Verifiability and effective persuasion

Noémie Cabau[1] Ming Li[2]

ESEM 2023-Barcelona School of Economics
2023.08.30.

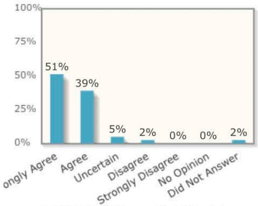
[1]Budapest University of Technology and Economics
[2]Concordia University, CIRANO, and CIREQ

Carbon Tax

DECEMBER 20, 2011 ADD TO POLL

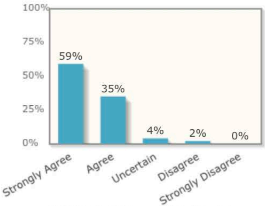
A tax on the carbon content of fuels would be a less expensive way to reduce carbon-dioxide emissions than would a collection of policies such as "corporate average fuel economy" requirements for automobiles.

Responses



Source: IGM Economic Experts Panel
www.igmchicago.org/igm-economic-experts-panel

Responses weighted by each expert's confidence



Source: IGM Economic Experts Panel
www.igmchicago.org/igm-economic-experts-panel

Recent Polls

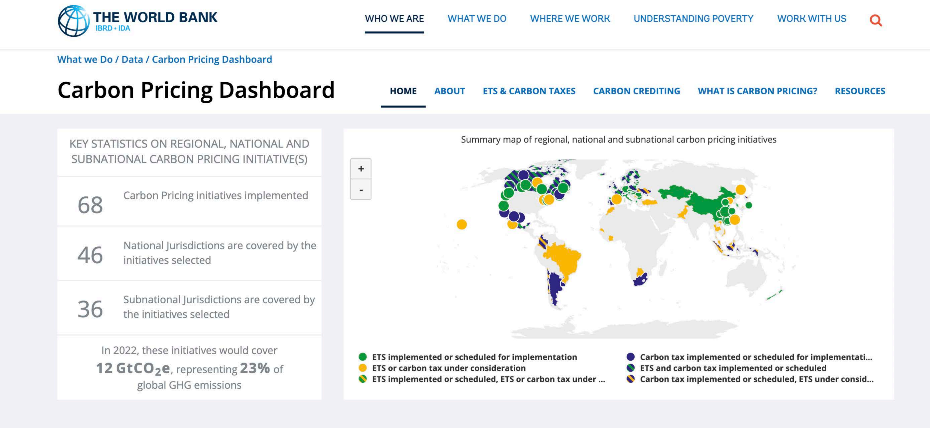

Stablecoins
May 25th, 2022


Stablecoins
May 25th, 2022


Energy Sanctions
May 11th, 2022


Energy Sanctions
May 11th, 2022


Ranked-Choice Voting
April 22nd, 2022


THE WORLD BANK
IBRD • IDA
WHO WE ARE
WHAT WE DO
WHERE WE WORK
UNDERSTANDING POVERTY
WORK WITH US
What we Do / Data / Carbon Pricing Dashboard
Carbon Pricing Dashboard
HOME
ABOUT
ETS & CARBON TAXES
CARBON CREDITING
WHAT IS CARBON PRICING?
RESOURCES
KEY STATISTICS ON REGIONAL, NATIONAL AND SUBNATIONAL CARBON PRICING INITIATIVE(S)
68 Carbon Pricing initiatives implemented
46 National Jurisdictions are covered by the initiatives selected
36 Subnational Jurisdictions are covered by the initiatives selected
In 2022, these initiatives would cover 12 GtCO2e, representing 23% of global GHG emissions
Summary map of regional, national and subnational carbon pricing initiatives
+
-
ETS implemented or scheduled for implementation
ETS or carbon tax under consideration
ETS implemented or scheduled, ETS or carbon tax under ...
Carbon tax implemented or scheduled for implementati...
ETS and carbon tax implemented or scheduled
Carbon tax implemented or scheduled, ETS under consid...

Introduction and Motivation

- Experts often fail to convince decision makers to select the right action appropriate for the situation.
  - medical experts on COVID-related health measures
  - economists on carbon tax
- Experts' interests are not perfectly aligned with the decision maker
  - experts consider what is optimal from their perspective, but decision maker has to consider other tradeoffs
  - experts may have biases: conflict of interest, strongly held opinion, etc.

Introduction and Motivation

Some casual empiricism:

- Judicial trials are very good at finding the facts; organizations' internal investigations in response to accusations usually less convincing.
- Using history to make an argument is easier than using forecasts to do so.

Observation:

- Information is not always fully verifiable.
    - Existence of accounting error is verifiable, but absence of one is not;
    - Possession of a qualification is verifiable, but lack of one is not.
- Verifiability matters to whether persuasion is effective.

When is effective persuasion possible and when is it not?

- Punchline: nature of verifiability matters.

Related literature

- Communication of hard information: Milgrom (1981), Milgrom and Roberts (1986), Grossman (1981);
- Multiple experts: Bhattacharya and Mukherjee (2013), Wolinsky (**2002**);
- Partial verifiability: Wolinsky (**2002**, 2003), Dziuda (2011);
- Information aggregation: Austen-Smith and Banks (1996), McLennan (1998).

Model setup

We follow closely the setup of Wolinsky (2002).

- Decision maker chooses action from $a \in \{L, H\}$.
- Experts $i \in \{1, 2, \dots, n\}$ each observes a signal $s_i \in \{0, 1\}$.
- Signal $s_i = 1$ is more favorable news for H than $s_i = 0$.
- State of the world: $S := \sum_j s_j$.

Preferences

There is a conflict of interest between experts and decision maker.

- *DM*'s payoff V:

$$\tilde{V}(s|a) = \begin{cases} V(S) & \textit{if } a = H \\ 0 & \textit{if } a = L \end{cases}$$

  where V is increasing.
- All experts have the same preferences, U, given by:

$$\tilde{U}(s|a) = \begin{cases} U(S) & \textit{if } a = H \\ 0 & \textit{if } a = L \end{cases}$$

  where U is increasing.
- $V(S-1) > U(S)$ for any $S \in \{1, \ldots, n\}$: panel has a *conservative* bias (think reopening)–less eager to take action H.
- $U(n) > 0 > V(0)$: agreement in extreme states.

Strategies

- Experts simultaneously report $r_i \in \{0, 1\}$;
- Decision maker chooses action H with probability $\rho(r_1, r_2, \ldots, r_n) \in [0, 1]$;
- Expert's strategy $y_i(0)$ and $y_i(1)$: Probability of telling the truth when signal is 0 or 1, respectively.

Verifiability

- **Definition.** A signal s is *verifiable* if the expert can prove with hard evidence that he has received signal s; signal s' is *unverifiable* if the expert cannot prove he has received s' (or equivalently, has not received $s \neq s'$).
- Possible verifiability scenarios:
  - Neither 0 nor 1 is verifiable $\Rightarrow$ Cheap-Talk;
  - Signal 1 is verifiable but 0 is not, verifiability opposite to experts' bias (organization scandals/accounting error);
  - Signal 0 is verifiable but 1 is not, verifiability towards experts' bias (employee competence);
  - Both signals are verifiable $\Rightarrow$ Mandatory disclosure (judicial trials).

Benchmark: Cheap talk

- Key observation: when an expert considers their strategy, they consider only the instances in which they are pivotal – changing a decision from L to H or vice versa.

Lemma 1. In the cheap-talk scenario, whenever pivotal, experts always tell the truth when their signal is 0: $y_i(0) = 1$, ignoring perverse equilibria.

- Intuition: in the pivotal instances, the DM is "indifferent" between H and L, but given that an expert is more conservative, they always prefer to implement L. Hence, they will not report 0 as 1.

Some unpleasant pivotal math

If i's report is pivotal given other experts' report r_{-i}, it must be that

$$\sum_{k=0}^{n} Pr_{DM}(S = k | r_{-i}, \mathbf{r_i} = \mathbf{0}, y) V(k) \leq 0$$

$$\leq \sum_{k=0}^{n} Pr_{DM}(S = k | r_{-i}, \mathbf{r_i} = \mathbf{1}, y) V(k).$$

Some unpleasant pivotal math

However, for expert i with signal $s_i = 0$, payoff from reporting 1 is

$$
\begin{aligned}
&\mathbb{E}_S[U(S)|r_i = 1, r_{-i}, y] \\
&= \sum_{k=0}^{n-1} Pr_i(S_{-i} = k|r_{-i}, y_{-i})U(k), \\
&= \sum_{k=0}^{n-1} Pr_{DM}(S = k|r_i = 0, \tilde{y}_i = 1, r_{-i}, y_{-i})U(k), \\
&< \sum_{k=0}^{n-1} Pr_{DM}(S = k|r_i = 0, \tilde{y}_i = 1, r_{-i}, y_{-i})V(k), \\
&\leq \sum_{k=0}^{n} Pr_{DM}(S = k|r_i = 0, y_i, r_{-i}, y_{-i})V(k) \leq 0.
\end{aligned}
$$

Some unpleasant pivotal math

1. First equality: DM infers that $s_i = 0$ if Expert i always tells the truth, so DM's posterior is the same as that of i;
2. Strict inequality: Assumption that experts are more conservative than the DM;
3. First weak inequality: Roughly speaking, if Expert i always reports the truth about 0, DM's inference about state is lower when receiving report $r_i = 0$ versus when Expert i sometimes potentially reports 0 as 1;
4. Last weak inequality: Expert i is pivotal.

Unpleasant pivotal math: Some details

Let

$$\omega_{DM}(k) := Pr_{DM}(S_{-i} = k | r_{-i}, y_{-i}),$$

Under $\tilde{y}_i = 1$,

$$Pr_{DM}(S = k | r_i = 0, r_{-i}, y_{-i}, \tilde{y}_i) = \omega(k).$$

Under y_i,

$$Pr_{DM}(S = k | r_i = 0, r_{-i}, y) = \begin{cases} \lambda_i \omega(0), & x = 0; \\ \lambda_i \omega(k) + (1 - \lambda_i)\omega(k-1), & 1 \leq x < n; \\ (1 - \lambda_i)\omega(n-1) & x = n, \end{cases}$$

where

$$\lambda_i := Pr_{DM}(s_i = 0 | r_i = 0, y_i) = \frac{(1-q)y_i}{(1-q)y_i + q(1-y_i)}.$$

Unpleasant pivotal math: Some details

Therefore,

$$\sum_{k=0}^{n-1} Pr_{DM}(S = k | r_i = 0, r_{-i}, y_{-i}, \tilde{y}_i) V(k) - \sum_{k=0}^{n} Pr_{DM}(S = k | r_i = 0, r_{-i}, y$$

$$= \sum_{k \in \{0,\dots,n-1\}} \omega(k) V(k) - \lambda_i \omega(0) V(0)$$

$$- \sum_{k \in \{1,\dots,n-1\}} [\lambda_i \omega(k) + (1 - \lambda_i)\omega(k-1)] V(k) - (1 - \lambda_i)\omega(n-1) V(n)$$

$$= (1 - \lambda_i) \sum_{k \in \{0,\dots,n-1\}} \omega(k) V(k) - (1 - \lambda_i) \sum_{k \in \{1,\dots,n\}} \omega(k-1) V(k)$$

$$= -(1 - \lambda_i) \sum_{k \in \{0,\dots,n-1\}} \omega(k) [V(k+1) - V(k)] \leq 0.$$

Last inequality: V is weakly increasing in k, for all $k \in \{0, \dots, n-1\}$.

Two corollaries

Corollary 1. If neither signals are verifiable, there is no informative equilibrium. DM takes the ex ante optimal action.

Corollary 2. (Wolinsky 2002) If 1 is verifiable but 0 is not, there is no informative equilibrium. DM takes the ex ante optimal action.

Intuition for both corollaries: given that 0 is always reported as 0, it will sway the decision towards L if equilibrium is informative. Thus, whenever expert is pivotal, they would lie about signal 1.

Déjà vu

Conditional on Expert i being pivotal, when $s_i = 1$,

$$
\begin{aligned}
&\mathbb{E}_S[U(S)|r_i = 1, r_{-i}, y] \\
&= \sum_{k=0}^{n-1} Pr(S_{-i} = k|r_{-i}, y_{-i})U(k+1) \\
&= \sum_{k=0}^{n-1} Pr_{DM}(S = k+1|r_i = 1, r_{-i}, y)U(k+1) \\
&< \sum_{k=0}^{n-1} Pr_{DM}(S = k+1|r_i = 1, r_{-i}, y)V(k) \\
&\leq \sum_{k=0}^{n} Pr_{DM}(S = k|r_i = 0, r_{-i}, y)V(k) \leq 0
\end{aligned}
$$

Result

Proposition. If signal 0 is verifiable but 1 is not, there exists a fully-revealing equilibrium.

Intuition: if an expert lies, they will switch the decision from L to H – bad pivotal math.

Asymmetric equilibria

An example. Suppose that $n = 3$, $q = 0.35$, $n_{DM} = 1$, and $n_E = 3$. The payoff from action H is

$$V(S) = S - 0.995$$

for DM and

$$U(S) = \begin{cases} -2 & \text{if } S \in \{0, 1, 2\} \\ 0.5 & \text{if } S = 3. \end{cases}$$

Expert suffers a big loss from action H taken when they do not think it is appropriate.
Asymmetric equilibrium: $(y_1, y_2, y_3) = (1, 0.7, 0)$ and the decision rule:

$$\rho(r) = \begin{cases} 0 & \text{if } r \in \{(0,0,0), (0,0,1), (0,1,0), (0,1,1)\} \\ 1 & \text{if } r \in \{(1,0,0), (1,1,0), (1,0,1), (1,1,1)\}. \end{cases}$$

Welfare

- DM: more information is always better. So full information equilibria are Pareto efficient.
- Experts: welfare comparison depends on comparison between "type-1" error and "type-2" error.

Example above. Experts 2 and 3 are not pivotal. When $S = 0$, H is not taken in the above equilibrium or the truthful equilibrium. When $S = 3$, H is always taken. However when $S = 1, 2$, H is taken less frequently, compared with the truthful equilibrium, which is a gain for the experts.

Asymmetric equilibria

Lemma

Suppose that $s_i = 0$ is verifiable. If (ρ^*, y^*) is an equilibrium and $y_i^* \in (0,1)$ for some $i \in E$, then the strategy profile (ρ', y') where $y_i' = 0$ and $y_j' = y_j^*$ for all $j \neq i$, and

$$\rho'(r) = \rho^*(r) \qquad \forall r \in \{0,1\}^n \tag{1}$$

is also an equilibrium. Furthermore, the two equilibria give the experts the same expected payoff.

Intuition.

Proposition

We may focus on equilibria where some experts always babble (bad cop) and some always tell the truth (good cop).

Mixed panel of experts

Proposition

Let C denote the number of conservatives experts, and L denote the number of liberal experts. (i) If no signal is verifiable, then no equilibrium is influential. (ii) Suppose that $s_i = 1$ is verifiable. If

$$\sum_{k=0}^{C} \binom{C}{k} q^k (1-q)^{C-k} V(k) \leq 0 \leq \sum_{k=0}^{C} \binom{C}{k} q^k (1-q)^{C-k} V(L+k),$$

then an influential equilibrium exists, where all liberal experts reveal their signals and all conservatives babble. (iii) Suppose that $s_i = 0$ is verifiable. A similar statement to (ii) follows under an analogous condition.

- **Intuition**. Same as above.

Experts-DM vs. Condorcet

- Differences between Experts-DM and Condorcet:
  - Decision rules are endogenous in ED, whereas it is exogenous in Condorcet.
  - Partial verifiability $\Rightarrow$ "vote" is restricted in ED, whereas it is not restricted in Condorcet.
- Consequences:
  - DM does not just count "votes"–in particular, not in asymmetric equilibria.
  - McLennan's (1998, *APSR*) insightful result does not apply.

Experts-DM vs. Condorcet

- McLennan (1998) shows that even though "sincere voting" is not necessarily an equilibrium, voters with common interests can achieve their ex ante optimal outcome through voting. So voters do not need commitment power.
- We have counterexample in which commitment benefits experts. ("Bayesian persuasion-ish"). Intuition: experts want to commit to misreporting "0" in a way that is not ex post incentive compatible.

Experts-DM vs. Condorcet: Example

In the Example above, $y = (1, 0.7, 0)$ is worse for the experts than $\hat{y} = (0.7, 0.7, 0)$, but the latter is not an equilibrium.

Table: Expected gain from $\hat{y} = \sum_{s \in \{0,1\}^n} Pr(s) \times \Delta\pi(y|s) = 0.10595$.

Signal profile	$Pr(H\|s, .)$		$\pi(\hat{y}\|s) - \pi(y\|s)$
	$\hat{y}$	y	
$s = (0,0,0)$	0.3^2	0	-0.18
$s = (0,0,1)$	0.3^2	0	-0.18
$s = (0,1,0)$	0.3	0	-0.6
$s = (1,0,0)$	0.3	1	$+1.4$
$s = (0,1,1)$	0.3	0	-0.6
$s = (1,0,1)$	0.3	1	$+1.4$
$s = (1,1,0)$	1	1	0
$s = (1,1,1)$	1	1	0

Conclusion

- Verifiability matters to effectiveness of persuasion.
- Verifiability of evidence in experts' favour: good for everyone.
- Verifiability of evidence against experts: difficulty of persuasion.