Privacy in Implementation

Ronen Gradwohl

Kellogg School of Management
Northwestern University
Mechanism Design with Information-Sensitive Agents: Privacy in Implementation

Ronen Gradwohl

Kellogg School of Management
Northwestern University
Information-sensitivity

You might be information-sensitive if...

▶ You care about what people think of you.
▶ You care about what people think of others.
▶ You care enough about these to incorporate them into your decision-making.

In game theory and mechanism design, agents typically have preferences only over outcomes. But what if agents are information-sensitive, and also care about the private information revealed in the course of an interaction? This could lead to entirely different strategic behavior...
Information-sensitivity

You might be information-sensitive if...

▶ You care about what people think of you.
Information-sensitivity

You might be information-sensitive if...

▶ You care about what people think of you.
▶ You care about what people think of others.
Information-sensitivity

You might be information-sensitive if...

▶ You care about what people think of you.
▶ You care about what people think of others.
▶ You care enough about these to incorporate them into your decision-making.
Information-sensitivity

You might be information-sensitive if...

▶ You care about what people think of you.
▶ You care about what people think of others.
▶ You care enough about these to incorporate them into your decision-making.

In game theory and mechanism design, agents typically have preferences only over outcomes.
Information-sensitivity

You might be information-sensitive if...

▶ You care about what people think of you.
▶ You care about what people think of others.
▶ You care enough about these to incorporate them into your decision-making.

In game theory and mechanism design, agents typically have preferences only over outcomes.

But what if agents are information-sensitive, and also care about the private information revealed in the course of an interaction?
Information-sensitivity

You might be information-sensitive if...

- You care about what people think of you.
- You care about what people think of others.
- You care enough about these to incorporate them into your decision-making.

In game theory and mechanism design, agents typically have preferences only over outcomes.

But what if agents are information-sensitive, and also care about the private information revealed in the course of an interaction?

This could lead to entirely different strategic behavior...
Information-sensitivity: Related literature

Psychological game theory: Utilities of agents depend on everybody's actions and beliefs. (Geanakoplos et. al, 1989)

Social image: Agents care about how they are perceived by others. (Bernheim, 1994; Glazer and Konrad, 1996)

The focus of these areas is the study of various phenomena resulting from such preferences.

▶ For example, Bernheim (1994) shows how such preferences can lead to conformity in behavior.

Our focus: Designing mechanisms for such agents.
Information-sensitivity: Related literature

Psychological game theory: Utilities of agents depend on everybody’s actions and beliefs. (Geanakoplos et. al, 1989)
Information-sensitivity: Related literature

Psychological game theory: Utilities of agents depend on everybody’s actions and beliefs. (Geanakoplos et. al, 1989)

Social image: Agents care about how they are perceived by others. (Bernheim, 1994; Glazer and Konrad, 1996)
Information-sensitivity: Related literature

Psychological game theory: Utilities of agents depend on everybody’s actions and beliefs. (Geanakoplos et. al, 1989)

Social image: Agents care about how they are perceived by others. (Bernheim, 1994; Glazer and Konrad, 1996)

The focus of these areas is the study of various phenomena resulting from such preferences.

- For example, Bernheim (1994) shows how such preferences can lead to conformity in behavior.
Information-sensitivity: Related literature

Psychological game theory: Utilities of agents depend on everybody’s actions and beliefs. (Geanakoplos et. al, 1989)

Social image: Agents care about how they are perceived by others. (Bernheim, 1994; Glazer and Konrad, 1996)

The focus of these areas is the study of various phenomena resulting from such preferences.

- For example, Bernheim (1994) shows how such preferences can lead to conformity in behavior.

Our focus: Designing mechanisms for such agents.
Privacy in implementation

We'll focus on a particular type of information-sensitivity: A predilection for privacy. For example, an agent may wish to prevent the disclosure of private information at all costs; wish to prevent disclosure unless he is sufficiently compensated; prefer different outcomes depending on what and how much private information is revealed.

We'll focus on a particular setting: Full implementation with complete information.

- Full implementation: All equilibria of a mechanism should correspond to socially optimal outcomes.
- Complete information: Simplest informational setting – all agents know each other's private information.
Privacy in implementation

We’ll focus on a particular type of information-sensitivity: A predilection for privacy.
Privacy in implementation

We’ll focus on a particular type of information-sensitivity: A predilection for 
privacy. For example, an agent may
Privacy in implementation

We’ll focus on a particular type of information-sensitivity: A predilection for privacy. For example, an agent may

- wish to prevent the disclosure of private information at all costs;
Privacy in implementation

We’ll focus on a particular type of information-sensitivity: A predilection for privacy. For example, an agent may

- wish to prevent the disclosure of private information at all costs;
- wish to prevent disclosure unless he is sufficiently compensated;
Privacy in implementation

We’ll focus on a particular type of information-sensitivity: A predilection for privacy. For example, an agent may

- wish to prevent the disclosure of private information at all costs;
- wish to prevent disclosure unless he is sufficiently compensated;
- prefer different outcomes depending on what and how much private information is revealed.
Privacy in implementation

We’ll focus on a particular type of information-sensitivity: A predilection for privacy. For example, an agent may

- wish to prevent the disclosure of private information at all costs;
- wish to prevent disclosure unless he is sufficiently compensated;
- prefer different outcomes depending on what and how much private information is revealed.

We’ll focus on a particular setting: Full implementation with complete information.
Privacy in implementation

We’ll focus on a particular type of information-sensitivity: A predilection for privacy. For example, an agent may

► wish to prevent the disclosure of private information at all costs;
► wish to prevent disclosure unless he is sufficiently compensated;
► prefer different outcomes depending on what and how much private information is revealed.

We’ll focus on a particular setting: Full implementation with complete information.

► Full implementation: All equilibria of a mechanism should correspond to socially optimal outcomes.
Privacy in implementation

We’ll focus on a particular type of information-sensitivity: A predilection for privacy. For example, an agent may

- wish to prevent the disclosure of private information at all costs;
- wish to prevent disclosure unless he is sufficiently compensated;
- prefer different outcomes depending on what and how much private information is revealed.

We’ll focus on a particular setting: Full implementation with complete information.

- Full implementation: All equilibria of a mechanism should correspond to socially optimal outcomes.
- Complete information: Simplest informational setting – all agents know each other’s private information.
Privacy in implementation

The main aspect of privacy – secrecy: Agents have private information that they may want to keep secret (Posner, 1981).
Privacy in implementation

The main aspect of privacy – secrecy: Agents have private information that they may want to keep secret (Posner, 1981).

Internal privacy: Keep information secret from other agents.
Privacy in implementation

The main aspect of privacy – secrecy: Agents have private information that they may want to keep secret (Posner, 1981).

Internal privacy: Keep information secret from other agents.

External privacy: Keep information secret from the planner and other outside observers.
Privacy in implementation

The main aspect of privacy – secrecy: Agents have private information that they may want to keep secret (Posner, 1981).

Internal privacy: Keep information secret from other agents.

External privacy: Keep information secret from the planner and other outside observers.
Privacy in implementation

The main aspect of privacy – secrecy: Agents have private information that they may want to keep secret (Posner, 1981).

Internal privacy: Keep information secret from other agents.

External privacy: Keep information secret from the planner and other outside observers.

There are two ways in which privacy is relevant for implementation:
Privacy in implementation

The main aspect of privacy – secrecy: Agents have private information that they may want to keep secret (Posner, 1981).

Internal privacy: Keep information secret from other agents.

External privacy: Keep information secret from the planner and other outside observers.

There are two ways in which privacy is relevant for implementation:

(i) Equilibria should pertain to *information-sensitive* preferences, not just preferences over outcomes.
Privacy in implementation

The main aspect of privacy – secrecy: Agents have private information that they may want to keep secret (Posner, 1981).

Internal privacy: Keep information secret from other agents.

External privacy: Keep information secret from the planner and other outside observers.

There are two ways in which privacy is relevant for implementation:

(i) Equilibria should pertain to information-sensitive preferences, not just preferences over outcomes.

(ii) Some equilibrium should “protect” privacy by not revealing too much private information.
Example: Policy recommendation

A local government is debating the details of a new policy. Several investors from the area have private information that is relevant, and the government would like to get a policy recommendation. The government may be able to design a mechanism that will elicit private information. However, the investors might have serious reservations about revealing their private information to the government...

Hence, (i) equilibria should pertain to information-sensitive preferences, not just preferences over outcomes.
Example: Policy recommendation

A local government is debating the details of a new policy.
Example: Policy recommendation

A local government is debating the details of a new policy.
Several investors from the area have private information that is relevant, and the government would like to get a policy recommendation.
Example: Policy recommendation

A local government is debating the details of a new policy.

Several investors from the area have private information that is relevant, and the government would like to get a policy recommendation.

The government may be able to design a mechanism that will elicit private information.
Example: Policy recommendation

A local government is debating the details of a new policy. Several investors from the area have private information that is relevant, and the government would like to get a policy recommendation.

The government may be able to design a mechanism that will elicit private information.

However, the investors might have serious reservations about revealing their private information to the government...
Example: Policy recommendation

A local government is debating the details of a new policy. Several investors from the area have private information that is relevant, and the government would like to get a policy recommendation.

The government may be able to design a mechanism that will elicit private information.

However, the investors might have serious reservations about revealing their private information to the government...

Hence,

(i) Equilibria should pertain to information-sensitive preferences, not just preferences over outcomes.
Example: Board of directors

The board of directors of a company is voting on an issue, say the appointment of a new CEO. Eventually, the board will have to relay its decision to the shareholders and CEO. However, they may not want them to know who was in favor, who was against, and who abstained...

The only information they'd like to pass on is the outcome of the vote. Everything else should be kept private.

Hence, some equilibrium should "protect" privacy by not revealing too much private information. Similar situation arises in cabinet meetings, faculty meetings, arbitration hearings...
Example: Board of directors

The board of directors of a company is voting on an issue, say the appointment of a new CEO.
Example: Board of directors

The board of directors of a company is voting on an issue, say the appointment of a new CEO.

Eventually, the board will have to relay its decision to the shareholders and CEO.
Example: Board of directors

The board of directors of a company is voting on an issue, say the appointment of a new CEO.

Eventually, the board will have to relay its decision to the shareholders and CEO.

However, they may not want them to know who was in favor, who was against, and who abstained...
Example: Board of directors

The board of directors of a company is voting on an issue, say the appointment of a new CEO.

Eventually, the board will have to relay its decision to the shareholders and CEO.

However, they may not want them to know who was in favor, who was against, and who abstained...

The only information they’d like to pass on is the outcome of the vote. Everything else should be kept private.
Example: Board of directors

The board of directors of a company is voting on an issue, say the appointment of a new CEO.

Eventually, the board will have to relay its decision to the shareholders and CEO.

However, they may not want them to know who was in favor, who was against, and who abstained...

The only information they’d like to pass on is the outcome of the vote. Everything else should be kept private.

Hence,

(ii) Some equilibrium should “protect” privacy by not revealing too much private information.
Example: Board of directors

The board of directors of a company is voting on an issue, say the appointment of a new CEO.

Eventually, the board will have to relay its decision to the shareholders and CEO.

However, they may not want them to know who was in favor, who was against, and who abstained...

The only information they’d like to pass on is the outcome of the vote. Everything else should be kept private.

Hence,

(ii) Some equilibrium should “protect” privacy by not revealing too much private information.

Similar situation arises in cabinet meetings, faculty meetings, arbitration hearings...
In this paper:

▶ A model of preferences that may include a desire for privacy.
▶ Recall:
  (i) Equilibria should pertain to information-sensitive preferences.
  (ii) Some equilibrium should “protect” privacy.
▶ Without restrictions on preferences, (i) may be impossible.
▶ A couple of restrictions on preferences.
▶ Even with restrictions, (ii) may be impossible.
▶ If we relax the model and allow extensive-form mechanisms, privacy-protecting implementation becomes possible,
In this paper:

- A model of preferences that may include a desire for privacy.
In this paper:

- A model of preferences that may include a desire for privacy.

- Recall:
  
  (i) Equilibria should pertain to *information-sensitive* preferences.
  
  (ii) Some equilibrium should “protect” privacy.
In this paper:

- A model of preferences that may include a desire for privacy.

- Recall:
  1. Equilibria should pertain to *information-sensitive* preferences.
  2. Some equilibrium should “protect” privacy.

- Without restrictions on preferences, (i) may be impossible.
In this paper:

- A model of preferences that may include a desire for privacy.

- Recall:
  (i) Equilibria should pertain to information-sensitive preferences.
  (ii) Some equilibrium should “protect” privacy.

- Without restrictions on preferences, (i) may be impossible.
In this paper:

- A model of preferences that may include a desire for privacy.

- Recall:
  1. Equilibria should pertain to information-sensitive preferences.
  2. Some equilibrium should “protect” privacy.

- Without restrictions on preferences, (i) may be impossible 😞
  
  - A couple of restrictions on preferences.
In this paper:

- A model of preferences that may include a desire for privacy.

- Recall:
  1. Equilibria should pertain to *information-sensitive* preferences.
  2. Some equilibrium should “protect” privacy.

- Without restrictions on preferences, (i) may be impossible 😞
  - A couple of restrictions on preferences.

- Even with restrictions, (ii) may be impossible 😞
In this paper:

- A model of preferences that may include a desire for privacy.

- Recall:
  
  (i) Equilibria should pertain to *information-sensitive* preferences.
  
  (ii) Some equilibrium should “protect” privacy.

- Without restrictions on preferences, (i) may be impossible 😞
  
  - A couple of restrictions on preferences.

- Even with restrictions, (ii) may be impossible 😞

- If we relax the model and allow extensive-form mechanisms, privacy-protecting implementation becomes possible 😊
Main take-away point(s) from today’s talk

1. Standard tools and mechanisms for implementation can break down if agents care about privacy.
2. The (credible) threat of revealing agents’ information can aid the design of mechanisms in the presence of privacy concerns.
Main take-away point(s) from today’s talk

1. Standard tools and mechanisms for implementation can break down if agents care about privacy.
Main take-away point(s) from today’s talk

1. Standard tools and mechanisms for implementation can break down if agents care about privacy.

2. The (credible) threat of revealing agents’ information can aid the design of mechanisms in the presence of privacy concerns.
Why model privacy concerns exogenously?

The crux of our model is that agents’ preferences over privacy are given *exogenously*.
Why model privacy concerns exogenously?

The crux of our model is that agents’ preferences over privacy are given *exogenously*.

Why not model these concerns as arising endogenously?
Why model privacy concerns exogenously?

The crux of our model is that agents’ preferences over privacy are given *exogenously*.

Why not model these concerns as arising endogenously?

One interpretation is that these are agents’ *deep* preferences.
Why model privacy concerns exogenously?

The crux of our model is that agents’ preferences over privacy are given *exogenously*.

Why not model these concerns as arising endogenously?

One interpretation is that these are agents’ deep preferences.

- Examples: fear of embarrassment; basic individual or social value like liberty, fairness, and justice.
Why model privacy concerns exogenously?

The crux of our model is that agents’ preferences over privacy are given *exogenously*.

Why not model these concerns as arising endogenously?

One interpretation is that these are agents’ *deep* preferences.

- Examples: fear of embarrassment; basic individual or social value like liberty, fairness, and justice.

Another interpretation is that agents are participating in multiple interactions, this being just one of them.
Why model privacy concerns exogenously?

The crux of our model is that agents’ preferences over privacy are given *exogenously*.

Why not model these concerns as arising endogenously?

One interpretation is that these are agents’ deep preferences.

► Examples: fear of embarrassment; basic individual or social value like liberty, fairness, and justice.

Another interpretation is that agents are participating in multiple interactions, this being just one of them.

► The agents or the planner do not fully understand future interactions.
Why model privacy concerns exogenously?

The crux of our model is that agents’ preferences over privacy are given *exogenously*.

Why not model these concerns as arising endogenously?

One interpretation is that these are agents’ **deep** preferences.

▶ Examples: fear of embarrassment; basic individual or social value like liberty, fairness, and justice.

Another interpretation is that agents are participating in multiple interactions, this being just one of them.

▶ The agents or the planner do not fully understand future interactions.

▶ The planner may have control only over this particular interaction.
Why can’t we use standard tools?

Why not apply standard tools to an extended type space?

In standard models, utilities of agents depend on outcomes. With information-sensitivity, however, utilities also depend on the mechanism.

Standard mechanisms often require revealing your type, which may be incompatible with a desire for privacy.
Why can’t we use standard tools?

Why not apply standard tools to an extended type space?
Why can’t we use standard tools?

Why not apply standard tools to an extended type space?

In standard models, utilities of agents depend on outcomes. With information-sensitivity, however, utilities also depend on the mechanism.
Why can’t we use standard tools?

Why not apply standard tools to an extended type space?

In standard models, utilities of agents depend on outcomes. With information-sensitivity, however, utilities also depend on the mechanism.

Standard mechanisms often require revealing your type, which may be incompatible with a desire for privacy.
What about using cryptography to provide privacy?

There are two ways to incorporate crypto into strategic interactions:

▶ As a black box: Participants sit at computer terminals or there is a trusted mediator.

▶ Cryptographic communication is part of participants' strategies.

We will assume away the former approach.

The latter approach poses two difficulties in the context of this paper:

1. We want a full implementation.
2. We probably want some form of sequential rationality.

Finally, a minor point to note: It will be useful to threaten to reveal an agent's information.
What about using cryptography to provide privacy?

There are two ways to incorporate crypto into strategic interactions:

- As a black box: Participants sit at computer terminals or there is a trusted mediator.
- Cryptographic communication is part of participants' strategies.

We will assume away the former approach. The latter approach poses two difficulties in the context of this paper:

1. We want a full implementation.
2. We probably want some form of sequential rationality.

Finally, a minor point to note: It will be useful to threaten to reveal an agent's information.
What about using cryptography to provide privacy?

There are two ways to incorporate crypto into strategic interactions:

- As a black box: Participants sit at computer terminals or there is a trusted mediator.
What about using cryptography to provide privacy?

There are two ways to incorporate crypto into strategic interactions:

- As a black box: Participants sit at computer terminals or there is a trusted mediator.
- Cryptographic communication is part of participants’ strategies.
What about using cryptography to provide privacy?

There are two ways to incorporate crypto into strategic interactions:

▶ As a black box: Participants sit at computer terminals or there is a trusted mediator.

▶ Cryptographic communication is part of participants’ strategies.

We will assume away the former approach.
What about using cryptography to provide privacy?

There are two ways to incorporate crypto into strategic interactions:

- As a black box: Participants sit at computer terminals or there is a trusted mediator.
- Cryptographic communication is part of participants’ strategies.

We will assume away the former approach.

The latter approach poses two difficulties in the context of this paper:
What about using cryptography to provide privacy?

There are two ways to incorporate crypto into strategic interactions:

- As a black box: Participants sit at computer terminals or there is a trusted mediator.
- Cryptographic communication is part of participants’ strategies.

We will assume away the former approach.

The latter approach poses two difficulties in the context of this paper:

1. We want a full implementation.
What about using cryptography to provide privacy?

There are two ways to incorporate crypto into strategic interactions:

- As a black box: Participants sit at computer terminals or there is a trusted mediator.
- Cryptographic communication is part of participants’ strategies.

We will assume away the former approach.

The latter approach poses two difficulties in the context of this paper:

1. We want a full implementation.
2. We probably want some form of sequential rationality.

Finally, a minor point to note: It will be useful to threaten to reveal an agent's information.
What about using cryptography to provide privacy?

There are two ways to incorporate crypto into strategic interactions:

- As a black box: Participants sit at computer terminals or there is a trusted mediator.
- Cryptographic communication is part of participants’ strategies.

We will assume away the former approach.

The latter approach poses two difficulties in the context of this paper:

1. We want a full implementation.
2. We probably want some form of sequential rationality.

Finally, a minor point to note: It will be useful to threaten to reveal an agent’s information.
Related literature – economics

Implementation:

Implementation with nonstandard preferences:

Dynamic mechanism design:
Related literature – economics

**Implementation:** Hurwicz (1972), Maskin (1999), Moore and Repullo (1988).
Related literature – economics

**Implementation:** Hurwicz (1972), Maskin (1999), Moore and Repullo (1988).

Related literature – economics

**Implementation:** Hurwicz (1972), Maskin (1999), Moore and Repullo (1988).


**Dynamic mechanism design:** Calzolari and Pavan (2006).
Related literature – computer science

Differential privacy as a tool for mechanism design: McSherry and Talwar (2007), Nissim et al. (2012)

Related literature – computer science

Differential privacy as a tool for mechanism design: McSherry and Talwar (2007), Nissim et al. (2012)
Related literature – computer science

**Differential privacy as a tool for mechanism design:** McSherry and Talwar (2007), Nissim et al. (2012)

**Explicit utility for privacy:** Miltersen et al. (2009), Xiao (2013), Ghosh and Roth (2011), Nissim et al. (2012), Chen et al. (2013).
Outline

Model and Results via an Example

Extensive-Form Mechanisms
An example

There are 2 possible outcomes \{a,b\} and 2 states \{t_a, t_b\}, specifying the "type" of an agent.

Type $t_a$ prefers $a$ over $b$ when the state is fully revealed: $(a, \{t_a\}) \succ_{t_a} (b, \{t_a\})$.

We call these the agent's intrinsic preferences.

Type $t_a$ has the following information-sensitive preferences: $(a, \{t_a, t_b\}) \succ_{t_a} (a, \{t_a\})$ and $(a, \{t_b\}) \succ_{t_a} (a, \{t_a\})$. 
An example

There are 2 possible outcomes \{a, b\} and 2 states \{t^a, t^b\}, specifying the “type” of an agent.
An example

There are 2 possible outcomes \( \{a, b\} \) and 2 states \( \{t^a, t^b\} \), specifying the “type” of an agent.

Type \( t^a \) prefers \( a \) over \( b \)
An example

There are 2 possible outcomes \{a, b\} and 2 states \{t^a, t^b\}, specifying the “type” of an agent.

Type \(t^a\) prefers \(a\) over \(b\)

- when the state is fully revealed:

\[(a, \{t^a\}) \succeq_{t^a} (b, \{t^a\})\]
An example

There are 2 possible outcomes \{a, b\} and 2 states \{t^a, t^b\}, specifying the “type” of an agent.

Type \(t^a\) prefers \(a\) over \(b\)

- when the state is fully revealed:

\[
(a, \{t^a\}) \succ_{t^a} (b, \{t^a\}).
\]

- We call these the agent’s intrinsic preferences.
An example

There are 2 possible outcomes \( \{a, b\} \) and 2 states \( \{t^a, t^b\} \), specifying the “type” of an agent.

Type \( t^a \) prefers \( a \) over \( b \)

- when the state is fully revealed:

\[ (a, \{t^a\}) \succ_{t^a} (b, \{t^a\}) . \]

- We call these the agent’s *intrinsic* preferences.

Type \( t^a \) has the following *information-sensitive* preferences:

\[ (a, \{t^a, t^b\}) \succ_{t^a} (a, \{t^a\}) \text{ and } (a, \{t^b\}) \succ_{t^a} (a, \{t^a\}) . \]
The model: Preferences

There is a set $\mathcal{O}$ of possible outcomes, a set $\Theta$ of outcome-states, and a set $\Psi$ of privacy-states.

▶ We call these $i$’s intrinsic preferences, and think of them as his preferences when the outcome-state $\theta$ is revealed.

Our setup: An outcome-state $\theta$, extended by a privacy-state $\psi$, yields the state $\theta \psi$.

▶ In every state $\theta \psi$, every agent has preferences over $\mathcal{O} \times 2$.

▶ The second element here is the set of outcome-states that an outside observer believes are possible.
The model: Preferences

There is a set \( \mathcal{O} \) of possible outcomes, a set \( \Theta \) of outcome-states, and a set \( \Psi \) of privacy-states.
The model: Preferences

There is a set $\mathcal{O}$ of possible outcomes, a set $\Theta$ of outcome-states, and a set $\Psi$ of privacy-states.

The usual setup: In every outcome-state $\theta \in \Theta$, every agent $i$ has preferences over $\mathcal{O}$.
The model: Preferences

There is a set $\mathcal{O}$ of possible outcomes, a set $\Theta$ of outcome-states, and a set $\Psi$ of privacy-states.

The usual setup: In every outcome-state $\theta \in \Theta$, every agent $i$ has preferences over $\mathcal{O}$.

▶ We call these $i$’s intrinsic preferences, and think of them as his preferences when the outcome-state $\theta$ is revealed.
The model: Preferences

There is a set $\mathcal{O}$ of possible outcomes, a set $\Theta$ of outcome-states, and a set $\Psi$ of privacy-states.

The usual setup: In every outcome-state $\theta \in \Theta$, every agent $i$ has preferences over $\mathcal{O}$.

- We call these $i$’s intrinsic preferences, and think of them as his preferences when the outcome-state $\theta$ is revealed.

Our setup: An outcome-state $\theta$, extended by a privacy-state $\psi$, yields the state $\theta^\psi$. 
The model: Preferences

There is a set $\mathcal{O}$ of possible outcomes, a set $\Theta$ of outcome-states, and a set $\Psi$ of privacy-states.

The usual setup: In every outcome-state $\theta \in \Theta$, every agent $i$ has preferences over $\mathcal{O}$.

- We call these $i$’s intrinsic preferences, and think of them as his preferences when the outcome-state $\theta$ is revealed.

Our setup: An outcome-state $\theta$, extended by a privacy-state $\psi$, yields the state $\theta\psi$.

- In every state $\theta\psi$, every agent has preferences over $\mathcal{O} \times 2^{\Theta}$.
The model: Preferences

There is a set $\mathcal{O}$ of possible outcomes, a set $\Theta$ of outcome-states, and a set $\Psi$ of privacy-states.

The usual setup: In every outcome-state $\theta \in \Theta$, every agent $i$ has preferences over $\mathcal{O}$.

- We call these $i$’s intrinsic preferences, and think of them as his preferences when the outcome-state $\theta$ is revealed.

Our setup: An outcome-state $\theta$, extended by a privacy-state $\psi$, yields the state $\theta^\psi$.

- In every state $\theta^\psi$, every agent has preferences over $\mathcal{O} \times 2^\Theta$.
- The second element here is the set of outcome-states that an outside observer believes are possible.
The model: Preferences

- In every state $\theta^\psi$, every agent has preferences over $\mathcal{O} \times 2^\Theta$.
- The second element here is the set of outcome-states that an outside observer believes are possible.
The model: Preferences

- In every state $\theta^\psi$, every agent has preferences over $\mathcal{O} \times 2^\Theta$.
- The second element here is the set of outcome-states that an outside observer believes are possible.

If $\psi, \phi \in \Psi$, then in states $\theta^\psi$ and $\theta^\phi$ agents have the same intrinsic preferences, but possibly different information-sensitive preferences.
The model: Preferences

- In every state $\theta^\psi$, every agent has preferences over $O \times 2^\Theta$.
- The second element here is the set of outcome-states that an outside observer believes are possible.

If $\psi, \phi \in \Psi$, then in states $\theta^\psi$ and $\theta^\phi$ agents have the same *intrinsic* preferences, but possibly different information-sensitive preferences.

We will (mostly) restrict ourselves to the case $|\Psi| = 1$. 
The model: Preferences

- In every state $\theta^\psi$, every agent has preferences over $\mathcal{O} \times 2^\Theta$.
- The second element here is the set of outcome-states that an outside observer believes are possible.

If $\psi, \phi \in \Psi$, then in states $\theta^\psi$ and $\theta^\phi$ agents have the same intrinsic preferences, but possibly different information-sensitive preferences.

We will (mostly) restrict ourselves to the case $|\Psi| = 1$.

- This is the simplest informational setting: Preferences for privacy are common knowledge, and only intrinsic preferences are unknown to the planner.
An example

There are 3 agents and outcomes \( \Omega = \{a, b\} \).

The possible states are \( \theta \in \{t_a, t_b\} \times \{t_a, t_b\} \times \{t_a, t_b\} \), specifying the "type" of each agent.

Agent 1 of type \( t_a \) prefers \( a \) over \( b \) when his type is fully revealed:

\[
\text{For any } S \subseteq \{t_a, t_b\} \times \{t_a, t_b\}, (a, \{t_a\} \times S) \succ_{t_a} (b, \{t_a\} \times S).
\]

Agent 1 of type \( t_a \) has the following information-sensitive preferences:

\[
(a, \{t_a, t_b\} \times S) \succ_{t_a} (a, \{t_a\} \times S) \text{ and } (a, \{t_b\} \times S) \succ_{t_a} (a, \{t_a\} \times S).
\]
An example

There are 3 agents and outcomes $\mathcal{O} = \{a, b\}$. 
An example

There are 3 agents and outcomes $\mathcal{O} = \{a, b\}$.

The possible states are $\theta \in \{t^a, t^b\} \times \{t^a, t^b\} \times \{t^a, t^b\}$, specifying the “type” of each agent.
An example

There are 3 agents and outcomes \( \mathcal{O} = \{a, b\} \).

The possible states are \( \theta \in \{t^a, t^b\} \times \{t^a, t^b\} \times \{t^a, t^b\} \), specifying the “type” of each agent.

Agent 1 of type \( t^a \) prefers \( a \) over \( b \) when his type is fully revealed:
An example

There are 3 agents and outcomes $\mathcal{O} = \{a, b\}$.

The possible states are $\theta \in \{t^a, t^b\} \times \{t^a, t^b\} \times \{t^a, t^b\}$, specifying the “type” of each agent.

Agent 1 of type $t^a$ prefers $a$ over $b$ when his type is fully revealed:

- For any $S \subseteq \{t^a, t^b\} \times \{t^a, t^b\}$,
An example

There are 3 agents and outcomes $\mathcal{O} = \{a, b\}$.

The possible states are $\theta \in \{t^a, t^b\} \times \{t^a, t^b\} \times \{t^a, t^b\}$, specifying the “type” of each agent.

Agent 1 of type $t^a$ prefers $a$ over $b$ when his type is fully revealed:

- For any $S \subseteq \{t^a, t^b\} \times \{t^a, t^b\}$,

\[(a, \{t^a\} \times S) \succ_{t^a} (b, \{t^a\} \times S).\]
An example

There are 3 agents and outcomes $\mathcal{O} = \{a, b\}$.

The possible states are $\theta \in \{t^a, t^b\} \times \{t^a, t^b\} \times \{t^a, t^b\}$, specifying the “type” of each agent.

Agent 1 of type $t^a$ prefers $a$ over $b$ when his type is fully revealed:

- For any $S \subseteq \{t^a, t^b\} \times \{t^a, t^b\}$,

  $$(a, \{t^a\} \times S) \succ_{t^a} (b, \{t^a\} \times S).$$

Agent 1 of type $t^a$ has the following information-sensitive preferences:

$$(a, \{t^a, t^b\} \times S) \succ_{t^a} (a, \{t^a\} \times S) \text{ and } (a, \{t^b\} \times S) \succ_{t^a} (a, \{t^a\} \times S).$$
An example

Suppose we wish to implement the majority function:

▶ The outcome that is intrinsically preferred by the majority of the agents.

We want to design a mechanism that, in equilibrium, yields the majority outcome.

But what is an equilibrium?
An example

Suppose we wish to implement the majority function:
An example

Suppose we wish to implement the majority function:

- The outcome that is *intrinsically* preferred by the majority of the agents.
An example

Suppose we wish to implement the majority function:

- The outcome that is *intrinsically* preferred by the majority of the agents.

We want to design a mechanism that, in equilibrium, yields the majority outcome.
An example

Suppose we wish to implement the majority function:

- The outcome that is *intrinsically* preferred by the majority of the agents.

We want to design a mechanism that, in equilibrium, yields the majority outcome.

But what is an equilibrium?
The model: Equilibrium

The usual setup: Each action profile corresponds to an outcome.
The model: Equilibrium

The usual setup: Each action profile corresponds to an outcome.

Our setup: Each action profile corresponds to a pair – an outcome $a$ and a set of possible states $S$. 
The model: Equilibrium

The usual setup: Each action profile corresponds to an outcome.

Our setup: Each action profile corresponds to a pair – an outcome $a$ and a set of possible states $S$.

An equilibrium is a strategy profile in which no agent can unilaterally deviate to obtain more favorable pair $(b, T)$.
The model: Equilibrium

The usual setup: Each action profile corresponds to an outcome.

Our setup: Each action profile corresponds to a pair – an outcome \( a \) and a set of possible states \( S \).

An equilibrium is a strategy profile in which no agent can unilaterally deviate to obtain more favorable pair \((b, T)\).

But where do the sets of possible states \( S \) and \( T \) come from?
The model: Equilibrium

The usual setup: Each action profile corresponds to an outcome.

Our setup: Each action profile corresponds to a pair – an outcome $a$ and a set of possible states $S$.

An equilibrium is a strategy profile in which no agent can unilaterally deviate to obtain more favorable pair $(b, T)$.

But where do the sets of possible states $S$ and $T$ come from?

Suppose that in some mechanism agents play a strategy profile $s$, and some action profile is realized.
The model: Equilibrium

The usual setup: Each action profile corresponds to an outcome.

Our setup: Each action profile corresponds to a pair – an outcome $a$ and a set of possible states $S$.

An equilibrium is a strategy profile in which no agent can unilaterally deviate to obtain more favorable pair $(b, T)$.

But where do the sets of possible states $S$ and $T$ come from?

Suppose that in some mechanism agents play a strategy profile $s$, and some action profile is realized.

The set of possible states at this action profile is

$$\{\theta : s(\theta) \text{ leads to the realized action profile}\}.$$
The model: Equilibrium

The set of possible states at this action profile is

\[ \{ \theta : s(\theta) \text{ leads to the realized action profile} \} . \]
The model: Equilibrium

The set of possible states at this action profile is

$$\{ \theta : s(\theta) \text{ leads to the realized action profile} \}.$$

What is the set of possible states at action profiles that cannot be reached by the strategy profile $s$?
The model: Equilibrium

The set of possible states at this action profile is

\[ \{ \theta : s(\theta) \text{ leads to the realized action profile} \} \, . \]

What is the set of possible states at action profiles that cannot be reached by the strategy profile \( s \)?

We will make one restriction on such “off-equilibrium beliefs” – one-deviation consistency:
The model: Equilibrium

The set of possible states at this action profile is

\[ \{ \theta : s(\theta) \text{ leads to the realized action profile} \} . \]

What is the set of possible states at action profiles that cannot be reached by the strategy profile \( s \)?

We will make one restriction on such “off-equilibrium beliefs” – one-deviation consistency:

- If there is no \( \theta \) such that \( s(\theta) \) leads to an action profile \( A \), but
The model: Equilibrium

The set of possible states at this action profile is

\[ \{ \theta : s(\theta) \text{ leads to the realized action profile} \}. \]

What is the set of possible states at action profiles that cannot be reached by the strategy profile \( s \)?

We will make one restriction on such “off-equilibrium beliefs” – one-deviation consistency:

- If there is no \( \theta \) such that \( s(\theta) \) leads to an action profile \( A \), but
- there is an \( i, s'_i, \theta \) such that \( (s_{-i} \circ s'_i)(\theta) \) does lead to \( A \), then
The model: Equilibrium

The set of possible states at this action profile is

\[ \{ \theta : s(\theta) \text{ leads to the realized action profile} \} . \]

What is the set of possible states at action profiles that cannot be reached by the strategy profile \( s \)?

We will make one restriction on such “off-equilibrium beliefs” – one-deviation consistency:

- If there is no \( \theta \) such that \( s(\theta) \) leads to an action profile \( A \), but
- there is an \( i, s'_i, \theta \) such that \( (s_{-i} \circ s'_i)(\theta) \) does lead to \( A \), then
- the set of possible states at \( A \) is a nonempty subset of
The model: Equilibrium

The set of possible states at this action profile is

\[ \{ \theta : s(\theta) \text{ leads to the realized action profile} \} . \]

What is the set of possible states at action profiles that cannot be reached by the strategy profile \( s \)?

We will make one restriction on such “off-equilibrium beliefs” – one-deviation consistency:

- If there is no \( \theta \) such that \( s(\theta) \) leads to an action profile \( A \), but
- there is an \( i, s'_i, \theta \) such that \( (s_{-i} \circ s'_i)(\theta) \) does lead to \( A \), then
- the set of possible states at \( A \) is a nonempty subset of

\[ \{ \theta : (s_{-i} \circ s'_i)(\theta) \text{ leads to } A \text{ for some } i \text{ and } s'_i \} . \]
Back to our example

Recall: Each of 3 agents is of type $t^a$ or $t^b$, where
Back to our example

Recall: Each of 3 agents is of type $t^a$ or $t^b$, where

- Type $t^a$ intrinsically prefers $a$ over $b$.
- Given outcome $a$, type $t^a$ prefers both privacy and deception over full revelation of his type.
Back to our example

Recall: Each of 3 agents is of type $t^a$ or $t^b$, where

- Type $t^a$ intrinsically prefers $a$ over $b$.
- Given outcome $a$, type $t^a$ prefers both privacy and deception over full revelation of his type.

Mechanism 1:

1. Each agent votes for the outcome that he intrinsically prefers.
2. Implement the majority.
Back to our example

Recall: Each of 3 agents is of type $t^a$ or $t^b$, where

- Type $t^a$ intrinsically prefers $a$ over $b$.
- Given outcome $a$, type $t^a$ prefers both privacy and deception over full revelation of his type.

Mechanism 1:

1. Each agent votes for the outcome that he intrinsically prefers.
2. Implement the majority.

Is this an equilibrium?
Back to our example

Recall: Each of 3 agents is of type $t^a$ or $t^b$, where

- Type $t^a$ intrinsically prefers $a$ over $b$.
- Given outcome $a$, type $t^a$ prefers both privacy and deception over full revelation of his type.

Mechanism 1:

1. Each agent votes for the outcome that he intrinsically prefers.
2. Implement the majority.

Is this an equilibrium?

No! Suppose all agents are of type $t^a$, which leads to $(a, \{(t^a, t^a, t^a)\})$. 
Back to our example

Recall: Each of 3 agents is of type \( t^a \) or \( t^b \), where

- Type \( t^a \) intrinsically prefers \( a \) over \( b \).
- Given outcome \( a \), type \( t^a \) prefers both privacy and deception over full revelation of his type.

Mechanism 1:

1. Each agent votes for the outcome that he intrinsically prefers.
2. Implement the majority.

Is this an equilibrium?

No! Suppose all agents are of type \( t^a \), which leads to \((a, \{(t^a, t^a, t^a)\})\).
But if agent 1 deviates to \( b \) we get \((a, \{(t^b, t^a, t^a)\})\), which he prefers.
An example

Complete information: Agents know each other's types.

Mechanism 1:
1. If majority is $a$, all agents vote for $a$.
2. If majority is $b$, all agents vote for $b$.
3. Implement the majority.

Is this an equilibrium? Yes, under arguably reasonable "off-equilibrium beliefs."
An example

Complete information: Agents know each other’s types.
An example

Complete information: Agents know each other’s types.

Mechanism 1’:

1. If majority is $a$, all agents vote for $a$.
2. If majority is $b$, all agents vote for $b$.
3. Implement the majority.

Is this an equilibrium? Yes, under arguably reasonable “off-equilibrium beliefs.”
An example

Complete information: Agents know each other’s types.

Mechanism 1’:

1. If majority is $a$, all agents vote for $a$.
2. If majority is $b$, all agents vote for $b$.
3. Implement the majority.

Is this an equilibrium?
An example

Complete information: Agents know each other’s types.

Mechanism 1’:

1. If majority is $a$, all agents vote for $a$.
2. If majority is $b$, all agents vote for $b$.
3. Implement the majority.

Is this an equilibrium?

Yes
An example

Complete information: Agents know each other’s types.

Mechanism 1’:

1. If majority is $a$, all agents vote for $a$.  
2. If majority is $b$, all agents vote for $b$.  
3. Implement the majority.

Is this an equilibrium?  
Yes, under arguably reasonable “off-equilibrium beliefs.”
An example

Complete information: Agents know each other’s types.

Mechanism 1’:

1. If majority is $a$, all agents vote for $a$.
2. If majority is $b$, all agents vote for $b$.
3. Implement the majority.

Is this an equilibrium?

Yes, under arguably reasonable “off-equilibrium beliefs.”

- Namely, that a deviation does not convey any new information.
An example

Mechanism 1’:

1. If majority is $a$, all agents vote for $a$.
2. If majority is $b$, all agents vote for $b$.
3. Implement the majority.
An example

Mechanism 1’:

1. If majority is \( a \), all agents vote for \( a \).
2. If majority is \( b \), all agents vote for \( b \).
3. Implement the majority.

Observe: In equilibrium, only information about outcome is conveyed.
An example

Mechanism 1’:

1. If majority is $a$, all agents vote for $a$.
2. If majority is $b$, all agents vote for $b$.
3. Implement the majority.

Observe: In equilibrium, only information about outcome is conveyed.

Problem: It is also an equilibrium for agents to always vote for $a$. 
An example

Mechanism 1’:

1. If majority is $a$, all agents vote for $a$.
2. If majority is $b$, all agents vote for $b$.
3. Implement the majority.

Observe: In equilibrium, only information about outcome is conveyed.

Problem: It is also an equilibrium for agents to always vote for $a$. It’s even possible that this equilibrium dominates the honest one.
We want *full* implementation: *All* equilibria yield “correct” outcome.
We want full implementation: All equilibria yield “correct” outcome. The following is the standard mechanism for achieving this.
We want *full* implementation: *All* equilibria yield “correct” outcome. The following is the standard mechanism for achieving this.

**Mechanism 2:**

1. Each agent submits a vector of *all* agents’ types.
2. If all agents agree, implement the majority.
3. If one agent disagrees, ignore him, unless the majority according to his report is different from his preferred outcome according to the others’ reports. In that case implement the majority according to the deviator.
4. Otherwise, play integer/modulo game.
We want full implementation: All equilibria yield “correct” outcome. The following is the standard mechanism for achieving this.

Mechanism 2:

1. Each agent submits a vector of all agents’ types.
2. If all agents agree, implement the majority.
3. If one agent disagrees, ignore him, unless the majority according to his report is different from his preferred outcome according to the others’ reports. In that case implement the majority according to the deviator.
4. Otherwise, play integer/modulo game.

With standard preferences this mechanism is a full implementation, since majority with strict preferences is Maskin monotonic.
An example

But what about our information-sensitive preferences?
An example

But what about our information-sensitive preferences?

Honesty is an equilibrium (when beliefs are one-deviation consistent).

But what about $(b, \{t_a, t_b\} \times S)$?
An example

But what about our information-sensitive preferences?

Honesty is an equilibrium (when beliefs are one-deviation consistent).

But do all equilibria lead to majority outcome under these information-sensitive preferences?
An example

But what about our information-sensitive preferences?

Honesty is an equilibrium (when beliefs are one-deviation consistent).

But do all equilibria lead to majority outcome under these information-sensitive preferences?

It depends:
An example

But what about our information-sensitive preferences?

Honesty is an equilibrium (when beliefs are one-deviation consistent).

But do all equilibria lead to majority outcome under these information-sensitive preferences?

It depends: We haven’t fully specified the information-sensitive preferences of the agents...
An example

But what about our information-sensitive preferences?

Honesty is an equilibrium (when beliefs are one-deviation consistent).

But do all equilibria lead to majority outcome under these information-sensitive preferences?

It depends: We haven’t fully specified the information-sensitive preferences of the agents...

Recall

\[(a, \{t^a, t^b\} \times S) \succ_{t_a} (a, \{t^a\} \times S) \quad \text{and} \quad (a, \{t^b\} \times S) \succ_{t_a} (a, \{t^a\} \times S).\]
An example

But what about our information-sensitive preferences?

Honesty is an equilibrium (when beliefs are one-deviation consistent).

But do all equilibria lead to majority outcome under these information-sensitive preferences?

It depends: We haven’t fully specified the information-sensitive preferences of the agents...

Recall

\[(a, \{t^a, t^b\} \times S) \succeq_{t_a} (a, \{t^a\} \times S) \quad \text{and} \quad (a, \{t^b\} \times S) \succ_{t_a} (a, \{t^a\} \times S).\]

But what about

\[(b, \{t^a, t^b\} \times S) \sim (a, \{t^a\} \times S)?\]
An example

But what about

$$(b, \{t^a, t^b\} \times S') \sim (a, \{t^a\} \times S')?$$
An example

If

\[(a, \{t^a, t^b\} \times S) \succ_t (b, \{t^a, t^b\} \times S) \succ_t (a, \{t^a\} \times S)\]
An example

If
\[(a, \{t^a, t^b\} \times S) \succ \tau_a (b, \{t^a, t^b\} \times S) \succ \tau_a (a, \{t^a\} \times S)\]

then all agents always submitting \((t^b, t^b, t^b)\) can also be an equilibrium:
An example

If
\[(a, \{t^a, t^b\} \times S) \succ_{t^a} (b, \{t^a, t^b\} \times S) \succ_{t^a} (a, \{t^a\} \times S)\]

then all agents always submitting \((t^b, t^b, t^b)\) can also be an equilibrium:

- On a deviation, fix the set of possible types so that deviator is believed to be of type \(t^a\). Then neither type will deviate.
An example

If

\[(a, \{t^a, t^b\} \times S) \succ_t (b, \{t^a, t^b\} \times S) \succ_t (a, \{t^a\} \times S)\]

then all agents always submitting \((t^b, t^b, t^b)\) can also be an equilibrium:

- On a deviation, fix the set of possible types so that deviator is believed to be of type \(t^a\). Then neither type will deviate.

In fact, for every set \(\mathcal{O}\) and intrinsic preferences \(\Theta\) there exists \(\Psi\) such that there is no full implementation of any non-constant SCF \(f : \Theta \mapsto \mathcal{O}\).
An example

Sufficient for full implementation: Lexicographic preferences.
An example

Sufficient for full implementation: Lexicographic preferences.

- Roughly, agents are willing to reveal their information to obtain an outcome they strictly prefer (intrinsically).
An example

Sufficient for full implementation: Lexicographic preferences.

- Roughly, agents are willing to reveal their information to obtain an outcome they strictly prefer (intrinsically).

A weaker condition: Minimal willingness to reveal (MWR).
An example

Sufficient for full implementation: Lexicographic preferences.

▶ Roughly, agents are willing to reveal their information to obtain an outcome they strictly prefer (intrinsically).

A weaker condition: Minimal willingness to reveal (MWR).

▶ Roughly, agents are willing to reveal their information to obtain the outcome they prefer most (intrinsically).
An example

Recall mechanism 2:

1. Each agent submits a vector of \textit{all} agents’ types.
2. ...
3. ...
4. ...
An example

Recall mechanism 2:

1. Each agent submits a vector of all agents’ types.
2. ...
3. ...
4. ...

Problem: In mechanism 2, all private information is revealed!
An example

Recall mechanism 2:

1. Each agent submits a vector of *all* agents’ types.
2. ...
3. ...
4. ...

Problem: In mechanism 2, all private information is revealed!

Question: Do there exist mechanisms for full implementation that do not reveal information beyond the outcome (as in mechanism 1’)?
An example

Recall mechanism 2:

1. Each agent submits a vector of *all* agents’ types.
2. ...
3. ...
4. ...

Problem: In mechanism 2, all private information is revealed!

Question: Do there exist mechanisms for full implementation that do not reveal information beyond the outcome (as in mechanism 1’)?

No!
An example

Recall mechanism 2:

1. Each agent submits a vector of all agents’ types.
2. ...
3. ...
4. ...

Problem: In mechanism 2, all private information is revealed!

Question: Do there exist mechanisms for full implementation that do not reveal information beyond the outcome (as in mechanism 1’)?

No!

- In fact, only the dictatorship function has a privacy-protecting implementation (under some weak conditions).
An example

Recall mechanism 2:

1. Each agent submits a vector of all agents’ types.
2. ...
3. ...
4. ...

Problem: In mechanism 2, all private information is revealed!

Question: Do there exist mechanisms for full implementation that do not reveal information beyond the outcome (as in mechanism 1’)?

No, unless we allow extensive-form mechanisms, in which case the answer is yes!
Outline

Model and Results via an Example

Extensive-Form Mechanisms
Full implementation – extensive-form mechanisms

Moore and Repullo (1988), Abreu and Sen (1990), and Vartiainen (2007) studied the extent of implementation in subgame perfect equilibrium (SPE).

Theorem
If \( f \) is implementable in SPE and preferences are lexicographic, then there is a privacy-protecting implementation of \( f \).

Theorem
If preferences satisfy MWR, then there is a privacy-protecting implementation of \( f \) under a somewhat stronger condition than above.
Full implementation – extensive-form mechanisms

Moore and Repullo (1988), Abreu and Sen (1990), and Vartiainen (2007) studied the extent of implementation in subgame perfect equilibrium (SPE).

Much less restrictive than Nash implementation.
Full implementation – extensive-form mechanisms

Moore and Repullo (1988), Abreu and Sen (1990), and Vartiainen (2007) studied the extent of implementation in subgame perfect equilibrium (SPE).

Much less restrictive than Nash implementation.

Theorem

If $f$ is implementable in SPE and preferences are lexicographic, then there is a privacy-protecting implementation of $f$. 
Moore and Repullo (1988), Abreu and Sen (1990), and Vartiainen (2007) studied the extent of implementation in subgame perfect equilibrium (SPE).

Much less restrictive than Nash implementation.

**Theorem**

If $f$ is implementable in SPE and preferences are lexicographic, then there is a privacy-protecting implementation of $f$.

**Theorem**

If preferences satisfy MWR, then there is a privacy-protecting implementation of $f$ under a somewhat stronger condition than above.
Full implementation – direct mechanisms

An extensive-form mechanism is direct if it has the following form:

1. A designated leader agent submits an outcome $a$ from $O$.
2. All other agents either object or not.
3. If no agent objects, then $a$ is implemented.
4. If some agent does object, then we continue (contingency plan).

Theorem

If $f$ is implementable in SPE, preferences are lexicographic, and $f$ satisfies an additional mild condition, then there is a direct privacy-protecting implementation of $f$.

The additional mild condition: For every outcome that is not socially-optimal, there exists an agent who strictly prefers the socially-optimal outcome.
Full implementation – direct mechanisms

An extensive-form mechanism is *direct* if it has the following form:

1. A designated leader agent submits an outcome \( a \) from \( O \).
2. All other agents either object or not.
3. If no agent objects, then \( a \) is implemented.
4. If some agent does object, then we continue (contingency plan).

**Theorem**

If \( f \) is implementable in SPE, preferences are lexicographic, and \( f \) satisfies an additional mild condition, then there is a direct privacy-protecting implementation of \( f \).

The additional mild condition: For every outcome that is not socially-optimal, there exists an agent who strictly prefers the socially-optimal outcome.
Full implementation – direct mechanisms

An extensive-form mechanism is *direct* if it has the following form:

1. A designated *leader* agent submits an outcome $a$ from $\mathcal{O}$. 

Theorem

If $f$ is implementable in SPE, preferences are lexicographic, and $f$ satisfies an additional mild condition, then there is a *direct* privacy-protecting implementation of $f$.

The additional mild condition: For every outcome that is not socially-optimal, there exists an agent who strictly prefers the socially-optimal outcome.
Full implementation – direct mechanisms

An extensive-form mechanism is *direct* if it has the following form:

1. A designated *leader* agent submits an outcome $a$ from $O$.
2. All other agents either object or not.

Theorem

If $f$ is implementable in SPE, preferences are lexicographic, and $f$ satisfies an additional mild condition, then there is a *direct* privacy-protecting implementation of $f$.

The additional mild condition: For every outcome that is not socially-optimal, there exists an agent who strictly prefers the socially-optimal outcome.
Full implementation – direct mechanisms

An extensive-form mechanism is *direct* if it has the following form:

1. A designated *leader* agent submits an outcome $a$ from $O$.
2. All other agents either object or not.
3. If no agent objects, then $a$ is implemented.
Full implementation – direct mechanisms

An extensive-form mechanism is *direct* if it has the following form:

1. A designated *leader* agent submits an outcome $a$ from $\mathcal{O}$.
2. All other agents either object or not.
3. If no agent objects, then $a$ is implemented.
4. If some agent does object, then we continue (contingency plan).

Theorem

If $f$ is implementable in SPE, preferences are lexicographic, and $f$ satisfies an additional mild condition, then there is a direct privacy-protecting implementation of $f$.

The additional mild condition: For every outcome that is not socially-optimal, there exists an agent who strictly prefers the socially-optimal outcome.
Full implementation – direct mechanisms

An extensive-form mechanism is *direct* if it has the following form:

1. A designated *leader* agent submits an outcome $a$ from $O$.
2. All other agents either object or not.
3. If no agent objects, then $a$ is implemented.
4. If some agent does object, then we continue (contingency plan).

**Theorem**

If $f$ is implementable in SPE, preferences are lexicographic, and $f$ satisfies an additional mild condition, then there is a *direct* privacy-protecting implementation of $f$. 
Full implementation – direct mechanisms

An extensive-form mechanism is \textit{direct} if it has the following form:

1. A designated \textit{leader} agent submits an outcome \( a \) from \( \mathcal{O} \).
2. All other agents either object or not.
3. If no agent objects, then \( a \) is implemented.
4. If some agent does object, then we continue (contingency plan).

\textbf{Theorem}

If \( f \) is implementable in SPE, preferences are lexicographic, and \( f \) satisfies an additional mild condition, then there is a \textit{direct} privacy-protecting implementation of \( f \).

The additional mild condition: For every outcome that is \textit{not} socially-optimal, there exists an agent who strictly prefers the socially-optimal outcome.
Theorem

If $f$ is implementable in SPE, preferences are lexicographic, and $f$ satisfies an additional mild condition, then there is a direct privacy-protecting implementation of $f$. 
Full implementation – direct mechanisms

Theorem

If $f$ is implementable in SPE, preferences are lexicographic, and $f$ satisfies an additional mild condition, then there is a direct privacy-protecting implementation of $f$.

$f$ is implementable in SPE, so we use (a variant of) the implementing mechanism as the contingency plan.
**Full implementation – direct mechanisms**

**Theorem**

If $f$ is implementable in SPE, preferences are lexicographic, and $f$ satisfies an additional mild condition, then there is a *direct* privacy-protecting implementation of $f$.

$f$ is implementable in SPE, so we use (a variant of) the implementing mechanism as the contingency plan.

- If someone objects to the leader’s submitted outcome, then this leads to a contingency plan in which all outcomes are socially-optimal.
Full implementation – direct mechanisms

Theorem
If \( f \) is implementable in SPE, preferences are lexicographic, and \( f \) satisfies an additional mild condition, then there is a \textit{direct} privacy-protecting implementation of \( f \).

\( f \) is implementable in SPE, so we use (a variant of) the implementing mechanism as the contingency plan.

- If someone objects to the leader’s submitted outcome, then this leads to a contingency plan in which all outcomes are socially-optimal.
- However, all information is revealed in the contingency plan.
Full implementation – direct mechanisms

**Theorem**

If $f$ is implementable in SPE, preferences are lexicographic, and $f$ satisfies an additional mild condition, then there is a *direct* privacy-protecting implementation of $f$.

$f$ is implementable in SPE, so we use (a variant of) the implementing mechanism as the contingency plan.

- If someone objects to the leader’s submitted outcome, then this leads to a contingency plan in which all outcomes are socially-optimal.
- However, all information is revealed in the contingency plan.
- If the leader submits an outcome $a$ that is not socially-optimal, then some agent would be willing to reveal all information to get a better outcome.
Recap

A model of preferences that may include a desire for privacy.

Impossibility result: Without restrictions on these preferences, implementation may be impossible.

Restrictions: Lexicographic and MWR preferences.

Impossibility result: A non-constant SCC has a privacy-protecting implementation if and only if it is a dictatorship.

Possibility result: Any SCF that can be implemented in SPE can be privately implemented in SPE.

Possibility result: Direct mechanisms,
Recap

- A model of preferences that may include a desire for privacy.
Recap

- A model of preferences that may include a desire for privacy.
- Impossibility result: Without restrictions on these preferences, implementation may be impossible 😞
Recap

- A model of preferences that may include a desire for privacy.

- Impossibility result: Without restrictions on these preferences, implementation may be impossible 😞

- Restrictions: Lexicographic and MWR preferences.
Recap

- A model of preferences that may include a desire for privacy.
- Impossibility result: Without restrictions on these preferences, implementation may be impossible 😞
- Restrictions: Lexicographic and MWR preferences.
- Impossibility result: A non-constant SCC has a privacy-protecting implementation if and only if it is a dictatorship 😊
Recap

- A model of preferences that may include a desire for privacy.

- Impossibility result: Without restrictions on these preferences, implementation may be impossible 😞

- Restrictions: Lexicographic and MWR preferences.

- Impossibility result: A non-constant SCC has a privacy-protecting implementation if and only if it is a dictatorship 😃

- Possibility result: Any SCF that can be implemented in SPE can be privately implemented in SPE 😊
Recap

- A model of preferences that may include a desire for privacy.
- Impossibility result: Without restrictions on these preferences, implementation may be impossible 😞
- Restrictions: Lexicographic and MWR preferences.
- Impossibility result: A non-constant SCC has a privacy-protecting implementation if and only if it is a dictatorship 😊
- Possibility result: Any SCF that can be implemented in SPE can be privately implemented in SPE 😊
- Possibility result: Direct mechanisms 😊
Thank You!

😊