Information Revelation and Consumer Privacy

Rossella Argenziano[1] Alessandro Bonatti[2] Gonzalo Cisternas[2]

[1]University of Essex

[2]MIT Sloan

FTC Microeconomics Conference

November 13th, 2015

Introduction

Wide collection and diffusion of personal data in online markets.

- Sources: recorded purchases, browser cookies, social media.
- Uses: customized search results, web content, targeted advertising, promotional offers.

Consumers have partial control over available information.

Introduction

Wide collection and diffusion of personal data in online markets.

- Sources: recorded purchases, browser cookies, social media.
- Uses: customized search results, web content, targeted advertising, promotional offers.

Consumers have partial control over available information.

This paper: equilibrium analysis with rational consumer.

- Consumer has no intrinsic value of privacy.
- Understands information collection (profile building) mechanism and its payoff consequences.

BIG DATA AND DIFFERENTIAL PRICING

February 2015

Research Questions

1. What determines the *equilibrium* amount of available information about consumers?
2. How valuable is this information to sellers?
3. What are the implications of information collection (vs. privacy regulation) for consumers' welfare?
4. How do these answers depend on the *source* of the information and on its *intended use*?

Just the first step...

Model: Approach

Noisy *ratchet-effect* model with details from online markets.

Consumers' preferences are private information.

Each consumer interacts with heterogeneous firms: **sellers** and **websites** (non-merchant content providers).

Each interaction generates information of **endogenous precision**.

Firms use information (from either source) in future interactions.

- Aligned interests: matching content to tastes.
- Conflict: matching products and prices to willingness to pay.

Model: Actions

Consumer meets two firms in each period $t = 1, 2$.

Think of them as distinct firms.

Seller offers a single product at a unit **price** p_t.

Consumer chooses **quantity** q_t (intensity of interaction).

Website offers tailored **content** w_t (e.g., news stories).

Consumer chooses which **page** z_t to access (e.g., read).

Model: Payoffs

Consumer has type (θ_q, θ_z) =(taste for product, taste for news).

Correlated "vertical" and "horizontal" and components.

Today: perfectly correlated $\theta_q = \theta_z = \theta$.

Model: Payoffs

Consumer has type (θ_q, θ_z) =(taste for product, taste for news).

Correlated "vertical" and "horizontal" and components.

Today: perfectly correlated $\theta_q = \theta_z = \theta$.

Linear-quadratic *flow* utility function

$$U(\theta, q, z; p, w) = (\theta - p) q - q^2/2 - (\theta - z)^2 - (w - z)^2 .$$

Sellers maximize profits

$$\Pi(p, q) = p \cdot q.$$

Websites want to match content to consumer's type

$$L(\theta, w) = -(w - \theta)^2 .$$

Model: Information

Prior distribution of consumer's type

$$\theta \sim N\left(\theta_0, 1/\tau_0\right).$$

Consumer's actions at $t = 1$ **recorded** with noise.

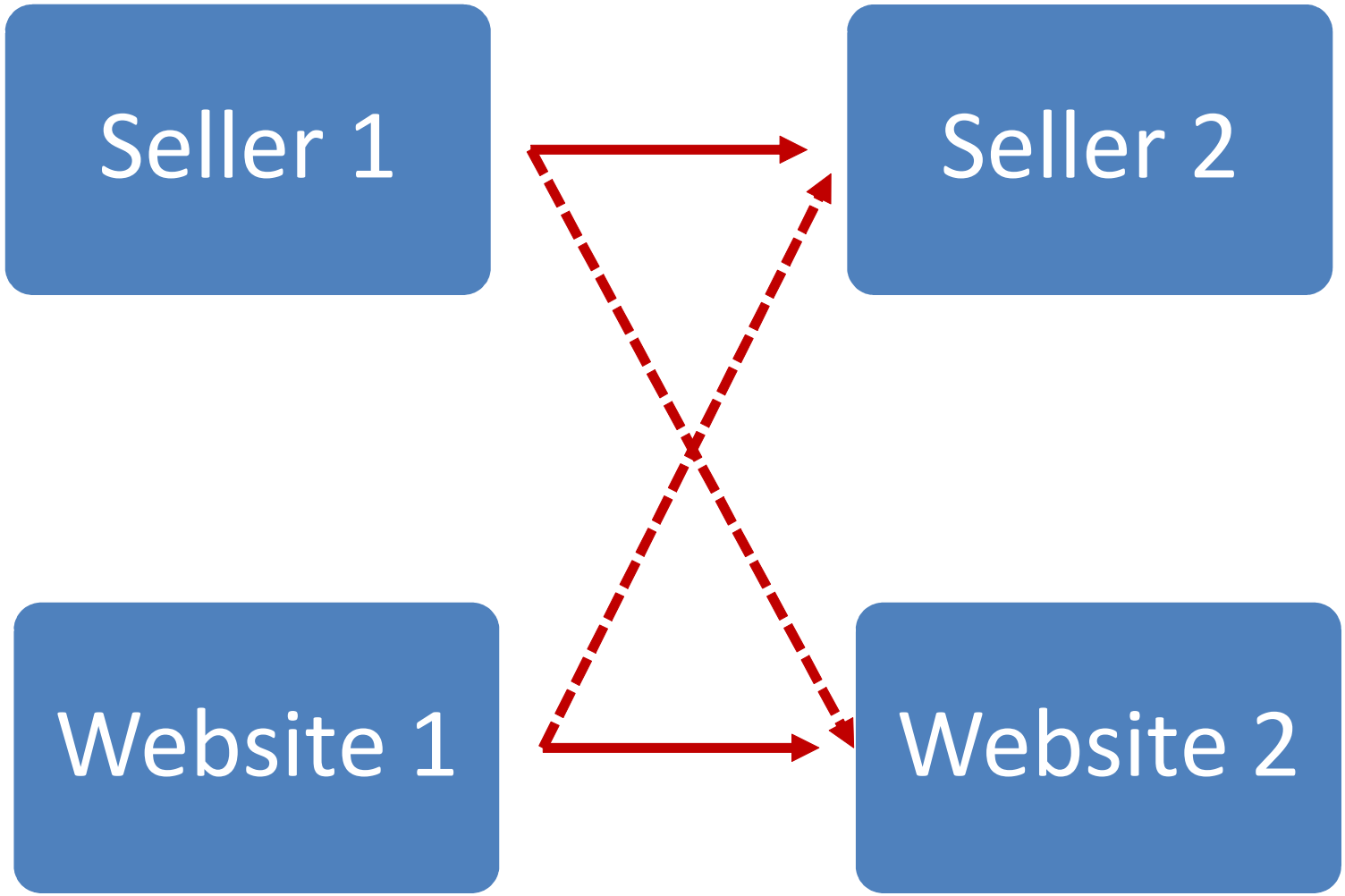
1. Browsing history
$$s_z \sim N\left(z, 1/\tau_z\right).$$
2. Purchase history
$$s_q \sim N\left(q, 1/\tau_q\right).$$

Firms at $t = 2$ observe (part of) the consumer's record.

Information set of firm $j \in \{W, S\}$

$$\mathcal{I}_j \subseteq \{s_q, s_z\}.$$

Signals



Model: Discussion

What is the source of conflict between consumers and firms?

Model: Discussion

What is the source of conflict between consumers and firms?

Here: firms customize prices based on purchase histories.

Richer model: firms sell multiple goods; searching is costly.

Customized search results steer high-value consumers to high-markup products.

Model: Discussion

What is the source of conflict between consumers and firms?

Here: firms customize prices based on purchase histories.

Richer model: firms sell multiple goods; searching is costly.

Customized search results steer high-value consumers to high-markup products.

Here: separate roles (players) for conflict and alignment.

Reality: additional product-quality dimension.

Consumer's Problem

Given firms' beliefs $m_j \triangleq \mathbb{E}\left[\theta \mid \mathcal{I}_j\right]$, continuation payoff:

$$V_2\left(\theta, m_S, m_W\right) = \frac{1}{2}\left(\theta - \frac{m_S}{2}\right)^2 - \frac{1}{2}\left(\theta - m_W\right)^2.$$

- Wants seller to under-estimate θ and website to learn θ.
- Ex-ante: $\mathrm{cov}(\theta, m_S)$ hurts, $\mathrm{cov}(\theta, m_W)$ benefits consumers.

Consumer's Problem

Given firms' beliefs $m_j \triangleq \mathbb{E}\left[\theta \mid \mathcal{I}_j\right]$, continuation payoff:

$$V_2\left(\theta, m_S, m_W\right) = \frac{1}{2}\left(\theta - \frac{m_S}{2}\right)^2 - \frac{1}{2}\left(\theta - m_W\right)^2.$$

- Wants seller to under-estimate θ and website to learn θ.
- Ex-ante: $\mathrm{cov}(\theta, m_S)$ hurts, $\mathrm{cov}(\theta, m_W)$ benefits consumers.

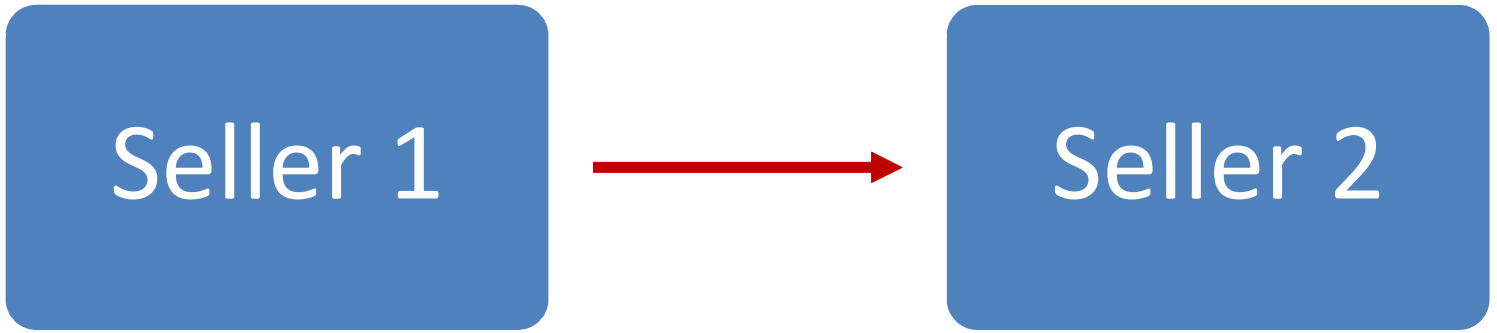
Period 1: consumer chooses (q_1, z_1) to maximize

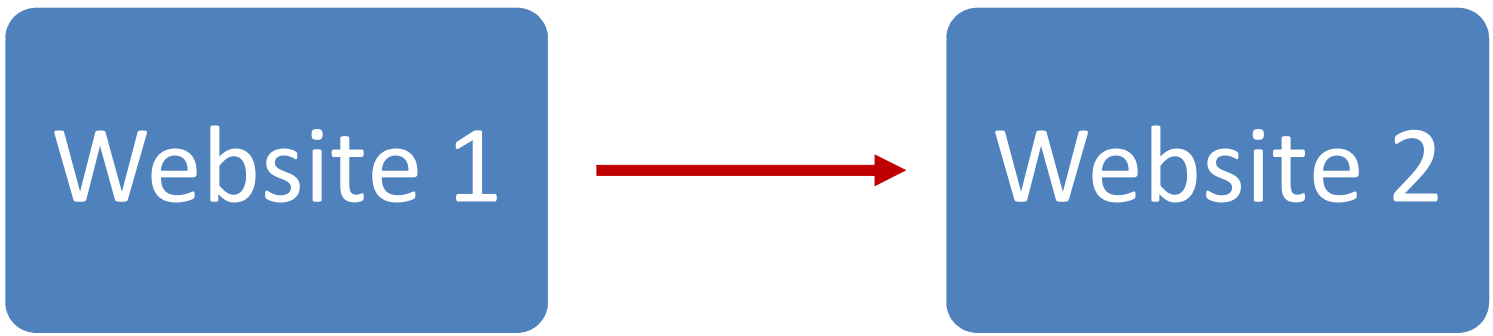
$$V_1\left(\theta, q, z\right) = U\left(\theta, q, z; p, w\right) + \mathbb{E}\left[V\left(\theta, m_S, m_W\right) \mid q, z\right].$$

Trades off flow utility vs. manipulating firms' beliefs.

Value of manipulating depends on type θ and information sets $\mathcal{I}_j$.

Private Signals





Firms' Beliefs

To form beliefs, firms conjecture linear strategies

$$\begin{aligned} q_1 &= \alpha_q \cdot \theta + \beta_q (p, w), \\ z_1 &= \alpha_z \cdot \theta + \beta_z (p, w). \end{aligned}$$

Linearity + normality $\Rightarrow$ consumer influences $\mathbb{E}[m_S], \mathbb{E}[m_W]$ only.

Firms' Beliefs

To form beliefs, firms conjecture linear strategies

$$\begin{aligned} q_1 &= \alpha_q \cdot \theta + \beta_q (p, w), \\ z_1 &= \alpha_z \cdot \theta + \beta_z (p, w). \end{aligned}$$

Linearity + normality $\Rightarrow$ consumer influences $\mathbb{E}[m_S], \mathbb{E}[m_W]$ only.

Consumer's first-period best replies (q_1, z_1) satisfy

$$\begin{aligned} q &= \theta - p - \frac{1}{2}\left(\theta - \frac{\mathbb{E}[m_S \mid q]}{2}\right)\frac{\partial \mathbb{E}[m_S \mid q]}{\partial q}, \\ z &= \frac{\theta + w}{2} + (\theta - \mathbb{E}[m_W \mid z])\frac{\partial \mathbb{E}[m_W \mid z]}{\partial z}. \end{aligned}$$

- Linear conjectures, linear replies, match coefficients.

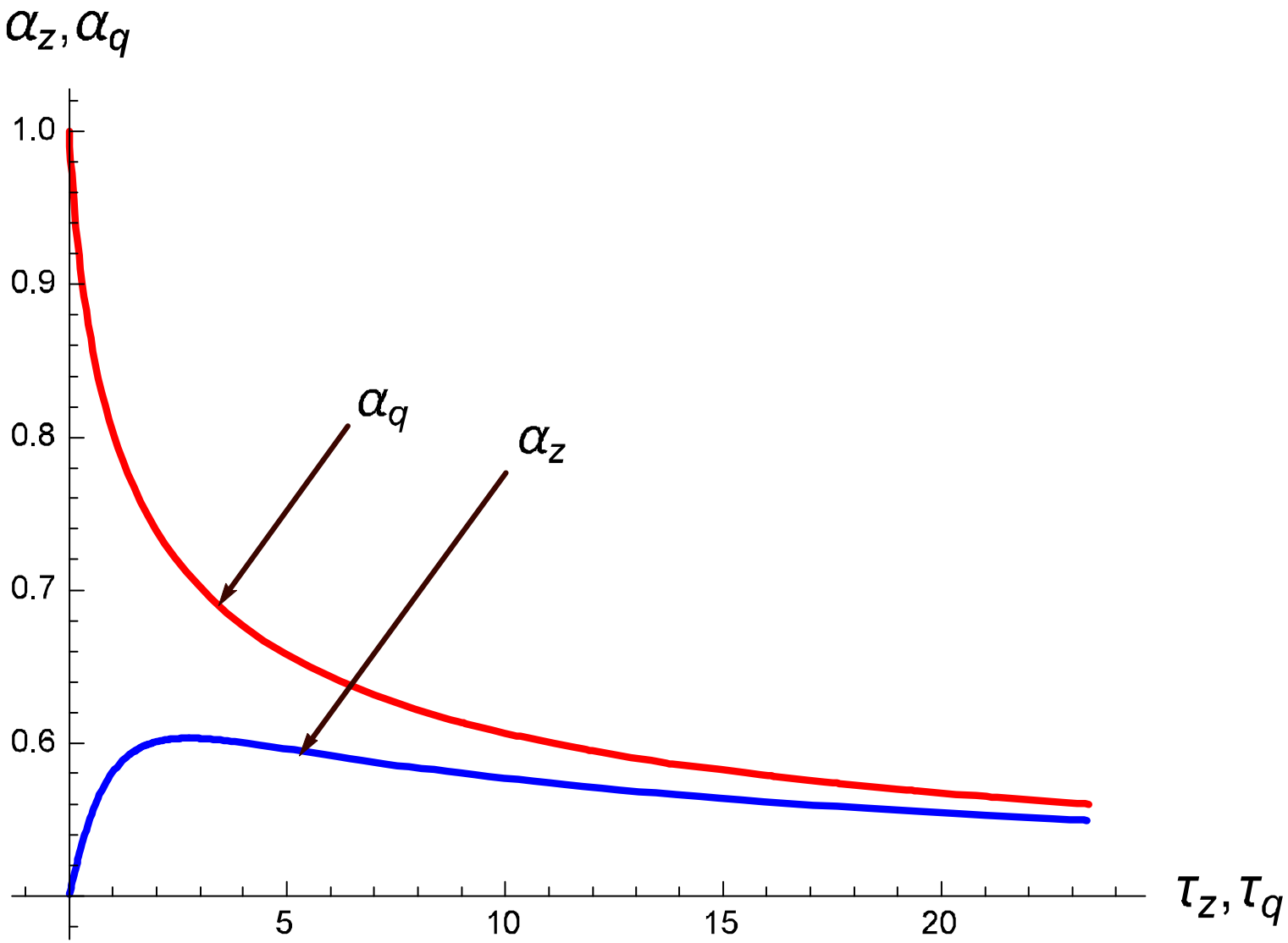
Proposition (Equilibrium with Private Signals)

There exists a unique equilibrium in linear strategies:

$$\begin{aligned} q_1^*(\theta, p) &= \alpha_q^*(\tau_q) \cdot \theta + \delta_q^*(\tau_q) \cdot \theta_0 - p, \\ z_1^*(\theta, w) &= \alpha_z^*(\tau_z) \cdot \theta + \delta_z^*(\tau_z) \cdot \theta_0 + w/2. \end{aligned}$$

1. *Signal precisions $\alpha_q^2 \tau_q$ and $\alpha_z^2 \tau_z$ are increasing in τ_q and τ_z.*
2. *The browsing weight $\alpha_z^* \geq 1/2$ is inverse-U shaped in τ_z.*
3. *The purchase weight $\alpha_q^* \leq 1$ is strictly decreasing in τ_q.*
4. *First-period price $p_1^* = \theta_0 \left(\alpha_q^* + \delta_q^*\right)/2$, decreasing in τ_q.*

Equilibrium Coefficients



Welfare with Private Signals

Consider *ex ante* welfare of consumers and *sellers*.

Decompose surplus into mean (θ_0) and variance (τ_θ) effects.

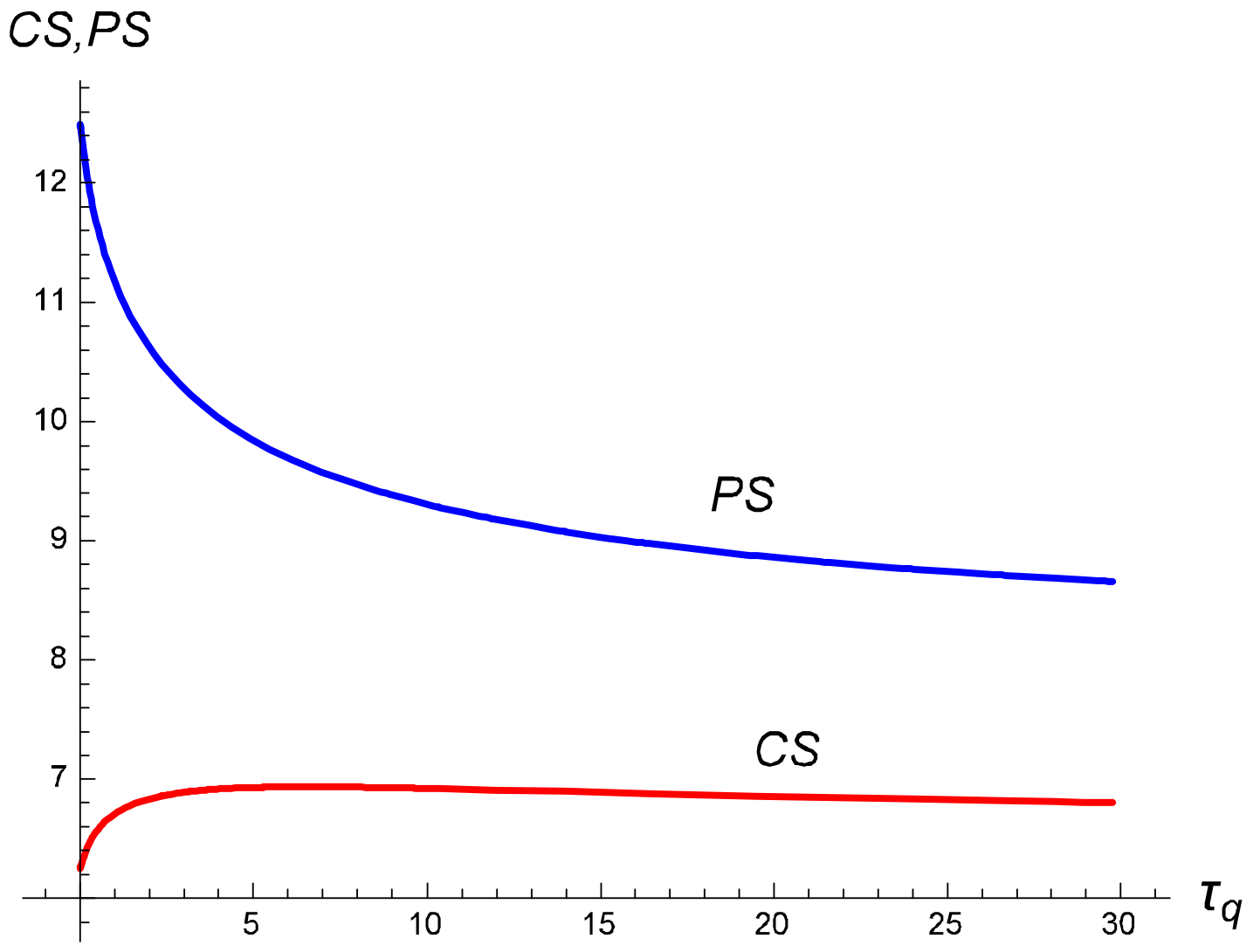
Welfare with Private Signals

Consider *ex ante* welfare of consumers and *sellers*.

Decompose surplus into mean (θ_0) and variance (τ_θ) effects.

Proposition (Welfare with Private Signals)

1. *Consumer surplus is increasing in* τ_z.
2. *Consumer surplus is inverse-U shaped in* τ_q *(strictly decreasing if* $\theta_0^2 \tau_\theta < 6$*).*
3. *Producer surplus is inverse-U shaped in* τ_q *(strictly decreasing if* $\theta_0^2 \tau_\theta > 2$*).*
4. *Total surplus is strictly decreasing in* τ_q.



Parameters: $(\theta_0, \tau_\theta, \tau_z) = (5, 1, 0)$

Intuition

Seller 1 anticipates consumer's concern over second-period price.

Expects lower demand, charges lower price.

Consumer buys fewer units at a lower price.

Intuition

Seller 1 anticipates consumer's concern over second-period price.

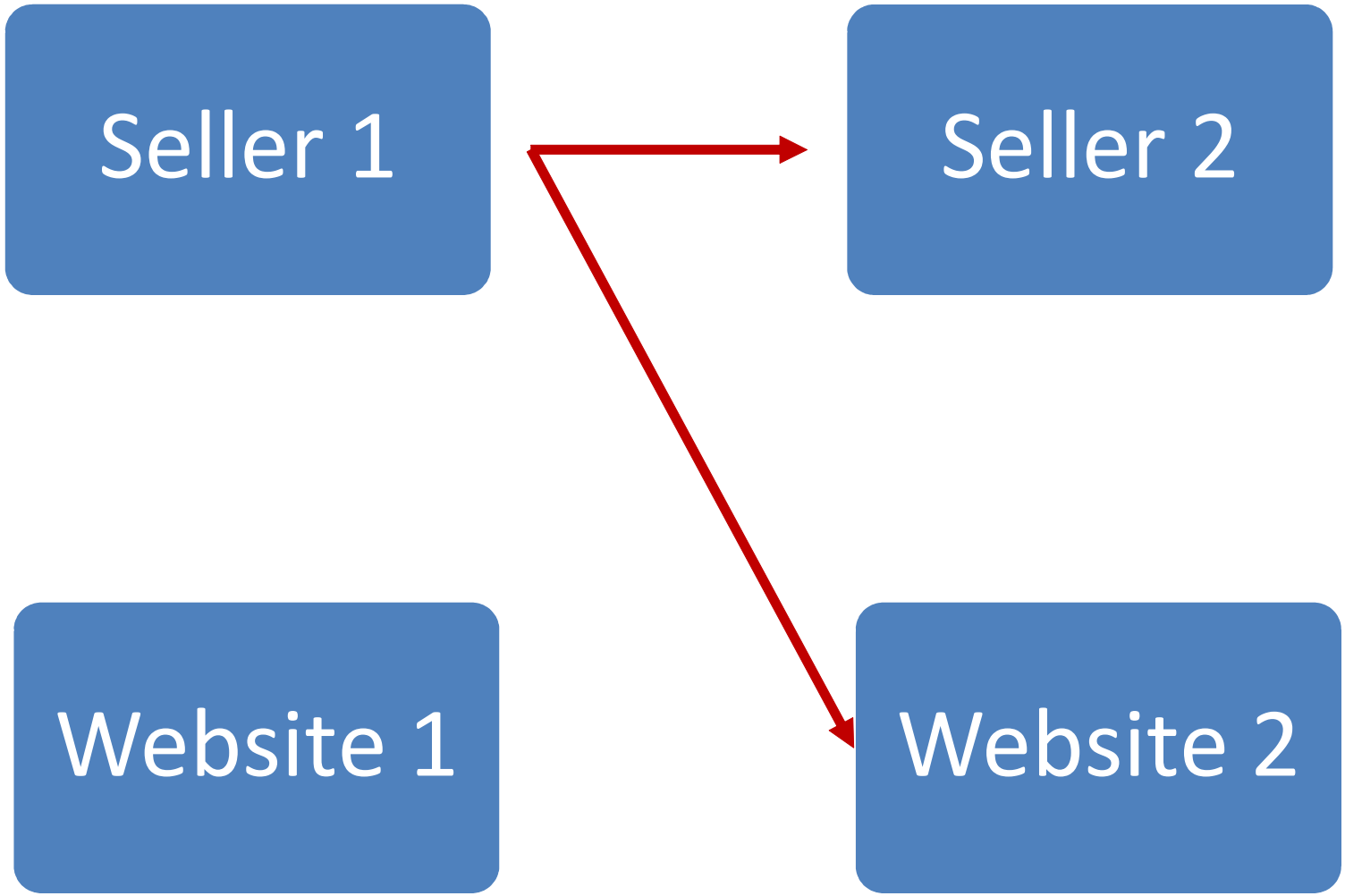
Expects lower demand, charges lower price.

Consumer buys fewer units at a lower price.

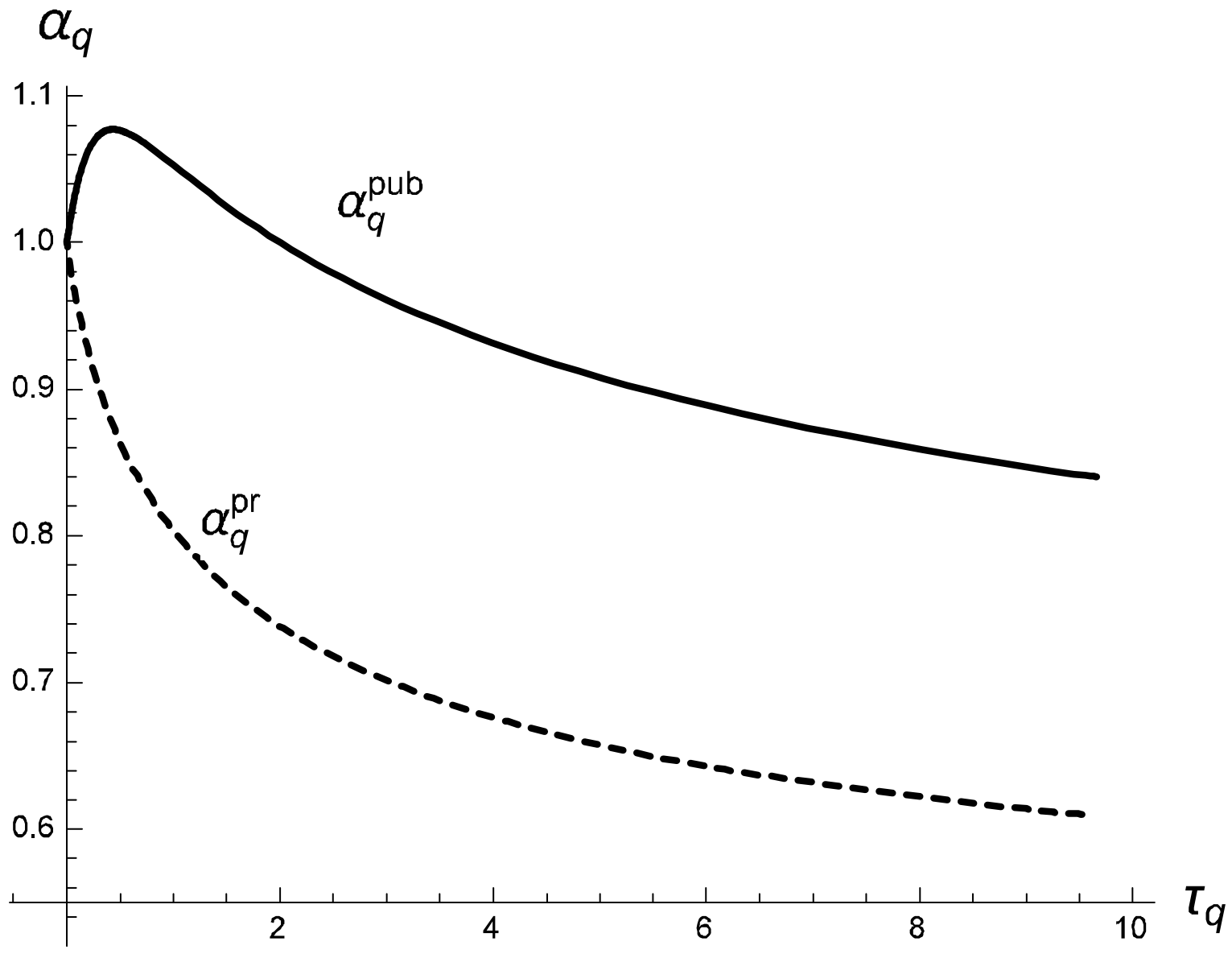
Period-1 gains can compensate loss in expected period-2 surplus.

Note: source of period-2 loss is irrelevant

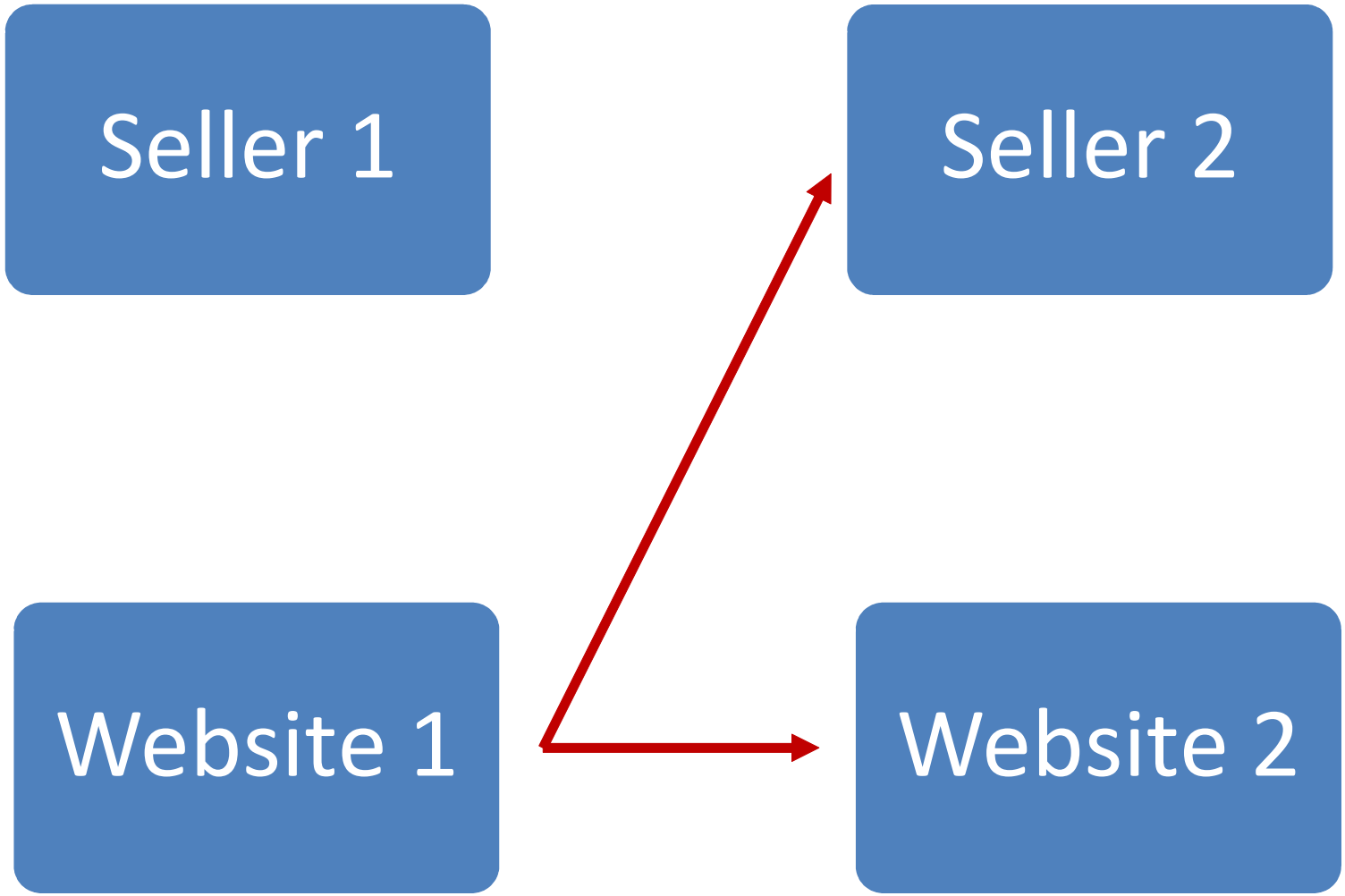
Public Signals



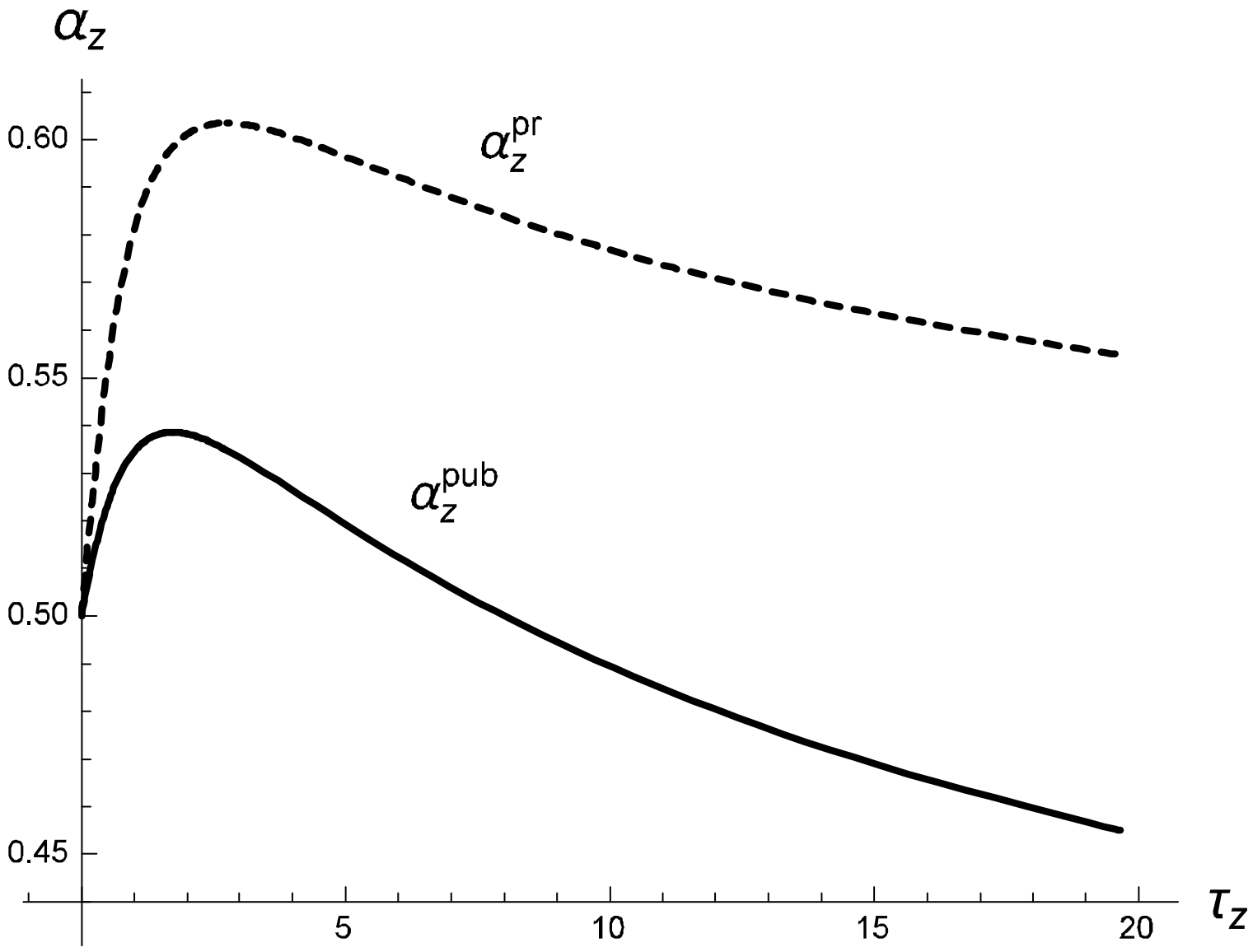
Public vs. Private Purchase Signal



Public Signals



Public vs. Private Browsing Signal



Consumer Surplus: Comparison

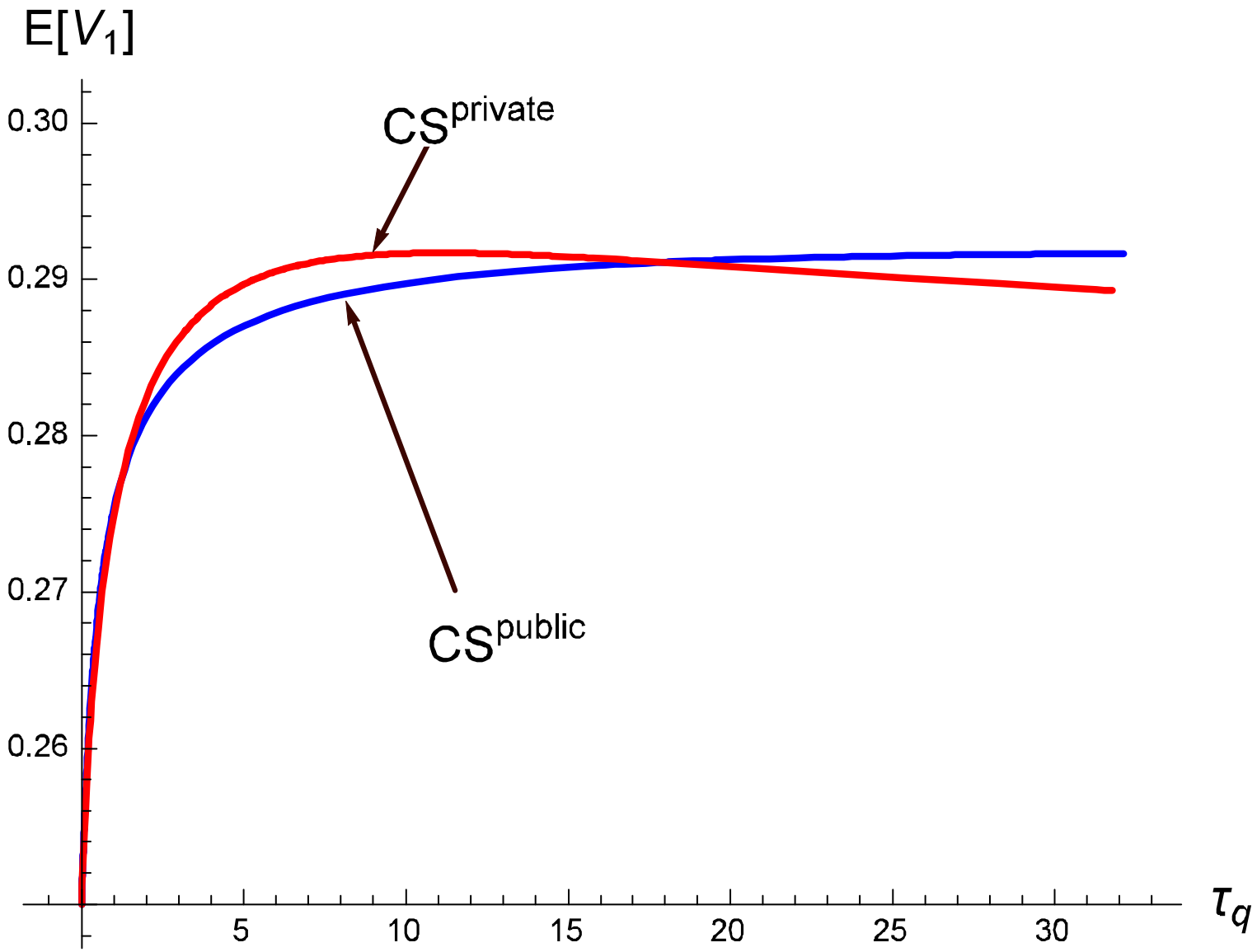
Proposition (Effect of Leakages)

1. *For all τ_z, ex ante consumer surplus is higher when browsing signals are private.*

2. *For all τ_q, period-1 prices are higher when purchase signals are public.*

3. *For sufficiently large τ_q, consumer surplus is higher when purchase histories are public.*

Payoff Comparison: Purchase Signal



Public vs. Private Purchases

Public purchase signals introduce signaling value through q_1.

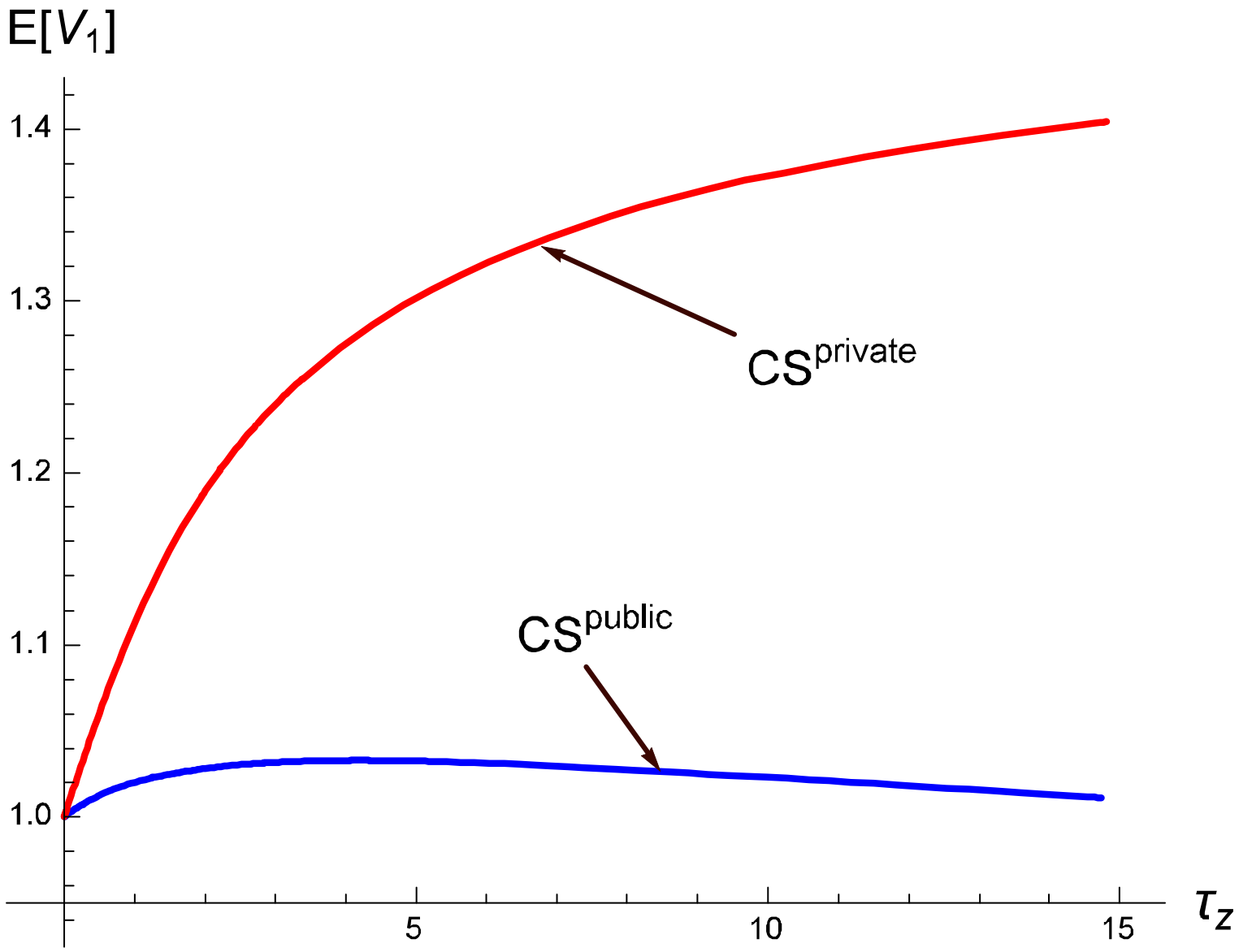
Trade-off:

- better match with period-2 website,
- *vs*. higher period-1 price.

With high precision τ_q, a small change in q_1 suffices to signal
$\Rightarrow$ public purchase signals are beneficial.

A closer look: public purchases increase consumer surplus restricted to her interaction with sellers.

Payoff Comparison: Browsing Signal



Public vs. Private Browsing

Public browsing signal $\approx$ exogenous signal for period-2 seller.

Reduces consumer surplus in period 2.

No balancing effect in period 1.

Implications for Consumers

Two kinds of information, very different effects.

"Compensation" for information revealed *within a transaction.*

Consumer cannot be compensated when "just browsing."

Benefits of cross-tracking of purchases.

Downsides to cross-tracking of "browsing" behavior.

Implications for Consumers

Two kinds of information, very different effects.

"Compensation" for information revealed *within a transaction.*

Consumer cannot be compensated when "just browsing."

Benefits of cross-tracking of purchases.

Downsides to cross-tracking of "browsing" behavior.

> *"A poorly thought out legislative solution would likely result in a very rigid framework that assigned individuals additional rights with respect to information about themselves, but did not allow for ways to sell such property rights in exchange for other considerations." (Varian, 1996)*

Implications for Firms

Fix signal technologies $\tau_q > 0,\ \tau_z > 0$.

Websites access *purchase* histories:

- informational content of *browsing* signal $\alpha_z^2 \tau_z$ decreases.

Implications for Firms

Fix signal technologies $\tau_q > 0,\ \tau_z > 0$.

Websites access *purchase* histories:

- informational content of *browsing* signal $\alpha_z^2 \tau_z$ decreases.

Sellers access *browsing* histories:

- informational content of *purchase* signal $\alpha_q^2 \tau_q$ increases;
- first-period price increases.

Strong incentives for sellers to purchase browsing signals.

Future Directions

Stationary (fully dynamic) model.

Externalities on other consumers ("Look-alikes").

Endogenous information structures.

Market structure in data sector.

Place in the Literature

"The Economics of Privacy" by Acquisti, Taylor, and Wagman (forthcoming *JEL*).

Behavior-based price discrimination: Fudenberg and Villas-Boas (2006, 2012).

Tracking and selling purchase histories: Taylor (2004), Acquisti and Varian (2005), Hermalin and Katz (2006).

Selling consumer-level information: Bergemann and Bonatti (2015), Bergemann, Bonatti and Smolin (2015).

Richer Information Structures

Numerical results show that

1. There exists $\bar{\tau}_q > 0$ such that it is optimal to keep browsing histories private, i.e., if $\tau_q < \bar{\tau}_q$ then $V^{\text{pr}}(\tau_z) > V^{\text{pub}}(\tau_z)$ for all $\tau_z \geq 0$.

2. If $V^{\text{pr}}(\tau_q) < V^{\text{pub}}(\tau_q)$ holds for $\tau_z = \bar{\tau}$, then it also holds for all $\tau_z > \bar{\tau}$. In the limit $\tau_z \longrightarrow \infty$ it is irrelevant whether purchase histories are private or public.

3. $\alpha_q(\tau_z)$ is increasing and $\alpha_z(\tau_q)$ is decreasing (both closer to myopic behavior).

Steady-State Model

Value of signaling is effectively exogenous.

Pure price discrimination. In each period, meet a different seller.

The consumer's type θ_t follows

$$\mathrm{d}\theta_t = -k(\theta_t - \theta_0)\mathrm{d}t + \sigma_\theta \mathrm{d}Z_t^\theta.$$

The purchases signal is given by

$$\mathrm{d}Y_t = \alpha\theta_t \mathrm{d}t + \sigma_Y \mathrm{d}Z_t^Y.$$

Steady-State of Learning

The market's posterior mean is m_t follows

$$\mathrm{d}m_t = -\kappa(m_t - \theta_0)\mathrm{d}t + \frac{\alpha_t \gamma_t}{\sigma_Y^2}(\mathrm{d}Y_t - \alpha m_t \mathrm{d}t),$$

The variance γ_t is deterministic, with

$$\dot{\gamma}_t = -2k\gamma_t + \sigma_\theta^2 - \left(\frac{\alpha\gamma_t}{\sigma_Y}\right)^2.$$

Equilibrium in the steady state of learning (α, γ) where $\dot{\gamma} = 0$.

The precision of beliefs is endogenous.

Proposition (Public Browsing Signal)

1. *The weight α_z on the consumer's type a is inverse-U shaped in τ_z and below its myopic level $(a = 1/2)$ for large enough τ_z.*

2. *The precision of the firms' beliefs is increasing in τ_z.*

3. *Consumer surplus is inverse-U shaped in τ_z.*

Proposition (Public Purchase Signal)

1. *The equilibrium weight on the consumer's type α is inverse-U shaped in τ_q and below the myopic level $(\alpha = 1)$ for large enough τ_q.*

2. *The precision of the firms' beliefs is increasing in τ_q.*

3. *The first-period equilibrium price is decreasing in τ_q.*

4. *Consumer surplus is inverse-U shaped in τ_q and maximized when $\alpha^*(\tau_q) < 1$.*

A Closer Look

Public signal $\approx$ commitment device to higher weight α_q.

As $\tau_q \to \infty$, seller 2 puts weight $=1$ on purchase signal:

$$m_S \approx \frac{s_q - \beta(p_1, w_1)}{\alpha}.$$

Higher α *reduces* sensitivity of beliefs, allows for higher consumption with less price impact $\partial \mathbb{E}[p_2] / \partial q_1$.

Continuation Game (Myopic Benchmark)

Consumer θ's best replies:

$$\begin{aligned} q_2 &= \theta - p, \\ z_2 &= (\theta + w)/2. \end{aligned}$$

Period-2 firms' beliefs:

$$\begin{aligned} m_S &\triangleq \mathbb{E}[\theta \mid \mathcal{I}_S] \\ m_W &\triangleq \mathbb{E}[\theta \mid \mathcal{I}_W] \end{aligned}$$

Seller and website choose

$$\begin{aligned} p_2^* &= m_S/2, \\ w_2^* &= m_W. \end{aligned}$$

Firms' Beliefs

To form beliefs, firms conjecture linear strategies

$$\begin{aligned} q_1 &= \alpha_q \cdot \theta + \beta_q (p, w), \\ z_1 &= \alpha_z \cdot \theta + \beta_z (p, w). \end{aligned}$$

Posterior means

$$\begin{aligned} m_S &= \frac{\alpha_q^2 \tau_q}{\alpha_q^2 \tau_q + \tau_0} \frac{s_q - \beta_q (p, w)}{\alpha_q} + \frac{\tau_0}{\alpha_q^2 \tau_q + \tau_0} \theta_0, \\ m_W &= \frac{\alpha_z^2 \tau_z}{\alpha_z^2 \tau_z + \tau_0} \frac{s_z - \beta_z (p, w)}{\alpha_z} + \frac{\tau_0}{\alpha_z^2 \tau_z + \tau_0} \theta_0. \end{aligned}$$

Signal-to-noise ratios depend on conjectured strategies.

Firms' Beliefs

To form beliefs, firms conjecture linear demands

$$\begin{aligned} q_1 &= \alpha_q \cdot \theta + \beta_q (p, w), \\ z_1 &= \alpha_z \cdot \theta + \beta_z (p, w). \end{aligned}$$

From the consumer's perspective,

$$\begin{aligned} \mathbb{E}[m_S \mid q] &= \frac{\alpha_q^2 \tau_q}{\alpha_q^2 \tau_q + \tau_0} \frac{q - \beta_q (p, w)}{\alpha_q} + \frac{\tau_0}{\alpha_q^2 \tau_q + \tau_0} \theta_0, \\ \mathbb{E}[m_W \mid z] &= \frac{\alpha_z^2 \tau_z}{\alpha_z^2 \tau_z + \tau_0} \frac{z - \beta_z (p, w)}{\alpha_z} + \frac{\tau_0}{\alpha_z^2 \tau_z + \tau_0} \theta_0. \end{aligned}$$

Note: consumer can influence posterior means (m_S, m_Z) only.

Equilibrium with Private Signals

Consumer's first-period best replies (q_1, z_1) satisfy

$$\begin{aligned} q &= \theta - p - \frac{1}{2}\left(\theta - \frac{\mathbb{E}\left[m_S \mid q\right]}{2}\right)\frac{\partial \mathbb{E}\left[m_S \mid q\right]}{\partial q}, \\ z &= \frac{\theta + w}{2} + \left(\theta - \mathbb{E}\left[m_W \mid z\right]\right)\frac{\partial \mathbb{E}\left[m_W \mid z\right]}{\partial z}. \end{aligned}$$

Equilibrium:

- Linear conjectures, linear replies, match coefficients.
- Amount of information conveyed by the signals is endogenous.