left(\tau_\theta+\tau_y\right)^2} \propto \frac{\partial\delta_1}{\partial c} < 0$$

2. $\tau_s^R = 0$ at the limit (19), as $0 \le \tau_s^R \le \tau_s = 0$. Combining this with (1.), we obtain that $\tau_s^R > 0$ local to this upper bound on costs.

3. $\tau_s^R \to \infty$ as $c \to 0$, as shown above.

4. For any c satisfying the existence of an interior equilibrium in the $\tau_p \to \infty$ limit (i.e. $\sqrt{c} \le \chi\frac{1-r}{\tau_\theta+\tau_y}$, see footnote 41), there exists a τ_p sufficiently large such that $\left(\tau_s^R\right)^{\text{FOC}} < 0$. To see this, pick an interior c. Then, because $\delta_1^\infty > 0$ and $\tau_p \to \infty$

$$\left(\tau_s^R\right)^{\text{FOC}} \to -\delta_1^2\tau_p < 0$$

This establishes nonmonotonicity, as τ_s^R is large for small c, zero for intermediate c, but nonzero local to $\sqrt{c} = \frac{1-r}{\tau_\theta+\tau_y}$.

To analyze the derivative in χ, we compute

$$\frac{d\tau_s^R}{d\chi} \propto r - 2\sqrt{c}\delta_1\tau_p$$

It is apparent that τ_s^R is decreasing for $r \le 0$ and that it is increasing as $r \to 1 - \sqrt{c}\left(\tau_\theta+\tau_y\right)$ when this is positive, as then $\delta_1 \to 0$. To see that we can have a hump shape, note that δ_1 is increasing in χ and hence

$$\frac{d^2\tau_s^R}{d\chi d\chi} \propto -\frac{d\delta_1}{d\chi} \le 0$$

around $\frac{d\tau_s^R}{d\chi} = 0$, which establishes a hump-shape (but, importantly, not necessarily concavity!). Clearly, all these comparative statics only apply for interior solutions, otherwise $\tau_s^R \equiv 0$ locally. □

Remark 1 (The Shrewd Agent in the Transparent Limit). Consider the limit as $\tau_p \to \infty$. Since $\delta_1^\infty > 0$ whenever a limit equilibrium exists, the rational agent can exactly infer the state. Therefore, he does not acquire or use private information[38] and relies solely

[38] Indeed, notice that the interval (60) vanishes as $\tau_p \to \infty$.

on the aggregative signal for information

$$\alpha_1^{\mathrm{R}} \to 0, \quad \tau_s^{\mathrm{R}} \to 0, \quad \alpha_2^{\mathrm{R}} \to \frac{-(1-r)\sqrt{c}\tau_y}{1-r-\sqrt{c}\left(\tau_\theta+\tau_y\right)} < 0, \quad \alpha_3^{\mathrm{R}} \to \frac{1-r-\sqrt{c}\left(\tau_\theta+\tau_y\right)r}{1-r-\sqrt{c}\left(\tau_\theta+\tau_y\right)} > 0 \tag{61}$$

The apparent anti-imitation in $\alpha_2^{\mathrm{R}} < 0$ allows the shrewd agent to filter out the over-reliance of the cursed crowd on the public signal.

Proof of Proposition 9: For χ, we plug the general δ into the welfare expression, take derivative w.r.t. χ, set $\chi = 1$, $\delta_1 = \delta_1^{\mathrm{FC}}$ and we get

$$\begin{aligned}\frac{\mathrm{d}W_\chi^{\mathrm{R}}}{\mathrm{d}\chi} &= -\frac{2\sqrt{c}(1-r)\tau_p\left(1-r-\sqrt{c}\left(\tau_\theta+\tau_y\right)\right)^2\left[r\left(\tau_\theta+\tau_y+\tau_s^{\mathrm{R}}+\tau_p\right)-\tau_p\left(1-\sqrt{c}\left(\tau_\theta+\tau_y\right)\right)\right]\left(1-r\left(1+\sqrt{c}\left(\tau_\theta+\tau_y\right)\right)\right)^2}{[\cdot]^2\left[1+\sqrt{c}\tau_p\left(1-\sqrt{c}\left(\tau_\theta+\tau_y\right)\right)^2+r^2\left(1+\sqrt{c}\left(\tau_\theta+\tau_y+\tau_p\right)\right)-r\left(2-2c\tau_p\left(\tau_\theta+\tau_y\right)+\sqrt{c}\left(\tau_\theta+\tau_y+2\tau_p\right)\right)\right]} \\ &\propto -\frac{2\sqrt{c}\tau_p\delta_1^2(1-r)^3\left[r\left(\tau_\theta+\tau_y+\tau_s^{\mathrm{R}}+\tau_p\delta_1\right)-\tau_p\delta_1\right]}{1+\sqrt{c}\tau_p\left(1-\sqrt{c}\left(\tau_\theta+\tau_y\right)\right)^2+r^2\left(1+\sqrt{c}\left(\tau_\theta+\tau_y+\tau_p\right)\right)-r\left(2-2c\tau_p\left(\tau_\theta+\tau_y\right)+\sqrt{c}\left(\tau_\theta+\tau_y+2\tau_p\right)\right)}\end{aligned} \tag{62}$$

Consider first the case of $r > 0$, which implies that $1-\sqrt{c}\left(\tau_\theta+\tau_y\right) > 0$. As $r \to 1-\sqrt{c}\left(\tau_\theta+\tau_y\right)$, we have for the above expression

$$\to -2\tau_p\left(\tau_\theta+\tau_y\right)\sqrt{c}\left(1-\sqrt{c}\left(\tau_\theta+\tau_y\right)\right)\delta_1^2 \leq 0$$

This converges to zero from below since $\delta_1 \to 0$ as $r \to 1-\sqrt{c}\left(\tau_\theta+\tau_y\right)$. Therefore, for large χ, the shrewd agent prefers a less cursed environment if r is big.

Consider now $r \leq 0$. Note that the numerator in (62) is always negative. Furthermore, note that the denominator $v(r)$ is positive for $r = 0$. We will show that it is positive for all $r \leq 0$ and hence the expression is positive for all $r \leq 0$. First, note that v is a convex quadratic function in r. Minimizing, we find that the minimizer and minimum, resp., are given by

$$\begin{aligned} r^* &\propto 2-2c\tau_p\left(\tau_\theta+\tau_y\right)+\sqrt{c}\left(\tau_\theta+\tau_y+2\tau_p\right) \\ v^* &\propto -1+4c\tau_p\left(\tau_\theta+\tau_y\right) \end{aligned}$$

If $v^* > 0$, we are done. If it is negative, it is easy see that r^* must be positive. But then, the fact that $v(0) > 0$ implies that $v(r) > 0$ for all $r \leq 0$. □

Proposition 11. *The welfare of a shrewd agent in a fully cursed population,* W_1^{R}*, has the following properties.*

1. *It is strictly increasing in* τ_p.

2. *It has ambiguous comparative statics with respect to τ_y. In particular,*

   - *At the boundary of the activity region, (i.e. $\tau_y = \frac{(1-r)\left(1-\frac{1}{\tau_p\sqrt{c}}\right)}{\sqrt{c}} - \tau_\theta$), we have $\frac{dW_1^R}{d\tau_y} < 0$,*
   - *for τ_y large (i.e. local to the nontriviality limit $\tau_y = \frac{(1-r)}{\sqrt{c}} - \tau_\theta$), we have $\frac{dW_1^R}{d\tau_y} > 0$.*

3. *It is decreasing in c whenever $\tau_s^R > 0$, but it has ambiguous comparative statics with respect to c if $\tau_s^R = 0$. In particular,*

   - *If $\tau_p \geq \frac{1}{\sqrt{c}}$ and r sufficiently negative, then $\frac{dW_1^R}{dc} > 0$.*

Proof. To obtain the welfare of the shrewd agent in the fully cursed equilibrium, we simply plug action rule (56)-(36) and the equilibrium δs into the welfare equation to obtain, for the unconstrained case

$$W_1^{R,\tau_s^R>0} = -2\sqrt{c} + \frac{-2c^{3/2}(1-r)\tau_p\left(\tau_\theta+\tau_y\right) + c^2\tau_p\left(\tau_\theta+\tau_y\right)^2 + c(1-r)\left(\tau_\theta+\tau_y+(1-r)\tau_p\right)}{(1-r)^2}$$

as well as for the constrained case, where we leave the expression in general form both for compactness and ease of analysis

$$W_1^{R,\tau_s^R=0} = -\frac{(1-r)r\delta_2^2}{\tau_y} - \frac{(1-r)r(1-\delta_1-\delta_2)^2}{\tau_\theta} - \frac{(1-(1-\delta_1)r)^2}{\tau_\theta+\tau_y+\delta_1^2\tau_p}$$

For transparency, we obtain by direct computation for an interior solution $\frac{\partial W_1^{R,\tau_s^R>0}}{\partial\tau_p} = c\frac{\left(1-r-\sqrt{c}(\tau_\theta+\tau_y)\right)^2}{(1-r)^2} = c\delta_1^2 > 0$, and through an envelope argument from the general expression, we get form corner solutions $\frac{\partial W_1^{R,\tau_s^R=0}}{\partial\tau_p} = \frac{\delta_1^2(1-(1-\delta_1)r)^2}{\left(\tau_\theta+\tau_y+\delta_1^2\tau_p\right)^2} > 0$.

For τ_θ,τ_y: Consider the derivative of $W_1^{R,\tau_s^R>0}$ and let $\tau_\theta+\tau_y \to \frac{(1-r)\left(1-\frac{1}{\tau_p\sqrt{c}}\right)}{\sqrt{c}}$. Then, we get

$$\frac{\partial W_1^{R,\tau_s^R>0}}{\partial\tau_y} \to c\frac{1-r-(1-r)2\sqrt{c}\tau_p+2c\tau_p\left(\tau_\theta+\tau_y\right)}{(1-r)^2} = \frac{c}{1-r}\left(1-2\sqrt{c}\tau_p\delta_1^{FC}\right) = -\frac{c}{1-r} < 0$$

Note that this result also holds for $W_1^{R,\tau_s^R=0}$, as the value function is $\mathcal{C}^1$. Taking instead the limit as $\tau_\theta+\tau_y \to \frac{(1-r)}{\sqrt{c}}$ (where we always are at an interior solution), we have

$$\frac{\partial W_1^{R,\tau_s^R>0}}{\partial\tau_y} \to \frac{c}{1-r}\left(1-2\sqrt{c}\tau_p\delta_1^{FC}\right) = \frac{c}{1-r}$$

For c, let us first demonstrate a setting where an increase in costs is beneficial for the shrewd agent. Consider constrained welfare, let $r \to -\infty$ and hence $\delta_1 = 1 - \frac{\sqrt{c}(\tau_\theta+\tau_y)}{(1-r)} \to 1$, $\delta_2 = \frac{\sqrt{c}}{1-r}\tau_y \to 0$. Then

$$W_1^{R,\tau_s^R=0} \to -\frac{rc\tau_y}{(1-r)} - \frac{r\tau_\theta c}{(1-r)} - \frac{\left(1 - \frac{\sqrt{c}(\tau_\theta+\tau_y)}{(1-r)}r\right)^2}{\tau_\theta+\tau_y+\tau_p} \to c\tau_\theta + \tau_y c - \frac{\left(1+\sqrt{c}\left(\tau_\theta+\tau_y\right)\right)^2}{\tau_\theta+\tau_y+\tau_p}$$

and

$$\frac{\partial}{\partial c}\left(c\tau_\theta + \tau_y c - \frac{\left(1+\sqrt{c}\left(\tau_\theta+\tau_y\right)\right)^2}{\tau_\theta+\tau_y+\tau_p}\right) = \tau_\theta + \tau_y - \frac{\left(\tau_\theta+\tau_y\right)}{\sqrt{c}}\frac{\left(1+\sqrt{c}\left(\tau_\theta+\tau_y\right)\right)}{\tau_\theta+\tau_y+\tau_p}$$
$$= \left(\tau_\theta+\tau_y\right)\left(\frac{\tau_p - \frac{1}{\sqrt{c}}}{\tau_\theta+\tau_y+\tau_p}\right) > 0$$

since we consider the case $\tau_s^R = 0$, and therefore $\tau_p > \frac{1}{\sqrt{c}}$.

Generally, we can also have $\frac{\partial}{\partial c}W_1^R < 0$. To see this, consider the case of $r = 0$. Then

$$W^R|_{r=0} = \max_{\tau_s^R \geq 0} -\frac{1}{\tau_\theta+\tau_y+\tau_s^R+\delta_1^2\tau_p} - c\tau_s^R$$

where we may or may not have a corner solution. In either case, welfare is decreasing in c because δ_1 is decreasing in c and an envelope argument.

At an interior solution (which occurs for $\tau_p < \frac{1}{\delta_1\sqrt{c}}$), we get

$$\frac{\partial}{\partial c}W_1^{R,\tau_s^R>0} = -\frac{1}{\sqrt{c}} + \frac{-3\sqrt{c}(1-r)\tau_p\left(\tau_\theta+\tau_y\right) + 2c\tau_p\left(\tau_\theta+\tau_y\right)^2 + (1-r)\left(\tau_\theta+\tau_y+(1-r)\tau_p\right)}{(1-r)^2}$$
$$= \frac{\left(\tau_\theta+\tau_y\right)}{(1-r)} - \frac{1}{\sqrt{c}} + \delta_1\tau_p\left(\delta_1 - \frac{\sqrt{c}\left(\tau_\theta+\tau_y\right)}{(1-r)}\right)$$

which is always negative by condition (19) if $\delta_1 \leq \frac{\sqrt{c}(\tau_\theta+\tau_y)}{(1-r)}$. If this is violated, we have

$$\frac{\left(\tau_\theta+\tau_y\right)}{(1-r)} - \frac{1}{\sqrt{c}} + \delta_1\tau_p\left(\delta_1 - \frac{\sqrt{c}\left(\tau_\theta+\tau_y\right)}{(1-r)}\right) \leq \frac{\left(\tau_\theta+\tau_y\right)}{(1-r)} - \frac{1}{\sqrt{c}} + \frac{1}{\sqrt{c}}\left(\delta_1 - \frac{\sqrt{c}\left(\tau_\theta+\tau_y\right)}{(1-r)}\right)$$
$$= -\frac{\delta_1}{\sqrt{c}} + \frac{1}{\sqrt{c}}\left(\delta_1 - \frac{\sqrt{c}\left(\tau_\theta+\tau_y\right)}{(1-r)}\right) = \frac{1}{\sqrt{c}}\left(-\frac{\sqrt{c}\left(\tau_\theta+\tau_y\right)}{(1-r)}\right) < 0$$

concluding the proof. □

E Additional Proofs and Results (For Online Publication)

E.1 Model With Exogenous Private Information τ_s

We start by analyzing the model with exogenous private information, some results from which are cited in the discussion after Propositions 1, 2 and 3 to clarify the impact of endogenous information acquisition. We first establish a helpful lemma.

Lemma 9. *In equilibrium, we have $f_\delta > 0$.*

Proof. Compute

$$\begin{aligned}
f_\delta &= \tau_\theta + \tau_y + 3\delta_1^2\tau_p - \chi\frac{\tau_s}{\tau_\theta + \tau_y + \tau_s}\left[2 + r(3\delta_1 - 2)\right]\delta_1\tau_p + (1-r)\tau_s \\
&= \left(\tau_\theta + \tau_y + 3\delta_1^2\tau_p\left[1 - \chi\frac{\tau_s}{\tau_\theta + \tau_y + \tau_s}\left[\frac{1 + r(\delta_1 - 1)}{\delta_1}\right]\right]\right) + \chi\frac{\tau_s}{\tau_\theta + \tau_y + \tau_s}\left[1 - r\right]\delta_1\tau_p + (1-r)\tau_s \\
&= \left(\tau_\theta + \tau_y + 3\delta_1^2\tau_p\left[1 - \frac{\chi\tau_\theta + \chi\tau_y + \chi\tau_s + \chi\delta_1^2\tau_p}{\tau_\theta + \tau_y + \tau_s + \chi\delta_1^2\tau_p}\right]\right) + \chi\frac{\tau_s}{\tau_\theta + \tau_y + \tau_s}\left[1 - r\right]\delta_1\tau_p + (1-r)\tau_s > 0
\end{aligned}$$

where in the final step we used (47) in the transformation:

$$1 - \chi\frac{\tau_s}{\tau_\theta + \tau_y + \tau_s}\left[\frac{1 + r(\delta_1 - 1)}{\delta_1}\right] = 1 - \frac{\chi\tau_\theta + \chi\tau_y + \chi\tau_s + \chi\delta_1^2\tau_p}{\tau_\theta + \tau_y + \tau_s + \chi\delta_1^2\tau_p} \geq 0. \qquad \square$$

Proposition 12. *The comparative statics of the model with exogenous τ_s are given by*

$$\frac{\partial\delta_1}{\partial\chi} \geq 0, \qquad \frac{\partial\delta_2}{\partial\chi} \geq 0, \qquad \frac{\partial\delta_3}{\partial\chi} \leq 0, \qquad \frac{\partial}{\partial\chi}\frac{\delta_1}{\delta_2} \leq 0 \tag{63}$$

$$\frac{\partial\delta_1}{\partial\tau_p} \leq 0, \qquad \frac{\partial\delta_2}{\partial\tau_p} \leq 0, \qquad \frac{\partial\delta_3}{\partial\tau_p} \geq 0, \qquad \frac{\partial}{\partial\tau_p}\frac{\delta_1}{\delta_2} \geq 0, \tag{64}$$

all inequalities being strict if $\tau_p \neq 0$ and $\chi \neq 1$. Furthermore,

$$\frac{\partial}{\partial\tau_p}\delta_1^2\tau_p > 0. \tag{65}$$

Proof. By implicit differentiation

$$\frac{\mathrm{d}\delta_1}{\mathrm{d}\chi} = -\frac{f_\chi}{f_\delta} \propto \frac{\tau_s}{\tau_\theta + \tau_y + \tau_s}\left[(1-r) + r\delta_1\right]\delta_1^2\tau_p \geq 0$$

and from (12), it is immediate that $\frac{\mathrm{d}\delta_2}{\mathrm{d}\chi} \propto \frac{\mathrm{d}\delta_1}{\mathrm{d}\chi} \geq 0$ and $\frac{\mathrm{d}\delta_3}{\mathrm{d}\chi} \propto -\frac{\mathrm{d}\delta_1}{\mathrm{d}\chi} \leq 0$.

Relative size of δ_1 and δ_2:

$$\frac{d}{d\chi}\frac{\delta_2}{\delta_1} = \frac{d}{d\chi}\frac{\delta_1\tau_y}{(1-r)\tau_s - \delta_1\left(\tau_\theta+\tau_y\right)} = \frac{\tau_y\left((1-r)\tau_s - \delta_1\left(\tau_\theta+\tau_y\right)\right) + \delta_1\tau_y\left(\tau_\theta+\tau_y\right)}{\left((1-r)\tau_s - \delta_1\left(\tau_\theta+\tau_y\right)\right)^2}\frac{d\delta_1}{d\chi} \geq 0$$

whence the result in the proposition follows.

τ_p comparative statics: Again, $\frac{d\delta_1}{d\tau_p} \propto -f_{\tau_p}$ and using (47), we have

$$f_{\tau_p} = \delta_1^3 - \chi\frac{\tau_s}{\tau_\theta+\tau_y+\tau_s}\left[(1-r)+r\delta_1\right]\delta_1^2 = \delta_1^3\left(1-\chi\frac{\tau_\theta+\tau_y+\tau_s+\delta_1^2\tau_p}{\tau_\theta+\tau_y+\tau_s+\chi\delta_1^2\tau_p}\right) > 0$$

From (12), the comparative statics are immediate. Finally, we have

$$\begin{aligned}
\frac{d\delta_1^2\tau_p}{d\tau_p} &= 2\delta_1\tau_p\frac{d\delta_1}{d\tau_p} + \delta_1^2 \\
&= -2\delta_1\tau_p\frac{\delta_1^3 - \chi\frac{\tau_s}{\tau_\theta+\tau_y+\tau_s}\left[(1-r)+r\delta_1\right]\delta_1^2}{\left(\tau_\theta+\tau_y+3\delta_1^2\tau_p\right) - \chi\frac{\tau_s}{\tau_\theta+\tau_y+\tau_s}\left[2+r(3\delta_1-2)\right]\delta_1\tau_p + (1-r)\tau_s} + \delta_1^2 \\
&= \frac{\delta_1^2}{f_\delta}\left\{-2\delta_1\tau_p\left(\delta_1 - \chi\frac{\tau_s}{\tau_\theta+\tau_y+\tau_s}\left[(1-r)+r\delta_1\right]\right) + \left(\left(\tau_\theta+\tau_y+3\delta_1^2\tau_p\right) - \chi\frac{\tau_s\left[2+r(3\delta_1-2)\right]}{\tau_\theta+\tau_y+\tau_s}\delta_1\tau_p + (1-r)\tau_s\right)\right\} \\
&\propto \left(\tau_\theta+\tau_y+\delta_1^2\tau_p\right) - \chi\frac{\tau_s}{\tau_\theta+\tau_y+\tau_s}r\delta_1^2\tau_p + (1-r)\tau_s \\
&= \frac{1}{\delta_1}\left((1-r)\tau_s + \chi\frac{\tau_s}{\tau_\theta+\tau_y+\tau_s}(1-r)\delta_1^2\tau_p\right) > 0
\end{aligned}$$

where we dropped $\frac{\delta_1^2}{f_\delta} > 0$ and, in the last step, we use $f(\delta_1) = 0$, which establishes the claim. □

Proposition 13. *In the model with exogenous τ_s, the weight on private information respond to parameter changes as follows:*

$$\frac{\partial\delta_1}{\partial\tau_s} \geq 0, \qquad \frac{\partial\delta_1}{\partial\tau_y} \leq 0, \qquad \frac{\partial\delta_1}{\partial r} \leq 0 \tag{66}$$

Proof. Since $f_\delta > 0$ (Lemma 9), for a generic parameter ν we get

$$\frac{d\delta_1}{d\nu} = -\frac{f_\nu}{f_\delta} \propto -f_\nu$$

And hence the comparative statics of δ_1 follow immediately, occasionally using (47), from

$$f_{\tau_y} = \delta_1 + \chi\left[(1-r)+r\delta_1\right]\delta_1^2\tau_p\frac{\tau_s}{\left(\tau_\theta+\tau_y+\tau_s\right)^2} > 0$$

$$f_{\tau_s} = -\chi \frac{\tau_\theta + \tau_y}{\left(\tau_\theta + \tau_y + \tau_s\right)^2} [(1-r) + r\delta_1] \delta_1^2 \tau_p - (1-r)(1-\delta_1) < 0$$

$$f_r = \chi \frac{\tau_s (1-\delta_1)}{\tau_\theta + \tau_y + \tau_s} \delta_1^2 \tau_p + (1-\delta_1)\tau_s = \tau_s (1-\delta_1) \left[\frac{\chi \delta_1^2 \tau_p}{\tau_\theta + \tau_y + \tau_s} + 1 \right] > 0$$

□

Notice the fundamental representation (24) is valid even in the model with exogenous τ_s as it does not make use of equation (17).

Proposition 14. *In the model with exogenous τ_s , the loadings in the fundamental representation (24) are*

$$\beta = 1 - \frac{\delta_1 \tau_\theta}{(1-r)\tau_s}, \qquad \gamma_2 = \frac{\delta_1 \tau_y}{(1-r)\tau_s}, \qquad \gamma_3 = \frac{1-\delta_1}{\delta_1} - \frac{\left(\tau_\theta + \tau_y\right)}{(1-r)\tau_s}. \tag{67}$$

Furthermore,

$$\frac{d\beta}{d\chi} < 0, \qquad \frac{d\beta}{d\tau_p} > 0, \qquad \frac{d\beta}{d\tau_s} > 0, \qquad \frac{d\gamma_2}{d\tau_y} > 0, \qquad \frac{d\gamma_2}{d\tau_s} < 0, \qquad \frac{d\gamma_2}{d\tau_p} < 0.$$

Proof. The result for $\frac{d\beta}{d\chi}, \frac{d\beta}{d\tau_p}, \frac{d\gamma_2}{d\tau_p}$ is immediate from the comparative statics of δ_1. For $\frac{d\gamma_2}{d\tau_y} \propto \frac{d\delta_1}{d\tau_y}\tau_y + \delta_1$, we get

$$\begin{aligned}
\frac{d\delta_1}{d\tau_y}\tau_y + \delta_1 &= -\frac{f_{\tau_y}}{f_\delta}\tau_y + \delta_1 \propto -f_{\tau_y}\tau_y + f_\delta \delta_1 \\
&= \frac{\delta_1}{\left(\tau_y + \tau_s + \tau_\theta\right)^2} \left[\left(3\delta_1^2\tau_p + \tau_\theta + (1-r)\tau_s\right)\left(\tau_y + \tau_s + \tau_\theta\right)^2 - \chi\delta_1\tau_p\tau_s\left((3-3r+4r\delta_1)\tau_y + (2-2r+3r\delta_1)(\tau_s+\tau_\theta)\right)\right] \\
&\propto \underbrace{\left(3\delta_1^2\tau_p + \tau_\theta + (1-r)\tau_s\right)\left(\tau_y + \tau_s + \tau_\theta\right)^2 - 3\chi\delta_1\tau_p\tau_s(1-r+r\delta_1)\left(\tau_y + \tau_s + \tau_\theta\right)}_{=:\mathrm{A}_1} + \mathrm{A}_2
\end{aligned}$$

where, using (47),

$$\mathrm{A}_1 = \left(\tau_y + \tau_s + \tau_\theta\right)^2 \left[\tau_\theta + (1-r)\tau_s + 3\frac{(1-\chi)\delta_1^2\tau_p\left(\tau_y + \tau_s + \tau_\theta\right)}{\tau_y + \tau_s + \tau_\theta + \chi\delta_1^2\tau_p} \right] > 0.$$

It remains to show that $\mathrm{A}_2 > 0$. We have

$$\begin{aligned}
\mathrm{A}_2 &\propto (1-r)(\tau_\theta + \tau_s) - \tau_y r \delta_1 > (1-r)(\tau_\theta + \tau_s) - \tau_y r \delta_1^{\mathrm{FC}} \\
&= (1-r)(\tau_\theta + \tau_s) - \tau_y r \frac{(1-r)\tau_s}{\tau_\theta + \tau_y + (1-r)\tau_s} = (1-r)\left(\tau_\theta + \tau_s\left(1 - r\frac{\tau_y}{\tau_\theta + \tau_y + (1-r)\tau_s}\right)\right) > 0.
\end{aligned}$$

Note that the result for $\frac{d\gamma_2}{d\tau_s}$ and $\frac{d\beta}{d\tau_s}$ follows when we establish $\frac{d}{d\tau_s}\frac{\delta_1}{\tau_s} < 0$:

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}\tau_s}\frac{\delta_1}{\tau_s} &= \frac{\frac{d\delta_1}{d\tau_s}\tau_s - \delta_1}{\tau_s^2} \propto \frac{d\delta_1}{d\tau_s}\tau_s - \delta_1 \propto -f_{\tau_s}\tau_s + f_\delta \delta_1 \\
&= (1-r)(1-\delta_1)\tau_s - \delta_1\left(\tau_y + \tau_\theta + (1-r)\tau_s\right) + 3\frac{\chi(1-r)\tau_s\delta_1^2\tau_p}{\tau_y+\tau_\theta+\tau_s} - \frac{\chi(1-r)\tau_s^2\delta_1^2\tau_p}{\left(\tau_y+\tau_\theta+\tau_s\right)^2} \\
&\quad - \delta_1^3\tau_p\left(3 - 4\frac{\chi r\tau_s}{\tau_y+\tau_\theta+\tau_s} + \frac{\chi r\tau_s^2}{\left(\tau_y+\tau_\theta+\tau_s\right)^2}\right)
\end{aligned}
$$

using f to replace $(1-r)(1-\delta_1)\tau_s$, we obtain a decomposition $\mathrm{B}_1 + \mathrm{B}_2$ where

$$
\begin{aligned}
\mathrm{B}_1 &= \frac{\chi\delta_1^2\tau_p\tau_s}{\left(\tau_y+\tau_\theta+\tau_s\right)^2}\left[\delta_1 r\left(\tau_y+\tau_\theta\right) - (1-r)\tau_s\right] \\
\mathrm{B}_2 &= -\delta_1\left(2\delta_1^2\tau_p + (1-r)\tau_s\right) + 2\frac{\chi\delta_1^2\left[1-r+\delta_1 r\right]\tau_p\tau_s}{\tau_y+\tau_\theta+\tau_s}
\end{aligned}
$$

Note that in B_1 the final term is negative if $r < 0$, otherwise, estimate $\delta_1 < \delta_1^{\mathrm{FC}}$ to arrive at

$$
\begin{aligned}
\delta_1 r\left(\tau_y+\tau_\theta\right) - (1-r)\tau_s &< \frac{(1-r)\tau_s}{\tau_\theta+\tau_y+(1-r)\tau_s} r\left(\tau_y+\tau_\theta\right) - (1-r)\tau_s \\
&= -\frac{(1-r)^2\tau_s\left(\tau_y+\tau_\theta+\tau_s\right)}{\tau_\theta+\tau_y+(1-r)\tau_s} < 0
\end{aligned}
$$

and therefore $\mathrm{B}_1 < 0$. Plugging (47) into B_2, we arrive at

$$
\mathrm{B}_2 = -\delta_1\left((1-r)\tau_s + 2\frac{(1-\chi)\delta_1^2\tau_p\left(\tau_y+\tau_\theta+\tau_s\right)}{\left(\tau_y+\tau_\theta+\tau_s+\chi\delta_1^2\tau_p\right)}\right) < 0,
$$

whence we have established $\frac{d}{d\tau_s}\frac{\delta_1}{\tau_s} < 0$ and therefore $\frac{d\gamma_2}{d\tau_s} < 0$ and $\frac{d\beta}{d\tau_s} > 0$. □

E.2 Informational Efficiency: Total Precision

The state-action coefficient β and the endogenous precision $\delta_1^2\tau_p$ represent natural metrics to measure the informational efficiency of markets and the information content of the aggregative signal, respectively. However, since they combine information acquisition, dissemination, and use, their comparative statics (contained in Proposition 1 and 2) give only a partial view of the impact of transparency on the efficiency of markets where agents underestimate the information content of aggregate statistics —

one of our key outcomes of interest.[39] To get a more complete understanding of this issue, we now study how τ_p affects $\tau_\Sigma := \tau_\theta + \tau_y + \tau_s + \delta_1^2\tau_p$, namely the total precision of information that agents possess about the state in the moment they choose their action (even though cursed agents do not use it efficiently — Section 6 studies an agent who can fully extrapolate from τ_Σ). We show that transparency has an ambiguous effect on τ_Σ as the dissemination improvement on $\delta_1^2\tau_p$ can exceed or fall short of the crowding out effect on τ_s. To see that, consider the following factorization of τ_Σ

$$\tau_\Sigma = \left(\frac{1 - r + r\delta_1}{\sqrt{c}}\right)\left(1 + \chi\frac{\delta_1^2\tau_p}{\tau_\theta + \tau_y + \tau_s}\right). \tag{68}$$

Since $\frac{d\delta_1}{d\tau_p} \geq 0$, the first factor in (68) is increasing in transparency if and only if $r < 0$, while the second factor is always increasing ($\frac{d\delta_1^2\tau_p}{d\tau_p} \geq 0$ and $\frac{d\tau_s}{d\tau_p} \leq 0$). Since $\chi = 0$ shuts down the second channel, transparency increases total precision in the rational benchmark if and only if actions are substitutes. In this case, transparency increases τ_Σ even for interior degrees of cursedness. In a game of complementarities instead, cursedness increases the range of parameters where transparency is desirable by scaling up the second factor. Hence, we obtain

Proposition 15. *The total precision available the agents is increasing in τ_p if and only if $r < \mathrm{R}(\chi)$, for a cutoff $\mathrm{R}(\chi)$, possibly trivial, with $\mathrm{R}(0) = 0$ and $\mathrm{R}' > 0$.*

Figure 9 gives a graphic representation of Proposition 15 (and shows that there is a similar pattern for the impact of public fundamental information τ_y): In the rational setting ($\chi = 0$, left panel), transparency increases total precision only in a game of strategic substitutes. In a game of strategic complements, by contrast, transparency increases total precision if and only if cursedness is large enough. A sufficiently cursed environment reduces the crowding out effect and therefore aggregative information becomes an effective tool to enhance total precision (right panel). In particular, in a fully cursed economy — where the crowding out effect is completely shut down — τ_Σ is always increasing in τ_p.

E.3 Proofs of the Lemmata in Appendix A (Learning Foundation)

Proof of Lemma 1: Let $\mathrm{W}^{(n)}$ denote the nth partial derivative of W with respect to τ, evaluated as above. We have

$\sigma\mathbb{E}[|\epsilon|]\Delta(\bar{\tau}, t) = \mathrm{W}^+(\bar{\tau}, t) - \mathrm{W}^-(\bar{\tau}, t)$

[39] A similar question is addressed in Morris and Shin (2005) who show that a central bank may inadvertently sabotage its own information collection from observing aggregate outcomes by providing precise information about these outcomes to the market.

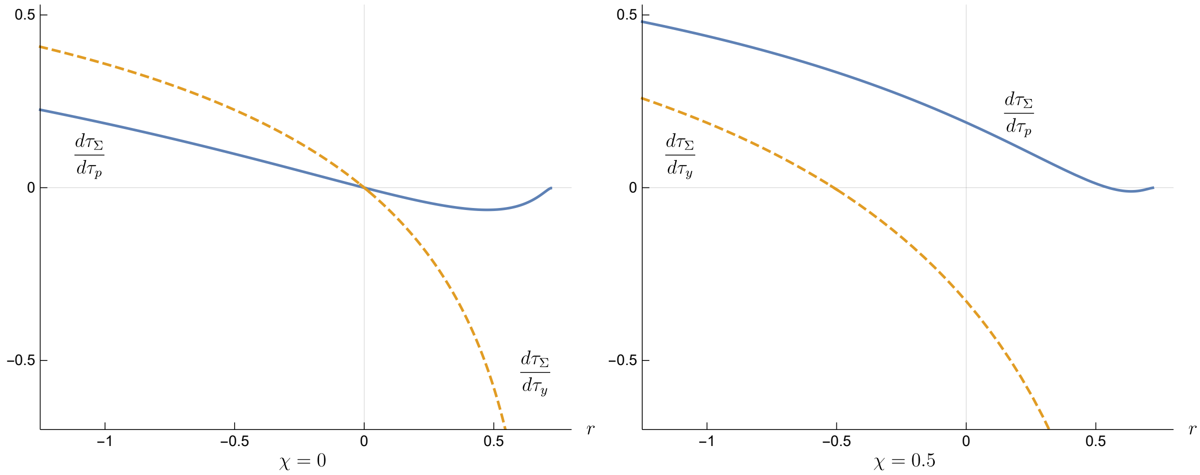

Figure 9: The effect of τ_p, τ_y on τ_Σ in the rational (left) and partially cursed ($\chi = 0.5$, right) model.

$$
\begin{aligned}
&\to \int_0^\infty W(\alpha(\delta,\bar\tau),\delta,\bar\tau+\sigma\epsilon)f(\epsilon)\,d\epsilon - \int_{-\infty}^0 W(\alpha(\delta,\bar\tau),\delta,\bar\tau+\sigma\epsilon)f(\epsilon)\,d\epsilon \\
&= \int_0^\infty \left(W(\alpha(\delta,\bar\tau),\delta,\bar\tau)+\sigma\epsilon W' + \frac{\sigma^2\epsilon^2}{2}W'' + o(\sigma^3\epsilon^3)\right)f(\epsilon)\,d\epsilon \\
&\quad - \int_{-\infty}^0 \left(W(\alpha(\delta,\bar\tau),\delta,\bar\tau)+\sigma\epsilon W' + \frac{\sigma^2\epsilon^2}{2}W'' + o(\sigma^3\epsilon^3)\right)f(\epsilon)\,d\epsilon \\
&= \int_0^\infty \left(\sigma\epsilon W' + \sigma^2\epsilon^2\frac{1}{2}W'' + o(\sigma^3\epsilon^3)\right)f(\epsilon)\,d\epsilon - \int_{-\infty}^0 \left(\sigma\epsilon W' + \sigma^2\epsilon^2\frac{1}{2}W'' + o(\sigma^3\epsilon^3)\right)f(\epsilon)\,d\epsilon \\
&= \sigma\mathbb{E}[|\epsilon|]W' + \frac{\sigma^2}{2}\mathbb{E}[\epsilon^2]W'' + o(\sigma^3)
\end{aligned}
$$

Where we use the fact that $W''' = \frac{\alpha_1^2}{\tau^4}$ is bounded for all $\tau+\sigma\epsilon$ since $\tau-\sigma\epsilon > \delta-\sigma > 0$ to ensure the order of convergence of the Taylor expansion. Dividing through by $\sigma\mathbb{E}[|\epsilon|]$ we obtain the result. □

Proof of Lemma 2: Suppose towards a contradiction that τ is a rest point with probability $p > 0$. Since $\Delta(\tau,t) \to \Delta(\tau)$ in probability, there exists a T such that for all $t \geq T$, $\mathbb{P}(|\Delta(\tau,t)-\Delta(\tau)| < ||\Delta(\tau)| - B(\sigma)|) > 1-p$. However, this implies that τ_T has moved away from τ with probability $1-p$, a contradiction to it being a rest point with probability p. □

Proof of Lemma 3: Since the reoptimization process is local, let us write everything in terms of periods spent with $\bar\tau_t = \tau$, denoted as $t = 1,2,\ldots,T,\ldots$ with slight abuse of notation. Note that the number of draws with $\tau_t > \tau$ at time t follows a binomial distribution $\text{Bin}(t,\frac{1}{2})$.

We need to show that with positive probability $\Delta(\tau,t) \in [-B(\sigma), B(\sigma)]$ for all $t > 2K$. We will proceed as follows. For every t, we provide a bound on the probability that

$|\Delta(t)| > B(\sigma)$. We then show that the sum of these bounds is smaller than 1, i.e., that with positive probability the process never leaves the bounds.

Let $\Psi_+ := \inf\{s > 0 : \mathbb{E}\left[e^{\frac{W_t}{s}} | \tau_t > 0, \bar{\tau}_t = \tau\right] \leq 2\}$ and $\Psi_- := \inf\{s > 0 : \mathbb{E}\left[e^{\frac{W_t}{s}} | \tau_t < 0, \bar{\tau}_t = \tau\right] \leq 2\}$ be the sub-exponential norms of realized welfare for positive and negative implementation errors, respectively. Note that these norms exist as W is the square of a Gaussian random variable and hence sub-exponential. Recall from Bernstein's inequality

$$\mathbb{P}\left\{\left|W_t^+ - \overline{W^+}\right| \geq \varepsilon\right\} \leq 2\exp\left\{-k\min\left\{\frac{\varepsilon^2 T^+(t)}{\Psi_+^2}, \frac{\varepsilon T^+(t)}{\Psi_+}\right\}\right\}$$

where $k > 0$ is a positive constant. For ε small, the first bound is the relevant one, and we will focus on this case.

Consider $\varepsilon = \sigma\mathbb{E}[|\epsilon|]\frac{B(\sigma)-|\Delta|}{2}$. Then, the probability that $|\Delta(t)| > B(\sigma)$ is bounded by $\mathbb{P}\left\{\left|W_t^+ - \overline{W^+}\right| \geq \varepsilon\right\} + \mathbb{P}\left\{\left|W_t^- - \overline{W^-}\right| \geq \varepsilon\right\}$, which is given by

$$\sum_{s=K}^{t-K} \frac{\binom{t}{s}\left(\frac{1}{2}\right)^t}{\sum_{s=K}^{t-K}\binom{t}{s}\left(\frac{1}{2}\right)^t}\left(2\exp\left\{-k\frac{\varepsilon^2 s}{\Psi_+^2}\right\} + 2\exp\left\{-k\frac{\varepsilon^2(t-s)}{\Psi_-^2}\right\}\right)$$

. Let $\psi_+ = k\frac{\varepsilon^2}{\Psi_+^2}$ and $\psi_- = k\frac{\varepsilon^2}{\Psi_-^2}$ denote the exponents. Note that

$$\sum_{s=K}^{t-K}\binom{t}{s}\left(\frac{1}{2}\right)^t \exp\{-\psi_+ s\} = \left(\frac{1}{2}\right)^t e^{-\psi_+ K}(1+e^{-\psi_+})^t - \left(\frac{1}{2}\right)^t e^{-\psi_+(1+t)}\binom{t}{1-K+t}\mathrm{Hyp}(1, 1-K, 2-K+t, -e^{-\psi_+}) \tag{69}$$

where Hyp denotes the Gaussian hypergeometric function. If a, c are positive integers, b is a negative integer and $z \in (-1, 0)$, the hypergeometric function evaluates to

$$\mathrm{Hyp}(a, b, c, z) = \sum_{n=0}^{|b|}\binom{|b|}{n}\frac{(a)_n}{(c)_n}|z|^n$$

where $(a)_n, (c)_n$ are the rising Pochhammer series, which are positive for all n. Therefore the second term in (69) is positive and

$$\sum_{s=K}^{t-K}\binom{t}{s}\left(\frac{1}{2}\right)^t \exp\{-\psi_+ s\} \leq \left(\frac{1}{2}\right)^t e^{-\psi_+ K}(1 + e^{-\psi_+})^t$$

Consider now the normalization factor $\sum_{s=K}^{t-K}\binom{t}{s}\left(\frac{1}{2}\right)^t$. Note that this is the probability that a $\mathrm{Bin}(t, \frac{1}{2})$ is not in the tails of length K. Clearly, this probability is increasing in t for fixed K. We can therefore estimate it by its value at $t = 2K$ and obtain $\sum_{s=K}^{t-K}\binom{t}{s}\left(\frac{1}{2}\right)^t >$

$\binom{2K}{K}\left(\frac{1}{2}\right)^K$. Using Stirling's approximation, we write[40]

$$\binom{2K}{K}\left(\frac{1}{2}\right)^K = \frac{\sqrt{2\pi 2K}\left(\frac{2K}{e}\right)^{2K}}{\left(\sqrt{2\pi K}\right)^2\left(\frac{K}{e}\right)^{2K}}\left(\frac{1}{2}\right)^K = \frac{1}{\sqrt{\pi K}}$$

Therefore

$$\sum_{s=K}^{t-K}\frac{\binom{t}{s}\left(\frac{1}{2}\right)^t}{\sum_{s=K}^{t-K}\binom{t}{s}\left(\frac{1}{2}\right)^t}\left(2\exp\left\{-k\frac{\varepsilon^2 s}{\Psi_+^2}\right\}+2\exp\left\{-k\frac{\varepsilon^2(t-s)}{\Psi_-^2}\right\}\right) \leq 2\sqrt{\pi K}\left(\frac{1}{2}\right)^t\left(e^{-\psi_+ K}(1+e^{-\psi_+})^t + e^{-\psi_- K}(1+e^{-\psi_-})^t\right)$$

Finally, consider the sum of these estimates

$$\sum_{t=2K}^{\infty} 2\sqrt{\pi K}\left(\frac{1}{2}\right)^t\left(e^{-\psi_+ K}(1+e^{-\psi_+})^t + e^{-\psi_- K}(1+e^{-\psi_-})^t\right) = 2\sqrt{\pi K}\left(e^{-\psi_+ K}\frac{\left(\frac{1+e^{-\psi_+}}{2}\right)^{2K}}{1-\frac{1+e^{-\psi_+}}{2}} + e^{-\psi_- K}\frac{\left(\frac{1+e^{-\psi_-}}{2}\right)^{2K}}{1-\frac{1+e^{-\psi_-}}{2}}\right)$$

Note that the RHS is eventually decreasing in K and goes to zero as $K \to \infty$. Therefore, we can pick a K sufficiently big such that the RHS is smaller than one. Then, the process stays within the bounds forever with positive probability. □

E.4 Proofs of the Lemmata in Appendix B (Proof Appendix)

Proof of Lemma 4: Towards a contradiction, let $\delta_1 < 0$. Then, $\tau_s = -\frac{\delta_1}{\sqrt{c}}$ and f reads

$$\left(\tau_\theta + \tau_y + \delta_1^2\tau_p\right) + \chi\frac{1}{\sqrt{c}\left(\tau_\theta+\tau_y\right)-\delta_1}\left[(1-r)+r\delta_1\right]\delta_1^2\tau_p + \frac{1}{\sqrt{c}}(1-r)(1-\delta_1) = 0$$

Clearly, for $\delta_1 = 0$, the expression is strictly positive. Furthermore, we have $\frac{\mathrm{d}}{\mathrm{d}\delta_1}f\left(\delta_1, -\frac{\delta_1}{\sqrt{c}}\right) < 0$, as

$$\begin{aligned}
&2\delta_1\tau_p + \chi\frac{\left[(1-r)+r\delta_1\right]\delta_1^2\tau_p}{\left(\sqrt{c}\left(\tau_\theta+\tau_y\right)-\delta_1\right)^2} + \chi\frac{\left[(1-r)+3r\delta_1^2\right]\tau_p}{\sqrt{c}\left(\tau_\theta+\tau_y\right)-\delta_1} - \frac{1}{\sqrt{c}}(1-r)\\
=&\frac{1}{\left(\sqrt{c}\left(\tau_\theta+\tau_y\right)-\delta_1\right)^2}\Big[2\delta_1\tau_p\left(\sqrt{c}\left(\tau_\theta+\tau_y\right)-\delta_1\right)^2 + \chi\left[(1-r)+r\delta_1\right]\delta_1^2\tau_p\\
&+\chi\left(\sqrt{c}\left(\tau_\theta+\tau_y\right)-\delta_1\right)\left[2(1-r)\delta_1+3r\delta_1^2\right]\tau_p - \frac{1}{\sqrt{c}}(1-r)\left(\sqrt{c}\left(\tau_\theta+\tau_y\right)-\delta_1\right)^2\Big]\\
\propto& 2\delta_1^3\tau_p - 4\delta_1\tau_p\sqrt{c}\left(\tau_\theta+\tau_y\right) + 2\delta_1\tau_p c\left(\tau_\theta+\tau_y\right)^2 + \chi\left[(1-r)+r\delta_1\right]\delta_1^2\tau_p\\
&+\chi\left(\sqrt{c}\left(\tau_\theta+\tau_y\right)-\delta_1\right)\left[2(1-r)\delta_1+3r\delta_1^2\right]\tau_p - \frac{1}{\sqrt{c}}(1-r)\left(\sqrt{c}\left(\tau_\theta+\tau_y\right)-\delta_1\right)^2
\end{aligned}$$

[40] Note that the approximation error is small and can be easily taken into account in the following arguments. We suppress it for brevity.

$$=2\delta_1^3\tau_p(1-\chi r)-\delta_1^2\left[\frac{1}{\sqrt{c}}(1-r)+\left(4\tau_p-3\chi r\tau_p\right)\sqrt{c}\left(\tau_\theta+\tau_y\right)+\chi(1-r)\tau_p\right]$$
$$+\delta_1 2c\left(\tau_\theta+\tau_y\right)\left[(1-r)\left(1+\chi\tau_p\sqrt{c}\right)\tau_p+\tau_p c\left(\tau_\theta+\tau_y\right)\right]-\sqrt{c}(1-r)\left(\tau_\theta+\tau_y\right)^2<0$$

Hence, we cannot have an solution as $f>0$ for all $\delta_1<0$ and there is no such equilibrium. □

Proof of Lemma 5: Suppose (19) holds. Then, using (47) and (17), we have

$$\frac{\mathrm{d}}{\mathrm{d}\delta_1}\tilde{f}(\delta_1)=\frac{(1-r)\left(\delta_1+\sqrt{c}\left(\tau_\theta+\tau_y\right)\right)+\sqrt{c}\delta_1^2\tau_p\chi\left[\frac{\delta_1+\sqrt{c}\left(\tau_\theta+\tau_y+\delta_1^2\tau_p\right)}{\delta_1+\sqrt{c}\left(\tau_\theta+\tau_y+\chi\delta_1^2\tau_p\right)}-r\right]}{\delta_1+\sqrt{c}\left(\tau_\theta+\tau_y\right)}+2\sqrt{c}\delta_1\tau_p\left((1-\chi)\frac{\delta_1+\sqrt{c}\left(\tau_\theta+\tau_y\right)}{\delta_1+\sqrt{c}\left(\tau_\theta+\tau_y+\chi\delta_1^2\tau_p\right)}\right)$$
$$\geq\frac{(1-r)\left(\delta_1+\sqrt{c}\left(\tau_\theta+\tau_y\right)\right)+\sqrt{c}\delta_1^2\tau_p\chi\left[\frac{\delta_1+\sqrt{c}\left(\tau_\theta+\tau_y+\delta_1^2\tau_p\right)}{\delta_1+\sqrt{c}\left(\tau_\theta+\tau_y+\chi\delta_1^2\tau_p\right)}-r\right]}{\delta_1+\sqrt{c}\left(\tau_\theta+\tau_y\right)}\geq\frac{\sqrt{c}\delta_1^2\tau_p\chi[1-r]}{\delta_1+\sqrt{c}\left(\tau_\theta+\tau_y\right)}>0$$

since the first term in the numerator is positive and the fraction in square brackets is greater than 1. □

Proof of Lemma 6: Note that $g_\delta=2\delta_1>0$, $g_{\tau_s}=-2c\tau_s<0$ and

$$f_\delta=\left(\tau_\theta+\tau_y+3\delta_1^2\tau_p\right)-\chi\frac{\tau_s}{\tau_\theta+\tau_y+\tau_s}\left[2+r(3\delta_1-2)\right]\delta_1\tau_p+(1-r)\tau_s$$
$$f_{\tau_s}=-\chi\frac{\tau_\theta+\tau_y}{\left(\tau_\theta+\tau_y+\tau_s\right)^2}\left[1+r(\delta_1-1)\right]\delta_1^2\tau_p-(1-r)(1-\delta_1).$$

By direct computation

$$\begin{aligned}
g_\delta f_{\tau_s}-g_{\tau_s}f_\delta=&2\delta_1\left(-\chi\frac{\tau_\theta+\tau_y}{\left(\tau_\theta+\tau_y+\tau_s\right)^2}\left[1+r(\delta_1-1)\right]\delta_1^2\tau_p-(1-r)(1-\delta_1)\right)\\
&-(-2c\tau_s)\left(\left(\tau_\theta+\tau_y+3\delta_1^2\tau_p\right)-\chi\frac{\tau_s}{\tau_\theta+\tau_y+\tau_s}\left[2+r(3\delta_1-2)\right]\delta_1\tau_p+(1-r)\tau_s\right)\\
=&2\delta_1\left(-\chi\frac{\tau_\theta+\tau_y}{\left(\tau_\theta+\tau_y+\tau_s\right)^2}\left[1+r(\delta_1-1)\right]\delta_1^2\tau_p-\left[1+r(\delta_1-1)\right]+\delta_1\right)\\
&+2\frac{\delta_1}{\tau_s}\left(\left(\tau_\theta+\tau_y+3\delta_1^2\tau_p\right)\delta_1-2\chi\frac{\tau_s}{\tau_\theta+\tau_y+\tau_s}\left[1+r(\delta_1-1)\right]\delta_1^2\tau_p+(1-r)\tau_s\delta_1-\chi\frac{\tau_s}{\tau_\theta+\tau_y+\tau_s}r\delta_1\delta_1^2\tau_p\right)\\
\overset{(f_\delta>0)}{\geq}&2\frac{\delta_1}{\tau_s}\left(-\chi\frac{\tau_\theta+\tau_y}{\tau_\theta+\tau_y+\tau_s}\frac{\left(\tau_\theta+\tau_y+\tau_s+\delta_1^2\tau_p\right)\delta_1}{\tau_\theta+\tau_y+\tau_s+\chi\delta_1^2\tau_p}\delta_1^2\tau_p-\frac{\delta_1\left(\tau_\theta+\tau_y+\tau_s\right)\left(\tau_\theta+\tau_y+\tau_s+\delta_1^2\tau_p\right)}{\tau_\theta+\tau_y+\tau_s+\chi\delta_1^2\tau_p}+\delta_1\tau_s\right)\\
&+2\frac{\delta_1}{\tau_s}\left(\left(\tau_\theta+\tau_y+3\left(1-\chi\frac{\tau_\theta+\tau_y+\tau_s+\delta_1^2\tau_p}{\tau_\theta+\tau_y+\tau_s+\chi\delta_1^2\tau_p}\right)\delta_1^2\tau_p\right)\delta_1+(1-r)\tau_s\delta_1+\chi\frac{\tau_s}{\tau_\theta+\tau_y+\tau_s}[1-r]\delta_1^2\tau_p\right)\\
=&2\delta_1^2\left\{-\frac{\tau_y}{\tau_s}-\frac{\tau_\theta}{\tau_s}+\delta_1^2\tau_p\left(-\frac{1}{\tau_s}+\frac{1}{\tau_\theta+\tau_y+\tau_s}-\frac{1-\chi}{\tau_\theta+\tau_y+\tau_s+\chi\delta_1^2\tau_p}\right)\right\}\\
&+2\frac{\delta_1}{\tau_s}\left(\left(\tau_\theta+\tau_y+3\left(1-\chi\frac{\tau_\theta+\tau_y+\tau_s+\delta_1^2\tau_p}{\tau_\theta+\tau_y+\tau_s+\chi\delta_1^2\tau_p}\right)\delta_1^2\tau_p\right)\delta_1+(1-r)\tau_s\delta_1+\chi\frac{\tau_s}{\tau_\theta+\tau_y+\tau_s}[1-r]\delta_1^2\tau_p\right)
\end{aligned}$$

$$\overset{(r<1\&(17))}{\geq} 2\delta_1^3\tau_p(1-\chi)\sqrt{c}\left\{\frac{2\delta_1^2+\sqrt{c}\chi\delta_1^3\tau_p+4\sqrt{c}\delta_1\left(\tau_\theta+\tau_y\right)+2c\left(\tau_\theta+\tau_y\right)^2}{\left[\delta_1+\sqrt{c}\left(\tau_\theta+\tau_y\right)\right]\left[\delta_1+\sqrt{c}\chi\delta_1^2\tau_p+\sqrt{c}\left(\tau_\theta+\tau_y\right)\right]}\right\}\geq 0$$

The last equality follows from lengthy but straightforward calculation. The inequality follows since the expression is decreasing in r and we hence set $r=1$ as a worst case, obtaining our result. □

Proof of Lemma 7: We have

$$\begin{aligned}
f_\delta^\star &= \frac{\left[2\left(\tau_\theta+\tau_y+\delta_1^2\tau_p\right)\delta_1\tau_p\right]\left(\tau_\theta+\tau_y+\tau_p\delta_1\right)-\tau_p\left(\tau_\theta+\tau_y+\delta_1^2\tau_p\right)^2}{\left(\tau_\theta+\tau_y+\tau_p\delta_1\right)^2}+(1-r)\frac{1}{\sqrt{c}}\\
&= \frac{\tau_p\left(\tau_\theta+\tau_y+\delta_1^2\tau_p\right)}{\left(\tau_\theta+\tau_y+\tau_p\delta_1\right)^2}\left[\frac{(1-r)}{\sqrt{c}}\frac{\left(\tau_\theta+\tau_y+\tau_p\delta_1\right)^2}{\tau_p\left(\tau_\theta+\tau_y+\delta_1^2\tau_p\right)}-[2\delta_1-1]\left(\tau_\theta+\tau_y\right)+\delta_1^2\tau_p\right]\\
&\geq \frac{\tau_p\left(\tau_\theta+\tau_y+\delta_1^2\tau_p\right)}{\left(\tau_\theta+\tau_y+\tau_p\delta_1\right)^2}\left[\frac{(1-r)}{\sqrt{c}}\left(\frac{\left(\tau_\theta+\tau_y+\tau_p\delta_1\right)^2}{\tau_p\left(\tau_\theta+\tau_y+\delta_1^2\tau_p\right)}-[2\delta_1-1]\right)+\delta_1^2\tau_p\right]
\end{aligned}$$

using that $\frac{1-r}{\tau_\theta+\tau_y}>\sqrt{c}$. Then, from

$$\begin{aligned}
\frac{\left(\tau_\theta+\tau_y+\tau_p\delta_1\right)^2}{\tau_p\left(\tau_\theta+\tau_y+\delta_1^2\tau_p\right)}-[2\delta_1-1] &\propto \tau_p\left(\tau_\theta+\tau_y+\delta_1^2\tau_p\right)+\left(\tau_\theta+\tau_y+\tau_p\delta_1\right)^2-2\delta_1\tau_p\left(\tau_\theta+\tau_y+\delta_1^2\tau_p\right)\\
&= \tau_p\left(\tau_\theta+\tau_y\right)+\left(\tau_\theta+\tau_y\right)^2+(2-2\delta_1)\tau_p\delta_1^2\tau_p>0
\end{aligned}$$

we obtain the result. □

Proof of Lemma 8: Taking the $\tau_p\to\infty$ limit in (18) we get the limit reliance on private information is[41]

$$\delta_1^\infty=\frac{\chi(1-r)-\left(\tau_\theta+\tau_y\right)\sqrt{c}}{(1-r\chi)}, \tag{70}$$

which we can plug in (28) to obtain that welfare in the limit equilibrium is given by

$$W^\infty=-\frac{2\sqrt{c}\chi(1-r)^2-c(1-2r+r\chi)\left(\tau_\theta+\tau_y\right)}{(1-r)(1-\chi r)}, \tag{71}$$

[41] Because the optimality condition (17) implies that τ_s^∞ solves $\tau_s\sqrt{c}=\frac{\chi\tau_s(1-r)}{\tau_\theta+\tau_y+\tau_s(1-r\chi)}$, a interior limit equilibrium exists if (and only if) $\chi>\sqrt{c}\frac{\tau_\theta+\tau_y}{(1-r)}$, a condition that we assume satisfied. Notice that this condition does not contradict the sufficiency of condition (19) which holds for any interior τ_p. The order of limits is relevant, as for every fixed τ_p, the influence of transparency vanishes as δ_1 and τ_s go to zero. All details about the derivations of the transparent limit are available upon request.

which we can directly differentiate. To see the comparative static, note that the coefficient of $\tau_\theta + \tau_y$ is $1 - 2r + r\chi$. Consider the case where $r < 0$, then this expression is negative only for $\chi > 2$, so this case is irrelevant. Instead, with $r > 0$, we get that the impact of τ_θ, τ_y is negative iff $\chi \leq 2 - \frac{1}{r}$.

The derivative w.r.t. costs is

$$\frac{\partial W^\infty}{\partial c} \propto -\frac{\chi}{\sqrt{c}}(1-r)^2 + (1 - 2r + r\chi)\left(\tau_\theta + \tau_y\right)$$

and

$$\begin{aligned} -\frac{\chi}{\sqrt{c}}(1-r)^2 + (1 - 2r + r\chi)\left(\tau_\theta + \tau_y\right) \geq 0 \\ \chi\left[(1-r)^2 - r\sqrt{c}\left(\tau_\theta + \tau_y\right)\right] \leq \sqrt{c}(1-2r)\left(\tau_\theta + \tau_y\right) \end{aligned}$$

Hence, there is two cases we need to consider. First, if $(1-r)^2 - r\sqrt{c}\left(\tau_\theta + \tau_y\right) < 0$: Note that this can only be the case if $r > \frac{1}{2}$, since otherwise by (19)

$$(1-r)^2 - r\sqrt{c}\left(\tau_\theta + \tau_y\right) \geq (1-r)^2 - r(1-r) = (1-2r)(1-r) > 0.$$

Then, we obtain a lower bound for χ:

$$\chi \geq \sqrt{c}(1-2r)\frac{\tau_\theta + \tau_y}{(1-r)^2 - r\sqrt{c}\left(\tau_\theta + \tau_y\right)}$$

but this bound rules out increasingness in costs, since

$$\sqrt{c}(1-2r)\frac{\tau_\theta + \tau_y}{(1-r)^2 - r\sqrt{c}\left(\tau_\theta + \tau_y\right)} \geq 1 \iff \sqrt{c}\left(\tau_\theta + \tau_y\right) \leq (1-r)$$

which is guaranteed by (19).

Second, if $(1-r)^2 - r\sqrt{c}\left(\tau_\theta + \tau_y\right) > 0$: we get an upper bound

$$\chi \leq \sqrt{c}(1-2r)\frac{\tau_\theta + \tau_y}{(1-r)^2 - r\sqrt{c}\left(\tau_\theta + \tau_y\right)}.$$

The resulting interval is nontrivial only if $r < 0$, as by (19)

$$\begin{aligned} \sqrt{c}\frac{\tau_\theta + \tau_y}{1-r} < \sqrt{c}(1-2r)\frac{\tau_\theta + \tau_y}{(1-r)^2 - r\sqrt{c}\left(\tau_\theta + \tau_y\right)} \\ r\left[\frac{\sqrt{c}\left(\tau_\theta + \tau_y\right)}{(1-r)}\right] > r \iff r < 0. \end{aligned}$$

□

E.5 Omitted Proofs

Proof of Proposition 15: Computing

$$\frac{\partial}{\partial \tau_p}\left(\tau_\theta + \tau_y + \tau_s + \tau_p\delta_1^2\right) = \frac{\partial \tau_s}{\partial \tau_p} + \delta_1^2 + 2\delta_1\tau_p\frac{\partial \delta_1}{\partial \tau_p} = \frac{-g_\delta f_{\tau_p}}{g_\delta f_{\tau_s} - g_{\tau_s} f_\delta} + \delta_1^2 + 2\delta_1\tau_p\frac{g_{\tau_s} f_{\tau_p}}{g_\delta f_{\tau_s} - g_{\tau_s} f_\delta}$$
$$\propto -g_\delta f_{\tau_p} + \delta_1^2\left(g_\delta f_{\tau_s} - g_{\tau_s} f_\delta\right) + 2\delta_1\tau_p g_{\tau_s} f_{\tau_p}$$

Plugging in and using (47) and (17) wherever apparent yields a linear equation in r given δ_1, which can be solved for the implicit equation

$$\mathrm{R}_p(\chi) = \chi\left(\frac{\delta_1 + \sqrt{c}\left(\tau_\theta + \tau_y + \delta_1^2\tau_p\right)}{\delta_1 + \sqrt{c}\left(\tau_\theta + \tau_y + \chi\delta_1^2\tau_p\right)}\right)^2 = \chi\left(\frac{\tau_\theta + \tau_y + \tau_s + \delta_1^2\tau_p}{\tau_\theta + \tau_y + \tau_s + \chi\delta_1^2\tau_p}\right)^2$$

such that $\frac{\partial}{\partial \tau_p}\left(\tau_\theta + \tau_y + \tau_s + \tau_p\delta_1^2\right) > 0$ if $r < \mathrm{R}_p$. To derive the properties of R_p, let us define

$$k(\chi, r) = \chi\left(\frac{\delta_1 + \sqrt{c}\left(\tau_\theta + \tau_y + \delta_1^2\tau_p\right)}{\delta_1 + \sqrt{c}\left(\tau_\theta + \tau_y + \chi\delta_1^2\tau_p\right)}\right)^2 - r = 0$$

To show that $\mathrm{R}_p(\chi)$ is increasing in χ, we need to establish that $\mathrm{R}_p'(\chi) = -\frac{k_\chi}{k_r} > 0$. By lengthy computation, it is easy to show that,

$$k_\chi \propto \left(\delta_1 + \sqrt{c}\left(\tau_\theta + \tau_y + \delta_1^2\tau_p\right)\right)\left(\delta_1 + \sqrt{c}\left(\tau_\theta + \tau_y - \chi\delta_1^2\tau_p\right)\right) + \frac{\partial \delta_1}{\partial \chi} 2\sqrt{c}(1-\chi)\tau_p\delta_1\left(\delta_1 + 2\sqrt{c}\left(\tau_\theta + \tau_y\right)\right)$$

Note that $\frac{\partial \delta_1}{\partial \chi} > 0$, and – if the first term is positive – we have $k_\chi > 0$. This is the case for valid parameters: Suppose towards a contradiction that it is not, i.e.

$$\delta_1 + \sqrt{c}\left(\tau_\theta + \tau_y - \chi\delta_1^2\tau_p\right) < 0 \iff \tau_p > \frac{\delta_1 + \sqrt{c}\left(\tau_\theta + \tau_y\right)}{\sqrt{c}\chi\delta_1^2}$$

But note that $r = \chi\left(\frac{\delta_1 + \sqrt{c}\left(\tau_\theta + \tau_y + \delta_1^2\tau_p\right)}{\delta_1 + \sqrt{c}\left(\tau_\theta + \tau_y + \chi\delta_1^2\tau_p\right)}\right)^2$ is increasing in τ_p (as a partial derivative), so this would imply that

$$r = \chi\left(\frac{\delta_1 + \sqrt{c}\left(\tau_\theta + \tau_y + \delta_1^2\tau_p\right)}{\delta_1 + \sqrt{c}\left(\tau_\theta + \tau_y + \chi\delta_1^2\tau_p\right)}\right)^2 > \chi\left(\frac{\delta_1 + \sqrt{c}\left(\tau_\theta + \tau_y + \delta_1^2\frac{\delta_1 + \sqrt{c}(\tau_\theta + \tau_y)}{\sqrt{c}\chi\delta_1^2}\right)}{\delta_1 + \sqrt{c}\left(\tau_\theta + \tau_y + \chi\delta_1^2\frac{\delta_1 + \sqrt{c}(\tau_\theta + \tau_y)}{\sqrt{c}\chi\delta_1^2}\right)}\right)^2$$
$$= \chi\left(\frac{\delta_1 + \sqrt{c}\left(\tau_\theta + \tau_y + \frac{\delta_1}{\chi\sqrt{c}} + \frac{1}{\chi}\left(\tau_\theta + \tau_y\right)\right)}{\delta_1 + \sqrt{c}\left(\tau_\theta + \tau_y + \frac{\delta_1}{\sqrt{c}} + \left(\tau_\theta + \tau_y\right)\right)}\right)^2 = \chi\left(\frac{\left(1 + \frac{1}{\chi}\right)\left(\delta_1 + \sqrt{c}\left(\tau_\theta + \tau_y\right)\right)}{2\delta_1 + 2\sqrt{c}\left(\tau_\theta + \tau_y\right)}\right)^2$$

$$= \frac{(\chi+1)^2}{4\chi} = 1 + \frac{(\chi-1)^2}{4\chi} > 1$$

a contradiction. Hence we require that τ_p is smaller, otherwise the cutoff is trivial (i.e. greater than one). Hence, whenever we have an interior cutoff, we have $k_\chi > 0$.

It remains to show (to get $\frac{dr}{d\chi} > 0$) that $k_r < 0$. (with linear costs), which is the case, as

$$k_r = \frac{\partial}{\partial r}\left(\chi\left(\frac{\delta_1 + \sqrt{c}\left(\tau_\theta + \tau_y + \delta_1^2\tau_p\right)}{\delta_1 + \sqrt{c}\left(\tau_\theta + \tau_y + \chi\delta_1^2\tau_p\right)}\right)^2 - r\right) = \underbrace{\chi(1-\chi)\,2\sqrt{c}\tau_p\delta_1 \frac{\left(\delta_1 + 2\sqrt{c}\left(\tau_\theta+\tau_y\right)\right)\left(\delta_1 + \sqrt{c}\left(\tau_\theta + \tau_y + \delta_1^2\tau_p\right)\right)}{\left[\delta_1 + \sqrt{c}\left(\tau_\theta + \tau_y + \chi\delta_1^2\tau_p\right)\right]^3}\frac{\partial\delta_1}{\partial r} - 1}_{\leq -1 < 0}$$

In addition, we have that at the solution to $k = 0$, we always have $k_r < 0$, whence there exists a unique solution and therefore a cutoff $\mathrm{R}(\chi)$, such that $k \geq 0$ iff $r \leq \mathrm{R}(\chi)$, as we wanted to show. In addition, $\mathrm{R}' > 0$, and $\mathrm{R}(0) = 0$.

To see that the cutoff can be trivial, note that

$$\frac{\mathrm{d}}{\mathrm{d}\tau_p}\left(\tau_\theta + \tau_y + \tau_s + \delta_1^2\tau_p\right) \to 1 - \frac{1-\chi}{2\sqrt{c}\left(\tau_\theta+\tau_y\right)}$$

as $r \to 1 - \sqrt{c}\left(\tau_\theta + \tau_y\right)$, which is of ambiguous sign. □

Lemma 10. *Welfare as a function of* δ *is given by (28).*

Proof. We have

$$\begin{aligned}
\mathrm{W}(\alpha,\delta) &= \mathbb{E}\left[-(1-r)(a_i-\theta)^2 - r(a_i-\bar{a})^2\right] \qquad (72)\\
&= \mathbb{E}\left[-(1-r)(\alpha_1 s_i + \alpha_2 y + \alpha_3 p - \theta)^2 - r(\alpha_1 s_i + \alpha_2 y + \alpha_3 p - \bar{a})^2\right]\\
&= \mathbb{E}\left[-(1-r)\left(\alpha_1(\theta+z_s) + \alpha_2\left(\theta+z_y\right) + \alpha_3\left(\frac{\delta_1+\delta_2}{1-\delta_3}\theta + \frac{\delta_2}{1-\delta_3}z_y + \frac{1}{1-\delta_3}z_p\right) - \theta\right)^2\right.\\
&\left. -r\left(\alpha_1(\theta+z_s) + \alpha_2\left(\theta+z_y\right) + \alpha_3\left(\frac{\delta_1+\delta_2}{1-\delta_3}\theta + \frac{\delta_2}{1-\delta_3}z_y + \frac{1}{1-\delta_3}z_p\right) - \left(\frac{\delta_1+\delta_2}{1-\delta_3}\theta + \frac{\delta_2}{1-\delta_3}z_y + \frac{\delta_3}{1-\delta_3}z_p\right)\right)^2\right]\\
&= -\frac{1}{(1-\delta_3)^2}\left\{\frac{(\alpha_2(1-\delta_3)+\alpha_3\delta_2)^2 + \delta_2(\delta_2 - 2\alpha_3\delta_2 - 2\alpha_2(1-\delta_3))r}{\tau_y} + \frac{\alpha_3^2 - 2\alpha_3\delta_3 r + \delta_3^2 r}{\tau_p} + \frac{\alpha_1^2(1-\delta_3)^2}{\tau_s}\right.\\
&\left. + \frac{(1-(1-\delta_3)(\alpha_1+\alpha_2) - \alpha_3(\delta_1+\delta_2) - \delta_3)^2 - (1 + (\delta_1+\delta_2)(1-2\alpha_3) - 2(\alpha_1+\alpha_2)(1-\delta_3) - \delta_3)(1-\delta_1-\delta_2-\delta_3)r}{\tau_\theta}\right\}
\end{aligned}$$

where the last step follows after lengthy but straightforward computation. Imposing $\alpha_i = \delta_i$, we get

$$\mathrm{W}(\delta) = -\frac{(1-r)}{(1-\delta_3)^2}\left\{\frac{\delta_2^2}{\tau_y} + \frac{\delta_3^2}{\tau_p} + \frac{(1-\delta_1-\delta_2-\delta_3)^2}{\tau_\theta}\right\} - \frac{\delta_1^2}{\tau_s}$$ □

Proof of Proposition 5. Plugging the equilibrium expressions (12) into W_{δ_1} and using $f(\delta) = 0$ yields

$$W_{\delta_1} = 2\frac{(1-\chi)\delta_1^3\tau_p\left(\tau_\theta + \tau_y + \tau_s\right)}{(1-r)\tau_s^2\left(\tau_\theta + \tau_y + \tau_s + \chi\delta_1^2\tau_p\right)}$$

which is zero for $\chi = 1$.

Plugging the equilibrium expressions for δ_2, δ_3 into W_{δ_2} yields

$$\begin{aligned} W_{\delta_2} =& 2\frac{\left[(1-r)\left(\tau_\theta + \tau_y + \tau_s + (1-\chi)\delta_1\tau_p\right) + (1-r\chi)\delta_1^2\tau_p\right]}{(1-r)\tau_\theta\left(\tau_\theta + \tau_y + \tau_s\right)\left(\tau_\theta + \tau_y + \tau_s + \delta_1^2\tau_p\right)^2} \\ &\cdot\left[(1-r)\chi\delta_1^2\tau_p\tau_s + (1-r)\tau_s\left(\tau_\theta + \tau_y + \tau_s\right) + \delta_1\left(\tau_\theta + \tau_y + \tau_s\right)\left(\tau_\theta + \tau_y + (1-r)\tau_s\right) + \delta_1^3\tau_p\left(\tau_\theta + \tau_y + (1-r\chi)\tau_s\right)\right] \end{aligned}$$

note that the final factor can be written as

$$\left(\tau_\theta + \tau_y + \tau_s\right)\left[\chi\frac{\tau_s}{\left(\tau_\theta + \tau_y + \tau_s\right)}\delta_1^2\tau_p(1-r+\delta_1 r) + \delta_1\left(\tau_\theta + \tau_y + \delta_1^2\tau_p\right) + (1-r)(1-\delta_1)\tau_s\right]$$

where we recognize the factor as $f(\delta_1) = 0$, whence $W_{\delta_2}(\delta^{EQ}) = 0$.

Plugging into W_{δ_3} and simplifying with heavy use of $f(\delta) = 0$, we get

$$W_{\delta_3} = -\frac{2\chi\left(\tau_\theta + \tau_y + \tau_s + \delta_1^2\tau_p\right)\left[\delta_1\left(\tau_\theta + \tau_y\right) - (1-r)\tau_s\right]^2}{(1-r)^2\tau_s^3\left(\tau_\theta + \tau_y + \tau_s + \chi\delta_1^2\tau_p\right)}$$

which clearly is zero in the rational case. □

Lemma 11. *Equation (34): if $\chi \in \{0, 1\}$, then* $W_\chi^{EQ} = -\sqrt{c}(1 + \delta_1)$.

Proof. Plugging the fully cursed equilibrium into (28), we obtain

$$W_1^{EQ} = -\frac{2(1-r)\sqrt{c} - c\left(\tau_\theta + \tau_y\right)}{1-r} = -\sqrt{c}(1 + \delta_1)$$

In the rational case, we instead have, using f

$$W_0^{EQ} = -\frac{2(1-r)\sqrt{c}\delta_1^R + c\left(\tau_\theta + \tau_y + \left(\delta_1^R\right)^2\tau_p\right)}{1-r} = -\frac{(1-r)\sqrt{c}\delta_1^R + (1-r)\sqrt{c}}{1-r} = -\sqrt{c}(1 + \delta_1).$$ □