The converse envelope theorem

Ludvig Sinander
University of Oxford

19 July 2023

paper: doi.org/10.3982/ECTA18119

Envelope theorem: optimal decision-making $\implies$ ⌧ formula.

Textbook intuition: ⌧ formula $\iff$ FOC.

Modern envelope theorem of MS02:* almost no assumptions.

$\hookrightarrow$ FOC ill-defined, so need different intuition.

My theorem: with almost no assumptions,
⌧ formula equivalent to generalised FOC.

- an envelope theorem: FOC $\implies$ ⌧
- *a converse:* ⌧ $\implies$ FOC.

Application to mechanism design.

*Milgrom, P., & Segal, I. (2002). Envelope theorems for arbitrary choice sets. *Econometrica*, *70*(2), 583–601. https://doi.org/10.1111/1468-0262.00296

Setting

Agent chooses action x from a set $\mathcal{X}$.
Objective $f(x,t)$, where $t \in [0,1]$ is a parameter.

No assumptions on $\mathcal{X}$, almost none on f:

(1) $f(x,\cdot)$ differentiable for each $x \in \mathcal{X}$

(2) $\{f(x,\cdot)\}_{x\in\mathcal{X}}$ absolutely equi-continuous. (def'n: slide 17)

– a sufficient condition (maintained by MS02):

  (a) $f(x,\cdot)$ absolutely continuous $\forall x \in \mathcal{X}$, and

  (b) $t \mapsto \sup_{x\in\mathcal{X}} |f_2(x,t)|$ dominated by an integrable f'n.

– a stronger sufficient condition: f_2 bounded.

Decision rule: a map $X : [0,1] \to \mathcal{X}$.

Associated *value function:* $V_X(t) := f(X(t),t)$.

Envelope theorem

X *satisfies the* $\boxtimes$ *formula* iff

$$V_X(t) = V_X(0) + \int_0^t f_2(X(s), s)\mathrm{d}s \quad \text{for every } t \in [0,1].$$

Equivalently: V_X is absolutely continuous and

$$V_X'(t) = f_2(X(t), t) \quad \text{for a.e. } t \in (0,1).$$

X is *optimal* iff for every t, $X(t)$ maximises $f(\cdot, t)$.

Modern envelope theorem (MS02).[†]
Any optimal decision rule satisfies the $\boxtimes$ formula.

[†]Really a slight refinement of MS02.

Textbook intuition

Differentiation identity for $V_X(t) := f(X(t),t)$:

$$V_X'(t) = \underbrace{\frac{\mathrm{d}}{\mathrm{d}m} f(X(t+m),t)\bigg|_{m=0}}_{\text{'indirect effect'}} + \underbrace{f_2(X(t),t)}_{\text{'direct effect'}}.$$

Indirect effect: t's gain from mimicking $t+m$ (for small m).

$$\begin{aligned} & \text{indirect effect} = 0 && \text{(FOC)} \\ \iff \quad & V_X'(t) = \text{direct effect} && (\boxtimes \text{ formula}). \end{aligned}$$

Problem: 'indirect effect' (hence FOC) ill-defined!

- $f(\cdot,t)$ & X need not be differentiable.
- actions $\mathcal{X}$ need have no convex or topological structure.

The outer first-order condition

Disjuncture: in general, $\boxtimes$ formula $\not\Leftrightarrow$ FOC.

- one solution: add strong 'classical' assumptions. (slide 18)
- my solution: find the correct FOC!

Decision rule X *satisfies the outer FOC* iff

$$\frac{\mathrm{d}}{\mathrm{d}m} \int_r^t f(X(s+m), s)\mathrm{d}s \bigg|_{m=0} = 0 \quad \text{for all } r, t \in (0,1).$$

Motivation: given decision rule $X : [0,1] \to \mathcal{X}$,

- type s can 'mimic' $s+m$ by choosing $X(s+m)$.
- oFOC: if types $s \in [r,t]$ do this,
  it's collectively unprofitable (to first order).

Housekeeping

Housekeeping lemma. Under classical assump'ns,
oFOC $\iff$ classical FOC.

(sketch proof: slide 19)

Necessity lemma. Any optimal decision rule X satisfies oFOC & has $V_X(t) := f(X(t), t)$ absolutely continuous.

(sketch proof: slide 20)

Main theorem

Envelope theorem & converse.
For a decision rule $X : [0,1] \to \mathcal{X}$, the following are equivalent:

(1) X satisfies the oFOC

$$\frac{\mathrm{d}}{\mathrm{d}m} \int_r^t f(X(s+m), s)\mathrm{d}s \bigg|_{m=0} = 0 \quad \text{for all } r, t \in (0,1),$$

and $V_X(t) \coloneqq f(X(t), t)$ is absolutely continuous.

(2) X satisfies the ⌧ formula

$$V_X(t) = V_X(0) + \int_0^t f_2(X(s), s)\mathrm{d}s \quad \text{for every } t \in [0,1].$$

Main theorem

Envelope theorem & converse. For $X : [0,1] \to \mathcal{X}$, TFAE:

(1) X satisfies the oFOC, & $V_X(t) \coloneqq f(X(t),t)$ is AC.

(2) X satisfies the $\boxtimes$ formula.

$\Longrightarrow$: an envelope theorem.
Implies the MS02 envelope theorem.

$\Longleftarrow$: *converse* envelope theorem.

The key lemma

Textbook intuition relied on differentiation identity

$$V_X'(s) = \underbrace{\frac{\mathrm{d}}{\mathrm{d}m} f(X(s+m), s)\bigg|_{m=0}}_{\text{`indirect effect'}} + \underbrace{f_2(X(s), s)}_{\text{`direct effect'}},$$

or (integrated & rearranged)

$$\int_r^t \frac{\mathrm{d}}{\mathrm{d}m} f(X(s+m), s)\bigg|_{m=0} \mathrm{d}s = V_X(t) - V_X(r) - \int_r^t f_2(X(s), s)\mathrm{d}s.$$

The 'outer' version is valid:

Identity lemma. If V_X is AC, then

$$\frac{\mathrm{d}}{\mathrm{d}m} \int_r^t f(X(s+m), s)\mathrm{d}s\bigg|_{m=0} = V_X(t) - V_X(r) - \int_r^t f_2(X(s), s)\mathrm{d}s.$$

(Where both sides are well-defined.)

(sketch proof: slide 21)

(proof of th'm: slide 24)

Application: environment

Agent with preferences $f(y, p, t)$ over
outcome $y \in \mathcal{Y}$ and payment $p \in \mathbf{R}$.

- $\mathcal{Y}$ partially ordered
- type $t \in [0, 1]$ is agent's private info
- assume single-crossing.

An *allocation* is $Y : [0, 1] \to \mathcal{Y}$.

Y is *implementable* iff $\exists$ payment rule $P : [0, 1] \to \mathbf{R}$
s.t. (Y, P) is incentive-compatible.

Application: goal

Classical result: implementable $\iff$ increasing.

'$\Longleftarrow$' is the substantial part. Versions: (lit: slide 25)

	literature	this paper
outcomes $\mathcal{Y}$	$\subseteq \mathbf{R}$	general
preferences f	quasi-linear	general.

Application: result

Implementability theorem. Under regularity assumptions, any increasing allocation is implementable.

Argument:

- fix an increasing allocation $Y : [0,1] \to \mathcal{Y}$
- choose a payment rule P so that $\boxtimes$ formula holds
- then by *converse envelope theorem*, oFOC holds $\iff$ mechanism (Y, P) is locally IC.
- finally, local IC $\implies$ global IC by single-crossing.

Application: example

Monopolist selling information.

Outcomes $\mathcal{Y}$:
distributions of posterior beliefs, ordered by Blackwell.

By the implementability theorem,
any Blackwell-increasing information allocation
can be implemented.

Application: details

Regular $\mathcal{Y}$: 'rich' & 'not too large'. (def'n: slide 26)

Examples:

- $\mathbf{R}^n$ ordered by 'coordinate-wise smaller'
- finite-expectation RVs ordered by 'a.s. smaller'
- distributions of posteriors updated from a given prior ordered by Blackwell.

Regular f: (def'n: slide 27)

(a) type derivative f_3 exists, bounded, continuous in p.

(b) f jointly continuous (when $\mathcal{Y}$ has order topology).

Single-crossing f: (def'n: slide 28)
if type t willing to pay to increase $y \in \mathcal{Y}$, then so is type $t' > t$.

Thanks!

Absolute equi-continuity

A family $\{\phi_x\}_{x\in\mathcal{X}}$ of functions $[0,1]\to\mathbf{R}$ is *AEC* iff the family

$$\left\{t \mapsto \sup_{x\in\mathcal{X}}\left|\frac{\phi_x(t+m)-\phi_x(t)}{m}\right|\right\}_{m>0} \quad \text{is uniformly integrable.}$$

Name inspired by the following (Fitzpatrick & Hunt, 2015):

AC–UI lemma. A continuous $\phi : [0,1] \to \mathbf{R}$ is AC iff

$$\left\{t \mapsto \frac{\phi(t+m)-\phi(t)}{m}\right\}_{m>0} \quad \text{is uniformly integrable.}$$

As name 'AEC' suggests, an AEC family

- is (uniformly) equi-continuous
- has AC functions as its members.

↪ back to slide 3

The classical approach

Classical assumptions:

– $\mathcal{X}$ is a convex subset of $\mathbf{R}^n$

– action derivative f_1 exists & is bounded

– only Lipschitz continuous decision rules X are considered.

(Bad for applications. Especially the Lipschitz restriction!)

$\Longrightarrow$ 'mimicking payoff' $\quad m \mapsto f(X(t+m), t) \quad$ diff'able a.e.

$\Longrightarrow$ FOC well-defined, differentiation identity valid.

Thus $\boxtimes$ formula $\iff$ FOC.

↪ back to slide 6

Sketch proof of the housekeeping lemma

Housekeeping lemma. Under classical assump'ns,
oFOC $\iff$ classical FOC.

Sketch proof. Fix a decision rule $X : [0,1] \to \mathcal{X}$.

Classical assump'ns & Vitali convergence theorem:

$$\underbrace{\frac{\mathrm{d}}{\mathrm{d}m} \int_r^t f(X(s+m), s)\mathrm{d}s \bigg|_{m=0}}_{\substack{= 0 \text{ for all } r,t \\ \text{iff oFOC holds}}} = \underbrace{\int_r^t \frac{\mathrm{d}}{\mathrm{d}m} f(X(s+m), s) \bigg|_{m=0} \mathrm{d}s}_{\substack{= 0 \text{ for all } r,t \\ \text{iff classical FOC holds.}}} .$$

■

↪ back to slide 7

Sketch proof of the necessity lemma

Necessity lemma. Any optimal decision rule X satisfies the oFOC & has $V_X(t) \coloneqq f(X(t),t)$ AC.

Sketch proof. X optimal & $\{f(x,\cdot)\}_{x\in\mathcal{X}}$ AEC $\implies$ V_X AC.

Since X optimal, have for any s and $m > 0 > m'$ that

$$\frac{f(X(s+m),s) - f(X(s),s)}{m} \leq 0 \leq \frac{f(X(s+m'),s) - f(X(s),s)}{m'}.$$

Integrating over (r,t) and letting $m, m' \to 0$,

both sides (in fact) converge to same limit:

$$\frac{\mathrm{d}}{\mathrm{d}m} \int_r^t f(X(s+m),s)\mathrm{d}s\bigg|_{m=0} \leq 0 \leq \frac{\mathrm{d}}{\mathrm{d}m} \int_r^t f(X(s+m),s)\mathrm{d}s\bigg|_{m=0}. \blacksquare$$

↪ back to slide 7

Sketch proof of the identity lemma I

$$\text{For } m>0 \text{, write} \quad \left.\frac{V_X(t+m)-V_X(t)}{m}\right\} =: \phi_m(t)$$

$$= \quad \left.\frac{f(X(t+m),t+m)-f(X(t+m),t)}{m}\right\} =: \psi_m(t)$$

$$+ \quad \left.\frac{f(X(t+m),t)-f(X(t),t)}{m}\right\} =: \chi_m(t).$$

$$\lim_{m\downarrow 0}\int_r^t \chi_m = \frac{\mathrm{d}}{\mathrm{d}m}\int_r^t f(X(s+m),s)\mathrm{d}s\Big|_{m=0} \quad \text{if limit exists.}$$

Must show: limit exists & equals

$$V_X(t)-V_X(r)-\int_r^t f_2(X(s),s)\mathrm{d}s.$$

Sketch proof of the identity lemma II

$$
\left.\begin{array}{lr}
 & \dfrac{V_X(t+m)-V_X(t)}{m}
\end{array}\right\} =: \phi_m(t)
$$

$$
= \left.\dfrac{f(X(t+m),t+m)-f(X(t+m),t)}{m}\right\} =: \psi_m(t)
$$

$$
+ \left.\dfrac{f(X(t+m),t)-f(X(t),t)}{m}\right\} =: \chi_m(t).
$$

$\{\psi_m\}_{m>0}$ need not converge a.e. (Even with strong assump'ns.)

But consider $\psi^\star_m(t) := \dfrac{f(X(t),t)-f(X(t),t-m)}{m}$.

$\{\psi^\star_m\}_{m>0}$ is UI & converges pointwise to $t \mapsto f_2(X(t),t)$. And

$$
\begin{aligned}
\int_r^t \psi_m = \int_{r+m}^{t+m} \psi^\star_m &= \int_r^t \psi^\star_m + \left(\int_t^{t+m} \psi^\star_m - \int_r^{r+m} \psi^\star_m\right) \\
&= \int_r^t \psi^\star_m + \mathrm{o}(1) \qquad \text{by UI.}
\end{aligned}
$$

Sketch proof of the identity lemma III

$$
\begin{aligned}
&\left.\frac{V_X(t+m)-V_X(t)}{m}\right\} =: \phi_m(t)\\
=\ &\left.\frac{f(X(t+m),t+m)-f(X(t+m),t)}{m}\right\} =: \psi_m(t)\\
+\ &\left.\frac{f(X(t+m),t)-f(X(t),t)}{m}\right\} =: \chi_m(t).
\end{aligned}
$$

$\int_r^t \psi_m = \int_r^t \psi_m^\star + \mathrm{o}(1)$, $\{\psi_m^\star\}_{m>0}$ UI & converges pointwise to $t \mapsto f_2(X(t),t)$.

V_X AC $\implies$ $\{\phi_m\}_{m>0}$ UI & converges a.e. to V_X'. So

$$
\begin{aligned}
\lim_{m\downarrow 0}\int_r^t \chi_m &= \lim_{m\downarrow 0}\int_r^t [\phi_m-\psi_m] = \lim_{m\downarrow 0}\int_r^t [\phi_m-\psi_m^\star]\\
\text{(Vitali)}\quad &= \int_r^t \lim_{m\downarrow 0}[\phi_m-\psi_m^\star] = \int_r^t \left[V_X'(s)-f_2(X(s),s)\right]\mathrm{d}s\\
\text{(FToC)}\quad &= V_X(t)-V_X(r)-\int_r^t f_2(X(s),s)\mathrm{d}s.
\end{aligned}
$$

↪ back to slide 10

Proof of the main theorem

Identity lemma. If V_X is AC, then

$$\underbrace{\frac{\mathrm{d}}{\mathrm{d}m}\int_r^t f(X(s+m),s)\mathrm{d}s\bigg|_{m=0}}_{\substack{=0 \text{ for all } r,t \\ \text{iff oFOC holds.}}} = \underbrace{V_X(t)-V_X(r)-\int_r^t f_2(X(s),s)\mathrm{d}s}_{\substack{=0 \text{ for all } r,t \\ \text{iff } \boxtimes \text{ formula holds.}}}\,.$$

Proof of main theorem. X satisfies the oFOC & V_X is AC $\implies$ identity lemma applies. So oFOC $\implies$ $\boxtimes$ formula.

X satisfies the $\boxtimes$ formula $\implies$ V_X is AC (by Lebesgue's FToC) $\implies$ identity lemma applies. So $\boxtimes$ formula $\implies$ oFOC. ■

↪ back to slide 10

Application: existing results

- $\mathcal{Y} \subseteq \mathbf{R}$
  - classical assump'ns — Mirrlees (1976), Spence (1974), Guesnerie and Laffont (1984)
  - no classical assump'ns — Nöldeke and Samuelson (2018)
- general $\mathcal{Y}$
  - quasi-linear f — Matthews and Moore (1987), García (2005)
  - general f — this paper.

↪ back to slide 12

Application: outcome regularity

A set $\mathcal{A}$ partially ordered by $\lesssim$ is

(1) *order-dense-in-itself* iff for any $a < a'$ in $\mathcal{A}$,
there is a $b \in \mathcal{A}$ such that $a < b < a'$,

(2) *chain-separable* iff for each chain $C \subseteq \mathcal{A}$,
there is a countable set $B \subseteq \mathcal{A}$ that is order-dense in C,[‡]

(3) *countably chain-complete* iff every countable chain in $\mathcal{A}$
with a lower (upper) bound in $\mathcal{A}$
has an infimum (a supremum) in $\mathcal{A}$.

(1) & (2): $\mathcal{A}$ 'rich'. (3): $\mathcal{A}$ 'not too large'.

Definition. $\mathcal{Y}$ is *regular* iff it satisfies properties (1)–(3).

↪ back to slide 15

[‡] $B \subseteq \mathcal{A}$ is *order-dense* iff for any $a < a'$ in $\mathcal{A}$, $\exists\, b \in B$ s.t. $a \lesssim b \lesssim a'$.

Application: preference regularity

Order topology on a set $\mathcal{A}$ partially ordered by $\lesssim$:
the topology generated by the open order rays

$$\{b \in \mathcal{A} : b < a\} \quad \text{and} \quad \{b \in \mathcal{A} : a < b\}.$$

Definition. f is *regular* iff

(a) type derivative f_3 exists & is bounded & continuous in p

(b) for any chain $\mathcal{C} \subseteq \mathcal{Y}$, f jointly continuous on $\mathcal{C} \times \mathbf{R} \times [0,1]$ when $\mathcal{C}$ has relative top'gy inherited from order top'gy on $\mathcal{Y}$.

↪ back to slide 15

Application: single-crossing

Definition. For $\phi : [0,1] \to \mathbf{R}$, upper & lower derivatives

$$\mathrm{D}^\star \phi(t) \coloneqq \limsup_{m \to 0} \frac{\phi(t+m) - \phi(t)}{m}$$
$$\mathrm{D}_\star \phi(t) \coloneqq \liminf_{m \to 0} \frac{\phi(t+m) - \phi(t)}{m}.$$

Partial upper/lower derivatives: $(\mathrm{D}^\star)_i$ & $(\mathrm{D}_\star)_i$.

Definition. f is *single-crossing* iff
for any increasing $Y : [0,1] \to \mathcal{Y}$ & any $P : [0,1] \to \mathbf{R}$,
mis-reporting payoff $U(r,t) \coloneqq f(Y(r), P(r), t)$ satisfies

$$\begin{aligned} & (\mathrm{D}^\star)_1 U(t,t) \geq 0 \quad \text{implies} \quad (\mathrm{D}_\star)_1 U(t,t') > 0 \quad \text{for } t' > t \\ \text{and} \quad & (\mathrm{D}_\star)_1 U(t,t) \leq 0 \quad \text{implies} \quad (\mathrm{D}^\star)_1 U(t,t') < 0 \quad \text{for } t' < t. \end{aligned}$$

↪ back to slide 15

References I

Fitzpatrick, P. M., & Hunt, B. R. (2015). Absolute continuity of a function and uniform integrability of its divided differences. *American Mathematical Monthly*, *122*(4), 362–366. https://doi.org/10.4169/amer.math.monthly.122.04.362

García, D. (2005). Monotonicity in direct revelation mechanisms. *Economics Letters*, *88*(1), 21–26. https://doi.org/10.1016/j.econlet.2004.12.022

Guesnerie, R., & Laffont, J.-J. (1984). A complete solution to a class of principal-agent problems with an application to the control of a self-managed firm. *Journal of Public Economics*, *25*(3), 329–369. https://doi.org/10.1016/0047-2727(84)90060-4

References II

Matthews, S., & Moore, J. (1987). Monopoly provision of quality and warranties: An exploration in the theory of multidimensional screening. *Econometrica*, *55*(2), 441–467. https://doi.org/10.2307/1913245

Milgrom, P., & Segal, I. (2002). Envelope theorems for arbitrary choice sets. *Econometrica*, *70*(2), 583–601. https://doi.org/10.1111/1468-0262.00296

Mirrlees, J. A. (1976). Optimal tax theory: A synthesis. *Journal of Public Economics*, *6*(4), 327–358. https://doi.org/10.1016/0047-2727(76)90047-5

Nöldeke, G., & Samuelson, L. (2018). The implementation duality. *Econometrica*, *86*(4), 1283–1324. https://doi.org/10.3982/ECTA13307

References III

Spence, M. (1974). Competitive and optimal responses to signals: An analysis of efficiency and distribution. *Journal of Economic Theory*, *7*(3), 296–332. https://doi.org/10.1016/0022-0531(74)90098-2