CSEF
Centre for Studies in Economics and Finance

# WORKING PAPER NO. 690

# *The Social Value of Overreaction to Information*

**Matteo Bizzarri and Daniele d'Arienzo**

**November 2023**

**University of Naples Federico II** **University of Salerno** **Bocconi University, Milan**

**CSEF - Centre for Studies in Economics and Finance**
**DEPARTMENT OF ECONOMICS AND STATISTICS – UNIVERSITY OF NAPLES FEDERICO II**
**80126 NAPLES - ITALY**
**Tel. and fax +39 081 675372 – e-mail: csef@unina.it**
**ISSN: 2240-9696**

CSEF
Centre for Studies in Economics and Finance

# WORKING PAPER NO. 690

# *The social value of overreaction to information*

**Matteo Bizzarri*[*] and Daniele d'Arienzo[†]**

**Abstract**
We study the welfare effects of overreaction to information in the form of diagnostic expectations in markets with asymmetric information, and the effect of a simple intervention in the form of a tax or a subsidy. A large enough level of overreaction is always welfare-decreasing and can rationalize a tax on financial transactions. A small degree of overreaction to private information can both increase or decrease welfare. This is because there are two competing externalities: an information externality, due to the informational role of prices, and a pecuniary externality, due to the allocative role of prices. When the information externality prevails on the pecuniary externality, the loading on private information in agents' trades is too small compared to the welfare optimum: in this case, a small degree of overreaction is welfare-improving.

**JEL Classification:** D82, D83, D91, G14, H23.

**Keywords:** Overreaction, Diagnostic Expectations, Non-Bayesian learning, Taxes on Financial Transactions, Asymmetric Information, Externalities.

**Acknowledgments:** This paper was initially part of a different project titled "Learning, overreaction, and the wisdom of the crowds". We wish to thank for useful comments Nicola Gennaioli, Philip Matejka, Filippo Massari, Fabrizio Panebianco, Antonio Rosato, Alex Teytelboym, Fernando Vega-Redondo, Giovanni Immordino, and participants to the 2019 SMYE, the 2021 ASSET, the 2021 GRASS, and seminar participants in Bocconi and Napoli.

[*] University of Naples Federico II and CSEF. E-mail: matteo.bizzarri@unina.it

[†] Capital Fund Management and Nova Business School.

# Introduction

Information aggregation is understood to be one of the fundamental roles of markets, and financial markets in particular. As a consequence, a large literature has studied the welfare properties and the social value of information in markets, from Hayek (1945) to, e.g., Angeletos and Pavan (2007). In doing so, it is crucial to understand how agents make inferences from the information they receive: for example, traders in financial markets constantly update their beliefs about valuations of financial assets, as a consequence of changes in market prices, fundamentals, and investment choices of other traders. There is growing evidence that agents' updating rules depart from Bayesian rationality in the form of under/overreaction to information (Benjamin, 2019, Bordalo et al., 2020).[1] In this paper, we ask: how do departures of individual updating rules from Bayesian rationality impact welfare and informational efficiency in financial markets? Can a simple intervention such as a tax or a subsidy mitigate inefficiencies?

To formalize departures from Bayesian rationality in a parsimonious way, we follow the logic of the diagnostic expectations model (Bordalo et al., 2016 and Bordalo et al., 2018) that formalizes overreaction (and underreaction) in beliefs as a parametric deviation from Bayesian updating. Biased agents depart from Bayesian rationality in computing posterior beliefs by under/over-reacting to recent information. We embed over-reacting agents in a market game in which agents submit conditional bids, or schedules, that depend on the market price and a private signal. We adopt the tractable linear-quadratic Gaussian setting from Vives (2017).

In this environment, there are two sources of information: the private signal and the (public) market price. We adopt the *diagnostic expectations equilibrium* of Bordalo et al. (2020), in which prices are formed in equilibrium given agents' trade choices, and agents correctly understand this mechanism, but their posterior expectation about the fundamental value is distorted due to over/under reaction to both private information (the private signal) and public information (the market price). In particular, in our context, the bias does not come from (possibly partially) failing to realize that other traders also understand the information contained in prices, as in the "cursed equilibrium" of Eyster et al. (2019) or the "partial equilibrium thinking" of Bastianello and Fontanier (2022). The main difference between Bordalo et al. (2020) and our work is that their focus is on bubbles rather than welfare and taxes.[2]

[1]Moreover, such departure can explain several facts about macro-financial variables, such as credit cycles (Bordalo et al., 2018), stock returns (Bouchaud et al., 2019 and Bordalo et al., 2018), interest rates (d'Arienzo, 2020) and even the likelihood of a financial crisis (Maxted, 2023).

[2]Moreover, they use a model with CARA utility and inelastic supply, whereas, for tractability, we follow Vives

In a version of this model with standard Bayesian agents, Vives (2017) highlights two competing externalities: a learning externality, due to the fact that agents do not internalize that their actions reveal information by changing the informativeness of the price as a signal of the underlying value; and a pecuniary externality, due to the fact that agents do not internalize that, conditioning their trade on the price, they also change how the price reacts to the underlying value. As a consequence, the loading on private information can be either too high with respect to the efficient benchmark (if the pecuniary externality prevails) or too low (if the learning externality prevails). Both cases are possible, for different values of the parameters.

We characterize the equilibrium in a tractable linear-quadratic setting. When agents display overreaction, agents trade more aggressively for the same private signal, because they overweight the information contained in it. As a consequence, they increase the informativeness of the price as a public signal of the value. However, this increase is not sufficient to offset the first-order effect, and so the loading on the private signal in agents' actions is larger than it would be for Bayesian agents. So, overreaction changes the relative importance of the learning and the pecuniary externality with respect to the benchmark model. As a consequence, the price reveals more information than in an economy with Bayesian agents.

Having characterized the equilibrium, we study the effect of overreaction on welfare. The externality that prevails in the Bayesian benchmark determines the sign of the welfare effect a small level of overreaction, and so it can be positive or negative. However, for a large enough level of overreaction, a further increase in the diagnostic bias is always decreasing welfare. Then, we explore whether introducing a small quadratic tax or subsidy can be optimal. We show that when the overreaction parameter is large enough, the introduction of a small tax is always welfare-improving. Such result can offer a rationalization of a Tobin-type tax (Tobin, 1978) on financial transactions, for reasons related to the interaction of a behavioral bias (diagnostic expectations) and informational efficiency, that are distinct both from arguments relating to curbing speculation (as in Stiglitz, 1989 and Summers and Summers, 1989), and arguments arising from disagreement in agents' evaluations such as in Dávila (2023), and thus can be seen as complementary to such arguments.[3] When overreaction is close to zero, the welfare effect of a tax depends on the balance between the learning and pecuniary externality in the Bayesian benchmark. So the model implications for the optimality of

(2017) using a model with elastic supply and quadratic utility.

[3]Such a tax has been the subject of a long debate and is still a important issue in economic policy: it has been first advocated by Keynes, is currently in place in multiple countries (such as UK and Sweden), and is the object of a European Commission official proposal since 2011.

a tax crucially depend on the degree of agents' overreaction to information.

Our work is related to three literatures: the literature on overreaction and related biases in information processing, the literature studying taxes in the presence of behavioral biases, especially on financial transactions, and the literature on the social value of information. Our contribution is to show how overreaction can be welfare improving via mitigating the learning externality: that, is, overreaction can have a "social value". However, when overreaction is large enough, it can rationalize a tax on financial transactions, even in the presence of the learning externality. The literature on overreaction in finance and macroeconomics has mostly focused on identifying and measuring overreaction and on its explanatory power for rationalizing various macroeconomic phenomena (Bordalo et al., 2022). Some papers have explored macroeconomic policy under overreaction or exuberance, such as Maxted (2023), which also finds a positive welfare effect, that does not work through the learning externality but a balance sheet mechanism. Walther (2020) explores macro-prudential policy implications with extrapolative beliefs. The fact that overreaction helps learning via revealing more information is similar to the effect of overconfidence in the social learning model of Bernardo and Welch (2001): they study a simple sequential learning model instead of a financial market and so, in their setting, only the learning externality is present, but not the pecuniary externality.

While the taxation literature has studied various behavioral biases, for example, related to attention and salience as in Goldin (2015), Moore and Slemrod (2021), Farhi and Gabaix (2020), the literature specifically on taxation *of financial trasactions* has mostly focused on rational models: Auerbach and Bradford (2004), Rochet and Biais (2023), Adam et al. (2017), Buss et al. (2016), at most with heterogeneous priors as in Dávila (2023). The literature on the social value of information has also mainly focused on Bayesian agents, e.g. Angeletos and Pavan (2007) and Angeletos and Pavan (2009). An exception is Ostrizek and Sartori (2021) that study a strategic setting in which agents follow the cursed equilibrium model of Eyster and Rabin (2005) and Eyster et al. (2019), showing that cursedness can improve welfare: their mechanism works through information acquisition and not via the pecuniary externality like ours.

The next section introduces the model, Section 2 describes the equilibrium characterization, Section 3 describes our results and Section 4 concludes. All the proofs are in the Online Appendix.

# 1 The model

Our model closely follows Vives (2017), in its financial market interpretation, except for the behavioral bias due to diagnostic expectations.[4] We consider a financial market populated by informed speculators and liquidity suppliers. There is only one asset traded.

**Informed agents** There is a continuum of informed speculators indexed by $i \in [0, 1]$ and represented with the density $f$. Informed speculators face quadratic transaction costs. Each of them can decide her position $D_i$ with respect to the only asset exchanged, where short sales are allowed ($D_i$ can be negative).

The profit of an informed agent $i$ holding $D_i$ units of the asset when the market price is $p$ is:

$$u_i = (V - p)D_i - \frac{1}{2}\gamma D_i^2$$

where $V$ is the (unobservable) fundamental value of the asset, and the quadratic term represents transaction costs. Equivalently, it can be considered a form of (non constant) risk aversion.[5] Informed speculators have a prior over the fundamental value $V$ that is Gaussian: $V \sim \mathcal{N}(0, \tau_0^{-1})$. They also have access to a private signal $s_i$ that, conditional on $V$, follows a Gaussian distribution: $s_i \mid V \sim \mathcal{N}\left(V, \tau_\varepsilon^{-1}\right)$. Moreover, $s_i$ is independent of $s_j$ for $i \neq j$, conditionally on $V$: $s_i \perp s_j \mid V$.

In the following, we are going to have repeatedly to integrate a continuum of random variables over $[0, 1]$. We follow the literature[6] *defining* the integral over a continuum of independent random variables $(X_i)_{i\in[0,1]}$ as $\int X_i \mathrm{d}i := \int \mathbb{E}(X_i)\mathrm{d}i$ whenever the map $\mathbb{E}(X_i)$ is integrable (that will always be the case in our setting). This implies that a form of the Law of Large numbers holds, so that, conditionally on $V$, we have $\int s_i \mathrm{d}i = V$. This is going to be the only property of such an integral we need.[7] We denote the total demand from all informed agents as $\overline{D} = \int D_i \mathrm{d}i$.

**Diagnostic expectations** Agents update over their prior using the private signal $s_i$ and also the information contained in the price $p$ but, crucially, not in a Bayesian way. If the price depends on

[4] Vives (2017) studies different interpretations of the same abstract model, one being agents in a financial market, another firms competing in schedules. For our purposes we stick to the interpretation of agents trading in a financial market.

[5] The quadratic functional form makes the model very tractable. A similar approached is followed in Vives (2014).

[6] See Vives (2010).

[7] There are various formalizations of such an integral that deliver such a property, discussed e.g. in Acemoglu and Jensen (2010). Since the only property we are going to need is the Law of Large numbers, we avoid this technical issues and directly assume it.

the fundamental $V$ and the noise $S$ according to $p = A + BV - CS$, then $(p - A)/B$ is a Gaussian random variable of mean $V$ and precision $B^2/C^2\tau_S$: the agents understand this dependence and use it for their updating. So, after observing private signal $s_i$ and the price $p$, the Bayesian posterior distribution of the belief on the fundamental $V$ is a Normal with parameters:

$$\begin{aligned}\mathbb{E}(V \mid s_i, p) &= \frac{\tau_\varepsilon}{\tau_\varepsilon + \tau_0 + B^2/C^2\tau_S} s_i + \frac{B^2/C^2\tau_S}{\tau_\varepsilon + \tau_0 + B^2/C^2\tau_S} \frac{p - A}{B} \\ Var(V \mid s_i, p) &= \left(\tau_0 + \tau_\varepsilon + B^2/C^2\tau_S\right)^{-1}\end{aligned}$$

Our informed agents do *not* hold these beliefs because we assume that they over/under-react to information according to the diagnostic expectations model of Bordalo et al. (2018) and Bordalo et al. (2020). Namely, their posterior beliefs follow a Gaussian with the same variance, but expectation equal to:

$$\mathbb{E}^{\theta,i}(V \mid s_i, p) := \mathbb{E}(V \mid s_i, p) + \theta(\mathbb{E}(V \mid s_i, p) - \mathbb{E}(V))$$

where $\theta \in (-1, +\infty)$ represents the strength of the diagnostic bias. When $\theta > 0$ agents *over-react* to the information: when the information leads them to revise their prior expectation upwards ($\mathbb{E}(V \mid s_i, p) > \mathbb{E}(V)$), they revise it upwards more than a Bayesian would: $\mathbb{E}^{\theta,i}(V \mid s_i, p) > \mathbb{E}(V \mid s_i, p)$; while if the information leads to a downward revision ($\mathbb{E}(V \mid s_i, p) < \mathbb{E}(V)$), they revise it downwards more than a Bayesian would: $\mathbb{E}^{\theta,i}(V \mid s_i, p) < \mathbb{E}(V \mid s_i, p)$. The case of Bayesian agents corresponds to $\theta = 0$. For $\theta < 0$ we instead obtain *under*-reaction: agents revise their priors *less* than a Bayesian would; for $\theta \to -1$, agents to not revise their prior at all.

**Liquidity suppliers** Liquidity suppliers trade according to the aggregate (inverse) supply function $p = -\mu_S - S + \beta\overline{D}$, where $\beta > 0$ and $S \sim \mathcal{N}(0, \tau_S^{-1})$ is a random component, while $\mu_S$ is a constant.[8] In the welfare measure (1), we include the surplus of the liquidity suppliers, defined as is standard as the area below the supply curve: $\int_0^{\overline{D}} p(q)\mathrm{d}q$.[9] An alternative interpretation that does not rely on the concept of noise traders (and so might have a clearer welfare interpretation) is that there is an entrepreneur that can issue equity yielding a dividend $V$, with a preference for retaining

[8] Noise traders as in Grossman and Stiglitz (1980) are a special case of this specification in which $\beta \to \infty$, $\tau_S \to \infty$ and $\tau_S\beta^2 = \tau_S' > 0$. In this case, the aggregate supply is independent of prices, and simply a random variable with precision $\tau_S'$.

[9] If we were to exclude the liquidity traders from welfare calculations, there would still be a scope for intervention, as even in the Bayesian benchmark Vives (2017) shows that the learning and pecuniary externality would still be present, even if the precise expression would change.

shares (control) of the firm measured by $S$. This is explored formally in the Appendix A.

**Equilibrium** We follow (Bordalo et al. (2020)) in looking for a *diagnostic expectations equilibrium*. Namely, we look for a pricing function $g(S, V)$ that satisfies:

1. Individual optimization: the demand $D_i$ maximizes trader $i$ (diagnostic) expected utility given the observation of the private signal $s_i$ and the price $p$, formally: $D_i(p, s_i) \in \arg\max\{\mathbb{E}^{\theta,i}[u_i \mid s_i, g(S, V) = p]\}$;
2. the price clears the market: $p = -\mu_S - S + \beta \int D_i(g(S, V), s_i)\mathrm{d}i$.

Similar to Vives (2017), the functional forms assumptions guarantee that the equilibrium pricing function is linear: so determining the equilibrium reduces to finding the coefficients $A$, $B$ and $C$ such that $p = A + BV - CS$ satisfies the conditions above.

**The welfare measure** We follow Vives (2017) in considering our welfare measure the total surplus, defined as informed trader surplus plus the surplus of the liquidity suppliers:

$$W = \mathbb{E}\left(\left(\mu_S + S - \beta\frac{1}{2}\overline{D}\right)\overline{D} + \int\left(VD_i - \frac{\gamma}{2}D_i^2\right)\mathrm{d}i\right) \tag{1}$$

In the alternative interpretation of the asset supply as arising from an entrepreneur issuing equity, this expression represents the surplus of the informed traders plus the profit of the entrepreneur, that is also equivalent to the utilitarian welfare in this economy. Note that the expectations that appear in the expression are all taken from the perspective of Bayesian agents. In doing this, we interpret the agents' deviation from the Bayesian benchmark as a proper "mistake", not as a taste or preference feature, following a standard approach in the behavioral economics literature, e.g.: O'Donoghue and Rabin (2006), Spinnewijn (2015), and the survey by Mullainathan et al. (2012).[10]

In this context the first best allocation, that would realize if agents could pool their information, would be the complete information allocation, since by the law of large numbers $\int s_i \mathrm{d}i = V$. It is convenient to study welfare in terms of *welfare loss* from such an allocation. The first best allocation solves:

$$\max_{D_i} W = \left(\mu_S + S - \beta\frac{1}{2}\overline{D}\right)\overline{D} + \int\left(VD_i - \frac{\gamma}{2}D_i^2\right)\mathrm{d}i$$

[10] There is another, more conceptual reason. To compute the ex-ante welfare from the perspective of a diagnostic decision maker would require to specify how the decision maker predicts her future behavior *once she receives the information*: is she aware of her bias or not? this would require considerably more assumptions than simply compute the welfare from the perspective of a Bayesian agent, so we follow the latter approach.

and since the agents are ex-ante identical is a symmetric allocation, that we denote $D^o$. Denote $W^o$ the aggregate welfare in such an allocation. Define the welfare loss of some allocation $(D_i)_{i\in[0,1]}$ from the first best as $WL = W^o - W$, where $W$ is the welfare in the allocation $(D_i)_{i\in[0,1]}$. The following lemma from Vives (2017) characterizes the welfare loss from the first best:

**Lemma 1.1.** *At the allocation $(D_i)_{i\in[0,1]}$ the welfare loss from the first best allocation $D^o$ is*

$$WL = \mathbb{E}(W^* - W) = (\beta+\gamma)\frac{1}{2}\mathbb{E}(\overline{D} - D^o)^2 + \frac{\gamma}{2}\mathbb{E}\int (D_i - \overline{D})^2 \mathrm{d}i$$

The interpretation of the above expression is that the welfare loss results from two parts, that Angeletos and Pavan (2007) name, respectively, "variance" and "dispersion": the first relative to the departure of the aggregate demand from its first best level, the second relative to the cross-sectional dispersion of trades across agents. The effect of information (and thus overreaction to information) results from this trade-off: precise information means a small aggregate deviation from the first best, but a large dispersion, because precise information means traders trade more aggressively. The welfare impact of overreaction will result from this fundamental trade-off.

## 2 Equilibrium characterization

In this section we illustrate the equilibrium and the welfare benchmark.

Define the *loading on private information* $\alpha = a(\theta+1)$, where $a$ is the solution of:

$$\gamma a = \frac{\tau_\varepsilon}{\tau_\varepsilon + \tau(a)} \tag{2}$$

where $\tau(a) = \tau_0 + \beta^2 a^2 \tau_S (\theta+1)^2$ represents, in equilibrium, the precision of public information.

**Proposition 1.** *The diagnostic expectation equilibrium of the model is such that the demand of assets of each agent $i$ is:*

$$D_i = \frac{1}{\gamma}(\mathbb{E}^{\theta,i}(V \mid s_i, p) - p)$$

*and the equilibrium price satisfies:*

$$p = A + BV - CS$$

*where:*

$$\mathbb{E}^{\theta,i}(V \mid s_i, p) = (\theta+1)(\gamma a s_i + (1-\gamma a)\mathbb{E}(V \mid p))$$

$$\mathbb{E}(V \mid p) = \frac{\beta^2 a^2 \tau_S (\theta+1)^2}{\tau_0 + \beta^2 a^2 \tau_S (\theta+1)^2}(p - A)/B$$

*and the price coefficients are:*

$$A = \frac{-\gamma \mu_S}{\gamma + \beta} \quad B = \beta \frac{(\theta+1)}{\gamma+\beta} \quad C = \frac{1}{(\gamma+\beta)a}$$

We underline some key positive properties of the equilibrium:

**Corollary 2.1.** *In equilibrium the following properties hold:*

1. *The sensitivity to private information $\alpha$ is increasing in $\theta$;*

2. *The precision of the price as a signal of the value $B^2/C^2\tau_S$ is increasing in $\theta$;*

3. *The volatility of the price $Var(p)$ is increasing in $\theta$.*

Point 1 yields the fundamental mechanism of what follows: overreaction increases the sensitivity to private information. This is immediate by construction when fixing the precision of the public signal but, in equilibrium, overreaction also affects such precision, because more information is revealed. This indirect effect on the precision of the price, though, is not strong enough to counteract the main effect, and so the loading $\alpha$ increases in $\theta$.

Point 2 shows that since with overreaction the sensitivity to private information is higher, the price reacts more to the true value than it would in the Bayesian case, and so the precision of the price as a signal of the value is higher. Point 3 shows that this effect is in place *despite* the fact that the price displays excess volatility under overreaction, a well known fact. The apparent contradiction is resolved noting that the higher precision of the price as a signal of the value causes indeed a higher variability when measured ex-ante, because when the price incorporates more information it moves more with respect to the prior.

# 3 The effect of overreaction

In this section we study the effect of overreaction. First, as a benchmark, we illustrate the welfare analysis of the Bayesian model with $\theta = 0$.

## 3.1 The Bayesian benchmark

In the following Proposition we summarize the characterization of the Bayesian case from Vives (2017)

**Proposition 2.** *Define $a^*$ as the loading on the private signal at the market solution in the Bayesian benchmark: that is the solution of equation* (2) *for $\theta = 0$. Define $a^T$ as the solution of:*

$$a^T = \frac{\tau_\varepsilon}{\gamma(\tau(a^T) + \tau_\varepsilon) + \beta\tau(a^T) - \Delta(a^T)}$$

*where $\Delta(a^T) = \frac{(1-\gamma a^T)^2\beta^2\tau_S\tau_\varepsilon}{\gamma\tau(a^T)}$.*

*The market solution is second-best efficient if and only if $a^* = a^T$: $\frac{\mathrm{d}WL}{\mathrm{d}a} > 0 \iff a^* > a^T$.*

The Proposition says that the loading on private information at the market equilibrium $a^*$ can either be too high or too low from a welfare perspective. This is because of the interplay between a *learning externality* and a *pecuniary externality.* The learning externality derives from the informational role of the price and is well understood: agents decisions to trade reveal information to other agents through the price, but agents do not internalize this effect in the market equilibrium. This force pushes the sensitivity $a^*$ to be too low with respect to the second best. The pecuniary externality derives from the allocative role of the price, and derives from the fact that agents decisions affect how the price correlates to the true value $V$, but do not internalize this in the market equilibrium. This externality pushes the sensitivity $a^*$ to be *too large.* Summing up:

1. if $a^T < a^*$, this means that the learning externality is stronger;
2. if $a^T > a^*$, this means that the pecuniary externality is stronger.
3. if $a^* = a^T$ the two externalities exactly balance each other and the market equilibrium maximizes welfare.

In the following, we study how overreaction to information changes this picture.

## 3.2 Welfare decomposition

The endogenous loading on private information $\alpha$ is crucial for the efficiency properties of the equilibrium. In the following Lemma, we provide a decomposition of the welfare loss that is going to be useful in the following.

**Lemma 3.1.** *In equilibrium, we can decompose the welfare loss* (1) *as* $WL = WL^B + WL^D$*:*

$$WL^B = \frac{1}{2}\frac{(1-\gamma\alpha)^2}{(\beta+\gamma)}\frac{1}{\tau} + \frac{\gamma\alpha^2}{2\tau_\varepsilon} \tag{3}$$

$$WL^D(\alpha) = \frac{1}{2}\theta^2\frac{(1-\gamma\alpha/(\theta+1))^2}{(\beta+\gamma)}\left(\frac{1}{\tau_0}-\frac{1}{\tau}\right) \tag{4}$$

The first term $WL^B$ is the welfare loss that would realize for Bayesian agents having loading on private information equal to $\alpha$. The second term $WL^D$ represents the additional bias that diagnostic expectations add *beyond* the change in $\alpha$. This is useful to separate the direct effect of overreaction from the effect on the loading $\alpha$.

### 3.3 Welfare effect

The following proposition characterizes the effect of overreaction on welfare.

**Proposition 3.** *In* $\theta = 0$ $\frac{\mathrm{d}WL}{\mathrm{d}\theta} > 0$ *if and only if* $a^* > a^T$*.*

*Moreover, for* $\theta$ *large enough* $\frac{\mathrm{d}WL}{\mathrm{d}\theta} > 0$*, and for* $\theta$ *small enough* $\frac{\mathrm{d}WL}{\mathrm{d}\theta} < 0$*.*

The proposition shows that, when overreaction is close to zero, its welfare impact depend solely on the balance of externalities in the Bayesian case: in particular, if $a^* < a^T$, so that the learning externality prevails, overreaction is welfare improving. The key mechanism driving the result is that overreaction increases the sensitivity to private information $\alpha = a(\theta+1)$. This has the effect of making the price more sensitive to the true value, that has two implications: first, this makes the price a better signal of the value, mitigating the information externality; second, it exhacerbates the pecuniary externality. Lemma 3.1 shows that the term $WL^D$ is second order in $\theta$, hence when $\theta = 0$ the first order effect is only the variation in $WL^B$ caused by the change in $\alpha$:

$$\frac{\mathrm{d}WL}{\mathrm{d}\theta} = \frac{\partial WL^B}{\partial\alpha}\frac{\mathrm{d}\alpha}{\mathrm{d}\theta}$$

So, since $\frac{\mathrm{d}\alpha}{\mathrm{d}\theta} > 0$ by Corollary 2.1, the sign of the welfare impact depends simply on which externality is stronger at the Bayesian benchmark, that determines the sign of $\frac{\partial WL^B}{\partial\alpha}\mid_{\theta=0} = \frac{\partial WL^B}{\partial a}\mid_{\theta=0}$ as in Proposition 2.

When the overreaction parameter is far from 0, the term $WL^D$ instead becomes important. Such a term incorporates the expected mistake the agents make overestimating (underestimating) $V$ when they get positive (negative) information. The second part of the Proposition says that if the overreaction parameter $\theta$, and the consequent expected error, has magnitude large enough in positive or negative, moving further from the Bayesian benchmark can only reduce welfare. To sum up: a limited amount of overreaction can have a positive effect, depending on the interplay of prediction error, information externality and pecuniary externality.

### 3.4 Policy

We have seen that in this economy there are multiple inefficiencies due to the fact that agents might trade too much or too little relative to what would be the optimum given their private signals. These inefficiencies are present already in the Bayesian case: moreover, the diagnostic bias can exhacerbate (or not) these inefficiencies. Since the inefficiencies stem from the departures of the amounts traded from the second best, now we explore whether a tax (or subsidy) on quantities exchanged can be used to correct the inefficiencies and provide higher welfare. Vives (2017) shows that, in the Bayesian case, a quadratic tax/subsidy can implement the second-best level of the loading on private information $a^T$. In this section we ask a related question, that is: when does the introduction of a small tax improves welfare, and when a small subsidy instead?[11]

A linear tax/subsidy here cannot improve welfare: it would simply shift uniformly all the demands, but would leave the loading on private and public information unaffected: so it would simply add an additional term $t^2$ to the welfare loss, contributing to the volatility term: so the introduction of such a linear tax/subsidy would never be optimal. So, the next natural step is to explore a quadratic tax/subsidy $\delta$.

Formally, we assume that when agents trade a volume $D_i$, they have to pay an additional amount$\frac{1}{2}\delta D_i^2$, where if $\delta < 0$ this is understood to be a subsidy. Both buyers and sellers have to pay the tax. So, the payoff of the informed speculators becomes:

$$u_i = (V - p)D_i - \frac{1}{2}(\gamma + \delta)D_i^2$$

Since the tax is levied also on the liquidity suppliers, the inverse demand becomes: $p = -\mu_S - S + (\beta + \delta)\overline{D}$. In the Online Appendix we show that the results are qualitatively the same if the tax is levied on informed speculators only.

We follow the assumption in Vives (2017) that the revenues/payments from this tax/subsidy

[11] Note that this is different from asking whether a tax/subsidy can implement the second best level of reaction to private information. This is possible, but such tax/subsidy level is not necessarily optimal even in the Bayesian case, because beyond changing the *relative* weight between private and public information the tax/subsidy also affects the *total* volume traded.

are rebated in a lump-sum amount $T$, to satisfy budget balance. The total amount paid from the informed speculators is $\frac{\delta}{2}\int D_i^2 \mathrm{d}i$ and the one paid by the entrepreneur is $\frac{\delta}{2}\overline{D}^2$, and the total revenues collected must equal the rebate, so: $T = \frac{\delta}{2}\int D_i^2 \mathrm{d}i + \frac{\delta}{2}\overline{D}^2$.

So the welfare loss with respect to the first best is:

$$W^* - \left(W - \frac{\delta}{2}\overline{D}^2 - \frac{\delta}{2}\int D_i^2 \mathrm{d}i + T\right) = W^* - W$$

because the additional terms cancel out thanks to budget balance. So we conclude that the welfare loss satisfies the same expression as in Lemma 1.1.

The next proposition studies the welfare effect of the introduction of a *small tax*, formally characterized as the derivative of the welfare loss, computed in $\delta = 0$: $\frac{\mathrm{d}WL}{\mathrm{d}\delta}\mid_{\delta=0}$. When $\frac{\mathrm{d}WL}{\mathrm{d}\delta}\mid_{\delta=0} < 0$ a small positive tax decreases the welfare loss, and so we say that *a small tax is welfare improving.* When the opposite is true, we say that *a small tax is welfare improving.*

**Proposition 4.** *The welfare loss from the introduction of a tax $\delta$ is:*

$$\begin{aligned} WL^\delta = {} & \frac{(1-(\gamma+\delta)\alpha)}{(\beta+\gamma+2\delta)^2\tau}\left((1-(\gamma+\delta)\alpha) + \frac{4\delta}{\beta+\gamma}(1+(\beta+\delta)\alpha)\right) + \frac{4\delta^2(\mu_S^2+\tau_S^{-1}+\tau_0^{-1})}{(\beta+\gamma)^2(\beta+\gamma+2\delta)^2} \\ & + \left(\frac{1}{\beta+\gamma+2\delta}\right)^2\left(\theta^2 + \frac{4\delta\theta}{\beta+\gamma}\left(1-\frac{1}{(\beta+\delta)\alpha\tau_S}\right)\right)\left(\frac{1}{\tau_0}-\frac{1}{\tau}\right) + \frac{\gamma\alpha^2}{\tau_\varepsilon} \end{aligned}$$

*Moreover:*

1. *If $\theta$ is large enough, the introduction of a small tax is welfare improving: $\frac{\mathrm{d}WL}{\mathrm{d}\delta}\mid_{\delta=0} < 0$;*
2. *If $\theta$ is small enough, the introduction of a small subsidy is welfare improving: $\frac{\mathrm{d}WL}{\mathrm{d}\delta}\mid_{\delta=0} > 0$;*
3. *If $\theta = 0$, a tax could be either welfare improving or decreasing depending on the parameters. For $a^* = a^T$ a small tax is welfare decreasing. If $a^*$ is sufficiently smaller than $a^T$, a small subsidy is welfare improving.*

A tax $\delta$ decreases the total amount traded, and in so doing it also changes the loading on private information $\alpha$: an increase in $\delta$ decreases $\alpha$. The expression of the welfare loss above sums up these direct and indirect effects. When $\delta = 0$ the indirect effect $\frac{\partial WL}{\partial \alpha}$ is the same as without the tax. But, contrary to Proposition 3, the direct effect of the tax is first order here, so it is not sufficient to look at the balance of the learning and pecuniary externality alone to understand the sign. However, when $\theta$ becomes large enough we obtain $\frac{\mathrm{d}WL}{\mathrm{d}\delta}\mid_{\delta=0} > 0$. This is because when $\theta$ goes to infinity, then also $\alpha$ does. So, the amount traded is larger than at the efficient level, and a tax becomes welfare improving. When $\theta$ is small enough the reasoning is analogous, obtaining a subsidy instead of a tax.

# 4 Conclusion

We show that overreaction to information in the form of diagnostic expectations can improve welfare in markets where there is a strong enough information externality. When the information externality is not strong enough, overreaction can rationalize a tax on financial transactions on efficiency grounds. These results highlight that understanding the *degree* of overreaction is crucial for understanding its welfare effect and the sign of the optimal intervention. The interactions of these effects with other rationales for trading, such as hedging or heterogeneity, and other biases such as cursedness, are interesting avenues for further research.

# References


Acemoglu, D. and M. K. Jensen (2010). Robust comparative statics in large static games. In *49th IEEE Conference on Decision and Control (CDC)*, pp. 3133–3139. IEEE.

Adam, K., A. Marcet, and J. Beutel (2017). Stock price booms and expected capital gains. *American Economic Review 107*(8), 2352–2408.

Angeletos, G.-M. and A. Pavan (2007). Efficient use of information and social value of information. *Econometrica 75*(4), 1103–1142.

Angeletos, G.-M. and A. Pavan (2009). Policy with dispersed information. *Journal of the European Economic Association 7*(1), 11–60.

Auerbach, A. J. and D. F. Bradford (2004). Generalized cash-flow taxation. *Journal of Public Economics 88*(5), 957–980.

Bastianello, F. and P. Fontanier (2022). Expectations and learning from prices.

Benjamin, D. J. (2019). Errors in probabilistic reasoning and judgment biases. In *Handbook of Behavioral Economics: Applications and Foundations 1*, Volume 2, pp. 69–186. Elsevier.

Bernardo, A. E. and I. Welch (2001). On the evolution of overconfidence and entrepreneurs. *Journal of Economics & Management Strategy 10*(3), 301–330.

Bordalo, P., K. Coffman, N. Gennaioli, and A. Shleifer (2016). Stereotypes. *The Quarterly Journal of Economics 131*(4), 1753–1794.

Bordalo, P., N. Gennaioli, S. Y. Kwon, and A. Shleifer (2020). Diagnostic bubbles. *Journal of Financial Economics*.

Bordalo, P., N. Gennaioli, Y. Ma, and A. Shleifer (2020). Overreaction in macroeconomic expectations. *American Economic Review 110*(9), 2748–2782.

Bordalo, P., N. Gennaioli, and A. Shleifer (2018). Diagnostic expectations and credit cycles. *The Journal of Finance 73*(1), 199–227.

Bordalo, P., N. Gennaioli, and A. Shleifer (2022). Overreaction and diagnostic expectations in macroeconomics. *Journal of Economic Perspectives 36*(3), 223–244.

Bouchaud, J.-P., P. Krueger, A. Landier, and D. Thesmar (2019). Sticky expectations and the profitability anomaly. *The Journal of Finance 74*(2), 639–674.

Buss, A., B. Dumas, R. Uppal, and G. Vilkov (2016). The intended and unintended consequences of financial-market regulations: A general-equilibrium analysis. *Journal of Monetary Economics 81*, 25–43.

d'Arienzo, D. (2020). Maturity increasing overreaction and bond market puzzles. *Available at SSRN 3733056*.

Dávila, E. (2023). Optimal financial transaction taxes. *The Journal of Finance 78*(1), 5–61.

Eyster, E. and M. Rabin (2005). Cursed equilibrium. *Econometrica 73*(5), 1623–1672.

Eyster, E., M. Rabin, and D. Vayanos (2019). Financial markets where traders neglect the informational content of prices. *The Journal of Finance 74*(1), 371–399.

Farhi, E. and X. Gabaix (2020). Optimal taxation with behavioral agents. *American Economic Review 110*(1), 298–336.

Goldin, J. (2015). Optimal tax salience. *Journal of Public Economics 131*, 115–123.

Grossman, S. J. and J. E. Stiglitz (1980). On the impossibility of informationally efficient markets. *The American economic review 70*(3), 393–408.

Hayek, F. A. (1945). The use of knowledge in society. *The American Economic Review 35*, 519–530.

Maxted, P. (2023). A macro-finance model with sentiment. *Review of Economic Studies*, rdad023.

Moore, D. T. and J. Slemrod (2021). Optimal tax systems with endogenous behavioral biases. *Journal of Public Economics 197*, 104384.

Mullainathan, S., J. Schwartzstein, and W. J. Congdon (2012). A reduced-form approach to behavioral public finance. *Annu. Rev. Econ. 4*(1), 511–540.

O'Donoghue, T. and M. Rabin (2006). Optimal sin taxes. *Journal of Public Economics 90*(10-11), 1825–1849.

Ostrizek, F. and E. Sartori (2021). Acquisition,(mis) use and dissemination of information: The blessing of cursedness and transparency. *Available at SSRN 3954947*.

Rochet, J.-C. and B. Biais (2023). Taxing financial transactions: A mirrleesian approach.

Spinnewijn, J. (2015). Unemployed but optimistic: Optimal insurance design with biased beliefs. *Journal of the European Economic Association 13*(1), 130–167.

Stiglitz, J. E. (1989). Using tax policy to curb speculative short-term trading. *Regulatory Reform of Stock and Futures Markets: A Special Issue of the Journal of Financial Services Research*, 3–17.

Summers, L. H. and V. P. Summers (1989). When financial markets work too well: A cautious case for a securities transactions tax. *Journal of financial services research 3*, 261–286.

Tobin, J. (1978). A proposal for international monetary reform. *Eastern economic journal 4*(3/4), 153–159.

Vives, X. (2010). *Information and learning in markets: the impact of market microstructure*. Princeton University Press.

Vives, X. (2014). On the possibility of informationally efficient markets. *Journal of the European Economic Association 12*(5), 1200–1239.

Vives, X. (2017). Endogenous public information and welfare in market games. *The Review of Economic Studies 84*(2), 935–963.

Walther, A. (2020). Prudential policy in an exuberant world.

# Appendix

## A Alternative interpretation for the liquidity suppliers

In this section we illustrate an alternative interpretation for the origin of the elastic inverse demand, originating from a simple reduced form model of an entrepreneur issuing equity. There is an entrepreneur that has a project with dividend value $V$, that is not ex-ante known. The entrepreneur has preferences for remaining in control of the firm, measured by the random variable $\mu_S + S$, that represents the disutility per share sold for the entrepreneur. If she sells an amount $\overline{D}$ of equity, she can raise $p\overline{D}$, at the utility cost $(\mu_S + S)\overline{D}$, paying the transaction costs $\frac{\beta}{2}\overline{D}^2$. So, in total, the profit of the entrepreneur is:

$$u_i^e = (p + \mu_S + S)\overline{D} - \frac{\beta}{2}\overline{D}^2$$

that gives rise exactly to the inverse demand in the main text.

# Online Appendix

## B Tax only on informed speculators

In this Appendix we explore a variation in which it is possible to levy the tax only on informed speculators, and we show that the qualitative results are very similar. The results are collected in the following Proposition.

**Proposition 5.** *The welfare loss from the introduction of a tax $\delta$ is:*

$$WL = \frac{1}{2}\left(\frac{(1-(\gamma+\delta)\alpha)}{(\beta+\gamma+\delta)^2\tau}\left((1-(\gamma+\delta)\alpha)(\beta+\gamma)+2\delta(1+\beta\alpha)\right)+\frac{\delta^2(\mu_S^2+\tau_S^{-1}+\tau_0^{-1})}{(\beta+\gamma)(\beta+\gamma+\delta)^2}\right.$$
$$\left.+\left(\frac{1}{\beta+\gamma+\delta}\right)^2\left(\theta^2(\beta+\gamma)+2\delta\theta\left(1-\frac{1}{\beta\alpha\tau_S}\right)\right)\left(\frac{1}{\tau_0}-\frac{1}{\tau}\right)+\frac{\gamma\alpha^2}{\tau_\varepsilon}\right)$$

1. *If $\theta$ is large enough, the introduction of a small tax is welfare improving.*

2. *If $\theta$ is small enough, the introduction of a small subsidy is welfare improving.*

3. *if $\theta = 0$, a tax could be either welfare improving or decreasing depending on the parameters. For $a^* = a^T$, welfare is at a local optimum.*

*Proof.* All the equilibrium expressions are analogous to what derived in Proposition 1, with $\gamma+\delta$ in place of $\gamma$.

The level of trade for agent $i$ is:

$$D_i = (\theta+1)as_i + \frac{(1-(\gamma+\delta)a)(\theta+1)\mathbb{E}(V \mid p) - p}{\gamma+\delta}$$

where $a$ solves:

$$(\gamma+\delta)a = \frac{\tau_\varepsilon}{\tau_\varepsilon+\tau(a)}$$

From Lemma 1.1, we know that the expression for the welfare loss is:

$$\frac{1}{2}\left((\beta+\gamma)\mathbb{E}(D^o-\overline{D})^2+\gamma\mathbb{E}Var(D_i)\right)$$

The first best solution $D^o$ is of course not affected by the tax. We have to compute the two terms using the individual demands under a tax $\delta$. The dispersion term has the same form as a function of $a$ as would without the tax.

Instead, for the volatility term:

$$\overline{D} = \frac{1}{\beta+\gamma+\delta}\left(S+\mu_S+(\gamma+\delta)a(\theta+1)V+(1-(\gamma+\delta)a)(\theta+1)\mathbb{E}(V \mid p)\right)$$

$$\begin{aligned}
\overline{D}^o - \overline{D} &= \frac{\mu_S+S+V}{\beta+\gamma} - \frac{1}{\beta+\gamma+\delta}\left(S+\mu_S+(\gamma+\delta)a(\theta+1)V+(1-(\gamma+\delta))(\theta+1)\mathbb{E}(V \mid p)\right) \\
&= \frac{(\beta+\gamma)((1-(\gamma+\delta)a(\theta+1)V)+(1-(\gamma+\delta)a)(\theta+1)\mathbb{E}(V \mid p))+\delta(\mu_S+S+V)}{(\beta+\gamma)(\beta+\gamma+\delta)} \\
&= \frac{1}{(\beta+\gamma)(\beta+\gamma+\delta)}\left((\beta+\gamma)(1-(\gamma+\delta)\alpha)(V-\mathbb{E}(V \mid p))+\right. \\
&\left.+(\beta+\gamma)\theta\mathbb{E}(V \mid p)+\delta(\mu_S+V+S)\right)
\end{aligned}$$

Taking the square and the expectation we get:

$$\begin{aligned}
\mathbb{E}(D^o-\overline{D})^2 = &\frac{(1-(\gamma+\delta)\alpha)}{(\beta+\gamma+\delta)^2\tau}\left((1-(\gamma+\delta)\alpha)+\frac{2\delta}{\beta+\gamma}(1+\beta\alpha)\right)+\frac{\delta^2(\mu_S^2+\tau_S^{-1}+\tau_0^{-1})}{(\beta+\gamma)^2(\beta+\gamma+\delta)^2} \\
&+\left(\frac{1}{\beta+\gamma+\delta}\right)^2\left(\theta^2+\frac{2\delta\theta}{\beta+\gamma}\left(1-\frac{1}{\beta\alpha\tau_S}\right)\right)\left(\frac{1}{\tau_0}-\frac{1}{\tau}\right)
\end{aligned}$$

So the total welfare loss is:

$$\begin{aligned}
WL = \frac{1}{2}&\left(\frac{(1-(\gamma+\delta)\alpha)}{(\beta+\gamma+\delta)^2\tau}\left((1-(\gamma+\delta)\alpha)(\beta+\gamma)+2\delta(1+\beta\alpha)\right)+\frac{\delta^2(\mu_S^2+\tau_S^{-1}+\tau_0^{-1})}{(\beta+\gamma)(\beta+\gamma+\delta)^2}\right. \\
&\left.+\left(\frac{1}{\beta+\gamma+\delta}\right)^2\left(\theta^2(\beta+\gamma)+2\delta\theta\left(1-\frac{1}{\beta\alpha\tau_S}\right)\right)\left(\frac{1}{\tau_0}-\frac{1}{\tau}\right)+\frac{\gamma\alpha^2}{\tau_\varepsilon}\right)
\end{aligned}$$

Using the implicit function theorem, the effect of $\delta$ on the loading on private information is:

$$\frac{\mathrm{d}\alpha}{\mathrm{d}\delta} = -\frac{\alpha}{(\gamma+\delta)\left(\frac{2\alpha^2\beta^2(\theta+1)^2\tau_S}{\alpha^2\beta^2(\theta+1)^2\tau_S+\tau_0+\tau_\epsilon}+1\right)} < 0$$

Calculating the derivatives in $\delta = 0$ we get:

$$\frac{\partial WL}{\partial \alpha}\Big|_{\delta=0} = \frac{\alpha\gamma}{\tau_\epsilon} + \frac{\gamma\tau_0(\alpha\gamma-1)+\alpha\beta^2\tau_S\left(\alpha\gamma+\theta^2-1\right)}{(\beta+\gamma)\left(\alpha^2\beta^2\tau_S+\tau_0\right)^2} \tag{5}$$

$$\frac{\partial WL}{\partial \delta}\Big|_{\delta=0} = -\frac{\alpha\beta\theta\left(\alpha\beta(\theta-1)\tau_S+1\right)}{\tau_0(\beta+\gamma)^2\left(\alpha^2\beta^2\tau_S+\tau_0\right)} \tag{6}$$

Now consider part 1 of the result. In the limit $\theta \to \infty$ the total derivative $\frac{\mathrm{d}WL}{\mathrm{d}\delta} \mid_{\delta=0} = \frac{\partial WL}{\partial \delta} \mid_{\delta=0} + \frac{\partial WL}{\partial \alpha} \frac{\mathrm{d}\alpha}{\mathrm{d}\delta} \mid_{\delta=0}$ goes to an indeterminate form. We notice that both summands have a factor of $\alpha$, and analyze $\frac{1}{\alpha} \frac{\mathrm{d}WL}{\mathrm{d}\delta} \mid_{\delta=0} = \frac{1}{\alpha} \frac{\partial WL}{\partial \delta} \mid_{\delta=0} + \frac{\partial WL}{\partial \alpha} \frac{1}{\alpha} \frac{\mathrm{d}\alpha}{\mathrm{d}\delta} \mid_{\delta=0}$. Then, note that:

$$\lim_{\theta \to \infty} \frac{1}{\alpha} \frac{\mathrm{d}\alpha}{\mathrm{d}\delta} = -\frac{1}{3\gamma}$$

Moreover, $\frac{\partial WL}{\partial \alpha} \mid_{\delta=0}$ goes to $+\infty$, because as before the only term surviving is $\alpha/\tau_\varepsilon$. Finally, the higher order term in $\frac{1}{\alpha} \frac{\partial WL}{\partial \alpha} \mid_{\delta=0}$ is $-\alpha\theta^2$ and in the denominator is $\alpha^2$: $\theta^2/\alpha$ is asymptotically equivalent to $\theta/a$, that diverges negatively. Hence the total derivative is negative for $\theta$ large enough.

Consider part 2. The total derivative goes to zero as $\theta \to -1$ (and $\alpha \to 0$). We can observe that both $\frac{\partial WL}{\partial \delta} \mid_{\delta=0}$ and $\frac{\mathrm{d}\alpha}{\mathrm{d}\delta} \mid_{\delta=0}$ have a factor of $\alpha$. So, we collect $\alpha$, and calculating we get:

$$\lim_{\theta \to -1} \frac{1}{\alpha} \frac{\mathrm{d}WL}{\mathrm{d}\delta} \mid_{\delta=0} = \lim_{\theta \to -1} \frac{1}{\alpha} \left( \frac{\partial WL}{\partial \delta} \mid_{\delta=0} + \frac{\partial WL}{\partial \alpha} \frac{\mathrm{d}\alpha}{\mathrm{d}\delta} \mid_{\delta=0} \right) = \frac{\tau_0(\beta+\gamma)+\beta}{\tau_0^2(\beta+\gamma)^2} > 0$$

Now consider part 3. If $\theta = 0$ expression 6 shows that $\frac{\partial WL}{\partial \delta} \mid_{\theta=\delta=0} = 0$. If $a^* = a^T$, by definition, $\frac{\partial WL}{\partial \alpha} \mid_{\delta=0} = 0$, hence also the total derivative is null: $\frac{\mathrm{d}WL}{\mathrm{d}\delta} \mid_{\theta=\delta=0} = 0$.

□

# C Proofs

## C.1 Proof of Proposition 1

If $p = A + BV - CS$ the optimal choice for agents is:

$$D_i = \frac{1}{\gamma} \left( \frac{(\theta+1)\tau_\varepsilon}{\tau_\varepsilon + \tau_0 + \tau_{p|V}} V + \frac{(\theta+1)\tau_{p|V}}{\tau_\varepsilon + \tau_0 + \tau_{p|V}} (p - A)/B - p \right)$$

Now if $p = A + BV - CS$ then $\tau_{p|V} = B^2/C^2\tau_S$. Define $a = \frac{1}{\gamma} \frac{\tau_\varepsilon}{\tau_\varepsilon + \tau_0 + B^2/C^2\tau_S}$. Then, by the LLN, $\overline{D} = (\gamma a(\theta+1)V + (1-\gamma a)(\theta+1)\mathbb{E}(V \mid p) - p)/\gamma$. So the market clearing reads:

$$p = -\mu_S - S + \beta/\gamma(\gamma a(\theta+1)V + (1-\gamma a)(\theta+1)(p-A)/B - p)$$

$$(1 - \beta/\gamma((1-\gamma a)(\theta_p+1)/B - 1))p = -\mu_S - S + \beta/\gamma(\gamma a(\theta+1)V + (1-\gamma a)(\theta+1)(-A)/B)$$

$$p = \frac{-\gamma\mu_S - \gamma S + \beta(\gamma a(\theta+1)V - (1-\gamma a)(\theta_p+1)A/B)}{\gamma + \beta(1 - (1-\gamma a)(\theta+1)/B)}$$

So:

$$A = \frac{-\gamma\mu_S - \beta((1-\gamma a)(\theta+1)A/B)}{\gamma + \beta(1-(1-\gamma a)(\theta+1)/B)}$$

$$A + \frac{\beta((1-\gamma a)(\theta+1)A/B)}{\gamma + \beta(1-(1-\gamma a)(\theta+1)/B)} = \frac{-\gamma\mu_S}{\gamma + \beta(1-(1-\gamma a)(\theta+1)/B)}$$

$$A = \frac{-\gamma\mu_S}{\gamma+\beta}$$

$$B = \frac{\beta\gamma a(\theta+1)}{\gamma + \beta(1-(1-\gamma a)(\theta+1)/B)}$$

$$1 = \frac{\beta\gamma a(\theta+1)}{B(\gamma+\beta) - \beta(1-\gamma a)(\theta+1))}$$

$$B = \beta\frac{(\theta+1)}{\gamma+\beta}$$

$$C = \frac{\gamma}{\gamma + \beta(1-(1-\gamma a)(\theta+1)/B)} = \frac{\gamma}{\gamma + \beta\,(1-(1-\gamma a)(\beta+\gamma)/\beta)} = \frac{1}{(\gamma+\beta)a}$$

so that: $B^2/C^2 = \beta^2 a^2(\theta+1)^2$. So $a$ must satisfy:

$$\gamma a = \frac{\tau_\varepsilon}{\tau_\varepsilon + \tau_0 + B^2/C^2\tau_S} = \frac{\tau_\varepsilon}{\tau_\varepsilon + \tau_0 + \beta^2 a^2(\theta+1)^2\tau_S}$$

Define $\tau(a) = \beta^2 a^2(\theta+1)^2\tau_S$ the precision of public information. Since the RHS is monotone decreasing and the LHS is monotone increasing (from 0 to $\infty$), there is a unique positive solution.

So, finally:

$$p = \frac{-\gamma\mu_S + (\gamma a(\theta+1) + (1-\gamma a)(\theta+1))(\beta V - S/a)}{(\gamma+\beta)}$$

$$\overline{D} = (\gamma a(\theta+1)V + (1-\gamma a)(\theta+1)\mathbb{E}(V \mid p))/\gamma - (-S - \mu_S + \beta\overline{D})/\gamma$$

$$(1+\beta/\gamma)\overline{D} = (\gamma a(\theta+1)V + (1-\gamma a)(\theta+1)\mathbb{E}(V \mid p))/\gamma - (-S - \mu_S)/\gamma$$

$$\overline{D} = \frac{\gamma a(\theta+1)V + (1-\gamma a)(\theta+1)\mathbb{E}(V \mid p) + S + \mu_S}{\beta+\gamma}$$

□

## C.2 Proof of Corollary 2.1

1. The first point follows from the implicit function theorem. Indeed, we have:

$$\frac{\mathrm{d}a}{\mathrm{d}\theta} = -\frac{\frac{2a^2\beta^2(\theta+1)\tau_S\tau_\epsilon}{(a^2\beta^2(\theta+1)^2\tau_S+\tau_0+\tau_\epsilon)^2}}{\frac{2a\beta^2(\theta+1)^2\tau_S\tau_\epsilon}{(a^2\beta^2(\theta+1)^2\tau_S+\tau_0+\tau_\epsilon)^2}+\gamma} = -\frac{2\gamma a(1-\gamma a)\frac{\tau-\tau_0}{\tau(\theta+1)}}{2\gamma(1-\gamma a)\frac{\tau-\tau_0}{\tau}+\gamma} = -\frac{2a(1-\gamma a)\frac{\tau-\tau_0}{\tau(\theta+1)}}{2(1-\gamma a)\frac{\tau-\tau_0}{\tau}+1}$$

so $\frac{\mathrm{d}a}{\mathrm{d}\theta} < 0$. But:

$$\frac{\mathrm{d}\alpha}{\mathrm{d}\theta} = \frac{\mathrm{d}a}{\mathrm{d}\theta}(\theta+1) + a = -a\frac{2\gamma(1-\gamma a)\frac{\tau-\tau_0}{\tau}}{2\gamma(1-\gamma a)\frac{\tau-\tau_0}{\tau}+\gamma} + a > 0$$

2. from the proof of Proposition 1 we get that $B^2/C^2 = a^2(\theta+1)^2\beta^2\tau_S = \alpha^2\beta^2\tau_S$, hence it is increasing in $\theta$.

3. the volatility of the price is given by:

$$Var(p) = B^2 + C^2 = \frac{1}{(\gamma+\beta)^2}\left(\beta^2(\theta+1)^2\frac{1}{\tau_0} + \frac{1}{a^2\tau_S}\right)$$

that is increasing in $\theta$.

$\square$

## C.3 Proof of Lemma 3.1

The expression for the welfare loss is:

$$WL = W^* - W = (\beta+\gamma)\frac{1}{2}\mathbb{E}(\overline{D} - \overline{D}^*)^2 + \frac{\gamma}{2}\mathbb{E}Var(D_i)$$

$$\mathbb{E}(VarD_i) = \mathbb{E}\int(-a(\theta+1)s_i + a(\theta+1)V)^2 = \frac{a^2(\theta+1)^2}{\tau_\varepsilon}$$

Instead:

$$D^o - \overline{D} = \frac{V+\mu_S+S}{\beta+\gamma} - \frac{1}{\beta+\gamma}\left(\mu_S + S + \gamma a(\theta+1)V + (1-\gamma a)(\theta+1)\mathbb{E}(V \mid p)\right)$$

$$= \frac{(1-\gamma a)}{\beta+\gamma}(V - \mathbb{E}(V \mid p)) - \frac{\gamma a\theta V + (1-\gamma a)\theta\mathbb{E}(V \mid p))}{\beta+\gamma}$$

Now we want to compute the expectation of the square. This is equivalent to the variance since all the variables involved have zero expectation. We are going to use some facts:

$$\mathbb{E}(\mathbb{E}(V \mid p)^2) = \frac{(\tau-\tau_0)^2}{\tau^2}\mathbb{E}\left(V - \frac{C}{B}S\right)^2 = \frac{(\tau-\tau_0)^2}{\tau^2}\left(\frac{1}{\tau_0} + \frac{C^2}{B^2}\frac{1}{\tau_S}\right) = \frac{\tau-\tau_0}{\tau\tau_0}$$

Moreover, by the LIE also: $\mathbb{E}(\mathbb{E}(V \mid p)V) = \mathbb{E}(\mathbb{E}(\mathbb{E}(V \mid p)V \mid p)) = \mathbb{E}(\mathbb{E}(V \mid p)^2) = \frac{\tau-\tau_0}{\tau\tau_0}$. So we have:

$$\mathbb{E}(V-\mathbb{E}(V\mid p))^2=\mathbb{E}(\mathbb{E}((V-\mathbb{E}(V\mid p))^2\mid p))=\mathbb{E}(Var(V\mid p))=\mathbb{E}\left(\frac{1}{\tau}\right)=\frac{1}{\tau}$$

$$Cov((V-\mathbb{E}(V\mid p))V)=Var(V)-\mathbb{E}(\mathbb{E}(V\mid p)V)=\frac{1}{\tau_0}-\frac{\tau-\tau_0}{\tau\tau_0}=\frac{1}{\tau}$$

$$Cov((V-\mathbb{E}(V\mid p))\mathbb{E}(V\mid p))=\mathbb{E}(\mathbb{E}(V\mid p)V)-\mathbb{E}(\mathbb{E}(V\mid p)^2)=0$$

So:

$$\begin{aligned}\mathbb{E}(D^o-\overline{D})^2&=\frac{(1-\gamma a(\theta+1))^2}{(\beta+\gamma)^2}\frac{1}{\tau}+\theta^2\frac{(1-\gamma a)^2}{(\beta+\gamma)^2}\left(\frac{1}{\tau}-\frac{1}{\tau_0}\right)\\&=\frac{(1-\gamma\alpha)^2}{(\beta+\gamma)^2}\frac{1}{\tau}+\theta^2\frac{(1-\gamma\alpha/(\theta+1))^2}{(\beta+\gamma)^2}\left(\frac{1}{\tau_0}-\frac{1}{\tau}\right)\end{aligned}$$

So, the total welfare loss is:

$$WL=\frac{1}{2}\frac{(1-\gamma\alpha)^2}{(\beta+\gamma)}\frac{1}{\tau}+\theta^2\frac{(1-\gamma\alpha/(\theta+1))^2}{(\beta+\gamma)}\left(\frac{1}{\tau_0}-\frac{1}{\tau}\right)+\frac{\gamma\alpha^2}{2\tau_\varepsilon}$$

and can be decomposed as:

$$WL^B(\alpha)=\frac{1}{2}\frac{(1-\gamma\alpha)^2}{(\beta+\gamma)}\frac{1}{\tau}+\frac{\gamma\alpha^2}{2\tau_\varepsilon}$$

$$WL^D=\frac{1}{2}\theta^2\frac{(1-\gamma\alpha/(\theta+1))^2}{(\beta+\gamma)}\left(\frac{1}{\tau_0}-\frac{1}{\tau}\right)$$

□

## C.4 Proof of Proposition 3

From Lemma 3.1, we have that the welfare loss has two components:

$$WL=WL^B(\alpha)+WL^D$$

where $WL^B$ depends on $\theta$ only via $\alpha$, and $WL^D$ is second order in $\theta$. Hence, in $\theta=0$:

$$\frac{\mathrm{d}WL}{\mathrm{d}\theta}\mid_{\theta=0}=\frac{\partial WL^B}{\partial\alpha}\frac{\mathrm{d}\alpha}{\mathrm{d}\theta}\mid_{\theta=0}$$

Moreover, from Corollary 2.1 we know that $\alpha$ is increasing in $\theta$, so we conclude that, in $\theta=0$, $\frac{\mathrm{d}WL}{\mathrm{d}\theta}$ this has the same sign as $\frac{\partial WL^B}{\partial\alpha}$. Since this is the Bayesian welfare loss, this is positive if and only if $a^*>a^T$.

In general:

$$\frac{\partial WL^D}{\partial \theta} = \frac{\theta(1-\gamma\alpha/(\theta+1))^2 + \theta^2(1-\gamma\alpha/(\theta+1))\gamma\alpha/(\theta+1)^2}{\beta+\gamma}\left(\frac{1}{\tau_0}-\frac{1}{\tau}\right)$$

$$= \theta(1-\gamma\alpha/(\theta+1))\frac{(1-\gamma\alpha/(\theta+1)) + \theta\gamma\alpha/(\theta+1)^2}{\beta+\gamma}\left(\frac{1}{\tau_0}-\frac{1}{\tau}\right)$$

$$= \theta(1-\gamma a)\frac{1-\gamma a/(\theta+1)}{\beta+\gamma}\left(\frac{1}{\tau_0}-\frac{1}{\tau}\right)$$

$$\frac{\partial WL^D}{\partial \alpha} = -\theta^2\frac{\gamma(1-\gamma\alpha/(\theta+1))/(\theta+1)}{(\beta+\gamma)^2}\left(\frac{1}{\tau}-\frac{1}{\tau_0}\right) - \theta^2\frac{(1-\gamma\alpha/(\theta+1))^2}{(\beta+\gamma)^2}\frac{\alpha\beta^2\tau_S}{\tau^2} < 0$$

Now we compute limits. For $\theta \to -1$ we have that $a$ goes to its maximum, $\underline{a} = \frac{\tau_\varepsilon}{\tau_\varepsilon+\tau_0}$, and $\alpha \to 0$, as $\tau$. For $\theta \to \infty$ instead we have $a \to 0$ but $\alpha \to \infty$. Indeed, both $a$ and $\alpha$ are monotonic so they have a limit. Indeed, if $\lim_{\theta\to\infty} a = a' > 0$ (possibly infinite) we would have:

$$\lim_{\theta\to\infty} a = \lim_{\theta\to\infty} \frac{\tau_\varepsilon}{\gamma(\tau_\varepsilon + \tau_0(a')^2\beta^2(\theta+1)^2)} = 0$$

and if $\lim_{\theta\to\infty} \alpha = \alpha' < \infty$ (possibly zero), we would have:

$$\lim_{\theta\to\infty} \alpha = \lim_{\theta\to\infty} \frac{\tau_\varepsilon(\theta+1)}{\gamma(\tau_\varepsilon + \tau_0(\alpha')^2\beta^2)} = \infty$$

that would be contradictions.

From these, it follows that for $\theta \to -1$ $\frac{\mathrm{d}\alpha}{\mathrm{d}\theta}$ goes to the finite value $\underline{a} > 0$, while for $\theta \to \infty$ it goes to zero.

Now we can compute the limit of the partial derivatives. $\frac{\partial WL^D}{\partial \theta}$ goes to infinity for $\theta \to \infty$, while for $\theta \to -1$ is an indeterminate form, equivalent to:

$$\lim_{\theta\to-1} \frac{\partial WL^D}{\partial \theta} = \lim_{\theta\to-1} -(1-\gamma a/(\theta+1))\frac{\beta^2 a^2(\theta+1)^2}{\tau\tau_0} = \lim_{\theta\to-1} \gamma\underline{a}\frac{\beta^2\underline{a}^2(\theta+1)}{\tau\tau_0} = 0$$

Instead, $\frac{\partial WL^D}{\partial \alpha}$ goes to a finite negative value for $\theta \to -1$. Hence $\frac{\mathrm{d}WL^D}{\mathrm{d}\theta} < 0$. For $\theta \to -1$ we also know that $\frac{\mathrm{d}WL^B}{\mathrm{d}\alpha} < 0$, hence we conclude that the welfare loss is decreasing: $\frac{\mathrm{d}WL}{\mathrm{d}\theta} < 0$. So $\theta = -1$ cannot be the optimal value of $\theta$.

Now for $\theta \to \infty$ the welfare loss diverges: hence the optimal value of $\theta$ has to be finite. (take any finite value $t = WL(\theta')$, there is a $\theta''$ such that $WL > t$ for all $\theta > \theta''$ and so the optimum is smaller than $\theta''$). Hence, for $\theta$ large enough, $\frac{\mathrm{d}WL}{\mathrm{d}\theta} > 0$.

$\square$

## C.5 Proof of Proposition 4

All the equilibrium expressions are analogous to what derived in Proposition 1, with $\gamma+\delta$ in place of $\gamma$ and $\beta+\delta$ in place of $\beta$.

The level of trade for agent $i$ is:

$$D_i = (\theta+1)as_i + \frac{(1-(\gamma+\delta)a)(\theta+1)\mathbb{E}(V \mid p) - p}{\gamma+\delta}$$

where $a$ solves:

$$(\gamma+\delta)a = \frac{\tau_\varepsilon}{\tau_\varepsilon + \tau(a)}$$

and $\tau(a) = a^2(\theta+1)^2(\beta+\delta)^2\tau_S$, and:

$$\overline{D} = \frac{1}{\beta+\gamma+2\delta}\left(S + \mu_S + (\gamma+\delta)a(\theta+1)V + (1-(\gamma+\delta)a)(\theta+1)\mathbb{E}(V \mid p)\right)$$

From Lemma 1.1, we know that the expression for the welfare loss is:

$$\frac{1}{2}\left((\beta+\gamma)\mathbb{E}(D^o - \overline{D})^2 + \gamma\mathbb{E}Var(D_i)\right)$$

this is not affected, because the lump-sum rebate means that the tax terms cancel out.

The first best solution $D^o$ is of course not affected by the tax. We have to compute the two terms using the individual demands under a tax $\delta$. The dispersion term has the same form as a function of $a$ as would without the tax:

$$\begin{aligned}\mathbb{E}Var(D_i) &= \mathbb{E}\int \alpha^2(s_i - V)^2\mathrm{d}i = \alpha^2\int \mathbb{E}(s_i - V)^2\mathrm{d}i \\ &= \alpha^2\int \mathbb{E}(\mathbb{E}((s_i - V)^2 \mid V))\mathrm{d}i = \frac{\alpha^2}{\tau_\varepsilon}\end{aligned}$$

Instead, for the volatility term:

$$
\begin{aligned}
\overline{D}^o - \overline{D} &= \frac{\mu_S + S + V}{\beta + \gamma} - \frac{1}{\beta + \gamma + 2\delta}\left(S + \mu_S + (\gamma + \delta)a(\theta + 1)V + (1 - (\gamma + \delta))(\theta + 1)\mathbb{E}(V \mid p)\right) \\
&= \frac{(\beta + \gamma)((1 - (\gamma + \delta)a(\theta + 1)V) + (1 - (\gamma + \delta)a)(\theta + 1)\mathbb{E}(V \mid p)) + 2\delta(\mu_S + S + V)}{(\beta + \gamma)(\beta + \gamma + 2\delta)} \\
&= \frac{1}{(\beta + \gamma)(\beta + \gamma + 2\delta)}\left((\beta + \gamma)(1 - (\gamma + \delta)\alpha)(V - \mathbb{E}(V \mid p)) + \right. \\
&\left. + (\beta + \gamma)\theta\mathbb{E}(V \mid p) + 2\delta(\mu_S + V + S)\right)
\end{aligned}
$$

Taking the square and the expectation we get:

$$
\begin{aligned}
\mathbb{E}(D^o - \overline{D})^2 = &\frac{(1 - (\gamma + \delta)\alpha)}{(\beta + \gamma + 2\delta)^2\,\tau}\left((1 - (\gamma + \delta)\alpha) + \frac{4\delta}{\beta + \gamma}(1 + (\beta + \delta)\alpha)\right) + \frac{4\delta^2(\mu_S^2 + \tau_S^{-1} + \tau_0^{-1})}{(\beta + \gamma)^2(\beta + \gamma + 2\delta)^2} \\
&+ \left(\frac{1}{\beta + \gamma + 2\delta}\right)^2\left(\theta^2 + \frac{4\delta\theta}{\beta + \gamma}\left(1 - \frac{1}{(\beta + \delta)\alpha\tau_S}\right)\right)\left(\frac{1}{\tau_0} - \frac{1}{\tau}\right)
\end{aligned}
$$

So the total welfare loss is:

$$
\begin{aligned}
WL = &\frac{(1 - (\gamma + \delta)\alpha)}{(\beta + \gamma + 2\delta)^2\,\tau}\left((1 - (\gamma + \delta)\alpha) + \frac{4\delta}{\beta + \gamma}(1 + (\beta + \delta)\alpha)\right) + \frac{4\delta^2(\mu_S^2 + \tau_S^{-1} + \tau_0^{-1})}{(\beta + \gamma)^2(\beta + \gamma + 2\delta)^2} \\
&+ \left(\frac{1}{\beta + \gamma + 2\delta}\right)^2\left(\theta^2 + \frac{4\delta\theta}{\beta + \gamma}\left(1 - \frac{1}{(\beta + \delta)\alpha\tau_S}\right)\right)\left(\frac{1}{\tau_0} - \frac{1}{\tau}\right) + \frac{\gamma\alpha^2}{\tau_\varepsilon}
\end{aligned}
$$

Using the implicit function theorem, the effect of $\delta$ on the loading on private information is:

$$
\frac{\mathrm{d}\alpha}{\mathrm{d}\delta} = -\frac{\alpha\left(\alpha^2(\beta + \delta)(\beta + \delta + 2)\tau_S + \tau_0 + \tau_\epsilon\right)}{(\gamma + \delta)\left(3\alpha^2(\beta + \delta)^2\tau_S + \tau_0 + \tau_\epsilon\right)} < 0
$$

Calculating the derivatives in $\delta = 0$ we get:

$$
\begin{aligned}
\frac{\partial WL}{\partial \alpha} &= \frac{\alpha\gamma}{\tau_\epsilon} + \frac{\left(\gamma\tau_0(\alpha\gamma - 1) + \alpha\beta^2\tau_S\left(\alpha\gamma + \theta^2 - 1\right)\right)}{(\beta + \gamma)\left(\alpha^2\beta^2\tau_S + \tau_0\right)^2}, \\
\frac{\partial WL}{\partial \delta} &= -\frac{\alpha\left(\tau_0(\alpha\gamma - 1)(\beta + \gamma) + 2\beta\theta\left(\alpha\beta(\theta - 1)\tau_S + 1\right)\right)}{\tau_0(\beta + \gamma)^2\left(\alpha^2\beta^2\tau_S + \tau_0\right)}
\end{aligned}
\tag{7}
$$

Now consider part 1 of the result. In the limit $\theta \to \infty$ the total derivative goes to an indeterminate form. Then, note that:

$$
\lim_{\theta \to \infty} \frac{1}{\alpha}\frac{\mathrm{d}\alpha}{\mathrm{d}\delta} = -\frac{2 + \beta}{3\gamma\beta}
$$

Moreover, $\frac{\partial WL}{\partial \alpha}$ goes to $+\infty$, because as before the only term surviving is $\alpha/\tau_\varepsilon$, and so the first part diverges negatively since $\frac{\mathrm{d}\alpha}{\mathrm{d}\delta} < 0$. Finally, the higher order term in $\frac{\partial WL}{\partial \alpha}$ is $-\alpha\theta^2$ and in the denominator is $\alpha^2$: $\theta^2/\alpha$ is asymptotically equivalent to $\theta/a$, that diverges. Hence the total derivative is negative for $\theta$ large enough.

Consider part 2. The total derivative goes to zero as $\theta \to -1$ (and $\alpha \to 0$). We can observe that both $\frac{\partial WL}{\partial \delta}$ and $\frac{\mathrm{d}\alpha}{\mathrm{d}\delta}$ have a factor of $\alpha$. So, we collect $\alpha$, and calculating we get:

$$\lim_{\theta \to -1} \frac{1}{\alpha} \frac{\mathrm{d}WL}{\mathrm{d}\delta} = \lim_{\theta \to -1} \frac{1}{\alpha} \left( \frac{\partial WL}{\partial \delta} + \frac{\partial WL}{\partial \alpha} \frac{\mathrm{d}\alpha}{\mathrm{d}\delta} \right) = \frac{2\left(\tau_0(\beta+\gamma)+\beta\right)}{\tau_0^2(\beta+\gamma)^2} > 0$$

Now consider part 3. If $\theta = 0$ expression 6 shows that $\frac{\partial WL}{\partial \delta} = -\frac{\alpha(\tau_0(1-a^*\gamma))}{\tau_0(\beta+\gamma)((a^*)^2\beta^2\tau_S+\tau_0)} > 0$. If $a^* = a^T$, by definition, $\frac{\partial WL}{\partial \alpha} = 0$, and $\frac{\mathrm{d}\alpha}{\mathrm{d}\delta}$ remains finite. Hence also the total derivative is positive: $\frac{\mathrm{d}WL}{\mathrm{d}\delta} > 0$.

□