***Константин Сонин***
***Российская экономическая школа***
***Волгоград, 22-25 апреля 2002 года***

# Лекции по ЭКОНОМИКЕ ОБЩЕСТВЕННОГО СЕКТОРА

Лекция 3
***Определение механизма. Схема Гровса***

## Основные понятия

- Конечное множество агентов (участников, игроков) $N = \{1, 2, ..., N\}$
- У каждого агента $i$ есть множество типов $\theta_i \in \Theta_i$, $\Theta = \times_i \Theta_i$
  - интерпретация: тип игрока отражает доступную ему (но не другим) информацию
- Множество возможных (общественных) *решений* $A$
- У каждого агента $i$ есть предпочтения относительно возможных решений; предпочтения зависят от типа

$$v_i : A \times \Theta_i \longrightarrow \mathbf{R}$$

  - мы предполагаем, что предпочтения агента $i$ зависят только от типа самого участника $\theta_i$, но не от типов остальных участников $\theta_{-i}$ ("индивидуальные ценности"="private values")
- *Функция общественного выбора*

$$\begin{aligned} f &: \Theta \longrightarrow A \\ &: \theta \longrightarrow f(\theta) \end{aligned}$$

- *Ex post эффективность* (функции общественного выбора)*:*

$$\sum_i v_i(f(\theta), \theta_i) \geq \sum_i v_i(f', \theta_i)$$

для всех $\theta \in \Theta$ и $f' \in A$

- *Трансферы*:

$$t : \Theta \longrightarrow \mathbf{R}^N$$

  – трансферы *доступны,* если для всех $\theta \in \Theta$

$$\sum_i t_i(\theta) \leq 0$$

  – трансферы *сбалансированы,* если для всех $\theta \in \Theta$

$$\sum_i t_i(\theta) = 0$$

- Трансферы входят в общую полезность квази-линейно: если "объявлен" профиль типов $\hat{\theta}$, а настоящие типы - $\theta$, то

$$u_i = v_i(f(\hat{\theta}), \theta_i) + t_i(\hat{\theta})$$

## Пример: Общественное благо

- $N$ агентов размышляют - не произвести ли какое общественное благо?
- Возможные исходы $(f, t) \in \{0, 1\} \times \mathbf{R}^N$
- Стоимость постройки равна $c \geq 0$
- Полезность агента $i$ есть $u_i = \theta_i f(\theta) + t_i$
- *Ex post эффективность:*

$$f(\theta) = \begin{cases} 1, & \text{если } \sum_i \theta_i \geq 0 \\ 0, & \text{если } \sum_i \theta_i < 0 \end{cases}$$

- Сбалансированность $\sum_i t_i(\theta) = -f(\theta)$
- Есть ли у агентов стимулы выявлять свои истинные предпочтения?

# Механизм

- *Механизм* - это пара $(M, g)$, где $M = M_1 \times ... \times M_N$ - пространство стратегий (действий, посланий), а $g$ - функция исходов

$$g : M \longrightarrow A \times \mathbf{R}^N,$$

то есть $g$ определяет решение (общественный выбор) и трансферы

- Зафиксируем механизм $(M, g)$. Стратегия $m_i \in M_i$ является *доминирующей* для типа $\theta_i$ агента $i$, если

$$v_i(g(m_i, m_{-i}), \theta_i) + t_i(m_i, m_{-i}) \geq v_i(g(m_i', m_{-i}), \theta_i) + t_i(m_i', m_{-i})$$

для всех $m_i' \in M_i$ и всех $m_{-i} \in M_{-i}$

- Механизм $(M, g)$ *реализует* функцию общественного выбора $f$ *в доминирующих стратегиях,* если существуют такие функции

$$m_i : \Theta_i \longrightarrow M_i,$$

что $m_i(\theta_i)$ является доминирующей стратегией для любого $i$ и $\theta_i \in \Theta_i$ и

$$g(m(\theta)) = f(\theta)$$

для всех $\theta \in \Theta$

# Прямые механизмы

- Механизм называется *прямым*, если для каждого $i$ пространство стратегий $M_i$ есть множество возможных типов агента $i$, и сама функция общественного выбора $f$ определяет соответствующее решение

- Прямой механизм $(f, t)$ *совместим со стимулами в доминирующих стратегиях (dominant strategy incentive compatible),* если для каждого типа $\theta_i$ каждого агента $i$ сообщение правды о своем типе является доминирующей стратегией:

$$v_i(f(\theta_i, \theta_{-i}), \theta_i) + t_i(\theta_i, \theta_{-i}) \geq v_i(f(\theta_i', \theta_{-i}), \theta_i) + t_i(\theta_i', \theta_{-i})$$

для всех $\theta_i' \in M_i = \Theta_i$ и всех $\theta_{-i} \in M_{-i} = \Theta_{-i}$

## Принцип выявления

- *Если механизм $(M, g)$ реализует функцию общественного выбора $(f, t)$ в доминирующих стратегиях, то соответствующий прямой механизм совместим со стимулами в доминирующих стратегиях*
- Это принцип (коллекция теорем), а не теорема - он верен всегда (неформально)!

## Пример: Частное благо

- Продавец продает один объект
  - (пока) считаем, что продавец ценит объект ровно в 0
- Для каждого игрока $i \in N$ объект имеет ценность $\theta_i \in \Theta_i$
- Каждый игрок знает ценность обьекта для себя, но не знает ценностей остальных (индивидуальные ценности)
  - $\theta_i$ получена из распределения $\Phi_i$ на множестве $\Theta_i$
- Пусть есть только продавец и 2 покупателя ($N = 3$)
  - продавец - игрок с номером 0
  - распределения $\Phi_i$ - равномерные на множестве $[0, 1]$ и независимые
- Рассмотрим такой пример механизма

## Плохой механизм (невыявляющий)

- Механизм
  - функция общественного выбора $f(\theta) = (f_0(\theta), f_1(\theta), f_2(\theta))$

$$f_1(\theta) = \begin{cases} 1, & \theta_1 \geq \theta_2 \\ 0, & \theta_1 < \theta_2 \end{cases}$$

$$f_2(\theta) = \begin{cases} 1, & \theta_1 < \theta_2 \\ 0, & \theta_1 \geq \theta_2 \end{cases}$$

$$f_0(\theta) = 0$$

- трансферы

$$
\begin{aligned}
t_1(\theta) &= -\theta_1 f_1(\theta) \\
t_2(\theta) &= -\theta_2 f_2(\theta) \\
t_0(\theta) &= -\left(t_1(\theta) + t_2(\theta)\right)
\end{aligned}
$$

- Игроки максимизируют ожидаемую полезность
- ВОПРОС: Если один из участников (пусть 2-ой) говорит правду, то имеет ли другой (1-ый) стимулы говорить правду?
  - Для каждого $\theta_1$, 1-ый решает задачу

$$\max_{\hat{\theta}_1}(\theta_1 - \hat{\theta}_1)\Pr(\theta_2 \leq \hat{\theta}_1) = \max_{\hat{\theta}_1}(\theta_1 - \hat{\theta}_1)\hat{\theta}_1$$

  - Решение этой задачи

$$\hat{\theta}_1 = \frac{1}{2}\theta_1,$$

    то есть ответ на вопрос - *нет*.

- У первого агента есть выбор при решении его оптимизационной задачи
  - назвать свою ценность побольше, что объект доставался ему чаще
  - назвать свою ценность поменьше, чтобы заплатить поменьше в тех случаях, когда объект достается ему

## Другой механизм (выявляющий)

- Функция $f(\theta) = (f_0(\theta), f_1(\theta), f_2(\theta))$ та же самая, а трансферы другие

$$
\begin{aligned}
t_1(\theta) &= -\theta_2 f_1(\theta) \\
t_2(\theta) &= -\theta_1 f_2(\theta) \\
t_0(\theta) &= -\left(t_1(\theta) + t_2(\theta)\right)
\end{aligned}
$$

- Этот механизм выявляющий (то есть у агентов есть стимулы говорить правду, если другой участник говорит правду - и даже если другой участник говорит неправду)!

## Аукцион первой цены

- По-прежнему есть только продавец и 2 покупателя ($N = 3$)
  - продавец - игрок с номером 0
  - распределения $\Phi_i$ - равномерные на множестве $[0, 1]$ и независимые
- Рассмотрим такой механизм (*аукцион первой цены)*
  - участники аукциона подают заявки $b_i \geq 0$ в конвертах
  - тот, чья заявка оказалась больше (после открытия конвертов) получает объект и платит то, что написано у него в конверте ($b_i$)
- Будем искать равновесие в виде $b_i(\theta_i) = \alpha_i \theta_i$, где $\alpha_i \in [0, 1]$
- Допустим, игрок 2 придерживается такой стратегии
  - для каждого $\theta_1$, 1-ый решает задачу

$$\max_{b_1} (\theta_1 - b_1) \Pr(b_2(\theta_2) \leq b_1) = \max_{b_1 \in [0, \alpha_2]} (\theta_1 - b_1) \frac{b_1}{\alpha_2}$$

  - решение

$$b_1(\theta_1) = \begin{cases} \frac{1}{2}\theta_1, \ \frac{1}{2}\theta_1 \leq \alpha_2 \\ \alpha_2, \ \frac{1}{2}\theta_1 > \alpha_2 \end{cases}$$

$$b_2(\theta_2) = \begin{cases} \frac{1}{2}\theta_2, \ \frac{1}{2}\theta_2 \leq \alpha_1 \\ \alpha_1, \ \frac{1}{2}\theta_2 > \alpha_1 \end{cases}$$

  - если $\alpha_1 = \alpha_2 = \frac{1}{2}$, то $(b_1(\theta_1), b_2(\theta_2))$ - равновесие (по Байесу-Нэшу)
- Соответствующий прямой механизм
  - функция общественного выбора $f(\theta) = (f_0(\theta), f_1(\theta), f_2(\theta))$

$$f_1(\theta) = \begin{cases} 1, \ \theta_1 \geq \theta_2 \\ 0, \ \theta_1 < \theta_2 \end{cases}$$

$$f_2(\theta) = \begin{cases} 1, \ \theta_1 < \theta_2 \\ 0, \ \theta_1 \geq \theta_2 \end{cases}$$

$$f_0(\theta) = 0$$

– трансферы

$$\begin{aligned} t_1(\theta) &= -\frac{1}{2}\theta_1 f_1(\theta) \\ t_2(\theta) &= -\frac{1}{2}\theta_2 f_2(\theta) \\ t_0(\theta) &= -(t_1(\theta)+t_2(\theta)) \end{aligned}$$

### Аукцион второй цены

- Еще раз - продавец (игрок с номером 0) и 2 покупателя ($N = 3$)
  - распределения $\Phi_i$ - равномерные на множестве $[0, 1]$ и независимые
- Рассмотрим такой механизм (*аукцион второй цены)*
  - участники аукциона подают заявки $b_i \geq 0$ в конвертах
  - тот, чья заявка оказалась больше (после открытия конвертов) получает объект и платит то, что написано у другого (если участников больше 2-х, то *второго*) в конверте
- В этой игре $b_i(\theta_i) = \theta_i$ является (слабо) доминирующей стратегией

## Схема Гровса

### Пример

- Три человека думают - не построить ли дорогу?
- Стоимость дороги - $c$, полезность каждого участника от построения дороги равна $\theta_i$
- Доброжелатель предлагает следующий механизм:
  - Каждый участник платит $b_i = b_i(\theta_i)$
  - Платежи агента $i$ в механизме равны

$$u_i(b_1, b_2, b_3, \theta_i) = \begin{cases} 0, & \text{если } b < c \\ \theta_i - (c - b_j - b_k), & \text{если } b \geq c \text{и } b_j + b_k < c \\ \theta_i, & \text{если } b < c \text{и } b_j + b_k \geq c \end{cases}$$

  - То есть дорога строится, если набирается нужное количество денег
- Для каждого участника его действие - $b_i = b_i(\theta_i)$ *не* влияет на то, сколько он платит - только на то, строится дорога или нет
- Сообщение правды является доминирующей стратегией

- Проблемы со схемой Гровса:
  - Сумма, выплаченная всеми участниками в случае построения дороги, равна $3c-2b \leq c$, а, значит, без доброжелателя не обойтись
  - Чем выше суммарная ценность дороги для участников, тем *больше* должен заплатить доброжелатель - у игроков есть стимулы сговориться

## Пример с 2 игроками

- Два участника решают начать проект $\in \{0, 1\}$
- Для каждого участника выигрыш от завершения проекта равен $\theta_i \in (-\infty, \infty)$
- Игроки знают только свой собственный выигрыш
- *Ex post* эффективность требует, чтобы проект предпринимался тогда и только тогда, когда $\theta_1 + \theta_2 \geq 0$, то есть функция общественного выбора есть

$$f^*(\theta) = \begin{cases} 1, & \text{если } \theta_1 + \theta_2 \geq 0 \\ 0, & \text{если } \theta_1 + \theta_2 < 0 \end{cases}$$

- Из принципа выявления следует, что можно ограничиться прямыми механизмами $(f, t)$, где $f$ отображает действия (объявления типа) в решения, а $t$ определяет трансферы

$$\begin{aligned} f &: \mathbf{R}^2 \longrightarrow \{0, 1\} \\ t &: \mathbf{R}^2 \longrightarrow \mathbf{R}^2 \end{aligned}$$

- Что требуется от механизма?
  - *Ex post* эффективность: $f(\theta) = f^*(\theta)$
  - Совместимость со стимулами и доминирование: для всех $\theta_{-i}$

$$\theta_i \in \arg\max_{\hat{\theta}_i} \{\theta_i f(\hat{\theta}_i, \hat{\theta}_{-i}) + t_i(\hat{\theta}_i, \hat{\theta}_{-i})\}$$

- Механизм Гровса:
  - если $\hat{\theta} = (\hat{\theta}_1, \hat{\theta}_2)$ - объявленные типы, то

$$f(\hat{\theta}) = \begin{cases} 1, & \text{если } \hat{\theta}_1 + \hat{\theta}_2 \geq 0 \\ 0, & \text{если } \hat{\theta}_1 + \hat{\theta}_2 < 0 \end{cases}$$

– $t_i(\hat{\theta}) = f(\hat{\theta})\hat{\theta}_{-i} + h_i(\hat{\theta}_{-i})$, где $h_i : \mathbf{R} \longrightarrow \mathbf{R}$ - некоторая функция

- Механизм Гровса соответствует требованиям - говорить правду является доминирующей стратегией

  – Заметим, что
  $$\theta_i \in \arg\max_{\hat{\theta}_i}\{\theta_i f(\hat{\theta}_i, \hat{\theta}_{-i}) + t_i(\hat{\theta}_i, \hat{\theta}_{-i})\}$$
  тогда и только тогда, когда
  $$\theta_i \in \arg\max_{\hat{\theta}_i}\{f(\hat{\theta}_i, \hat{\theta}_{-i})(\theta_i + \hat{\theta}_{-i})\}$$

  – $\hat{\theta}_i = \theta_i$ действительно максимизирует последнее выражение, так как если $\hat{\theta}_i = \theta_i$, то $f(\hat{\theta}) = 1$ тогда и только тогда, когда $\theta_i + \hat{\theta}_{-i} \geq 0$

- Трансферы $t(\hat{\theta})$ устроены так, что задача, которую решает каждый игрок, совпадает с задачей социального планировщика

## Теорема Гровса - Грина - Лаффона

**Теорема.** *(I) Пусть $f$ - ex post эффективная функция общественного выбора, $t$ - функция трансферов и для каждого $i \in N$ существует такая функция $x_i : \times_{j \neq i}\Theta_j \longrightarrow \mathbf{R}$, что*
$$t_i(\theta) = x_i(\theta_{-i}) + \sum_{j \neq i} v_j(f(\theta), \theta_j). \qquad (*)$$
*Тогда прямой механизм $(f, t)$ реализует $f$ в доминирующих стратегиях.*
*(II) Обратно, пусть $f$ - ex post эффективная функция общественного выбора, прямой механизм $(f, t)$ реализует $f$ в доминирующих стратегиях и, кроме того,*
$$\{v_i(\cdot, \theta_i) | \theta_i \in \Theta_i\} = \{v : D \longrightarrow \mathbf{R}\}$$
*для каждого $i \in N$. Тогда для каждого $i \in N$ существует функция $x_i : \times_{j \neq i}\Theta_j \longrightarrow \mathbf{R}$,* удовлетворяющая $(*)$.

### Доказательство

- (I) Допустим, что прямой механизм $(f, t)$ **не** реализует $f$ в доминирующих стратегиях
- Существуют такие $i$, $\theta$, $\hat{\theta}_i$, что
$$v_i(f(\hat{\theta}_i, \theta_{-i}), \theta_i) + t_i(\hat{\theta}_i, \theta_{-i}) > v_i(f(\theta_i, \theta_{-i}), \theta_i) + t_i(\theta_i, \theta_{-i})$$

- То есть, используя $(*)$ :

$$v_i(f(\hat\theta_i,\theta_{-i}),\theta_i)+x_i(\theta_{-i})+\sum_{j\neq i}v_j(f(\hat\theta_i,\theta_{-i}),\theta_j)>v_i(f(\theta),\theta_i)+x_i(\theta_{-i})+\sum_{j\neq i}v_j(f(\theta),\theta_j)$$

- $x_i(\theta_{-i})$ сокращается, а

$$v_i(f(\hat\theta_i,\theta_{-i}),\theta_i)+\sum_{j\neq i}v_j(f(\hat\theta_i,\theta_{-i}),\theta_j)>v_i(f(\theta),\theta_i)+\sum_{j\neq i}v_j(f(\theta),\theta_j),$$

  что противоречит *ex post* эффективности функции $f$.

- (II) Существует такая функция $x_i : \times_j\Theta_j \longrightarrow \mathbf{R}$, что

$$t_i(\theta)=x_i(\theta)+\sum_{j\neq i}v_j(f(\theta),\theta_j).$$

- Надо лишь доказать, что $x_i(\theta)$ не зависит от $\theta_i$

- От противного: предположим, что существуют такие $i$, $\theta$, $\hat\theta_i$, что

$$x_i(\theta)>x_i(\hat\theta_i,\theta_{-i})$$

- Положим

$$\varepsilon=\frac{1}{2}\Big(x_i(\theta)-x_i(\hat\theta_i,\theta_{-i})\Big)$$

- Из реализуемости функции $f$ в доминирующих стратегиях следует, что $f(\theta)\neq f(\hat\theta_i,\theta_{-i})$

  – Иначе:

$$x_i(\theta)=v_i(f(\theta),\theta_i)+t_i(\theta_i,\theta_{-i})>v_i(f(\theta),\theta_i)+t_i(\hat\theta_i,\theta_{-i})=x_i(\hat\theta_i,\theta_{-i})$$

  и, следовательно,

$$t_i(\theta_i,\theta_{-i})>t_i(\hat\theta_i,\theta_{-i})$$

  – Теперь

$$v_i(f(\theta_i,\theta_{-i}),\hat\theta_i)+t_i(\theta_i,\theta_{-i})>v_i(f(\hat\theta_i,\theta_{-i}),\hat\theta_i)+t_i(\hat\theta_i,\theta_{-i}),$$

  то есть если истинный тип есть $\hat\theta_i$, то выгоднее сообщать $\theta_i$ (ведь $f(\theta)=f(\hat\theta_i,\theta_{-i})$!), что противоречит определению доминирующих стратегий

- Поскольку множество всех возможных предпочтений $\Theta_i$ содержит все возможные предпочтения, найдется такой тип $\tilde{\theta}_i$, что

$$v_i(f(\hat{\theta}_i, \theta_{-i}), \tilde{\theta}_i) + \sum_{j \neq i} v_j(f(\hat{\theta}_i, \theta_{-i}), \theta_j) = \varepsilon,$$

и, кроме того,

$$v_i(f') + \sum_{j \neq i} v_j(f') = 0$$

для всех $f' \neq f(\hat{\theta}_i, \theta_{-i})$

- Поскольку $f$ - *ex post* эффективно, $f(\tilde{\theta}_i, \theta_{-i}) = f(\hat{\theta}_i, \theta_{-i})$

  – Иначе получилось бы, что

$$v_i(f(\tilde{\theta}_i, \theta_{-i})) + \sum_{j \neq i} v_j(f(\tilde{\theta}_i, \theta_{-i})) = 0 < v_i(f(\hat{\theta}_i, \theta_{-i}), \tilde{\theta}_i) + \sum_{j \neq i} v_j(f(\hat{\theta}_i, \theta_{-i}), \theta_j) = \varepsilon$$

- Из реализуемости функции $f$ в доминирующих стратегиях следует, что $t_i(\tilde{\theta}_i, \theta_{-i}) = t_i(\hat{\theta}_i, \theta_{-i})$

  – Иначе (если $t_i(\tilde{\theta}_i, \theta_{-i}) > t_i(\hat{\theta}_i, \theta_{-i})$) получилось бы, что

$$v_i(f(\tilde{\theta}_i, \theta_{-i}), \hat{\theta}_i) + t_i(\tilde{\theta}_i, \theta_{-i}) > v_i(f(\hat{\theta}_i, \theta_{-i}), \hat{\theta}_i) + t_i(\hat{\theta}_i, \theta_{-i}),$$

  – Аналогично рассматривается случай $t_i(\tilde{\theta}_i, \theta_{-i}) < t_i(\hat{\theta}_i, \theta_{-i})$

- Если истинный тип есть $\tilde{\theta}_i$, то

  – правдивое объявление дает

$$\begin{aligned} u_i &= v_i(f(\tilde{\theta}_i, \theta_{-i}), \tilde{\theta}_i) + t_i(\tilde{\theta}_i, \theta_{-i}) \\ &= v_i(f(\hat{\theta}_i, \theta_{-i}), \tilde{\theta}_i) + t_i(\tilde{\theta}_i, \theta_{-i}) \\ &= \varepsilon - \sum_{j \neq i} v_j(f(\hat{\theta}_i, \theta_{-i}), \theta_j) + t_i(\tilde{\theta}_i, \theta_{-i}) \\ &= \varepsilon - \sum_{j \neq i} v_j(f(\hat{\theta}_i, \theta_{-i}), \theta_j) + t_i(\hat{\theta}_i, \theta_{-i}) \\ &= \varepsilon + x_i(\hat{\theta}_i, \theta_{-i}) \end{aligned}$$

  – ложь ($\theta_i$ вместо $\tilde{\theta}_i$) дает

$$u_i = v_i(f(\theta_i, \theta_{-i}), \tilde{\theta}_i) + t_i(\theta_i, \theta_{-i}) = x_i(\theta)$$

- Это противоречит совместимости со стимулами (реализуемости с помощью прямого механизма в доминирующих стратегиях), так как

$$x_i(\theta) > \varepsilon + x_i(\hat{\theta}_i, \theta_{-i}).$$

- Значит, наше предположение о том, что $x_i(\cdot)$ зависит от $\theta_i$ было неверно. ■

## Механизм Кларка (решающего голоса)

- Механизм "решающего голоса" - частный случай схемы Гровса (придуманный независимо Кларком)
- Положим

$$x_i(\theta_{-i}) = -\max_{a \in A} \sum_{j \neq i} v_j(a, \theta_j)$$

- Тогда функции трансферов заданы как

$$t_i(\theta) = \sum_{j \neq i} v_j(f(\theta), \theta_j) - \max_{a \in A} \sum_{j \neq i} v_j(a, \theta_j)$$

- $\sum_{i \in N} t_i(\theta) \leq 0$, так что этот механизм всегда доступен
- Интуиция:
  - если присутствие (вклад) агента $i$ не влияет на выбор $a$, то $t_i(\theta) = 0$
  - иначе, $t_i(\theta)$ есть потери остальных игроков, появляющиеся из-за присутствия агента $i$
  - когда агент $i$ решает свою собственную задачу и учитывает размеры трансфера, его задача в точности совпадает с задачей планировщика - исход эффективен

## Проблемы баланса

- Вернемся к задаче с двумя игроками
- Если механизм сбалансирован, то

$$t_1(\theta) + t_2(\theta) = f(\theta_1, \theta_2)(\theta_1 + \theta_2) + h_1(\theta_2) + h_2(\theta_1) = 0$$

- Отсюда

$$h_1(\theta_2) + h_2(\theta_1) = \begin{cases} -(\theta_1 + \theta_2), \text{ если } \theta_1 + \theta_2 \geq 0 \\ 0, \text{иначе}, \end{cases}$$

что невозможно, если $h_i$ не зависит от $\theta_i$

- Что можно сделать? Ослабим требования к равновесию

## Реализация в виде байесовского равновесия

- Вместо "$\theta_i \in \arg\max_{\hat{\theta}_i}\{\theta_i f(\hat{\theta}_i, \hat{\theta}_{-i}) + t_i(\hat{\theta}_i, \hat{\theta}_{-i})\}$ для всех $\hat{\theta}_{-i}$" (реализация в доминирующих стратегиях), потребуем

$$\theta_i \in \arg\max_{\hat{\theta}_i} E_{\hat{\theta}_{-i}}[\theta_i f(\hat{\theta}_i, \hat{\theta}_{-i}) + t_i(\hat{\theta}_i, \hat{\theta}_{-i})|\theta_i]$$

- Требование эффективности $f(\theta) = f^*(\theta)$ остается в силе
- Кроме этого, мы хотим, чтобы механизм был сбалансирован

$$t_1(\theta) + t_2(\theta) = 0$$

- Эрроу (1979) и д'Аспремон и Жерар-Варет (1979) показали, что можно построить механизм, соответствующий всем трем требованиям
- Рассмотрим байесовскую игру $\Gamma = \{\hat{\Theta}_1, \hat{\Theta}_2; \Theta_1, \Theta_2; p_1, p_2; u_1, u_2\}$
  - $\hat{\Theta}_i = \Theta_i = \mathbf{R}$,
  - $u_i(\hat{\theta}, \theta) = \theta_i f(\hat{\theta}) + t_i(\hat{\theta})$
  - $p_i(\theta_{-i}|\theta_i)$ не зависит от $\theta_i$ (то есть типы независимы)
- Механизм
  - если $\hat{\theta} = (\hat{\theta}_1, \hat{\theta}_2)$ - объявленные типы, то

$$f(\hat{\theta}) = \begin{cases} 1, & \text{если } \hat{\theta}_1 + \hat{\theta}_2 \geq 0 \\ 0, & \text{если } \hat{\theta}_1 + \hat{\theta}_2 < 0 \end{cases}$$

  - $t_i(\hat{\theta}) = g_i(\hat{\theta}_i) - g_{-i}(\hat{\theta}_{-i})$, где

$$g_i(\hat{\theta}_i) = E_{\theta_{-i}}\theta_{-i} f(\hat{\theta}_i, \theta_{-i})$$

- Очевидно - механизм сбалансирован

- Стимулы:
  - $$\hat{\theta}_i \in \arg\max_{\tilde{\theta}_i} E_{\theta_{-i}}[\theta_i f(\tilde{\theta}_i, \theta_{-i}) + t_i(\tilde{\theta}_i, \theta_{-i})]$$

    ($\theta_i$ - истинный тип, $\tilde{\theta}_i$ - возможный объявляемый тип, $\hat{\theta}_i$ - выбранный объявляемый тип)
  - Эквивалентно,
    $$\hat{\theta}_i \in \arg\max_{\tilde{\theta}_i} E_{\theta_{-i}}(\theta_i + \theta_{-i}) f(\tilde{\theta}_i, \theta_{-i})$$
  - Эквивалентно,
    $$\hat{\theta}_i \in \arg\max_{\tilde{\theta}_i} \int_{-\tilde{\theta}_i}^{\infty} (\theta_i + \theta_{-i}) d\Phi_{-i},$$
    где $\Phi_{-i}$ - функция распределения $\theta_{-i}$
  - Следовательно, $\hat{\theta}_i = \theta_i$
- К сожалению, и с этим механизмом есть проблемы
  - не исключено, что какой-то игрок, узнав свой тип, воздержится от участия в механизме - потому что механизм принесет ему, в ожидании, меньше, чем статус-кво (ноль)