



Public and Private Sphere and the Authentic Self

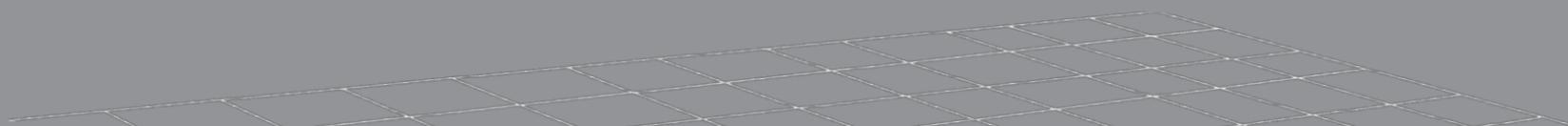
Jean Tirole

Toulouse School of Economics

Markus Brunnermeier

Princeton

24. Sept. 2020



Every Thursday at 12:30 (New York Time)

■ Upcoming



- Esther Duflo
“COVID”



- Esteban Rossi-Hansberg
“Climate Change”



- Chris Sims
“How to worry about
government debt”



- Bengt Holmstrom
“Seasonal COVID”

Markus' intro on **Privacy**

- Information rent/consumer rent vs. better recommender system
- In the private sphere one can one's authentic self
- Privacy and Brainstorming
 - Private sphere involve people who know me well already
(who have a tight prior and hence their posterior doesn't shift)
 - “free to think/bold brainstorm openly”
fosters creativity without distorting self-image
 - Overcome “cancel culture”
- “Free to plot a revolt/revolution”
 - Is privacy regulation (privacy of letters) helping/harming democracy?

Public Sphere

- Observable by public/a large group (social media)
- Jürgen Habermas' definition:
 - Institutional arrangement (democratic process) determines how people can interact and arrive at consensus (truth/beliefs)
 - Connection to cooperative game theory?
 - The way public sphere (coffee houses in 19th century) is designed determines how people interact and how networks are formed
 - Should this be decided by Facebook/Twitter/...?

The Shift from Private to Public Sphere

evolution

- Voice/sound/language
- Script allows interaction across generations
- Print press (Johannes Gutenberg)
- Social media
- Face recognition
- Brain-to-brain link (NeuraLink)
 - 86bn neurons vs. few thousand/million “connectors”
- Are we still “free to think”?
(Brainstorm with ourselves, without worrying which self-image we project to society/public sphere)

“Tech elite” can measure our inner feelings

- Tech elites 2 advantages
 - Informational advantage
 - Due designs public sphere

- Feelings
 - Beyond our control (limbic system)
 - In our control (prefrontal cortex)

Tech elites advantage
info advantage
change our behavior

- Jean Tirole will address the most important questions of our times.

Poll Results

1. Is there a trade-off between privacy and influence
 1. Yes, more privacy comes at the cost of less influence
 2. No
2. Human behavior is mostly driven by (multiple choice possible)
 1. Extrinsic motivation (monetary incentives)
 2. Intrinsic motivation
 3. Signaling intrinsic motivation (reputation), which distorts “authenticity”
3. With the emergence of social media, I distort my “self-authenticity”
 1. more (to present a favorable self-image is more important)
 2. less
4. Rules of public discourse and privacy should be decided by
 1. Facebook/Twitter/...
 2. Parliaments/governments
 3. Society at large through public interlocutor debate

Thank you!

markus@princeton.edu

PUBLIC AND PRIVATE SPHERES AND THE AUTHENTIC SELF

Work in progress!!!

Princeton University Bendheim Center

Markus' Academy

September 24, 2020

Jean Tirole

INTRODUCTION

Two observations

- Technological revolution [AI, ratings, facial recognition, recording of online and public space interactions, data externalities...] \Rightarrow our life is more and more exposed to public view
- Experimental literature: we change our behavior when observed by others whose judgment we value (more signaling = less authenticity)

How will this alter the nature of our social relationships, and is this desirable?
Failing empirical evidence on the matter, can theory shed some light on the design of privacy policies?

Inflation of the public sphere entails both a *level* and a *composition* effect:

- (1) Behaviors that once belonged to the private sphere become observable to a larger audience. How will private sphere be affected?
- (2) Expansion of the public sphere is not random: the selective relationships of our private sphere are biased towards like-minded and caring individuals. Behavioral implications?

Considerable lab-and-field evidence:

“Giving a socially-valued behavior more visibility makes it more prevalent”

This paper: making a behavior more visible may not increase its frequency in environments that diverge from the ones used in existing experiments.

Two apparent caveats in above statement:

- (1) *Single activity*. Consider signaling environment in which the individual's trait to be signaled (say, prosociality) is correlated or the same across *multiple activities*: increased visibility of an activity increases signaling incentive on this activity but alters it in others.
- (2) *Socially valued activity*. Corrida or boxing attendance, religious slaughtering of an animal: A fraction of the population may infer good traits (attachment to cultural roots, conviviality) while others will find the behavior repulsive. Politics, religion, sexuality...

⇒ Two topics

- (1) multitasking (public and private sphere)
- (2) divisive behaviors.

Add the normative dimension to the picture

We will focus on diagonal entries in:

Consumption Heterogenous trait	exerts an externality	is victimless
is a vertical attribute (altruism)	consensual consumption (prosocial behavior) Part 1	status consumption
is an horizontal attribute (taste)	mixed case	divisive consumption (religious, politics, sexuality...) Part 2

Upper-right corner: conspicuous consumption [Veblen, Pesendorfer, Bagwell-Bernheim...]

Lower-left corner: Two-dimensional type (altruism, taste)?

Less frequent (externality, yet not consensual)?

Outline

(1) Consensual behavior

- Single task: classic model, behavior driven by reputational concerns. Giving the behavior more visibility indeed makes it more prevalent.
- Multiple actions, some in private sphere, others public. Presumably better behavior in public sphere. Impact of an expansion of the public sphere on private sphere behavior? Is expansion overall desirable?

(2) Divisive behavior

- Impact of privacy on authenticity?
- Public policy?

(Some) related literature

- Modeling: that of signaling prosocial behavior [e.g. Bénabou-Tirole 2006]
Review of empirical evidence [Ashraf-Bandierra 2018, Bursztyn-Jensen 2017].
- Posturing and privacy: effect of visibility [e.g. Ali-Bénabou 2019], chilling effect [e.g. Jann-Schottmüller 2020], under/over signaling [e.g. Ball 2020, Bonatti-Cisternas 2020, Bernheim-Bodoh-Creed 2019]. Literature varies number of actions or their visibility. In part 1, multiple audiences with different information structures.
- Signaling to multiple audiences [e.g. Gertner et al 1988, Spiegel-Spulber 1997, Austen-Smith-Fryer 2005, Bar-Isaac-Deb 2014, Bouvard-Levy 2017, Bursztyn et al 2017]. In part 2, endogeneity of information is key.
- Philosophy: authenticity (associated with privacy) often synonym of emancipation.

1. CONSENSUAL BEHAVIOR

Single-task benchmark

Agent i selects $a_i \in \{0, 1\}$

$a_i = 1$

- is pro-social (generates some externality $e > 0$)
- is costly for agent i ($c > 0$)
- intrinsic motivation $v_i \sim dF(v_i)$ on \mathbb{R}^+

Payoff

$$u_i = (v_i - c)a_i + \theta \hat{v}_i(a_i)$$

intrinsic motivation extrinsic motivation intensity of image concerns image

Welfare: $(v_i - c + e)a_i$ (image is positional/zero sum)

Note: θ may be an average (e.g. $\theta = \int_0^\infty \theta(v)dF(v)$). “Privacy” \longleftrightarrow low θ .

Equilibrium: selects $a_i = 1$ iff $v_i \geq v^*$, where

$$v^* - c + \theta \Delta(v^*) = 0$$

and

$$\Delta(v^*) \equiv M^+(v^*) - M^-(v^*) \equiv E[v|v \geq v^*] - E[v|v < v^*].$$

Assume $v^* - c + \theta \Delta(v^*)$ increasing in v^* : Guarantees uniqueness of v^* .

Authenticity (flip side of signaling): $A \equiv \frac{v^*}{c} = 1 - \frac{\theta \Delta(v^*(\theta))}{c}$ decreases with θ .

An increase in authenticity is $\begin{cases} \text{desirable ("emancipating")} & \text{if } e < \theta \Delta(v^*) \\ \text{undesirable} & \text{otherwise} \end{cases}$

Multi-tasking

Same social preferences, but multiple actions a_{ij} , each with some counterparty j

- s (for “silo”) in private sphere: only j learns a_{ij}
- t (for “transparency”) in public sphere: everyone learns a_{ij} .

Assume a continuum of actions (could be finite number, even only 2)

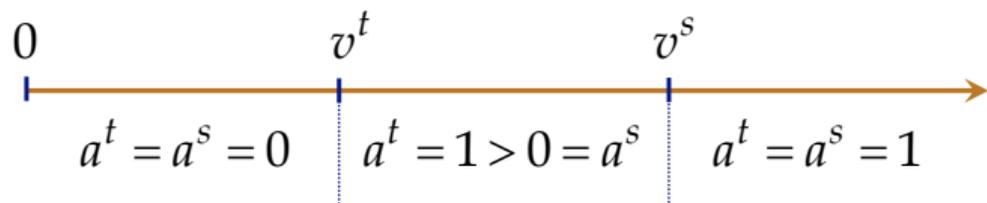
Focus on deterministic, symmetric behavior

$$a_{ij} = \begin{cases} a^s & \text{in private sphere} \\ a^t & \text{in public sphere} \end{cases}$$

Agent i solves

$$\max_{(a^t, a^s) \in \{0,1\}^2} \{ (v_i - c)(ta^t + sa^s) + \theta[t\hat{v}_i(a^t) + s\hat{v}_i(a^t, a^s)] \}$$

Look for two cutoffs: $v^t < v^s$.



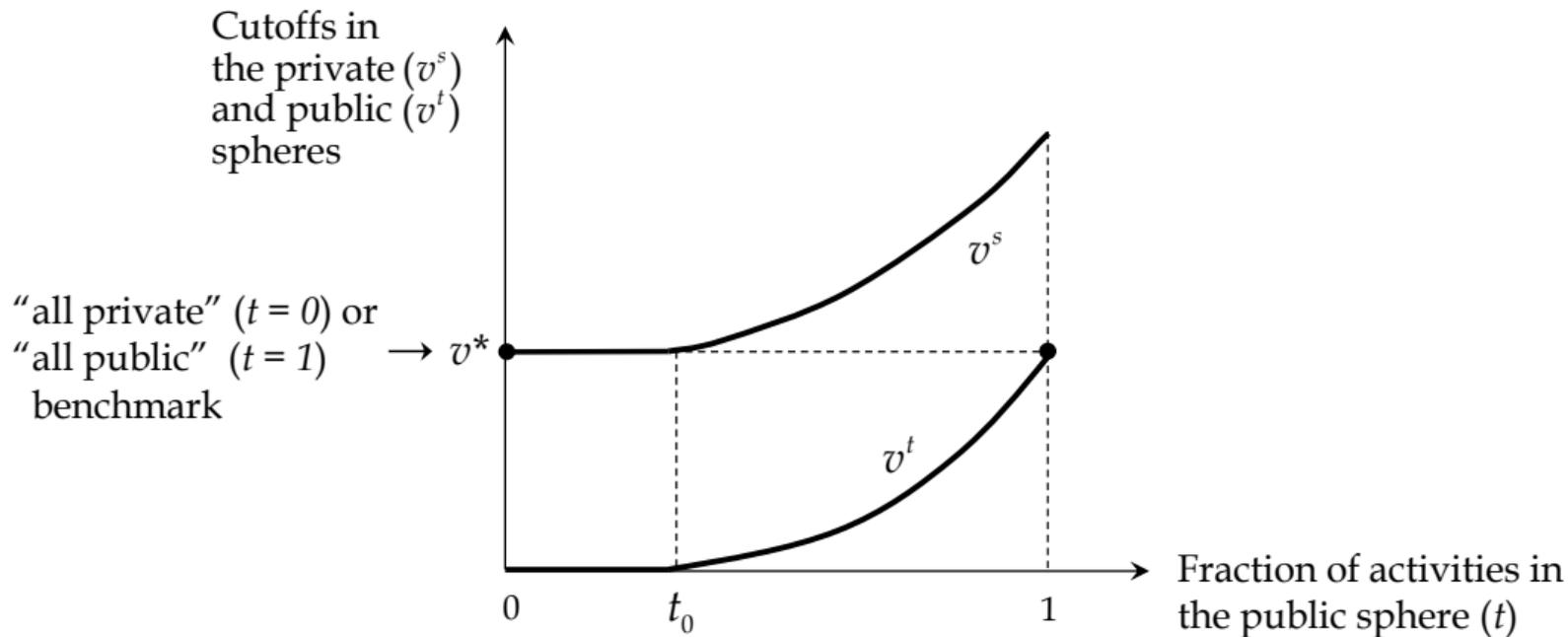
First-order condition if interior, letting $M(v_0, v_1)$ denote the mean of v conditional on $v \in [v_0, v_1]$, in the private sphere:

$$v^s - c + \theta[M^+(v^s) - M(v^t, v^s)] = 0$$

and (public sphere: scores double)

$$t[v^t - c] + \theta [t[M^+(v^t) - M^-(v^t)] + s[M(v^t, v^s) - M^-(v^t)]] = 0$$

Properties of equilibrium



Equilibrium contributions (where $v^* - c + \theta\Delta(v^*) = 0$)

Insights

- Misallocation of effort ($v^s > v^t$)
- Public sphere reduces prosociality in private sphere ($v^s \geq v^*$); already apart from the chaff
- Crowding out by public sphere
 - An increase in t reduces pro-sociality in both spheres
 - An increase in t reduces \bar{a} over some range
($\bar{a}(0) = \bar{a}(1)$ and \bar{a} initially increasing in t)
 - Uniform distribution: $\bar{a}(t)$ hump-shaped with a peak at t_0 .

Bottom line: an increase in transparency may not be desirable even if there is too little prosocial behavior (in contrast with single action).

2. DIVISIVE BEHAVIOR

People disagree as to what is “moral” or “immoral”, “right” or “wrong”.

To capture this, we assume that

- $v \sim F(v)$ on \mathbb{R} (not \mathbb{R}^+)
- F symmetric around 0

Measuring lack of consensus

- (a) *Symmetry* of F in itself captures a strong form of disagreement
- (b) *Polarization* (comparison among symmetric distributions)

$F(v, \rho)$ where ρ is a *rotation* index ($F_\rho \geq 0$ for $v \leq 0$ and $F_\rho \leq 0$ for $v \geq 0$).

Social approval: $\theta(v_j)$ such that $\theta(0) = 0$ and θ symmetric and increasing
Image remains a positional (zero-sum) good

(Single) action: $a_i \in \{-1, 0, 1\}$

“Acting” ($a_i \in \{-1, 1\}$) entails a cost $c \geq 0$.

Agent i 's payoff function

$$v_i a_i - c|a_i| + \int_{-\infty}^{+\infty} \theta(v_j) \hat{v}_{ji} dj$$

Assumed choice: victim-less behavior ($e = 0$).

(a) *Privacy as activity-contingent visibility*

Action $a_i = +1$ (resp. $a_i = -1$) is observed by agents who select the same action.

Non-action $a_i = 0$ is observed by no-one.

- safe space frequented only by peers (church, corrido, club, political party)
- captures in a stark form idea that social relationships are biased toward like-minded individuals.

Incentive compatibility \Rightarrow monotonicity/thresholds $v^- \leq v^+$.

Let $\Theta(v^+) \equiv \int_{v^+}^{+\infty} \theta(v) dF(v)$ (odd function, while Δ is even): weight put on image/in-group (out-group unable to tell $a_i = 1$ and $a_i = 0$ apart)

Equilibrium

- Must be symmetric: $v^+ = -v^- = v^s \geq 0$
- $v^s - c + \Theta(v^s)\Delta(v^s) = 0$ $\left(A = 1 - \frac{\Theta(v^s)\Delta(v^s)}{c} \right)$ if $v^- < 0 < v^+$
(otherwise $v^s = 0$ and non-action is not selected).

Contrast with transparency (a_i observed by all)

Then no image concerns (what is approved by some is disapproved by others)

$$v^t - c = 0 \quad (A = 1)$$

Proposition (divisive behaviors)

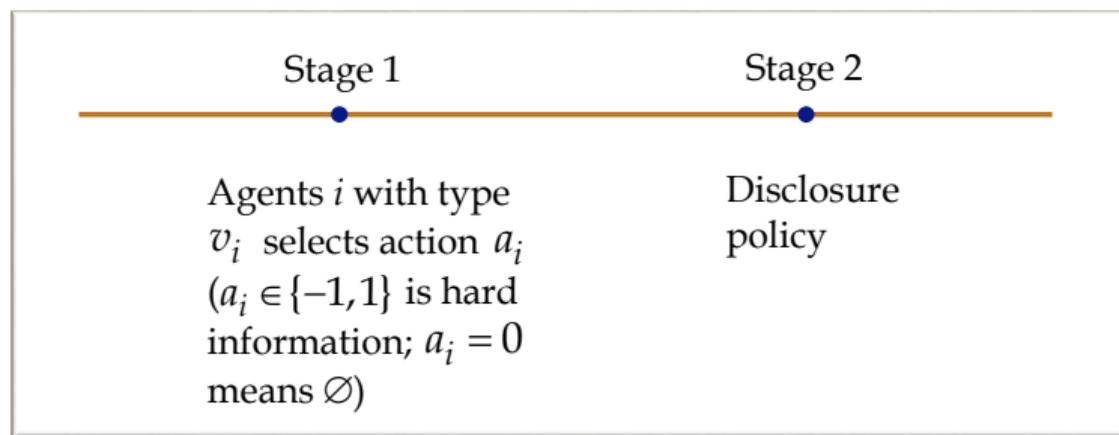
- (i) Authenticity. Transparency yields authentic behavior while privacy, by encouraging agents to impress like-minded peers, does not.
Authenticity under privacy decreases both with the importance of social approval and with polarization.
- (ii) Welfare. Welfare is higher under transparency.

Turns consensual-behavior benchmark on its head

Privacy { (1) encourages signaling (at the expense of moderates)
(2) is welfare-reducing despite absence of externality

(b) *Endogenous visibility*

(i) *Status-quo = full privacy* (work in progress)



Lexicographic preferences for simplicity:

- (1) Extrinsic, intrinsic & image motivation
- (2) If information available about others, consult it.

Repository = two-sided platform/club defined by information that is required to have access to repository's information

Left and right repositories: recreate safe spaces (visibility only to ingroups/peers).

Proposition (disclosing behavior)

The outcome under activity-contingent visibility is still an equilibrium when privacy is full, but agents can share their information within repositories.

(ii) *Status-quo = full transparency*

Outgroup transparency

Hiding cost

- sexual minorities not enjoying public space together
- not using freedom of speech
- resorting to costly or untrustworthy supplier

Hiding cost $h > 0$ to be seen only by peers.

Then $v^* - c + \Theta(v^*)\Delta(v^*) = h$

if $h \leq \Theta(v^*)[M^+(v^*) - M^+(-v^*)]$

Deadweight loss

COMPETITION FOR ATTENTION

Back to consensual behavior, but what is unknown by audience is *relative* preferences, as opposed to unknown absolute empathy. Unknown relative empathy model (link to Maskin-Tirole 2019)

- $F(v_z)$ where $z \in [0, 1]$ is rank-order (v_z increasing in z)
- all permutations $j \rightarrow \omega_i(j) = z$ equally likely

$$v^s - c + \theta \Delta(v^s) = 0 \quad \text{and} \quad v^t = c$$

Proposition (unknown relative empathy)

Transparency reduces prosocial behavior when the individual's relative (rather than absolute) empathy is unknown. Authenticity is maximal ($A = 1$) under transparency.

Analogy with divisive issues: what is good news for someone is bad news for others \Rightarrow when \bar{a} large, favors become suspicious.

3. CONCLUSION

Conventional wisdom

- (Positive) Increased visibility of a consensual action \Rightarrow more prevalence (lower authenticity)
- (Normative) Increased visibility is good if sizeable externality (accountability), bad if small one

Conclusions may be turned on their head in broader contexts

- Multitasking
 - Increased visibility for a task crowds out signaling incentive for less visible task
 - Increased visibility may thereby reduce average prosocial contribution
- Divisive behaviors
 - Privacy (safe space) \Rightarrow less authenticity
 - But making privacy the default avoids hiding costs.

Future research: broader views of privacy.

Demand for privacy

- emancipation from the need to posture if consensual behavior, avoid revealing to out-group if divisive behaviors (escape from oppressive government or society, form safe space to escape censorship)
- alternative reasons
 - letting off steam, throwing ideas around without antagonizing others

Cost of privacy

- literature: privacy \Rightarrow lack of accountability
- costly restriction to specific safe spaces, fake identity, etc. in this work; may also reduce authenticity
- privacy may lead to “ghettoisation of thinking”.