Dynamic Mechanism Design: Robustness and Endogenous Types

Alessandro Pavan

Northwestern University

November 2015
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


- **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...) 


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

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

- Most work:
  - Myersonian/first-order approach ("global" IC slack)
  - exogenous types
Dynamic Mechanism Design

- Most work:
  - Myersonian/first-order approach ("global" IC slack)
  - exogenous types

- Robust Predictions (to binding global IC constraints)?
Dynamic Mechanism Design

- Most work:
  - Myersonian/first-order approach ("global" IC slack)
  - exogenous types

- Robust Predictions (to binding global IC constraints)?

- Novel effects due to endogeneity of types?
Plan

- Robust Predictions
Plan

- Robust Predictions
- Endogenous Types
Plan

1. Robust Predictions
2. Endogenous Types
3. Conclusions
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 \( \theta \) 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 \( \theta \) drawn from \( F \) (density \( f \)), privately observed by Buyer

- Incentive compatibility:

\[ V^A(\theta) \equiv \theta q(\theta) - p(\theta) = \sup_{\hat{\theta}} \left\{ \theta q(\hat{\theta}) - p(\hat{\theta}) \right\} \]
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 \( \theta \) drawn from \( F \) (density \( f \)), privately observed by Buyer
- Incentive compatibility:
  \[ V^A(\theta) = \theta q(\theta) - p(\theta) = \sup_{\hat{\theta}} \{ \theta q(\hat{\theta}) - p(\hat{\theta}) \} \]
- 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)

\[ p(\theta) = \theta q(\theta) - \left\{ V^A(\theta) + \int_\theta^\theta q(s)ds \right\} \]
Static example

- Transfer (revenue equivalence)
  \[ p(\theta) = \theta q(\theta) - \left\{ V^A(\theta) + \int_\theta^\theta q(s)ds \right\} \]

- 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)

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

- 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: \( V^A(\theta) = 0 \)
  2. no distortion at the top: \( q(\bar{\theta}) = q^{FB}(\bar{\theta}) \)
  3. downward distortions elsewhere: \( q(\theta) < q^{FB}(\theta) \)
Static example

- Transfer (revenue equivalence)

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

- Optimal quantity schedule maximizes expected "virtual surplus"

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

- **Robust predictions** (e.g., Hellwig, 2010):
  1. participation constraint binds only for lowest type: \( V^A(\theta) = 0 \)
  2. no distortion at the top: \( q(\bar{\theta}) = q^{FB}(\bar{\theta}) \)
  3. downward distortions elsewhere: \( q(\theta) < q^{FB}(\theta) \)

- **Binding (M): “ironing” (just more pooling)**
Dynamic Environment

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

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

- $\theta_t$ privately observed by agent at beginning of period $t$
Type process

- type $\theta_t$ drawn from (exogenous) Markov chain on $\Theta = [\underline{\theta}, \bar{\theta}] \subseteq \mathbb{R}_+$

- transition probability kernels $F \equiv (F_t)$

- $F_t(\cdot \mid \theta)$: cdf of $\theta_t$, given $\theta_{t-1} = \theta$

- $F_1$: cdf of initial distribution; density $f_1$

- **stochastic monotonicity (FOSD):** $\theta' > \theta \Rightarrow F_t(\cdot \mid \theta') \succ_{FOSD} F_t(\cdot \mid \theta)$

- **ergodicity:** $\exists!$ invariant distribution $\pi$ s.t., for all $\theta \in \Theta$

\[
\sup_{A \in \mathcal{B}(\Theta)} |F^n(A, \theta) - \pi(A)| \rightarrow 0 \text{ as } n \rightarrow \infty
\]

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

Principal’s problem

- Principal designs $\chi = \langle q, p \rangle$ 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
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:**

\[
I_t(\theta_1, \varepsilon) = \frac{\partial}{\partial \theta_1} \theta_t = \frac{\partial Z_t(\theta_1, \varepsilon)}{\partial \theta_1}
\]
Examples

- AR(1):

\[ \theta_t = \gamma \theta_{t-1} + \varepsilon_t \]
\[ = Z_t(\theta_1, \varepsilon) = \gamma^{t-1} \theta_1 + \gamma^{t-2} \varepsilon_2 + \cdots + \gamma \varepsilon_{t-1} + \varepsilon_t \]
\[ \rightarrow I_t(\theta_1, \varepsilon) = \gamma^{t-1} \]
Examples

- AR(1):
  \[ \theta_t = \gamma \theta_{t-1} + \varepsilon_t \]
  \[ = Z_t(\theta_1, \varepsilon) = \gamma^{t-1} \theta_1 + \gamma^{t-2} \varepsilon_2 + \cdots + \gamma \varepsilon_{t-1} + \varepsilon_t \]
  \[ \rightarrow I_t(\theta_1, \varepsilon) = \gamma^{t-1} \]

- ARIMA:
  \[ \theta_t = Z_t(\theta_1, \varepsilon) = a_{t,1} \theta_1 + a_{t,2} \varepsilon_2 + \cdots + a_{t,t-1} \varepsilon_{t-1} + \varepsilon_t \]
  \[ \rightarrow I_t(\theta_1, \varepsilon) = a_{t,1} \]
Examples

- **AR(1):**
  \[
  \theta_t = \gamma \theta_{t-1} + \varepsilon_t \\
  = Z_t(\theta_1, \varepsilon) = \gamma^{t-1} \theta_1 + \gamma^{t-2} \varepsilon_2 + \ldots + \gamma \varepsilon_{t-1} + \varepsilon_t \\
  \rightarrow I_t(\theta_1, \varepsilon) = \gamma^{t-1}
  \]

- **ARIMA:**
  \[
  \theta_t = Z_t(\theta_1, \varepsilon) = a_{t,1} \theta_1 + a_{t,2} \varepsilon_2 + \ldots + a_{t,t-1} \varepsilon_{t-1} + \varepsilon_t \\
  \rightarrow I_t(\theta_1, \varepsilon) = a_{t,1}
  \]

- **Multiplicative shocks**
  \[
  \theta_t = Z_t(\theta_1, \varepsilon) = \theta_1 \times \varepsilon_2 \times \ldots \times \varepsilon_t \\
  \rightarrow I_t(\theta_1, \varepsilon) = \varepsilon_2 \times \ldots \times \varepsilon_t
  \]
Examples

- **AR(1):**
  \[ \theta_t = \gamma \theta_{t-1} + \epsilon_t \]
  \[ = Z_t(\theta_1, \epsilon) = \gamma^{t-1} \theta_1 + \gamma^{t-2} \epsilon_2 + \cdots + \gamma \epsilon_{t-1} + \epsilon_t \]
  \[ \rightarrow I_t(\theta_1, \epsilon) = \gamma^{t-1} \]

- **ARIMA:**
  \[ \theta_t = Z_t(\theta_1, \epsilon) = a_{t,1} \theta_1 + a_{t,2} \epsilon_2 + \cdots + a_{t,t-1} \epsilon_{t-1} + \epsilon_t \]
  \[ \rightarrow I_t(\theta_1, \epsilon) = a_{t,1} \]

- **Multiplicative shocks**
  \[ \theta_t = Z_t(\theta_1, \epsilon) = \theta_1 \times \epsilon_2 \times \cdots \times \epsilon_t \]
  \[ \rightarrow I_t(\theta_1, \epsilon) = \epsilon_2 \times \cdots \times \epsilon_t \]

- **Continuous-time** (Bergemann and Strack, 2015)
Examples

- **AR(1):**
  \[
  \theta_t = \gamma \theta_{t-1} + \epsilon_t \\
  = Z_t(\theta_1, \epsilon) = \gamma^{t-1} \theta_1 + \gamma^{t-2} \epsilon_2 + \cdots + \gamma \epsilon_{t-1} + \epsilon_t \\
  \rightarrow I_t(\theta_1, \epsilon) = \gamma^{t-1}
  \]

- **ARIMA:**
  \[
  \theta_t = Z_t(\theta_1, \epsilon) = a_{t,1} \theta_1 + a_{t,2} \epsilon_2 + \cdots + a_{t,t-1} \epsilon_{t-1} + \epsilon_t \\
  \rightarrow I_t(\theta_1, \epsilon) = a_{t,1}
  \]

- **Multiplicative shocks**
  \[
  \theta_t = Z_t(\theta_1, \epsilon) = \theta_1 \times \epsilon_2 \times \cdots \times \epsilon_t \\
  \rightarrow I_t(\theta_1, \epsilon) = \epsilon_2 \times \cdots \times \epsilon_t
  \]

- **Continuous-time (Bergemann and Strack, 2015)**

- **More generally,**
  \[
  I_t = \prod_{s \leq t} \frac{\partial}{\partial \theta} F_s^{-1}(\epsilon_s \mid \theta_{s-1})
  \]
Local IC – heuristics

- Assume $T = 2$

- Fix period-1 report, $\hat{\theta}_1$, and period-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 $\chi = \langle q, p \rangle$ is IC, then
  
  $$V^A_1(\theta_1) = \sup_{\hat{\theta}_1, \sigma} U^A(\theta_1, \hat{\theta}_1; \sigma)$$

- Envelope theorem
  
  $$\frac{\partial V^A_1}{\partial \theta_1} = \frac{\partial}{\partial \theta_1} U^A(\theta_1, \theta_1; \sigma^{truth}) = q_1(\theta_1) + \delta \mathbb{E} \left[ \frac{\partial Z_2(\theta_1, \varepsilon)}{\partial \theta_1} q_2(\theta_1, \varepsilon) \right]$$
  
  $$= \mathbb{E} \left[ \sum_{s \geq 1} \delta^{s-1} I_s q_s \mid \theta_1 \right]$$
Local IC – general case

More generally,

**Theorem (Pavan, Segal, Toikka, 2014)**

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

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

where $I_{t \rightarrow s} = \frac{d}{d\theta_t} \theta_s$ (with $I_t \equiv I_{1 \rightarrow t}$)
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 \( \chi = \langle q, p \rangle \) is IC iff, for all \( t \geq 0 \),

\[
\frac{\partial V^A_t}{\partial \theta_t} = \mathbb{E} \left[ \sum_{s \geq t} \delta_{s}^{s-1} I_{t \rightarrow s} q_s \mid \theta^t \right] \quad \text{a.e.,} \tag{ICFOC}
\]

and, for all \( \theta^t \) and \( \hat{\theta}_t \),

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

where

\[
D_t(\theta^t; \hat{\theta}_t) = \mathbb{E} \left[ \sum_{s \geq t} \delta_{s}^{s-1} I_{t \rightarrow s} q_s(\theta^t_s, \hat{\theta}_t) \mid \theta^t \right]
\]
Sufficiency and Integral Monotonicity

- Envelope conditions necessary but not sufficient
- Appropriate monotonicity conditions

**Theorem (PST, 2014)**

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

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

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

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

where

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

- Int-M $\rightarrow$ one-stage deviations suboptimal
Sufficiency and Integral Monotonicity

- Envelope conditions necessary but not sufficient
- Appropriate monotonicity conditions

**Theorem (PST, 2014)**

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

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

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

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

where

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

- Int-M → one-stage deviations suboptimal
- Int-M + Markov + OSDP → all deviations suboptimal
Stronger sufficient conditions

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

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

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 \( \theta^t \)), but that’s stronger than necessary.
Full program

- Principal’s **full program**

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

subject to

\(\text{IR:} \quad V_1^A(\theta_1) \geq 0 \text{ all } \theta_1\)

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

\(\text{Int-M:} \quad \int_{\hat{\theta}_t}^{\theta_t} \left[ D_t((\theta^{t-1}, x); x) - D_t((\theta^{t-1}, x); \hat{\theta}_t) \right] dx \geq 0.\)
Relax program – Myersonian/First-Order Approach

- Principal’s relaxed program

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

subject to

\text{IR:} \quad V_1^A(\theta_1) \geq 0 \text{ all } \theta_1 \quad \rightarrow \quad V_1^A(\theta) \geq 0

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

\text{Int-M:} \quad \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 objective as "Dynamic Virtual Surplus"

\[
\max_{q} \mathbb{E} \left[ \sum_{t} \delta^{t-1} \left( \theta_t - \frac{1-F_1(\theta_1)}{f_1(\theta_1)} I_t \right) q_t - c(q_t) \right]
\]
Relax program – Myersonian/First-Order Approach

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

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

- Pointwise maximization:

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

\[\Rightarrow \text{distortions driven by impulse responses } I_t\]
Validity of First-Order-Approach

- Remaining IR constraints slack under FOSD and $q \geq 0$
Validity of First-Order-Approach

- **Remaining IR constraints** slack under FOSD and \( q \geq 0 \)

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

\[
\mathbb{E} \left[ \sum_{s \geq t} \delta^t I_{t \rightarrow s} q_s (\theta_{s \leftarrow t}, \hat{\theta}_t) \mid \theta^t \right] \text{ nondecreasing in } \hat{\theta}_t \text{ all } t
\]
Validity of First-Order-Approach

- **Remaining IR constraints** slack under FOSD and \( q \geq 0 \)

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

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

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

\[
q_t = \theta_t - \frac{1-F_1(\theta_1)}{f_1(\theta_1)} I_t
\]

Monotone enough?
Validity of First-Order-Approach

- **Remaining IR constraints** slack under FOSD and \( q \geq 0 \)

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

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

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

\[
q_t = \theta_t - \frac{1-F_1(\theta_1)}{f_1(\theta_1)} I_t
\]

Monotone enough?

**Example (AR-1)**

\[
q_t = \theta_t - \frac{1-F_1(\theta_1)}{f_1(\theta_1)} \phi^{t-1} \Rightarrow \text{suffices that } F_1 \text{ log-concave}
\]
Robust predictions in Dynamic Screening
Garrett-Pavan-Toikka (2015)

- Predictions that do **not** hinge on FOA
Robust predictions in Dynamic Screening
Garrett-Pavan-Toikka (2015)

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

- Full program: **hard to solve**
Robust predictions in Dynamic Screening
Garrett-Pavan-Toikka (2015)

- Predictions that do not hinge on FOA

- Full program: hard to solve

- Idea: Let $q$ be optimal allocation process. Any perturbation preserving (Int-M) and IR constraints must be suboptimal
Robust predictions in Dynamic Screening
Garrett-Pavan-Toikka (2015)

- Predictions that do not hinge on FOA

- Full program: hard to solve

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

- Variational approach → robust predictions for average distortions
Robust predictions

- Assume IR binds only at $\theta_1 = \theta$ (always under FOSD and $q \geq 0$) and interior solutions.
Robust predictions

- Assume IR binds only at $\theta_1 = \theta$ (always under FOSD and $q \geq 0$) 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 $\theta_1 = \theta$ (always under FOSD and $q \geq 0$) 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 $\theta_1 = \theta$ (always under FOSD and $q \geq 0$) 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

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

  \[ \mathbb{E}[\text{period-$t$ distortion}] \equiv \mathbb{E}[\theta_t - c'(q_t)] \]

  \[ \equiv \mathbb{E} \left[ \frac{1-F_1(\theta_1)}{f_1(\theta_1)} I_t \right] \]
Theorem (Garrett, Pavan, Toikka, 2015)

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 $t$. 
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 general bounds

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

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

- More generally,
More general bounds

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

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

- More generally,

**Theorem (Garrett-Pavan-Toikka, 2015)**

*If $F$ is ergodic, then*

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

*If, in addition, $q$ eventually strictly interior, then*

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

*Finally, if distortions are eventually downward, then*

$$q_t \xrightarrow{p} q_{FB}^\infty$$
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, 2015)**

*If $F$ is ergodic, then*

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

*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*

$$
q_t \overset{p}{\to} q_t^{FB}
$$

---

**Corollary**

*Failure to converge $\to$ over-consumption and exclusion eventually infinitely often.*
Suppose now

\[ U^P \equiv \sum_{t=1}^{T} \delta^{t-1} [y_t - c_t] \]

\[ U^A \equiv \sum_{t=1}^{T} \delta^{t-1} \left[ v^A(c_t) - \psi(y_t, \theta_t) \right] \]

with

\[ \psi(y_t, \theta_t) = \frac{1}{2} (y_t - \theta_t)^2 \text{ and } \theta_t = \rho \theta_{t-1} + \epsilon_t \]

Non quasi-linear environment

Theorem (Garrett and Pavan, JET 2015)

When the agent’s risk aversion and productivity persistence are low, distortions decrease, on average, over time. Opposite is true when they are sufficiently high.

- distortions increasing over time reduce compensation risk
Managerial Turnover in a Changing World
Garrett and Pavan (JPE, 2013)

- Under risk neutrality and declining impulse responses
  
  $\rightarrow$ Effort gradually converges to FB

- Suppose firm can replace its managers and that each new employment relationship is affected by same information frictions as with incumbent

**Theorem (Garrett and Pavan, JPE 2013)**

*Firm’s optimal contract*

1. either induces excessive retention (inefficiently low turnover) at all tenure levels

2. or excessive firing early on followed by excessive retention in the long run.
Plan

1. Robust Predictions

2. Endogenous Types

3. Conclusions
Motivation/Perspective

- Most of DMD literature: exogenous types/information
Motivation/Perspective

- Most of DMD literature: exogenous types/information

- Many problems of interest: information/types *endogenous*
  - bandit auctions for sale of experience goods – Pavan, Segal, Toikka, 2014
  - dynamic matching w. unknown values – Fershtman and Pavan, 2015
  - habit formation/addiction
  - **taxation under learning-by-doing** (Makris and Pavan, 2015)
Incentives for Endogenous Types: Makris and Pavan (2015)

Design of incentive schemes in Markovian setting with

- endogenous types

- non-quasilinear payoffs

- (imperfect) type persistence
Incentives for Endogenous Types: Makris and Pavan (2015)

- Design of incentive schemes in Markovian setting with
  - endogenous types
  - non-quasilinear payoffs
  - (imperfect) type persistence

- Flexible framework embedding
  - taxation under LBD
  - managerial compensation
  - ...

Theory + quantitative analysis
Incentives for Endogenous Types: Makris and Pavan (2015)

- Design of incentive schemes in Markovian setting with
  - endogenous types
  - non-quasilinear payoffs
  - (imperfect) type persistence

- Flexible framework embedding
  - taxation under LBD
  - managerial compensation
  - ...

- Theory + quantitative analysis
Questions

- Effect of **endogeneity** of types on **distortions** and their **dynamics**?
Questions

- Effect of **endogeneity** of types on **distortions** and their **dynamics**?

- (Mirallesian) taxation:
  - Effect of LBD on **progressivity** and **dynamics of marginal taxes**
  - Welfare gains of tax reforms
  - Loss from neglecting LBD
**Related literature (incomplete)**

- **Dynamic Mechanism Design**
  ...

- **New Dynamic Public Finance**
  Farhi and Werning (2012) ←
  Kapicka (Restud 2013)
  Golosov, Troshkin, Tsyvinski (2014)
  ...

- **Taxation w. Human Capital Accumulation / LBD**
  Kapicka (2014)
  Kapicka and Neira (2014) ←
  Stantcheva (2014) ←
  Krause (2009)
  Best and Kleven (2013) ←
  Heathcote, Storesletten and Violante (2014)
  ...
Design problem

- Agent’s productivity

\[ \theta_t = z_t(\theta_{t-1}, y_{t-1}, \epsilon_t) \]

with \( z_t \) increasing
Design problem

- Agent’s productivity

\[ \theta_t = z_t(\theta_{t-1}, y_{t-1}, \varepsilon_t) \]

with \( z_t \) increasing

- Principal’s (dual) problem: choose \( \chi = \langle y, c \rangle \) to maximize

\[
\mathbb{E} \left[ \sum_{t=1}^{T} \delta^t (y_t - c_t) \right]
\]

s.t. \( \chi \) be IC and

\[
\int_{\theta_1}^{\bar{\theta}_1} q \left( V_1^A(\theta_1) \right) dF_1(\theta_1) \geq \kappa
\]

where \( q \left( V_1^A(\theta_1) \right) \) are non-linear Pareto weights
Design problem

- Agent’s productivity
  \[ \theta_t = z_t(\theta_{t-1}, y_{t-1}, \varepsilon_t) \]
  with \( z_t \) increasing

- Principal’s (dual) problem: choose \( \chi = \langle y, c \rangle \) to maximize
  \[
  \mathbb{E} \left[ \sum_{t=1}^{T} \delta^t (y_t - c_t) \right]
  \]
  s.t. \( \chi \) be IC and
  \[
  \int_{\theta_1}^{\bar{\theta}_1} q(V_A^1(\theta_1)) dF_1(\theta_1) \geq \kappa
  \]
  where \( q(V_A^1(\theta_1)) \) are non-linear Pareto weights

- Rawlsian: \( q(V_A^1(\theta_1)) = 0 \) all \( \theta_1 > \theta \)
Incentive Compatibility

- Analysis similar to exogenous case BUT

\[ I_{t \rightarrow s}(\theta^s, y^{s-1}) \]

now depend on past income
Wedges

- Period-\(t\) labor wedge (equivalently, marginal tax rate)

\[
[1 + LD^F_B; \chi_t^B]\left[1 - W_t\right] = \frac{\psi_y(y_t, \theta_t)}{vA'(c_t)}
\]

where \(LD^F_B; \chi_t^B(\theta_t)\) are FB effects of learning-by-doing
Wedges

- Period-\(t\) labor wedge (equivalently, marginal tax rate)

\[
\left[1 + LD_t^{FB;\chi}\right][1 - W_t] = \frac{\psi_y(y_t, \theta_t)}{\nu^{A'}(c_t)}
\]

where \(LD_t^{FB;\chi}(\theta^t)\) are FB effects of learning-by-doing

- Interested in how LBD affects progressivity and dynamics of \(W_t\)
Period-\(t\) labor wedge (equivalently, marginal tax rate)

\[
\left[1 + LD_t^{FB;\chi}\right][1 - W_t] = \frac{\psi_y(y_t, \theta_t)}{\nu^{A'}(c_t)}
\]

where \(LD_t^{FB;\chi}(\theta^t)\) are FB effects of learning-by-doing

Interested in how LBD affects progressivity and dynamics of \(W_t\)

We study

\[
\hat{\hat{W}}_t(\theta^t) \equiv \frac{W_t(\theta^t)}{1 - W_t(\theta^t)}
\]
Wedges: Two-period Rawlsian

- Period-2 wedge

\[ \hat{W}_2(\theta) = RA_2(\theta) \left[ \hat{W}_2^{RN}(\theta) \right] \]

where

- \( \hat{W}_2^{RN}(\theta) \) is wedge under risk neutrality and no LBD
- \( RA_2(\theta) > 0 \) is correction due to risk aversion
Wedges: Two-period Rawlsian

- Period-1 wedge under LBD

\[ \hat{W}_1(\theta_1) = RA_1(\theta_1) \left( \hat{W}_{1RN}(\theta_1) + \Omega_1(\theta_1) \right) \]

where \( \Omega_1(\theta_1) \) is **effect of LBD on period-1 wedge**
Period-1 wedge under LBD

\[ \hat{W}_1(\theta_1) = RA_1(\theta_1) [\hat{W}_1^{RN}(\theta_1) + \Omega_1(\theta_1)] \]

where \( \Omega_1(\theta_1) \) is effect of LBD on period-1 wedge

\( \Omega_1(\theta_1) \) takes into account how variation in \( y_1 \) affects cost of future incentives through

- distribution of future productivity \( F_2(\theta_2|\theta_1, y_1) \)
- impulse responses (handicaps): \( I_2(\theta, y_1) \)
Wedges: General case

Theorem (Makris and Pavan, 2015)

At any history $\theta^t$, optimal wedge given by

$$\hat{W}_t(\theta^t) = RA_t(\theta^t) \left[ \hat{W}^R_{RN} + \Omega_t(\theta^t) \right]$$
Literature: special cases

- **Static RN (Mirrlees-Diamond-Saez)**

\[
\hat{W}_1(\theta_1) = \hat{W}^{RN}_1(\theta_1) = \frac{1 - F_1(\theta_1)}{f_1(\theta_1)} \frac{1}{\theta_1} \varepsilon_{\theta}^{\psi_y}(y_1(\theta_1), \theta_1)
\]

where \( \varepsilon_{\theta}^{\psi_y}(y_1, \theta_1) \) is elasticity of marginal disutility of effort.
Literature: special cases

- **Static RA (Murryles-Saez):**

\[
\hat{W}_1(\theta_1) = RA_1(\theta_1)\hat{W}_1^{RN}(\theta_1)
\]

\[
= \left[ \int_{\theta_1}^{\overline{\theta}} \frac{1}{v'(c_1(s))} \frac{dF_1(s)}{1 - F_1(\theta_1)} v'(c_1(\theta_1)) \right] \frac{1 - F_1(\theta_1)}{f_1(\theta_1)} \frac{1}{\theta_1} \epsilon_{\theta y}^y(y_1(\theta_1), \theta_1)
\]
Literature: special cases

- **Dynamic RN (exogenous types):**

\[
\hat{W}_t(\theta^t) = \hat{W}_t^{RN}(\theta^t) = \frac{1 - F_1(\theta_1)}{f_1(\theta_1)} \frac{I_t(\theta^t)}{\theta^t} e^{\psi_y(y_t(\theta^t), \theta_t)}
\]
Literature: special cases

- Dynamic RA (exogenous types; Farhi and Werning, 2013):

\[ \hat{W}_t(\theta^t) = RA_t(\theta^t)[\hat{W}_t^{RN}(\theta^t)] \]


Literature: special cases

- Dynamic RA (exogenous types with training; Stantcheva, 2014):

\[
\hat{W}_t(\theta^t) = RA_t(\theta^t) \left[ \hat{W}_t^{RN}(\theta^t, b^t(\theta^{t-1})) \right]
\]

where

\[
\hat{W}_t^{RN}(\theta^t, b^t(\theta^{t-1})) = \frac{1 - F_1(\theta_1)}{f_1(\theta_1)} \frac{I_t(\theta^t)}{\theta_t} \varepsilon_{\theta^t}(y_t(\theta^t), \theta_t, b^t(\theta^{t-1}))
\]
General case

- More generally:

\[ \hat{W}_t(\theta^t) = RA_t(\theta^t) [\hat{W}^{RN}_t(\theta^t) + \Omega_t(\theta^t)] \]

- Key: endogeneity of type process
Example of effects of endogenous types on progressivity/dynamics

- \( \psi(y, \theta) = \frac{1}{1+\phi}(\frac{y}{\theta})^{1+\phi} \) where \( 1/\phi > 0 \) is Frisch elasticity
Example of effects of endogenous types on progressivity/dynamics

- $\psi(y, \theta) = \frac{1}{1+\phi} \left( \frac{y}{\theta} \right)^{1+\phi}$ where $1/\phi > 0$ is Frisch elasticity

- Pareto $F_1$ : $\frac{f_1(\theta_1)}{1-F_1(\theta_1)} \theta_1 = \lambda$
Example of effects of endogenous types on progressivity/dynamics

- $\psi(y, \theta) = \frac{1}{1+\phi} \left( \frac{y}{\theta} \right)^{1+\phi}$ where $1/\phi > 0$ is Frisch elasticity

- Pareto $F_1$: $\frac{f_1(\theta_1)}{1-F_1(\theta_1)} \theta_1 = \lambda$

- Ben-Porath specification:
  
  $\theta_2 = z_2(\theta_1, y_1, \varepsilon_1) = \theta_1^\rho \cdot y_1^\zeta \cdot \varepsilon_2$

  where

  - $\rho$ controls for exogenous type persistence
  - $\zeta$ controls for **intensity of LBD**
Example (cont’d)

- Consider economy described above, and assume agent is risk-neutral and principal Rawlsian
Example (cont’d)

- Consider economy described above, and assume agent is risk-neutral and principal Rawlsian

- Without LBD

\[ \hat{W}_t = (1 + \phi)/\lambda \text{ all } t \]
Example (cont’d)

- Consider economy described above, and assume agent is risk-neutral and principal Rawlsian

- Without LBD

\[ \hat{W}_t = (1 + \phi)/\lambda \text{ all } t \]

- With LBD

\[ \hat{W}_1(\theta_1) > \hat{W}_2(\theta) \text{ all } \theta \]

with \( \hat{W}_1(\theta_1) \) strictly increasing in \( \theta_1 \) (progressivity)
Marginal Wedges for Frisch elasticity 0.5 under Pareto-Lognormal Distribution with Pareto tail of alpha=5 at skill on 0.85 percentile (see Diamond, 1998, AER)

Wedges are decreasing for x<0.4
Intuition for progressivity and wedge dynamics

- **Positive effect of LBD on** $W_t$

  - reduction in $y_t$ reduces future cost of incentives through

    (a) **impulse responses** ($I_{t\rightarrow s}$ increasing in $y_t$ under Ben-Porath specification)

    (b) **distribution of future types** (handicaps increasing in future types)

  - importance of future cost of incentives declines with $t \rightarrow W_t$ *decline with t*

- **Progressivity**

  - handicaps increasing in types $\rightarrow$ reduction in cost of incentives most pronounced for higher types
Conclusions

- Dynamic Mechanism Design
  - useful tool for positive and normative analysis of environments in which
    - information arrives over time
    - sequence of decisions
Conclusions

- Dynamic Mechanism Design
  useful tool for positive and normative analysis of environments in which
  - information arrives over time
  - sequence of decisions

- Myersonian and Variational Approach
  - robust predictions
Conclusions

- Dynamic Mechanism Design
  useful tool for positive and normative analysis of environments in which
  - information arrives over time
  - sequence of decisions

- Myersonian and Variational Approach
  - robust predictions

- **Endogeneity of private information**
  → important qualitative and quantitative implications
  - taxation
  - dynamic matching (Fershtman and Pavan, 2015)
Conclusions

- **Dynamic Mechanism Design**
  - useful tool for positive and normative analysis of environments in which
    - information arrives over time
    - sequence of decisions

- **Myersonian and Variational Approach**
  - robust predictions

- **Endogeneity of private information**
  - important qualitative and quantitative implications
    - taxation
    - dynamic matching (Fershtman and Pavan, 2015)

- **(Much) more remains to be done**
  - partial commitment
  - interaction of time-varying information w. population dynamics (e.g., Garrett 2013)
  - endogenous disclosure (Calzolari and Pavan (2006a,b), Kamenica and Gentzkow, 2011,...)
  - empirics (Handel, Hendel, Whinston, 2014, Einav et al., 2015...)

Useful tool for positive and normative analysis of environments in which information arrives over time and sequence of decisions. Robust predictions considering endogeneity of private information.

(Much) more remains to be done, including partial commitment, interaction of time-varying information with population dynamics, endogenous disclosure, and empirics.
Thank You!
Mechanisms and Principal’s problem

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

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

subject to

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

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

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

- 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 \( \varepsilon \) publicly observable \( \Rightarrow \) Eso-Szentes’ irrelevance result
Integral Monotonicity: Proof sketch

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

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

- Equivalently,

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

- 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).
Existence

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

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

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

- Assume \( c(q) \geq q^2 \) for \(|q| > \bar{q}\), for some \( \bar{q} \)

- Then \( g(q) \to -\infty \) as \( \|q\| \to \infty \).

- Moreover, \( g \) is concave and Gateaux 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 Gateaux-differentiable functional that is “minus infinite at infinity” achieves its maximum on a closed convex set.
Recall that $E[I_t | \theta_1] = \frac{d}{d\theta_1} E[\theta_t | \theta_1]$.

Thus,

$$E \left[ \frac{1-F_1(\theta_1)}{f_1(\theta_1)} I_t \right] = E \left[ \frac{1-F_1(\theta_1)}{f_1(\theta_1)} E[I_t | \theta_1] \right] = \int_{\theta} (1 - F_1(\theta_1)) E[I_t | \theta_1] d\theta_1$$

$$= (1 - F_1(\theta_1)) E[\theta_t | \theta_1] |_{\theta_1 = \bar{\theta}} + \int_{\theta} f_1(\theta_1) E[\theta_t | \theta_1] d\theta_1$$

$$= E[\theta_t] - E[\theta_t | \theta] \to 0$$

by ergodicity.

- If $F$ monotone (FOSD),

$$E[\theta_t] - E[\theta_t | \theta] \geq 0$$

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

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

for $t > s$. 