The Dynamic Pivot Mechanism

Dirk Bergemann and Juuso Välimäki

January 22, 2012
Bandit Problems

recall the setting:

- $n$ alternatives
- each alternative $i$ has a state $\theta_{it}$
- when activated, alternative $i$ yields reward $r_i(\theta_{it})$
- exactly one alternative activated at each $t$, $q_t \in \{1, \ldots, n\}$
- $\theta_{it+1} \sim F_i(\cdot | \theta_{it})$ if $i$ is activated
- $\theta_{it+1} + \theta_{it}$ if $i$ is not activated
- objective function:

$$\mathbb{E} \sum_{t=0}^{\infty} \delta^t r_{it}(\theta_{it}) 1\{q_t=i\}$$
for each $i$, compute

$$\gamma_i (\theta_{it}) = \max_{\tau} \frac{\mathbb{E} \sum_{t=s}^{\tau} \delta^{s-t} r_{is} (\theta_{is})}{\mathbb{E} \sum_{t=s}^{\tau} \delta^{s-t}},$$

where $\tau$ is a stopping time determined by $(\theta_{i0}, ..., \theta_{it})$. 
Theorem (Gittins Index Theorem)

A policy \( \{ q_t \} \) solves

\[
\max_{(q_t)} \mathbb{E} \sum_{t=0}^{\infty} \delta^t r_{it}(\theta_{it}) 1_{\{q_t=i\}}
\]

if and only if for all \( t \),

\[ q_t \in \arg \max_i \gamma_i(\theta_{it}). \]

Alternative formulation for \( m_i(\theta_{it}) \) due to Whittle: Let

\[
V_i(m, \theta_{it}) = \max\{ r_{it}(\theta_{it}) + \delta \mathbb{E} V_i(m, \theta_{it+1}), \frac{m}{1 - \delta} \}.
\]

\[
\gamma_i(\theta_{it}) = \min\{ m \mid V_i(m, \theta_{it}) = m \}.
\]
General Dynamic Model with Private Values

- Model
- time horizon $t = 0, 1, \ldots$
- agent $i \in \{1, 2, \ldots, I\}$ in period $t \in \mathbb{N}$
- allocation $q_t \in Q$, $q$ finite
- transfer $p_{i,t} \in \mathbb{R}$ and type of agent $i$
- $\theta_{i,t} \in \Theta_i$.

von Neumann Morgenstern utility function $u_i$

$$u_i(q_t, p_{i,t}, \theta_{i,t}) \triangleq v_i(q_t, \theta_{i,t}) - p_{i,t}.$$  

- type of agent $\theta_{i,t}$ is a general Markov process on $\Theta_i$.
- type vector $\theta_t = (\theta_{1,t}, \ldots, \theta_{I,t})$ with $\Theta = \times_{i=1}^I \Theta_i$. 

Dirk Bergemann and Juuso Välimäki

The Dynamic Pivot Mechanism
common prior $F_i(\theta_{i,0})$

- current state $\theta_{i,t}$ and the current action $q_t$ define a probability distribution $F_i(\theta_{i,t+1}; \theta_{i,t}, q_t)$
- utility functions $u_i(\cdot)$ and the probability transition functions $F_i(\cdot; q_t, \theta_{i,t})$ are common knowledge at $t = 0$.
- $F_i(\theta_{i,0})$ and $F_i(\theta_{i,t+1}; \theta_{i,t}, q_t)$ independent
- each agent $i$ observes $\theta_{i,t}$ privately.
- an action $q_t \in Q$ is chosen and payoffs for period $t$ are realized
- There is a $K < \infty$ such that for every $i, q, \theta$ and allocation plan $q' : \Theta \rightarrow Q$:

$$
\int |v_i(q'(\theta'), \theta_i)| \ dF(\theta'; q, \theta) < K.
$$

- The nature of the state space $\Theta$ depends on the application at hand.
Efficient solution

- common discount factor $\delta$, $0 < \delta < 1$
- socially efficient policy:

$$W(\theta_t) \triangleq \max_{\{q_s\}_{s=t}^{\infty}} \mathbb{E}\left[ \sum_{s=t}^{\infty} \delta^{s-t} \sum_{i=1}^{l} v_i(q_s, \theta_{i,s}) \right].$$

- or with DP:

$$W(\theta_t) = \max_{q_t} \mathbb{E}\left[ \sum_{i=1}^{l} v_i(q_t, \theta_{i,t}) + \delta \mathbb{E}W(\theta_{t+1}) \right].$$

- socially efficient policy is denoted by $q^* = \{q^*_t\}_{t=0}^{\infty}$. The social externality cost of agent $i$ is determined by the social value in the absence of agent $i$:

$$W_{-i}(\theta_t) \triangleq \max_{\{q_s\}_{s=t}^{\infty}} \mathbb{E}\left[ \sum_{s=t}^{\infty} \delta^{s-t} \sum_{j\neq i} v_j(q_s, \theta_{j,s}) \right].$$
Direct Mechanism

- direct mechanisms that implement \( q^* \).
- at each \( t \) each \( i \) reports \( \theta_{i,t} \).
- public history in \( t : h_t = (r_0, q_0, r_1, q_1, \ldots r_{t-1}, q_{t-1}) \in H^t \), where each \( r_s = (r_{1,s}, \ldots, r_{i,s}) \).
- private history of \( i \):
  \[ h_{i,t} = (\theta_{i,0}, r_0, \theta_{i,1}, r_1, q_1, \ldots, \theta_{i,t-1}, r_{t-1}, q_{t-1}, \theta_{i,t}) \in H_{i,t}. \]
- an (efficient) dynamic direct mechanism is represented by a family of allocations and monetary transfers, \( \{ q_t^*, p_t \}_{t=0}^\infty : q_t^* : \Theta \to Q, \text{ and } p_t : H_t \times \Theta \to \mathbb{R}^I. \)
- a (pure) reporting strategy for agent \( i \) in period \( t \):
  \( r_{i,t} : H_{i,t} \to \Theta_i \)
- for a given mechanism, the expected payoff of agent \( i \) from reporting strategy \( r_i = \{ r_{i,t} \}_{t=0}^\infty \) given the strategies \( r_{-i} = \{ r_{-i,t} \}_{t=0}^\infty \) is:

\[
\mathbb{E} \sum_{t=0}^\infty \delta^t \left[ v_i (q^* (r_t), \theta_{i,t}) - p_i (h_t, r_t) \right].
\]
a dynamic direct mechanism is *interim incentive compatible*, if for every agent and every history, truthtelling is a best response given that all other agents report truthfully.

a dynamic direct mechanism is *periodic ex-post incentive compatible* if truthtelling is a best response regardless of the history and the current state of the other agents.

*periodic ex-post* participation constraints:

\[ V_i(h_{i,t}) \geq O_i(h_{i,t}), \]

where \( O_i(h_{i,t}) \) is determined by payoffs to \( i \) along \( q^*_{-i} \).

a mechanism is *ex-post incentive compatible and individually rational* if it satisfies the periodic ex-post incentive and participation constraints.
Marginal Contribution

- *flow marginal contribution* \( m_i(\theta_t) \)

\[
M_i(\theta_t) = m_i(\theta_t) + \delta \mathbb{E} M_i(\theta_{t+1}) ,
\]

or

\[
m_i(\theta_t) \triangleq W(\theta_t) - W_{-i}(\theta_t) - \delta \mathbb{E} [W(\theta_{t+1}) - W_{-i}(\theta_{t+1})].
\]

or

\[
m_i(\theta_t) = \sum_{j=1}^{l} v_j(q_i^*, \theta_{j,t}) - \sum_{j \neq i} v_j(q_{-i,t}^*, \theta_{j,t}) + \\
\delta \left[ \mathbb{E} W_{-i}(\theta_{t+1} | q_i^*, \theta_t) - \mathbb{E} W_{-i}(\theta_{t+1} | q_{-i,t}^*, \theta_t) \right].
\]
Efficient transfer

- transfer $p_i^* (\theta_t)$ equates flow payoff to flow marginal contribution if:

  $$p_i^* (\theta_t) \triangleq v_i (q_t^*, \theta_{i,t}) - m_i (\theta_t). \quad (1)$$

- We refer to $\{q^* (\theta_t) \ p^* (\theta_t)\}$ the dynamic pivot mechanism.

  $$p_i^* (\theta_t) = \sum_{j \neq i} [v_j (q_{-i,t}^*, \theta_{j,t}) - v_j (q_t^*, \theta_{j,t})] + \delta [W_{-i} (\theta_{t+1} \mid q_{-i,t}^*, \theta_t) - W_{-i} (\theta_{t+1} \mid q_{-i,t}^*, \theta_t)].$$

- The transfer $p_i^* (\theta_t) \geq 0$.

**Theorem (Dynamic Pivot Mechanism)**

$\{q^* (\theta_t) \ p^* (\theta_t)\}$ is ex-post incentive compatible and individually rational.
- scheduling tasks
- discrete time, infinite horizon: \( t = 0, 1, \ldots \)
- common discount factor \( \delta \)
- finite number of agents: \( i \in \{0, 1, \ldots, I\} \)
- each agent \( i \) has a single task
- value of task for \( i \) is: \( v_i > 0 \)

- quasilinear utility: \( v_i - p_i \)
Assignment

- values are given wlog in descending order:

\[ v_0 > v_1 > \cdots > v_i > 0 \]

- marginal contribution of task \( i \): difference in welfare with \( i \) and without \( i \)

- efficient task assignment policy:
Marginal Contribution

- insert valuable task \( i \):
- raise the value of all future tasks: \( t > i \)
- marginal contribution \( M_i \):

\[
M_i = \sum_{t=0}^{l} \delta^t v_t - \left( \sum_{t=0}^{i-1} \delta^t v_t + \sum_{t=i+1}^{l} \delta^{t-1} v_t \right)
\]

or

\[
M_i = \sum_{t=i}^{l} \delta^t (v_t - v_{t+1}) \geq 0
\]
Externality

- from marginal contribution to externality pricing:
  \[ M_i = v_i - p_i \]

- externality cost of task \( i \) is:
  \[ p_i = v_{i+1} - \sum_{t=i+1}^{l} \delta^{t-i}(v_t - v_{t+1}) > 0 \]

- task \( i \) directly replaces task \( i + 1 \), but also:
- task \( i \) raises the value of all future tasks
Incomplete Information

- $v_i$ is private information to agent $i$ at $t = 0$
- incentive compatibility and efficient sorting: when would $i$ like to win against $j$ versus $j + 1$:

$$
(v_i - v_j) - \sum_{t=j}^{l} \delta^{t-(j-1)} (v_t - v_{t+1})
$$

$$
\geq \delta (v_i - v_{j+1}) - \sum_{t=j+1}^{l} \delta^{t-j} (v_t - v_{t+1})
$$

- reduces to cost of delay:

$$
(1 - \delta) v_i \geq (1 - \delta) v_j.
$$

- report truthful if others report truthful: ex post equilibrium
single unit auction, repeated over time

- discrete time, infinite horizon: \( t = 0, 1, \ldots \)
- finite number of bidders: \( i \in \{1, \ldots, I\} \)
- realized value of object for winning bidder in period \( t \) is

\[
V_{i,t} = \omega_i + \varepsilon_{i,t}
\]

- \( \varepsilon_{i,t} \) is i.i.d. over time with \( \mathbb{E}[\varepsilon_{i,t}] = 0 \)
- \( \omega_i \) is true value of object, \( \varepsilon_{i,t} \) is random noise
Information Flow

- at $t = 0$: common prior distribution $\pi_i(\omega_i)$ for each agent $i$
- winning bidder $i$ receives informative signal $v_{i,t}$:
  \[ v_{i,t} = \omega_i + \epsilon_{i,t} \]
- loosing bidder $j$ doesn't receive additional information:
  \[ v_{j,t} = 0 \]
- realized value is private information for winning bidder:
  \[ \theta_{i,t} \triangleq P_i [\omega_i | v_{i,0}, \ldots, v_{i,t-1}] \]
  with
  \[ \theta_{i,t} \in \Theta_i = \Delta(\Omega_i) \]
bidder $i$ reports $r_{i,t} \in \Theta_i$ in every period $t$

inductively, a history of reports:

$$h_t = (h_{t-1}, r_{1,t}, ..., r_{I,t})$$

allocation rule:

$$q_t : H_t \rightarrow [0, 1]^I$$

transfer (or pricing) rule is given by:

$$p_t : H_t \rightarrow \mathbb{R}^I$$
Strategies

- private history of bidder $i$:
  \[ h_{i,t} = (r_0, v_{i,0}, \ldots, r_{t-1}, v_{i,t-1}) \]

- reporting strategy for agent $i$:
  \[ r_{i,t} : H_{i,t} \rightarrow \Theta_i \]

- expected payoff for bidder $i$:
  \[ \mathbb{E} \sum_{t=0}^{\infty} \delta^t \left[ q_{i,t}(r_t) v_i(\theta_{i,t}) - p_{i,t}(h_t) \right] . \]

- reporting strategy of $i$ solves sequential optimization problem $V_i(h_{i,t})$:
  \[ \max_{r_{i,t} \in \Theta_i} \mathbb{E} \left\{ q_{i,t}(r_t) v_{i,t}(\theta_{i,t}) - p_{i,t}(r_t) + \delta V_i(h_{i,t+1}) \right\} \]
denote by \( \theta_{i,t} \triangleq \theta_t \setminus \theta_{i,t} \)

Bayesian incentive compatible if \( r_{i,t} = \theta_{i,t} \) solves

\[
\max_{\theta_{i,t} \in \Theta_i} \mathbb{E} \left\{ q_{i,t} \left( r_{i,t}, \theta_{-i,t} \right) v_i (\theta_{i,t}) - p_{i,t} \left( r_{i,t}, \theta_{-i,t} \right) + \delta V_i (h_{i,t+1}) \right\}
\]

periodic ex post: with respect to all the information available at period \( t \)

(periodic) ex post incentive compatible if \( r_{i,t} = \theta_{i,t} \) solves

\[
\max_{\theta_{i,t} \in \Theta_i} \left\{ q_{i,t} \left( r_{i,t}, \theta_{-i,t} \right) v_i (\theta_{i,t}) - p_{i,t} \left( \theta_{i,t}, \theta_{-i,t} \right) + \delta V_i (h_{i,t+1}) \right\}
\]

for all \( \theta_{-i,t} \in \Theta_{-i} \)
Social Efficiency

- socially efficient assignment policy

\[ W(\theta_u) = \max_{\{q_t\}_{t=u}^{\infty}} \mathbb{E} \sum_{t=u}^{\infty} \sum_{i=1}^{I} \delta^{t-u} q_{i,t}(\theta_t) v_i(\theta_{i,t}) \]

- optimal assignment is a multi–armed bandit problem
- optimal policy is an index policy:

\[ \gamma_i(\theta_{i,u}) = \max_{\tau} \mathbb{E} \left\{ \frac{\sum_{t=0}^{\tau} \delta^t v_i(\theta_{i,u+t})}{\sum_{t=0}^{\tau} \delta^t} \right\} \]

- socially efficient allocation policy \( q^* = \{q_{i,t}^*\}_{t=0}^{\infty} : \)

\[ q_{i,t}^* > 0 \text{ if } \gamma_i(\theta_{i,t}) \geq \gamma_j(\theta_{j,t}) \text{ for all } j. \]
Marginal Contribution

- value of social program after removing bidder $i$

\[ W_{-i} (\theta_u) = \max_{\{q_{-i,t}\}_{t=u}} \mathbb{E} \sum_{t=u}^{\infty} \sum_{j \neq i} \delta^{t-u} q_{j,t} (\theta_t) v_j (\theta_{j,t}) \]

- marginal contribution $M_i (\theta_t)$ of bidder $i$ at state $\theta_t$ is:

\[ M_i (\theta_t) = W (\theta_t) - W_{-i} (\theta_t) \]

- value $M$ conditional on state $\theta_{t+1}$ and allocation $q_t$:

\[ M (\theta_{t+1} | q_t) \]

- flow marginal contribution $m_i (\theta_t)$:

\[ M_i (\theta_t) = m_i (\theta_t) + \delta M_i (\theta_{t+1} | q_t^*) \]
Flow Marginal Contribution

- flow marginal contribution:

\[ m_i(\theta_t) = M_i(\theta_t) - \delta M_i(\theta_t, q_t^*) \]

- expanding flow term with respect to time

\[ m_i(\theta_t) = \left( W(\theta_t) - W_{-i}(\theta_t) \right) - \delta \left( W(\theta_{t+1} | q_t^*) - W_{-i}(\theta_{t+1} | q_t^*) \right) \]

- expanding flow term with respect to identity (rearranging)

\[ m_i(\theta_t) = \left( W(\theta_t) - \delta W(\theta_{t+1} | q_t^*) \right) - \left( W_{-i}(\theta_t) - \delta W_{-i}(\theta_{t+1} | q_t^*) \right) \]
Efficient Assignment

\[ m_i (\theta_t) = (W(\theta_t) - \delta W(\theta_{t+1} | q_t^*)) - (W_{-i}(\theta_t) - \delta W_{-i}(\theta_{t+1} | q_t^*)) \]

- consider efficient assignment \( q_t^* = i \):
- information about \( q_t^* = i \) is worthless without \( i \):
  \[ W_{-i}(\theta_{t+1} | i) = W_{-i}(\theta_t) \]
  leads to
  \[ m_i (\theta_t) = v_i(\theta_{i,t}) - (1 - \delta) W_{-i}(\theta_t) \]

- consider inefficient bidder: \( q_t^* \neq j \):
  \[ q_{-j,t}^* = q_t^* \]
  leads to
  \[ m_j (\theta_t) = 0 \]

Dirk Bergemann and Juuso Välimäki

The Dynamic Pivot Mechanism
Dynamic Second Price Auction

- match net payoff to flow marginal contribution
- for winner $i$:
  \[ m_i(\theta_t) = v_i(\theta_t) - p_i(\theta_t) \]
- for losers, $j \neq i$:
  \[ m_j(\theta_t) = -p_j(\theta_t) \]
Theorem (Dynamic Second Price Auction)

The socially efficient allocation rule \( a^* \) satisfies ex post incentive and participation constraints with payment \( p^* \):

\[
p^*_j (\theta_t) = \begin{cases} 
(1 - \delta) W_{-j} (\theta_t) & \text{if } a^*_{j,t} = 1, \\
0 & \text{if } a^*_{j,t} = 0.
\end{cases}
\]

- price equals intertemporal opportunity cost
- delay \((1 - \delta)\) of the optimal program for all but \( j \)
with private values, static mechanism satisfies incentive compatibility in weakly dominant strategies

in dynamic mechanism, dominant incentive compatibility fails to hold in private value environment

truth telling after all histories fails to be a weakly dominant strategy as it removes the ability to respond to past announcements

yet ex post incentive compatibility can be satisfied in dynamic mechanism
many transfer rules support ex post incentive and ex post participation constraints in dynamic setting

temporal separation between allocative influence and monetary payments may be undesirable for many reasons:

- agent $i$ could be tempted to leave the mechanisms and break her commitment \textit{after} she ceases to have a pivotal role but \textit{before} her payments come due
- if arrival and departure of agents is random, then an agent could falsely claim to depart to avoid future payments

in intertemporal environment if agent $i$ ceases to influence current or future allocative decisions in period $t$, then she also ceases to have monetary obligations
Efficient Exit

- presence of \( i \) is not influential for the future at \( \tau \) and \( \theta_\tau \) if

\[
q_\tau^* (\theta_t) = q_{i,t}^* (\theta_t),
\]

for all \( t \geq \tau \) and for all \( \theta_t \)

**Definition (Efficient Exit)**

A mechanism satisfies the efficient exit condition if \( \forall i, \tau, \theta_\tau : \)

\[
p_{i,t} (\theta_t) = 0, \text{ for all } t \geq \tau.
\]

- weak online requirement: decisions regarding agent \( i \) have to be made in the presence of agent \( i \)
Diverse Valuations

- diverse flow valuations:

**Definition (Diverse Valuations)**

The valuations are diverse if

1. $\forall i \exists \theta_i = \theta^0_i$ s. th. for $\forall a_t$

   $$v_{i,t} \left( a_t, \theta^0_i \right) = 0 \text{ and } P_i \left( \theta^0_i; a_t, \theta^0_i \right) = 1;$$

2. $\forall i, a, x \in \mathbb{R}_+, \exists \theta_i$ s.th.

   $$v_{i,t} \left( a_t, \theta_i \right) = \begin{cases} x \text{ if } a_t = a, \\ 0 \text{ if } a_t \neq a, \end{cases} \text{ and } P_i \left( \theta_i; a_t, \theta_i \right) = 1.$$
Theorem (Uniqueness of Payoffs)

With diverse valuations, if a mechanism satisfies ex post incentive and participation constraints and the efficient exit condition, then the ex-ante expected value of agent $i$ is her marginal contribution.

- proof by contradiction: violation of EPIC or EPIR
- Green and Laffont (1977) provide characterization of Groves mechanism
- Moulin (1986) provides characterization of Pivotal mechanism
Diverse Departure

- departure probability $\lambda_i$: probability that agent $i$ will have constant valuation (at 0) starting tomorrow

**Definition (Diverse Departure)**

The departures are diverse if $\forall i, \theta_i, a_t$:

$$P_i \left( \theta_i^0; a_t, \theta_i \right) > 0$$

and if $\forall \theta_i, \lambda \in (0, 1), \exists \theta_i^\lambda$ s.th.:

$$v_{i,t} (\cdot, \theta_i) = v_{i,t} (\cdot, \theta_i^\lambda),$$

and

$$P_i \left( \cdot; a_t, \theta_i^\lambda \right) = \lambda P_i (\cdot; a_t, \theta_i) + (1 - \lambda) \mathbb{I} \left( \theta_i^0 \right).$$
Theorem (Uniqueness of Mechanism)

*With diverse valuations and departures, if a mechanism satisfies ex post incentive and participation constraints and the efficient exit condition, then it is the dynamic pivotal mechanism.*

- show that the flow transfers cannot depend on the departure probability
- analyze incentive constraints for types who only differ in their departure probability
- incentive compatibility and earlier result of marginal contribution payoff can only be reconciled with transfers of the dynamic pivotal mechanism
direct dynamic mechanism in private value environments with transferable utility

design of monetary transfers relies on notions of marginal contribution and flow marginal contribution

transfer the insights of pivotal mechanism from static to dynamic settings

interesting question: interdependent value environments