Robust Predictions in Dynamic Screening

Daniel Garrett (TSE), Alessandro Pavan (NU), JuusoToikka (MIT)

January 2017

Mechanism Design

- Mechanism Design: auctions, regulation, taxation, political economy, etc...

Mechanism Design

- Mechanism Design: auctions, regulation, taxation, political economy, etc...

- Standard model: one-time information, one-time decisions

Mechanism Design

- Mechanism Design: auctions, regulation, taxation, political economy, etc...
- Standard model: one-time information, one-time decisions
- Many settings

  - **information arrives over time** (serially correlated, possibly endogenous)

  - **sequence of decisions**

Long-Term Contracting

- Long-term contracting

Long-Term Contracting

- Long-term contracting
  - Trade

Long-Term Contracting

- Long-term contracting
  - Trade
  - Employment

Long-Term Contracting

- Long-term contracting
  - Trade
  - Employment
  - Taxation

Long-Term Contracting

- Long-term contracting
  - Trade
  - Employment
  - Taxation
  - Regulation

Long-Term Contracting

- Long-term contracting
  - Trade
  - Employment
  - Taxation
  - Regulation
  - Financing

Long-Term Contracting

- Long-term contracting
  - Trade
  - Employment
  - Taxation
  - Regulation
  - Financing
  - etc.

Long-Term Contracting

- Value of relationship changes over time

Long-Term Contracting

- Value of relationship changes over time
- "Shocks" to:

Long-Term Contracting

- Value of relationship changes over time
- “Shocks” to:
  - valuations

Long-Term Contracting

- Value of relationship changes over time
- "Shocks" to:
  - valuations
  - costs

Long-Term Contracting

- Value of relationship changes over time
- "Shocks" to:
  - valuations
  - costs
  - productivity

Long-Term Contracting

- Value of relationship changes over time
- “Shocks” to:
  - valuations
  - costs
  - productivity
  - outside options

Long-Term Contracting

- Value of relationship changes over time
- "Shocks" to:
  - valuations
  - costs
  - productivity
  - outside options
  - etc.

Long-Term Contracting

- Value of relationship changes over time
- "Shocks" to:
  - valuations
  - costs
  - productivity
  - outside options
  - etc.
- Changes often anticipated albeit not necessarily commonly observed

Questions

- Structure of optimal LT contract in changing environments

Questions

- Structure of optimal LT contract in changing environments

- Dynamics of distortions — convergence to FB?

Dynamic Mechanism Design

- Applications:

  - **revenue management** (Courty and Li, 2000, Battaglini 2005, Boleslavsky and Said, 2013, Ely, Garrett and Hinnosaar, 2014, Board and Skrzypacz, 2015, Akan, Ata, and Dana, 2015,...)

  - **disclosure in auctions** (Eso and Szentes, 2007, Bergemann and Wambach (2015), Li and Shi (2015)...)

  - **experimentation** (Bergemann and Välimäki, 2010, Pavan, Segal, and Toikka, 2014, Fershtman and Pavan, 2015...)

  - **taxation** (Farhi and Werning, 2012, Kapicka, 2013, Stantcheva, 2014, Makris and Pavan,2015,...)

  - **managerial compensation** (Garrett and Pavan, 2012, 2014,...)

  - **insurance** (Hendel and Lizzeri, 2003, Handel, Hendel, Whinston, 2015,...)

Static example

- Price discrimination (Mussa-Rosen, Maskin & Riley, Myerson)

Static example

- Price discrimination (Mussa-Rosen, Maskin & Riley, Myerson)
- Principal: seller

Static example

- Price discrimination (Mussa-Rosen, Maskin & Riley, Myerson)
- Principal: seller
- Agent: buyer

Static example

- Price discrimination (Mussa-Rosen, Maskin & Riley, Myerson)
- Principal: seller
- Agent: buyer
- Quasilinear payoffs

$$U^P = p - c(q) \quad \text{and} \quad U^A = \theta q - p$$

with θ drawn from F (density f), privately observed by Buyer

Static example

- Price discrimination (Mussa-Rosen, Maskin & Riley, Myerson)
- Principal: seller
- Agent: buyer
- Quasilinear payoffs

$$U^P = p - c(q) \quad \text{and} \quad U^A = \theta q - p$$

  with θ drawn from F (density f), privately observed by Buyer

- Incentive compatibility:

$$V^A(\theta) \equiv \theta q(\theta) - p(\theta) = \sup_{\hat{\theta}} \{\theta q(\hat{\theta}) - p(\hat{\theta})\}$$

Static example

- Price discrimination (Mussa-Rosen, Maskin & Riley, Myerson)
- Principal: seller
- Agent: buyer
- Quasilinear payoffs

$$U^P = p - c(q) \quad \text{and} \quad U^A = \theta q - p$$

with θ drawn from F (density f), privately observed by Buyer

- Incentive compatibility:

$$V^A(\theta) \equiv \theta q(\theta) - p(\theta) = \sup_{\hat{\theta}} \left\{ \theta q(\hat{\theta}) - p(\hat{\theta}) \right\}$$

- Envelope Th.

$$V^A(\theta) = V^A(\underline{\theta}) + \int_{\underline{\theta}}^{\theta} q(s)ds \quad \text{with} \quad q(\cdot) \text{ nondecreasing}$$

Static example

- Transfer (revenue equivalence)

$$p(\theta) = \theta q(\theta) - \left\{ V^A(\underline{\theta}) + \int_{\underline{\theta}}^{\theta} q(s)ds \right\}$$

Static example

- Transfer (revenue equivalence)

$$p(\theta) = \theta q(\theta) - \left\{ V^A(\underline{\theta}) + \int_{\underline{\theta}}^{\theta} q(s)ds \right\}$$

- Optimal quantity schedule maximizes expected "**virtual surplus**"

$$\mathbb{E}\left[\left(\theta - \frac{1-F(\theta)}{f(\theta)}\right) q(\theta) - c(q(\theta))\right] \quad \text{s.t. } q(\cdot) \text{ nondecreasing (M)}$$

Static example

- Transfer (revenue equivalence)

$$p(\theta) = \theta q(\theta) - \left\{ V^A(\underline{\theta}) + \int_{\underline{\theta}}^{\theta} q(s)ds \right\}$$

- Optimal quantity schedule maximizes expected "**virtual surplus**"

$$\mathbb{E}\left[\left(\theta - \frac{1-F(\theta)}{f(\theta)}\right) q(\theta) - c(q(\theta))\right] \quad \text{s.t. } q(\cdot) \text{ nondecreasing (M)}$$

- **Robust predictions** (e.g., Hellwig, 2010):

1. participation constraint binds only for lowest type: $V^A(\underline{\theta}) = 0$
2. no distortion at the top: $q(\bar{\theta}) = q^{FB}(\bar{\theta})$
3. downward distortions elsewhere: $q(\theta) < q^{FB}(\theta) \quad \forall \theta < \bar{\theta}$

Static example

- Transfer (revenue equivalence)

$$p(\theta) = \theta q(\theta) - \left\{ V^A(\underline{\theta}) + \int_{\underline{\theta}}^{\theta} q(s)ds \right\}$$

- Optimal quantity schedule maximizes expected "**virtual surplus**"

$$\mathbb{E}\left[\left(\theta - \frac{1-F(\theta)}{f(\theta)}\right) q(\theta) - c(q(\theta))\right] \quad \text{s.t. } q(\cdot) \text{ nondecreasing (M)}$$

- **Robust predictions** (e.g., Hellwig, 2010):

1. participation constraint binds only for lowest type: $V^A(\underline{\theta}) = 0$
2. no distortion at the top: $q(\bar{\theta}) = q^{FB}(\bar{\theta})$
3. downward distortions elsewhere: $q(\theta) < q^{FB}(\theta) \quad \forall \theta < \bar{\theta}$

- Binding (M): "ironing" (just more pooling)

Dynamic Environment

- $t = 1, \ldots, T$ (possibly infinite)

- Intertemporal payoffs

$$U^P = \sum_t \delta^{t-1}(p_t - c(q_t)) \quad \text{and} \quad U^A = \sum_t \delta^{t-1}(\theta_t q_t - p_t)$$

- θ_t privately observed by agent at beginning of period t

Type process

- type θ_t drawn from (exogenous) Markov chain on $\Theta = [\underline{\theta}, \overline{\theta}] \subseteq \mathbb{R}_+$
- transition probability kernels $F \equiv (F_t)$
- $F_t(\cdot \mid \theta)$: cdf of θ_t, given $\theta_{t-1} = \theta$
- F_1 : cdf of initial distribution; density f_1
- *stochastic monotonicity (FOSD):* $\theta' > \theta \Rightarrow F_t(\cdot \mid \theta') \succsim_{FOSD} F_t(\cdot \mid \theta)$
- *ergodicity:* $\exists!$ invariant distribution π s.t., for all $\theta \in \Theta$

$$\sup_{A \in \mathscr{B}(\Theta)} \left| F^t(A, \theta) - \pi(A) \right| \to 0 \text{ as } t \to \infty$$

- *stationarity:* $F_1 = \pi$ and $F_t = F_s$ all $t, s > 1$.

Principal's problem

- Principal designs $\chi = \langle \mathbf{q}, \mathbf{p} \rangle$ to maximize

$$\mathbb{E}\left[\sum_t \delta^{t-1}(p_t(\theta^t) - c(q_t(\theta^t)))\right]$$

subject to IR-1 and IC-t, all $t \geq 1$

- Stronger (periodic) IR

- Complexity:

  - different types have different beliefs about future

  - multi-period deviations

IC-IR-extended

State representation and impulse responses
Eso-Szentes (2007), Pavan, Segal, Toikka (2014)

- Auxiliary shocks, orthogonal to initial private information
- $\theta_t = Z_t(\theta_1, \varepsilon)$ where $\varepsilon \equiv (\varepsilon_t)$ are iid r.v.s
- Integral-transform-probability theorem (F_t^{-1} inductively)
- **Impulse responses**:

$$\boxed{I_t(\theta) = \frac{\partial}{\partial \theta_1}\theta_t = \left.\frac{\partial Z_t(\theta_1,\varepsilon)}{\partial \theta_1}\right|_{Z^t(\theta_1,\varepsilon)=\theta^t}}$$

Examples

- AR(1):

$$\begin{aligned}\theta_t &= \gamma\theta_{t-1} + \varepsilon_t \\ &= Z_t(\theta_1, \varepsilon) = \gamma^{t-1}\theta_1 + \gamma^{t-2}\varepsilon_2 + \cdots + \gamma\varepsilon_{t-1} + \varepsilon_t \\ &\rightarrow I_t(\theta_1, \varepsilon) = \gamma^{t-1}\end{aligned}$$

Examples

- AR(1):

$$\begin{aligned} \theta_t &= \gamma\theta_{t-1} + \varepsilon_t \\ &= Z_t(\theta_1, \varepsilon) = \gamma^{t-1}\theta_1 + \gamma^{t-2}\varepsilon_2 + \cdots + \gamma\varepsilon_{t-1} + \varepsilon_t \\ &\rightarrow I_t(\theta_1, \varepsilon) = \gamma^{t-1} \end{aligned}$$

- ARIMA:

$$\begin{aligned} \theta_t &= Z_t(\theta_1, \varepsilon) = a_{t,1}\theta_1 + a_{t,2}\varepsilon_2 + \cdots + a_{t,t-1}\varepsilon_{t-1} + \varepsilon_t \\ &\rightarrow I_t(\theta_1, \varepsilon) = a_{t,1} \end{aligned}$$

Examples

- AR(1):

$$\begin{aligned}\theta_t &= \gamma\theta_{t-1} + \varepsilon_t \\ &= Z_t(\theta_1, \varepsilon) = \gamma^{t-1}\theta_1 + \gamma^{t-2}\varepsilon_2 + \cdots + \gamma\varepsilon_{t-1} + \varepsilon_t \\ &\rightarrow I_t(\theta_1, \varepsilon) = \gamma^{t-1}\end{aligned}$$

- ARIMA:

$$\begin{aligned}\theta_t &= Z_t(\theta_1, \varepsilon) = a_{t,1}\theta_1 + a_{t,2}\varepsilon_2 + \cdots + a_{t,t-1}\varepsilon_{t-1} + \varepsilon_t \\ &\rightarrow I_t(\theta_1, \varepsilon) = a_{t,1}\end{aligned}$$

- Multiplicative shocks

$$\begin{aligned}\theta_t &= Z_t(\theta_1, \varepsilon) = \theta_1 \times \varepsilon_2 \times \cdots \times \varepsilon_t \\ &\rightarrow I_t(\theta_1, \varepsilon) = \varepsilon_2 \times \cdots \times \varepsilon_t\end{aligned}$$

Examples

- AR(1):

$$\begin{aligned} \theta_t &= \gamma\theta_{t-1} + \varepsilon_t \\ &= Z_t(\theta_1, \varepsilon) = \gamma^{t-1}\theta_1 + \gamma^{t-2}\varepsilon_2 + \cdots + \gamma\varepsilon_{t-1} + \varepsilon_t \\ &\rightarrow I_t(\theta_1, \varepsilon) = \gamma^{t-1} \end{aligned}$$

- ARIMA:

$$\begin{aligned} \theta_t &= Z_t(\theta_1, \varepsilon) = a_{t,1}\theta_1 + a_{t,2}\varepsilon_2 + \cdots + a_{t,t-1}\varepsilon_{t-1} + \varepsilon_t \\ &\rightarrow I_t(\theta_1, \varepsilon) = a_{t,1} \end{aligned}$$

- Multiplicative shocks

$$\begin{aligned} \theta_t &= Z_t(\theta_1, \varepsilon) = \theta_1 \times \varepsilon_2 \times \cdots \times \varepsilon_t \\ &\rightarrow I_t(\theta_1, \varepsilon) = \varepsilon_2 \times \cdots \times \varepsilon_t \end{aligned}$$

- More generally,

$$I_t = \prod_{s \le t} \frac{\partial}{\partial \theta} F_s^{-1}(\varepsilon_s \mid \theta_{s-1})$$

Examples

- AR(1):

$$\begin{aligned}\theta_t &= \gamma\theta_{t-1} + \varepsilon_t \\ &= Z_t(\theta_1, \varepsilon) = \gamma^{t-1}\theta_1 + \gamma^{t-2}\varepsilon_2 + \cdots + \gamma\varepsilon_{t-1} + \varepsilon_t \\ &\rightarrow I_t(\theta_1, \varepsilon) = \gamma^{t-1}\end{aligned}$$

- ARIMA:

$$\begin{aligned}\theta_t &= Z_t(\theta_1, \varepsilon) = a_{t,1}\theta_1 + a_{t,2}\varepsilon_2 + \cdots + a_{t,t-1}\varepsilon_{t-1} + \varepsilon_t \\ &\rightarrow I_t(\theta_1, \varepsilon) = a_{t,1}\end{aligned}$$

- Multiplicative shocks

$$\begin{aligned}\theta_t &= Z_t(\theta_1, \varepsilon) = \theta_1 \times \varepsilon_2 \times \cdots \times \varepsilon_t \\ &\rightarrow I_t(\theta_1, \varepsilon) = \varepsilon_2 \times \cdots \times \varepsilon_t\end{aligned}$$

- More generally,

$$I_t = \prod_{s \le t} \frac{\partial}{\partial \theta} F_s^{-1}(\varepsilon_s \mid \theta_{s-1})$$

- Continuous-time (Bergemann and Strack, 2015)

Local IC – heuristics

- Assume $T = 2$
- Fix period-1 report, $\hat{\theta}_1$, and priod-2 reporting strategy, $\sigma(\varepsilon)$
- Agent's payoff
$$U^A(\theta_1, \hat{\theta}_1; \sigma) = \theta_1 q_1(\hat{\theta}_1) - p_1(\hat{\theta}_1) + \delta \mathbb{E}\left[Z_2(\theta_1, \varepsilon) q_2(\hat{\theta}_1, \sigma(\varepsilon)) - p_2(\hat{\theta}_1, \sigma(\varepsilon))\right]$$
- If $\chi = \langle \mathbf{q}, \mathbf{p} \rangle$ is IC, then
$$V_1^A(\theta_1) = \sup_{\hat{\theta}_1; \sigma} U^A(\theta_1, \hat{\theta}_1; \sigma)$$
- Envelope theorem
$$\frac{\partial V_1^A}{\partial \theta_1} = \frac{\partial}{\partial \theta_1} U^A(\theta_1, \theta_1; \sigma^{truth}) = q_1(\theta_1) + \delta \mathbb{E}\left[\frac{\partial Z_2(\theta_1, \varepsilon)}{\partial \theta_1} q_2(\theta_1, \varepsilon))\right]$$
$$= \mathbb{E}\left[\sum_{s \geq 1} \delta^{s-1} I_s q_s \mid \theta_1\right]$$

Local IC – general case

- More generally,

Theorem (Pavan, Segal, Toikka, 2014)

If $\chi = \langle \mathbf{q}, \mathbf{p} \rangle$ is IC, then, for every truthful history θ^{t-1}, $t \geq 0$, V_t^A is equi-Lipschitz-continuous in θ_t and

$$\frac{\partial V_t^A}{\partial \theta_t} = \mathbb{E}\left[\sum_{s \geq t} \delta^{s-1} I_{t \to s} q_s \mid \theta^t\right] \quad a.e., \qquad \text{(ICFOC)}$$

where $I_{t \to s} = \frac{d}{d\theta_t} \theta_s$ (with $I_t \equiv I_{1 \to t}$)

▸ ICFOC-proof

Sufficiency and Integral Monotonicity

- Envelope conditions necessary but not sufficient

Sufficiency and Integral Monotonicity

- Envelope conditions necessary but not sufficient
- Appropriate monotonicity conditions

Sufficiency and Integral Monotonicity

- Envelope conditions necessary but not sufficient
- Appropriate monotonicity conditions

Theorem (PST, 2014)

Mechanism $\chi = \langle \mathbf{q}, \mathbf{p} \rangle$ *is IC iff, for all* $t \geq 0$,

$$\frac{\partial V_t^A}{\partial \theta_t} = \mathbb{E}\left[\sum_{s \geq t} \delta^{s-1} I_{t \to s} q_s \mid \theta^t\right] \text{ a.e.,} \qquad \text{(ICFOC)}$$

and, for all θ^t *and* $\hat{\theta}_t$,

$$\int_{\hat{\theta}_t}^{\theta_t} [D_t((\theta^{t-1}, x); x) - D((\theta^{t-1}, x); \hat{\theta}_t)] dx \geq 0 \qquad \text{(INT-M)}$$

where

$$D_t(\theta^t; \hat{\theta}_t) \equiv \mathbb{E}\left[\sum_{s \geq t} \delta^{s-1} I_{t \to s} q_s(\theta^s_{-t}, \hat{\theta}_t) \mid \theta^t\right]$$

Sufficiency and Integral Monotonicity

- Envelope conditions necessary but not sufficient
- Appropriate monotonicity conditions

Theorem (PST, 2014)

Mechanism $\chi = \langle \mathbf{q}, \mathbf{p} \rangle$ is IC iff, for all $t \geq 0$,

$$\frac{\partial V_t^A}{\partial \theta_t} = \mathbb{E}\left[\sum_{s \geq t} \delta^{s-1} I_{t \to s} q_s \mid \theta^t\right] \text{ a.e.,} \qquad \text{(ICFOC)}$$

and, for all θ^t and $\hat{\theta}_t$,

$$\int_{\hat{\theta}_t}^{\theta_t} [D_t((\theta^{t-1}, x); x) - D((\theta^{t-1}, x); \hat{\theta}_t)] dx \geq 0 \qquad \text{(INT-M)}$$

where

$$D_t(\theta^t; \hat{\theta}_t) \equiv \mathbb{E}\left[\sum_{s \geq t} \delta^{s-1} I_{t \to s} q_s(\theta^s_{-t}, \hat{\theta}_t) \mid \theta^t\right]$$

- Int-M $\rightarrow$ one-stage deviations suboptimal

Sufficiency and Integral Monotonicity

- Envelope conditions necessary but not sufficient
- Appropriate monotonicity conditions

Theorem (PST, 2014)

Mechanism $\chi = \langle \mathbf{q}, \mathbf{p} \rangle$ is IC iff, for all $t \geq 0$,

$$\frac{\partial V_t^A}{\partial \theta_t} = \mathbb{E}\left[\sum_{s \geq t} \delta^{s-1} I_{t \to s} q_s \mid \theta^t\right] \text{ a.e.,} \qquad \text{(ICFOC)}$$

and, for all θ^t and $\hat{\theta}_t$,

$$\int_{\hat{\theta}_t}^{\theta_t} [D_t((\theta^{t-1}, x); x) - D((\theta^{t-1}, x); \hat{\theta}_t)] dx \geq 0 \qquad \text{(INT-M)}$$

where

$$D_t(\theta^t; \hat{\theta}_t) \equiv \mathbb{E}\left[\sum_{s \geq t} \delta^{s-1} I_{t \to s} q_s(\theta^s_{-t}, \hat{\theta}_t) \mid \theta^t\right]$$

- Int-M $\rightarrow$ one-stage deviations suboptimal
- Int-M + Markov + continuity at ∞ $\rightarrow$ all deviations suboptimal

Int-M-Proof

Stronger sufficient conditions

- Int-M holds if expected future output, *discounted by impulse responses*

$$D_t(\theta^t;\hat{\theta}_t) = \mathbb{E}\left[\sum_{s\geq t} \delta^{s-1} I_{t\to s} q_s(\theta^s_{-t},\hat{\theta}_t) \mid \theta^t\right]$$

  is *nondecreasing* in current report $\hat{\theta}_t$.

- Output need not be monotone history by history, enough to have monotonicity **on average** over time and states.

- Literature typically checks "strong monotonicity" (i.e., $q_t(\theta^t)$ nondecreasing in θ^t), but that's stronger than necessary.

Full program

- Principal's **full program**

$$\max_{\chi=\langle \mathbf{q},\mathbf{p}\rangle} \mathbb{E}\left[\sum_t \delta^{t-1}(p_t - c(q_t))\right]$$

subject to

IR: $V_1^A(\theta_1) \geq 0$ all θ_1

ICFOC-(t): $$\frac{\partial V_t^A(\theta^t)}{\partial \theta_t} = D_t(\theta^t;\theta_t)$$

Int-M: $$\int_{\hat{\theta}_t}^{\theta_t} [D_t((\theta^{t-1},x);x) - D_t((\theta^{t-1},x);\hat{\theta}_t)]dx \geq 0.$$

Relax program – Myersonian/First-Order Approach

- Principal's **relaxed program**

$$\max_{\chi=\langle \mathbf{q},\mathbf{p}\rangle} \mathbb{E}\left[\sum_t \delta^{t-1}(p_t - c(q_t))\right]$$

subject to

IR: $V_1^A(\theta_1) \geq 0$ all θ_1 $\rightarrow$ $V_1^A(\underline{\theta}) \geq 0$

ICFOC-(t): $\dfrac{\partial V_t^A(\theta^t)}{\partial \theta_t} = D_t(\theta^t;\theta_t)$ $\rightarrow$ $\dfrac{\partial V_1^A(\theta_1)}{\partial \theta_1} = D_1(\theta_1;\theta_1)$

Int-M: $\int_{\hat{\theta}_t}^{\theta_t} [D_t((\theta^{t-1},x);x) - D_t((\theta^{t-1},x);\hat{\theta}_t)]dx \geq 0$ $\rightarrow$ $\varnothing$

Relax program – Myersonian/First-Order Approach

- Principal's objective as "**Dynamic Virtual Surplus**"

$$\max_{\mathbf{q}} \mathbb{E}\left[\sum_t \delta^{t-1}\left(\theta_t - \frac{1-F_1(\theta_1)}{f_1(\theta_1)} I_t\right) q_t - c(q_t))\right]$$

Relax program – Myersonian/First-Order Approach

- Principal's objective as "**Dynamic Virtual Surplus**"

$$\max_{\mathbf{q}} \mathbb{E}\left[\sum_t \delta^{t-1}\left(\theta_t - \frac{1-F_1(\theta_1)}{f_1(\theta_1)} I_t\right) q_t - c(q_t))\right]$$

- Pointwise maximization:

$$\text{period-}t \text{ virtual value } = \theta_t - \frac{1-F_1(\theta_1)}{f_1(\theta_1)} I_t = c'(q_t) = \text{ marginal cost}$$

$\Rightarrow$ **distortions** driven by **impulse responses** I_t

Validity of First-Order-Approach

- **Remaining IR constraints** slack under FOSD and $\mathbf{q} \geq 0$

$$V_1^A(\theta_1) = V_1^A(\underline{\theta}) + \int_{\underline{\theta}}^{\theta_1} \mathbb{E}\left[\sum_t \delta^t I_{1\to t} q_t(\theta^t) \mid x\right] dx \geq 0$$

Validity of First-Order-Approach

- **Remaining IR constraints** slack under FOSD and $\mathbf{q} \geq 0$

$$V_1^A(\theta_1) = V_1^A(\underline{\theta}) + \int_{\underline{\theta}}^{\theta_1} \mathbb{E}\left[\sum_t \delta^t I_{1\to t} q_t(\theta^t) \mid x\right] dx \geq 0$$

- **Remaining IC constraints** (equivalently, Int-M) slack if

$$\mathbb{E}\left[\sum_{s\geq t} \delta^t I_{t\to s} q_s(\theta^t_{-s}, \hat{\theta}_t) \mid \theta^t\right] \text{ nondecreasing in } \hat{\theta}_t \text{ all } t$$

Validity of First-Order-Approach

- **Remaining IR constraints** slack under FOSD and $\mathbf{q} \geq 0$

$$V_1^A(\theta_1) = V_1^A(\underline{\theta}) + \int_{\underline{\theta}}^{\theta_1} \mathbb{E}\left[\sum_t \delta^t I_{1\to t} q_t(\theta^t) \mid x\right] dx \geq 0$$

- **Remaining IC constraints** (equivalently, Int-M) slack if

$$\mathbb{E}\left[\sum_{s\geq t} \delta^t I_{t\to s} q_s(\theta^t_{-s}, \hat{\theta}_t) \mid \theta^t\right] \text{ nondecreasing in } \hat{\theta}_t \text{ all } t$$

- Suppose $c(q) = \frac{1}{2}q^2$. Solution to relaxed program

$$q_t = \max\left\{\theta_t - \frac{1-F_1(\theta_1)}{f_1(\theta_1)} I_{1\to t}; 0\right\}$$

Monotone enough?

Validity of First-Order-Approach

- **Remaining IR constraints** slack under FOSD and $\mathbf{q} \geq 0$

$$V_1^A(\theta_1) = V_1^A(\underline{\theta}) + \int_{\underline{\theta}}^{\theta_1} \mathbb{E}\left[\sum_t \delta^t I_{1\to t} q_t(\theta^t) \mid x\right] dx \geq 0$$

- **Remaining IC constraints** (equivalently, Int-M) slack if

$$\mathbb{E}\left[\sum_{s\geq t} \delta^t I_{t\to s} q_s(\theta^t_{-s}, \hat{\theta}_t) \mid \theta^t\right] \text{ nondecreasing in } \hat{\theta}_t \text{ all } t$$

- Suppose $c(q) = \frac{1}{2}q^2$. Solution to relaxed program

$$q_t = \max\left\{\theta_t - \frac{1-F_1(\theta_1)}{f_1(\theta_1)} I_{1\to t}; 0\right\}$$

Monotone enough?

Example (AR-1)

$q_t = \theta_t - \frac{1-F_1(\theta_1)}{f_1(\theta_1)} \phi^{t-1} \Rightarrow$ suffices that F_1 log-concave

Robust predictions in Dynamic Screening

Garrett-Pavan-Toikka

- Predictions that do **not** hinge on FOA

Robust predictions in Dynamic Screening

Garrett-Pavan-Toikka

- Predictions that do **not** hinge on FOA

- Full program: **hard to solve**

Robust predictions in Dynamic Screening

Garrett-Pavan-Toikka

- Predictions that do **not** hinge on FOA
- Full program: **hard to solve**
- Idea: Let $\mathbf{q}$ be optimal allocation process. Any perturbation preserving (Int-M) and IR constraints must be suboptimal

Robust predictions in Dynamic Screening

Garrett-Pavan-Toikka

- Predictions that do **not** hinge on FOA

- Full program: **hard to solve**

- Idea: Let $\mathbf{q}$ be optimal allocation process. Any perturbation preserving (Int-M) and IR constraints must be suboptimal

- **Variational approach** $\rightarrow$ robust predictions for **average distortions**

▸ Existence

Robust predictions

- Assume IR binds only at $\theta_1 = \underline{\theta}$ (always under FOSD and $\mathbf{q} \geq 0$) and interior solutions.

Robust predictions

- Assume IR binds only at $\theta_1 = \underline{\theta}$ (always under FOSD and $\mathbf{q} \geq 0$) and interior solutions.

- Simple perturbation: add constant $a \in \mathbb{R}$ to period-t allocation (equivalently, Gateux derivative in direction $(0,\ldots,0,1,0,\ldots)$)

Robust predictions

- Assume IR binds only at $\theta_1 = \underline{\theta}$ (always under FOSD and $\mathbf{q} \geq 0$) and interior solutions.

- Simple perturbation: add constant $a \in \mathbb{R}$ to period-t allocation (equivalently, Gateux derivative in direction $(0,\ldots,0,1,0,\ldots)$)

- FOC for optimum at $a = 0$:

$$\mathbb{E}\left[\theta_t - \frac{1-F_1(\theta_1)}{f_1(\theta_1)} I_t\right] = \mathbb{E}\left[c'(q_t)\right]$$

$\Rightarrow$ *average* virtual value equals *average* marginal cost

Robust predictions

- Assume IR binds only at $\theta_1 = \underline{\theta}$ (always under FOSD and $\mathbf{q} \geq 0$) and interior solutions.

- Simple perturbation: add constant $a \in \mathbb{R}$ to period-t allocation (equivalently, Gateux derivative in direction $(0, \ldots, 0, 1, 0, \ldots)$)

- FOC for optimum at $a = 0$:

$$\mathbb{E}\left[\theta_t - \tfrac{1-F_1(\theta_1)}{f_1(\theta_1)} I_t\right] = \mathbb{E}\left[c'(q_t)\right]$$

$\Rightarrow$ *average* virtual value equals *average* marginal cost

- Same prediction as under FOA, **but only in expectation**!

$$\begin{aligned} \mathbb{E}[\text{period-}t \text{ distortion}] &\equiv \mathbb{E}[\theta_t - c'(q_t)] \\ &= \mathbb{E}\left[\tfrac{1-F_1(\theta_1)}{f_1(\theta_1)} I_t\right] \end{aligned}$$

Handicap Dynamics

Theorem (Garrett, Pavan, Toikka)

Assume F is ergodic. Then

$$\mathbb{E}\left[\frac{1-F_1(\theta_1)}{f_1(\theta_1)}I_t\right] \to 0.$$

Moreover, if F satisfies FOSD, then convergence is from above.

If, in addition, F is stationary, then convergence is monotone in t.

Handicap-proof

More general bounds

- When IR binds only at bottom and $\mathbf{q}$ interior

$$\mathbb{E}[\text{distortion}] = \mathbb{E}[\text{handicap}] = \mathbb{E}\left[\frac{1-F_1(\theta_1)}{f_1(\theta_1)} I_t\right]$$

More general bounds

- When IR binds only at bottom and $\mathbf{q}$ interior

$$\mathbb{E}[\text{distortion}] = \mathbb{E}[\text{handicap}] = \mathbb{E}\left[\frac{1-F_1(\theta_1)}{f_1(\theta_1)} I_t\right]$$

- More generally,

More general bounds

- When IR binds only at bottom and $\mathbf{q}$ interior

$$\mathbb{E}[\text{distortion}] = \mathbb{E}[\text{handicap}] = \mathbb{E}\left[\frac{1-F_1(\theta_1)}{f_1(\theta_1)} I_t\right]$$

- More generally,

Theorem (Garrett-Pavan-Toikka)

If F is ergodic, then

$$\limsup_{t\to\infty} \mathbb{E}[\theta_t - c'(q_t)] \leq 0 \qquad \text{(limit upward distortions)}$$

If, in addition, q eventually strictly interior, then

$$\lim_{t\to\infty} \mathbb{E}[\theta_t - c'(q_t)] = 0$$

Finally, if distortions are eventually downward, then

$$q_t \xrightarrow{p} q_t^{FB}$$

More general bounds

- When IR binds only at bottom and **q** interior

$$\mathbb{E}[\text{distortion}] = \mathbb{E}[\text{handicap}] = \mathbb{E}\left[\frac{1-F_1(\theta_1)}{f_1(\theta_1)} I_t\right]$$

- More generally,

Theorem (Garrett-Pavan-Toikka)

If F is ergodic, then

$$\limsup_{t\to\infty} \mathbb{E}[\theta_t - c'(q_t)] \leq 0 \qquad \textit{(limit upward distortions)}$$

If, in addition, q eventually strictly interior, then

$$\lim_{t\to\infty} \mathbb{E}[\theta_t - c'(q_t)] = 0$$

Finally, if distortions are eventually downward, then

$$q_t \xrightarrow{p} q_t^{FB}$$

Corollary

Failure to converge → over-consumption and exclusion eventually infinitely often.

Summary

- Optimal dynamic screening contracts:
  output maximizes expected dynamic virtual surplus s.t. integral monotonicity, ICFOC-(t) and period-0 IR.

Summary

- Optimal dynamic screening contracts:
  output maximizes expected dynamic virtual surplus s.t. integral monotonicity, ICFOC-(t) and period-0 IR.

- FOA $\rightarrow$ path-wise predictions

Summary

- Optimal dynamic screening contracts:
  output maximizes expected dynamic virtual surplus s.t. integral monotonicity, ICFOC-(t) and period-0 IR.

- FOA $\longrightarrow$ path-wise predictions

- **Variational approach**: robust predictions for **average dynamics**

Summary

- Optimal dynamic screening contracts:
  output maximizes expected dynamic virtual surplus s.t. integral monotonicity, ICFOC-(t) and period-0 IR.

- FOA $\longrightarrow$ path-wise predictions

- **Variational approach**: robust predictions for **average dynamics**
  - ergodicity $\longrightarrow$ average distortions eventually upwards (weakly)

Summary

- Optimal dynamic screening contracts:
  output maximizes expected dynamic virtual surplus s.t. integral monotonicity, ICFOC-(t) and period-0 IR.

- FOA $\rightarrow$ path-wise predictions

- **Variational approach**: robust predictions for **average dynamics**
  - ergodicity $\rightarrow$ average distortions eventually upwards (weakly)
  - ergodicity + interiority $\rightarrow$ vanishing distortions (in expectations)

Summary

- Optimal dynamic screening contracts:
  output maximizes expected dynamic virtual surplus s.t. integral monotonicity, ICFOC-(t) and period-0 IR.

- FOA $\longrightarrow$ path-wise predictions

- **Variational approach**: robust predictions for **average dynamics**
  - ergodicity $\longrightarrow$ average distortions eventually upwards (weakly)
  - ergodicity + interiority $\longrightarrow$ vanishing distortions (in expectations)
  - ergodicity + interiority + FOSD $\longrightarrow$ expected distortions are downward in all periods

Summary

- Optimal dynamic screening contracts:
  output maximizes expected dynamic virtual surplus s.t. integral monotonicity, ICFOC-(t) and period-0 IR.

- FOA $\rightarrow$ path-wise predictions

- **Variational approach**: robust predictions for **average dynamics**
  - ergodicity $\rightarrow$ average distortions eventually upwards (weakly)
  - ergodicity + interiority $\rightarrow$ vanishing distortions (in expectations)
  - ergodicity + interiority + FOSD $\rightarrow$ expected distortions are downward in all periods
  - ergodicity + interiority + FOSD + stationarity $\rightarrow$ convergence to FB monotone in t

Summary

- Optimal dynamic screening contracts:
  output maximizes expected dynamic virtual surplus s.t. integral monotonicity, ICFOC-(t) and period-0 IR.

- FOA → path-wise predictions

- **Variational approach**: robust predictions for **average dynamics**
  - ergodicity → average distortions eventually upwards (weakly)
  - ergodicity + interiority → vanishing distortions (in expectations)
  - ergodicity + interiority + FOSD → expected distortions are downward in all periods
  - ergodicity + interiority + FOSD + stationarity → convergence to FB monotone in t
  - ergodicity + downward distortions → convergence in probability

Summary

- Optimal dynamic screening contracts:
  output maximizes expected dynamic virtual surplus s.t. integral monotonicity, ICFOC-(t) and period-0 IR.

- FOA $\rightarrow$ path-wise predictions

- **Variational approach**: robust predictions for **average dynamics**
  - ergodicity $\rightarrow$ average distortions eventually upwards (weakly)
  - ergodicity + interiority $\rightarrow$ vanishing distortions (in expectations)
  - ergodicity + interiority + FOSD $\rightarrow$ expected distortions are downward in all periods
  - ergodicity + interiority + FOSD + stationarity $\rightarrow$ convergence to FB monotone in t
  - ergodicity + downward distortions $\rightarrow$ convergence in probability

- ...more remains to be done

Thank You!

Mechanisms and Principal's problem

- direct mechanism $\chi = \langle \mathbf{q}, \mathbf{p} \rangle$, with $q_t : \Theta^t \to \mathscr{Q}$ and $p_t : \Theta^t \to \mathbb{R}$
- principal designs χ to maximize

$$\mathbb{E}\left[\sum_t \delta^{t-1}(p_t - c(q_t))\right]$$

subject to

$$\mathbb{E}\left[\sum_t \delta^{t-1}(\theta_t q_t - p_t) \mid \theta_1\right] \geq 0 \quad \text{for all } \theta_1 \in \Theta \qquad \text{(IR-1)}$$

$$\mathbb{E}\left[\sum_{s \geq t} \delta^{s-1}(\theta_s q_s - p_s) \mid \theta^t\right] \geq \mathbb{E}\left[\sum_{s \geq t} \delta^{s-1}(\theta_s q_s^\sigma - p_s^\sigma) \mid \theta^t\right] \text{(IC-t)}$$

$$\text{for all } \sigma, \text{ all } \theta^t = (\theta_1, \ldots, \theta_t) \in \Theta^t$$

▸ IC-IR-simple

ICFOC: Proof Sketch

- Agent's payoff in terms of state representation:

$$\mathbb{E}\left[\sum_t \delta^{t-1}(\theta_t q_t - p_t) \mid \theta_1\right] = \tilde{\mathbb{E}}\left[\sum_t \delta^{t-1}(\tilde{q}_t(\theta_1,\varepsilon^t)Z_t(\theta_1,\varepsilon^t) - \tilde{p}_t(\theta_1,\varepsilon^t)) \mid \theta_1\right]$$

- Thus,

$$V_1(\theta) = \max_{\hat{\theta}} U(\hat{\theta};\theta)$$

where

$$U(\hat{\theta};\theta) \equiv \tilde{\mathbb{E}}\left[\sum_t \delta^{t-1}\left(\tilde{q}_t(\hat{\theta},\varepsilon^t)Z_t(\theta_1,\varepsilon^t) - \tilde{p}_t(\hat{\theta},\varepsilon^t)\right) \mid \theta\right]$$

- For fixed $\hat{\theta}$,

$$\frac{d}{d\theta}U(\hat{\theta};\theta) = \tilde{\mathbb{E}}\left[\sum_t \delta^{t-1}\tilde{q}_t(\hat{\theta},\varepsilon^t)I_t \mid \theta\right]$$

- Envelope theorem then gives result
- Corollary: q pins down V_1 up to constant even if ε publicly observable $\Rightarrow$ Eso-Szentes' irrelevance result ICFOC

Integral Monotonicity: Proof sketch

- Fix t and θ^{t-1}.
- Let $U(\hat{\theta};\theta)$ = continuation utility of period-t type θ from one-stage deviation to $\hat{\theta}$.
- Markov and full support → IC equivalent to

$$V(\theta) \equiv U(\theta;\theta) = \max_{\hat{\theta}} U(\hat{\theta};\theta) \quad \text{all } \theta \in \Theta.$$

- Equivalently,

$$\hat{\theta} \in \arg\max_{\theta} \{U(\hat{\theta};\theta) - V(\theta)\} \quad \text{for all } \hat{\theta} \in \Theta.$$

- ICFOC implies that, for $\hat{\theta}$ fixed, $g(\theta) = U(\hat{\theta},\theta) - V(\theta)$ is Lipschitz with $g'(\theta) = U_2(\hat{\theta},\theta) - V'(\theta) = U_2(\hat{\theta},\theta) - U_2(\theta,\theta)$ a.e., so

$$g(\hat{\theta}) - g(\theta) = \int_{\theta}^{\hat{\theta}} [U_2(\hat{\theta},x) - U_2(x,x)]dx,$$

- Because $U_2(\hat{\theta},x) = D_t((\theta^{t-1},x);\hat{\theta})$, $\hat{\theta}$ maximizes $g(\theta)$ iff (Int-M).

Int-M

Existence

- Let $g(\mathbf{q}) = \mathbb{E}\left[\sum_t \delta^{t-1}\left(q_t \cdot \left(\theta_t - \frac{1-F_1(\theta_1)}{f_1(\theta_1)} I_t\right) - c(q_t)\right)\right]$ and consider

$$\sup_{\mathbf{q} \in L_2} g(\mathbf{q}) \quad \text{s.t. (Int-M)}$$

where $L_2 = L_2(\mathbb{R}^T)$ is space of square integrable processes with discounted measure, $\mathbf{q} \in L_2$ iff $\|\mathbf{q}\| = \mathbb{E}\left[\sum_t \delta^{t-1} q_t^2\right] < \infty$.

- Assume $c(q) \geq q^2$ for $|q| > \bar{q}$, for some $\bar{q}$
- Then $g(\mathbf{q}) \to -\infty$ as $\|\mathbf{q}\| \to \infty$.
- Moreover, g is concave and Gateux differentiable, and feasible set is closed, convex, and nonempty since defined by bounded linear operators.
- So supremum is achieved, because in a Hilbert space, every concave Gateux-differentiable functional that is "minus infinite at infinity" achieves its maximum on a closed convex set.
- robust

Handicap Dynamics – Proof sketch

- Recall that $\mathbb{E}[I_t \mid \theta_1] = \frac{d}{d\theta_1}\mathbb{E}[\theta_t \mid \theta_1]$.
- Thus,

$$\mathbb{E}\left[\frac{1-F_1(\theta_1)}{f_1(\theta_1)}I_t\right] = \mathbb{E}\left[\frac{1-F_1(\theta_1)}{f_1(\theta_1)}\mathbb{E}[I_t \mid \theta_1]\right] = \int_{\underline{\theta}}^{\bar{\theta}}(1-F_1(\theta_1))\mathbb{E}[I_t \mid \theta_1]d\theta_1$$
$$= (1-F_1(\theta_1))\mathbb{E}[\theta_t \mid \theta_1]|_{\theta_1=\underline{\theta}}^{\theta_1=\bar{\theta}} + \int_{\underline{\theta}}^{\bar{\theta}} f_1(\theta_1)\mathbb{E}[\theta_t \mid \theta_1]d\theta_1$$
$$= \mathbb{E}[\theta_t] - \mathbb{E}[\theta_t \mid \underline{\theta}] \to 0$$

by ergodicity.

- If F monotone (FOSD),

$$\mathbb{E}[\theta_t] - \mathbb{E}[\theta_t \mid \underline{\theta}] \geq 0$$

- If, in addition, $F_1 = \pi$, then

$$\mathbb{E}\left[\frac{1-F_1(\theta_1)}{f_1(\theta_1)}I_t\right] - \mathbb{E}\left[\frac{1-F_1(\theta_1)}{f_1(\theta_1)}I_s\right] = \mathbb{E}[\theta_s \mid \underline{\theta}] - \mathbb{E}[\theta_t \mid \underline{\theta}] \leq 0$$

for $t > s$.

handicap-dynamics