Brown University 250th Anniversary Symposium
John von Neumann Lecture on Economics

Efficient Mechanisms for Level-*k* Bilateral Trading

12 May 2015

Vincent P. Crawford
University of Oxford, All Souls College, and University of California, San Diego

Thanks to many people for valuable conversations and advice; and to Rustu Duran for outstanding research assistance.

The research leading to these results received primary funding from the European Research Council under the European Union’s Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement no. 339179. The contents reflect only the author’s views and not the views of the ERC or the European Commission, and the European Union is not liable for any use that may be made of the information contained therein. The University of Oxford and All Souls College also provided important funding.

Revised 18 May 2015

Advances in the Social Sciences
A Conference in Honor of Peyton Young

Efficient Mechanisms for Level-*k* Bilateral Trading

University of Oxford, 27 June 2015

Vincent P. Crawford
University of Oxford, All Souls College,
and University of California, San Diego

Thanks to many people for valuable conversations and advice; and to Rustu Duran for outstanding research assistance.

The research leading to these results received primary funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement no. 339179. The contents reflect only the author's views and not the views of the ERC or the European Commission, and the European Union is not liable for any use that may be made of the information contained therein. The University of Oxford and All Souls College also provided important funding.

Revised 18 May 2015

Economic Design: The Economist as Engineer
A Conference in Honor of Alvin Roth

Efficient Mechanisms for Level-*k* Bilateral Trading
University of Exeter, 14-15 July 2015

Vincent P. Crawford
University of Oxford, All Souls College,
and University of California, San Diego

Thanks to many people for valuable conversations and advice; and to Rustu Duran for outstanding research assistance.

The research leading to these results received primary funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement no. 339179. The contents reflect only the author's views and not the views of the ERC or the European Commission, and the European Union is not liable for any use that may be made of the information contained therein. The University of Oxford and All Souls College also provided important funding.

Revised 18 May 2015

Introduction

The paper revisits Myerson and Satterthwaite's 1983 *JET* ("MS") classic analysis of the design of incentive-efficient mechanisms for bilateral trading with independent private values, inspired by Chatterjee and Samuelson's 1983 *OR* ("CS") positive analysis.

Introduction

The paper revisits Myerson and Satterthwaite's 1983 *JET* ("MS") classic analysis of the design of incentive-efficient mechanisms for bilateral trading with independent private values, inspired by Chatterjee and Samuelson's 1983 *OR* ("CS") positive analysis.

MS assumed that traders will play any desired Bayesian Nash equilibrium in the game created by the chosen mechanism.

I focus on direct mechanisms; and replace MS's equilibrium assumption with a structural nonequilibrium "level-*k*" model of strategic thinking, meant to describe initial responses to games.

Introduction

The paper revisits Myerson and Satterthwaite's 1983 *JET* ("MS") classic analysis of the design of incentive-efficient mechanisms for bilateral trading with independent private values, inspired by Chatterjee and Samuelson's 1983 *OR* ("CS") positive analysis.

MS assumed that traders will play any desired Bayesian Nash equilibrium in the game created by the chosen mechanism.

I focus on direct mechanisms; and replace MS's equilibrium assumption with a structural nonequilibrium "level-*k*" model of strategic thinking, meant to describe initial responses to games.

To focus on nonequilibrium thinking, I maintain standard rationality assumptions regarding decisions and probabilistic judgment.

Motivation

- Mechanism design often creates novel games, weakening the usual learning justification for equilibrium; yet the design may need to work well the first time.

- Even if learning is possible, design may create games complex enough that convergence to equilibrium is behaviorally unlikely.

Motivation

- Mechanism design often creates novel games, weakening the usual learning justification for equilibrium; yet the design may need to work well the first time.

- Even if learning is possible, design may create games complex enough that convergence to equilibrium is behaviorally unlikely.

We usually assume equilibrium anyway, perhaps because:

- We doubt we can identify a credible basis for analysis among the enormous number of possible nonequilibrium models.

- We doubt that any nonequilibrium model could systematically out-predict a rational-expectations notion such as equilibrium.

But…

- There is now a large body of experimental research that studies strategic thinking by eliciting subjects' initial responses to games (surveyed in Crawford, Costa-Gomes, and Iriberri 2013 *JEL*).

- The evidence suggests that people's thinking in novel or complex games does not follow the fixed-point or indefinitely iterated dominance reasoning that equilibrium often requires.

But…

- There is now a large body of experimental research that studies strategic thinking by eliciting subjects' initial responses to games (surveyed in Crawford, Costa-Gomes, and Iriberri 2013 *JEL*).

- The evidence suggests that people's thinking in novel or complex games does not follow the fixed-point or indefinitely iterated dominance reasoning that equilibrium often requires.

  (Learning can still make people converge to something that *we* need fixed-point reasoning to characterize; the claim is that fixed-point reasoning doesn't *directly* describe people's thinking.)

But…

- There is now a large body of experimental research that studies strategic thinking by eliciting subjects' initial responses to games (surveyed in Crawford, Costa-Gomes, and Iriberri 2013 *JEL*).

- The evidence suggests that people's thinking in novel or complex games does not follow the fixed-point or indefinitely iterated dominance reasoning that equilibrium often requires.

  (Learning can still make people converge to something that *we* need fixed-point reasoning to characterize; the claim is that fixed-point reasoning doesn't *directly* describe people's thinking.)

- To the extent that people do not follow equilibrium logic, they must find another way to think about the game.

- Much evidence points to a class of nonequilibrium level-*k* or "cognitive hierarchy" models of strategic thinking.

Level-*k* models

In a level-*k* model people follow rules of thumb that:

- Anchor their beliefs in a naïve model of others' response to the game, called *L0*, often uniform random over feasible decisions;

and

- Adjust their beliefs via a small number (*k*) of iterated best responses, so *L1* best responds to *L0*, *L2* to *L1*, and so on.

Level-*k* models

In a level-*k* model people follow rules of thumb that:

- Anchor their beliefs in a naïve model of others' response to the game, called *L0*, often uniform random over feasible decisions;

and

- Adjust their beliefs via a small number (*k*) of iterated best responses, so *L1* best responds to *L0*, *L2* to *L1*, and so on.

People's levels are usually heterogeneous, and the population level frequencies are treated as behavioral parameters and either estimated from the data or calibrated from previous estimates.

Level-*k* models

In a level-*k* model people follow rules of thumb that:

- Anchor their beliefs in a naïve model of others' response to the game, called *L0*, often uniform random over feasible decisions;

and

- Adjust their beliefs via a small number (*k*) of iterated best responses, so *L1* best responds to *L0*, *L2* to *L1*, and so on.

People's levels are usually heterogeneous, and the population level frequencies are treated as behavioral parameters and either estimated from the data or calibrated from previous estimates.

Estimates vary with the setting and population, but normally the estimated frequency of *L0* is small or zero and the distribution of levels is concentrated on *L1*, *L2*, and *L3*.

- Lk (for $k > 0$) is decision-theoretically rational, with an accurate model of the game; it departs from equilibrium only in deriving its beliefs from an oversimplified model of others' responses.

- *Lk* (for $k > 0$) is decision-theoretically rational, with an accurate model of the game; it departs from equilibrium only in deriving its beliefs from an oversimplified model of others' responses.

- *Lk* (for $k > 0$) respects *k*-rationalizability (Bernheim 1984 *ECMA*), hence in two-person games its decisions survive *k* rounds of iterated elimination of strictly dominated strategies.

- Thus *Lk* mimics equilibrium decisions in *k*-dominance-solvable games, but may deviate systematically in more complex games.

- *Lk* (for $k > 0$) is decision-theoretically rational, with an accurate model of the game; it departs from equilibrium only in deriving its beliefs from an oversimplified model of others' responses.

- *Lk* (for $k > 0$) respects *k*-rationalizability (Bernheim 1984 *ECMA*), hence in two-person games its decisions survive *k* rounds of iterated elimination of strictly dominated strategies.

- Thus *Lk* mimics equilibrium decisions in *k*-dominance-solvable games, but may deviate systematically in more complex games.

- A level-*k* model (with zero weight on *L0*) can be viewed as a heterogeneity-tolerant refinement of *k*-rationalizability.

- *Lk* (for $k > 0$) is decision-theoretically rational, with an accurate model of the game; it departs from equilibrium only in deriving its beliefs from an oversimplified model of others' responses.

- *Lk* (for $k > 0$) respects *k*-rationalizability (Bernheim 1984 *ECMA*), hence in two-person games its decisions survive *k* rounds of iterated elimination of strictly dominated strategies.

- Thus *Lk* mimics equilibrium decisions in *k*-dominance-solvable games, but may deviate systematically in more complex games.

- A level-*k* model (with zero weight on *L0*) can be viewed as a heterogeneity-tolerant refinement of *k*-rationalizability.

- But unlike *k*-rationalizability, a level-*k* model makes precise predictions, given the population level frequencies: not only that deviations from equilibrium will sometimes occur, but also which settings evoke them and which forms they are likely to take.

- Level-*k* models share the generality and much of the tractability of equilibrium models (contrast *k*-rationalizability's set-valued predictions or quantal response equilibrium's noisy predictions); thus they can clarify the role of the equilibrium assumption.

- Level-*k* models share the generality and much of the tractability of equilibrium models (contrast *k*-rationalizability's set-valued predictions or quantal response equilibrium's noisy predictions); thus they can clarify the role of the equilibrium assumption.

- A level-*k* analysis can identify settings where conclusions based on equilibrium are robust to likely deviations from equilibrium.

- Level-*k* models share the generality and much of the tractability of equilibrium models (contrast *k*-rationalizability's set-valued predictions or quantal response equilibrium's noisy predictions); thus they can clarify the role of the equilibrium assumption.

- A level-*k* analysis can identify settings where conclusions based on equilibrium are robust to likely deviations from equilibrium.

- A level-*k* analysis could identify settings in which mechanisms that yield superior outcomes in equilibrium are worse in practice than others whose performance is less sensitive to deviations: an evidence-disciplined approach to robustness.

- Level-*k* models share the generality and much of the tractability of equilibrium models (contrast *k*-rationalizability's set-valued predictions or quantal response equilibrium's noisy predictions); thus they can clarify the role of the equilibrium assumption.

- A level-*k* analysis can identify settings where conclusions based on equilibrium are robust to likely deviations from equilibrium.

- A level-*k* analysis could identify settings in which mechanisms that yield superior outcomes in equilibrium are worse in practice than others whose performance is less sensitive to deviations: an evidence-disciplined approach to robustness.

- A level-*k* analysis might reduce optimal mechanisms' sensitivity to distributional and other details that real mechanisms seldom depend on, as advocated by Robert Wilson (1987) and others.

Antecedents

- Crawford and Iriberri's (2007 *ECMA*) level-*k* analysis of bidding behavior in sealed-bid independent-private-value and common-value auctions, which builds on Milgrom and Weber's 1982 *ECMA* equilibrium analysis.
- Crawford, Kugler, Neeman, and Pauzner's (2009 *JEEA*;"CKNP") level-*k* analysis of optimal independent-private-value auctions, which builds on Myerson's (1981 *MathOR*) equilibrium analysis.
- Saran's (2011 *GEB*) analysis of MS's design problem with a known population frequency of truthful traders.
- Kneeland's (2013) analysis of level-*k* implementation, with illustrations including bilateral trading.
- Gorelkina's (2015) level-*k* analysis of the expected externality mechanism.
- de Clippel, Saran, and Serrano's (2015) analysis of design with bounded depth of reasoning, and with small errors.

Outline

- CS's equilibrium analysis of bilateral trading via double auction.
- MS's analysis of equilibrium-incentive-efficient mechanisms.

Outline

- CS's equilibrium analysis of bilateral trading via double auction.
- MS's analysis of equilibrium-incentive-efficient mechanisms.
- A level-*k* model for direct games with asymmetric information.

Outline

- CS's equilibrium analysis of bilateral trading via double auction.
- MS's analysis of equilibrium-incentive-efficient mechanisms.
- A level-*k* model for direct games with asymmetric information.
- Level-*k* analysis of the double auction: *L1* optimism, aggressiveness, and incentive-*in*effiency; *L2* pessimism, *un*agressiveness, and incentive-*super*effiency.

Outline

- CS's equilibrium analysis of bilateral trading via double auction.
- MS's analysis of equilibrium-incentive-efficient mechanisms.
- A level-*k* model for direct games with asymmetric information.
- Level-*k* analysis of the double auction: *L1* optimism, aggressiveness, and incentive-*in*effiency; *L2* pessimism, *un*agressiveness, and incentive-*super*effiency.
- "Level-*k* menu effects" and the revelation principle.

Outline

● CS's equilibrium analysis of bilateral trading via double auction.

● MS's analysis of equilibrium-incentive-efficient mechanisms.

● A level-*k* model for direct games with asymmetric information.

● Level-*k* analysis of the double auction: *L1* optimism, aggressiveness, and incentive-*in*effiency; *L2* pessimism, *un*agressiveness, and incentive-*super*effiency.

● "Level-*k* menu effects" and the revelation principle.

● Mechanisms that are efficient in the set of level-*k*-incentive-compatible mechanisms for known populations with one level.

Outline

● CS’s equilibrium analysis of bilateral trading via double auction.

● MS’s analysis of equilibrium-incentive-efficient mechanisms.

● A level-*k* model for direct games with asymmetric information.

● Level-*k* analysis of the double auction: *L1* optimism, aggressiveness, and incentive-*in*effiency; *L2* pessimism, *un*agressiveness, and incentive-*super*effiency.

● “Level-*k* menu effects” and the revelation principle.

● Mechanisms that are efficient in the set of level-*k*-incentive-compatible mechanisms for known populations with one level.

● Relaxing level-*k*-incentive-compatibility (as will be explained, assuming it is *not* neutral for level-*k* traders).

● Relaxing concentration of the population on one level.

CS’s equilibrium analysis of bilateral trading via double auction

CS’s model has a potential seller and buyer of an indivisible object, in exchange for money.

CS's equilibrium analysis of bilateral trading via double auction

CS's model has a potential seller and buyer of an indivisible object, in exchange for money.

Their von Neumann-Morgenstern utility functions are quasilinear in money: risk-neutral, with money values for the object.

CS's equilibrium analysis of bilateral trading via double auction

CS's model has a potential seller and buyer of an indivisible object, in exchange for money.

Their von Neumann-Morgenstern utility functions are quasilinear in money: risk-neutral, with money values for the object.

Denote the buyer's value V and the seller's value C (for "cost").

V and C are independent, with positive densities f(V) and g(C) on their supports and distribution functions F(V) and G(C).

CS's equilibrium analysis of bilateral trading via double auction

CS's model has a potential seller and buyer of an indivisible object, in exchange for money.

Their von Neumann-Morgenstern utility functions are quasilinear in money: risk-neutral, with money values for the object.

Denote the buyer's value V and the seller's value C (for "cost").

V and C are independent, with positive densities f(V) and g(C) on their supports and distribution functions F(V) and G(C).

CS and MS allowed the densities to have any bounded overlapping supports, but without important loss of generality I take the supports to be identical and normalize them to [0, 1].

In the double auction:

- If the buyer's money bid $b \geq$ the seller's money ask a, the seller exchanges the object for a given weighted average of b and a.

- CS allowed any weights between 0 and 1, but I take the weights to be equal, so the buyer acquires the object at price $(a + b)/2$, the seller's utility is $(a + b)/2$, and the buyer's is $V - (a + b)/2$.

- If $b < a$, the seller retains the object, no money changes hands, the seller's utility is C, and the buyer's utility is 0.

- I ignore the possibility that $a = b$, which will have 0 probability.

The double auction has many Bayesian equilibria.

When f(V) and g(C) are uniformly distributed, CS identify a linear equilibrium, which also plays a central role in MS's analysis.

Denote the buyer's bidding strategy b(V) and the seller's asking strategy a(C), with * subscripts for the equilibrium strategies.

In the linear equilibrium, with value densities supported on [0, 1],

$$b_*(V) = 2V/3 + 1/12$$

unless V < ¼, when $b_*(V)$ can be anything that precludes trade;

and

$$a_*(C) = 2C/3 + ¼$$

unless C > ¾, when $a_*(C)$ can be anything that precludes trade.

In the linear equilibrium, with value densities supported on [0, 1],

$$b_*(V) = 2V/3 + 1/12$$

unless V < ¼, when $b_*(V)$ can be anything that precludes trade;

and

$$a_*(C) = 2C/3 + ¼$$

unless C > ¾, when $a_*(C)$ can be anything that precludes trade.

Trade occurs if and only if 2V/3 + 1/12 ≥ 2C/3 + ¼, or V ≥ C + ¼: with positive probability the outcome is ex post inefficient.

The ex ante probability of trade 9/32 ≈ 28% and the expected total surplus 9/64 ≈ 0.14, less than the maximum interim individually rational probability of trade 50% and expected surplus 1/6 ≈ 0.17.

MS's equilibrium analysis of incentive-efficient mechanisms for bilateral trading

MS characterized ex ante incentive-efficient mechanisms in CS's trading environment, requiring interim individual rationality.

MS, like CS, allowed general, independent value distributions with strictly positive densities on ranges that overlap for the buyer and seller; but I will continue to take both value supports to be [0, 1].

MS assumed that traders will play any desired Bayesian equilibrium in the game created by the chosen mechanism.

A *direct* (or direct-revelation) mechanism is one in which players' decisions are conformable to estimates of their values, with the outcome a function of the reported values.

A *direct* (or direct-revelation) mechanism is one in which players' decisions are conformable to estimates of their values, with the outcome a function of the reported values.

When traders are risk-neutral in money, denoting their value reports v and c (distinct from true values V and C), the payoff-relevant aspects of an outcome are determined by two functions:

- p(v, c), the probability that the object is transferred, and
- x(v, c), the expected monetary payment from buyer to seller.

Although these outcome functions depend only on reported values, traders' utilities are determined by their true values.

A *direct* (or direct-revelation) mechanism is one in which players' decisions are conformable to estimates of their values, with the outcome a function of the reported values.

When traders are risk-neutral in money, denoting their value reports v and c (distinct from true values V and C), the payoff-relevant aspects of an outcome are determined by two functions:

- $p(v, c)$, the probability that the object is transferred, and
- $x(v, c)$, the expected monetary payment from buyer to seller.

Although these outcome functions depend only on reported values, traders' utilities are determined by their true values.

A direct mechanism with outcome functions $p(\cdot, \cdot)$, $x(\cdot, \cdot)$ is *incentive-compatible* iff it makes truthful reporting an equilibrium; and is *(interim) individually rational* iff it yields buyer and seller expected utility ≥ 0 for every possible realization of their values.

The revelation principle shows that if traders can be counted on to play any desired equilibrium in the game created by the designer's chosen mechanism, there is no loss of generality in restricting attention to incentive-compatible direct mechanisms:

The revelation principle shows that if traders can be counted on to play any desired equilibrium in the game created by the designer's chosen mechanism, there is no loss of generality in restricting attention to incentive-compatible direct mechanisms:

"We can, without any loss of generality, restrict our attention to incentive-compatible direct mechanisms. This is because, for any Bayesian equilibrium of any bargaining game, there is an equivalent incentive-compatible direct mechanism that always yields the same outcomes (when the individuals play the honest equilibrium)….[w]e can construct [such a] mechanism by first asking the buyer and seller each to confidentially report his valuation, then computing what each would have done in the given equilibrium strategies with these valuations, and then implementing the outcome (transfer of money and object) as in the given game for this computed behavior. If either individual had any incentive to lie to us in this direct mechanism, then he would have had an incentive to lie to himself in the original game, which is a contradiction of the premise that he was in equilibrium in the original game." (MS, pp. 267-268)

MS's Theorem 1 uses the conditions for incentive-compatibility and individual rationality to derive an "incentive budget constraint" (my term, not theirs), subject to which, for traders with quasilinear utility functions, incentive-efficient outcome functions $p(\cdot, \cdot)$ and $x(\cdot, \cdot)$ must maximize the sum of traders' ex ante expected utilities.

MS's Theorem 2 uses Theorem 1's conditions to characterize the outcome functions associated with incentive-efficient mechanisms.

MS's Corollary 1 shows that no incentive-compatible individually rational mechanism is ex post Pareto-efficient with probability one.

The level-*k* counterparts of MS's results, explained below, will allow a more detailed explanation of MS's analysis as well.

In CS's example with uniform value densities, MS's Theorem 2 yields a closed-form solution for the incentive-compatible form of their incentive-efficient mechanism, which transfers the object when the reported values satisfy $v \geq c + \frac{1}{4}$, at price $(v + c + \frac{1}{2})/3$.

In CS's example with uniform value densities, MS's Theorem 2 yields a closed-form solution for the incentive-compatible form of their incentive-efficient mechanism, which transfers the object when the reported values satisfy $v \geq c + 1/4$, at price $(v + c + 1/2)/3$.

The linear equilibrium of the double auction with uniform value densities transfers the object when the *true* values satisfy $V \geq C + 1/4$.

In CS's example with uniform value densities, MS's Theorem 2 yields a closed-form solution for the incentive-compatible form of their incentive-efficient mechanism, which transfers the object when the reported values satisfy $v \geq c + \frac{1}{4}$, at price $(v + c + \frac{1}{2})/3$.

The linear equilibrium of the double auction with uniform value densities transfers the object when the *true* values satisfy $V \geq C + \frac{1}{4}$.

Thus, even though the double auction is not incentive-compatible, its equilibrium outcome is incentive-efficient: Equilibrium bidding strategies shade to mimic the effects of truthful reporting in the incentive-compatible form of MS's incentive-efficient mechanism.

In CS's example with uniform value densities, MS's Theorem 2 yields a closed-form solution for the incentive-compatible form of their incentive-efficient mechanism, which transfers the object when the reported values satisfy $v \geq c + 1/4$, at price $(v + c + 1/2)/3$.

The linear equilibrium of the double auction with uniform value densities transfers the object when the *true* values satisfy $V \geq C + 1/4$.

Thus, even though the double auction is not incentive-compatible, its equilibrium outcome is incentive-efficient: Equilibrium bidding strategies shade to mimic the effects of truthful reporting in the incentive-compatible form of MS's incentive-efficient mechanism.

(Satterthwaite and Williams 1989 *JET* showed, however, that for generic densities CS's double auction is *not* incentive-efficient. In that sense, MS's result for this example is coincidental.)

A level-*k* model for direct games with asymmetric information

- Recall that in a level-*k* model people anchor their beliefs in a naïve model of others' responses, *L0*, and adjust their beliefs via iterated best responses: *L1* best responds to *L0*, and so on.

- In complete-information games *L0* is usually assumed to make decisions uniformly distributed over the feasible decisions.

A level-*k* model for direct games with asymmetric information

- Recall that in a level-*k* model people anchor their beliefs in a naïve model of others' responses, *L0*, and adjust their beliefs via iterated best responses: *L1* best responds to *L0*, and so on.

- In complete-information games *L0* is usually assumed to make decisions uniformly distributed over the feasible decisions.

- Following Camerer, Ho, and Chong (2004 *QJE*), Crawford and Iriberri (2007 *ECMA*), and CKNP, I take *L0*'s decisions to be uniform over the feasible decisions, *independent of own value*.

A level-*k* model for direct games with asymmetric information

- Recall that in a level-*k* model people anchor their beliefs in a naïve model of others' responses, *L0*, and adjust their beliefs via iterated best responses: *L1* best responds to *L0*, and so on.

- In complete-information games *L0* is usually assumed to make decisions uniformly distributed over the feasible decisions.

- Following Camerer, Ho, and Chong (2004 *QJE*), Crawford and Iriberri (2007 *ECMA*), and CKNP, I take *L0*'s decisions to be uniform over the feasible decisions, *independent of own value*.

- Specifically, in a direct mechanism for the bilateral trading setting, I take *L0*'s decisions to be uniform over the range [0, 1].

  (If I allowed bounded overlapping supports, as in CS and MS, this corresponds to assuming that *L0* is uniform on the overlap, which traders have enough information to identify.)

A level-*k* model for direct games with asymmetric information

- Recall that in a level-*k* model people anchor their beliefs in a naïve model of others' responses, *L0*, and adjust their beliefs via iterated best responses: *L1* best responds to *L0*, and so on.

- In complete-information games *L0* is usually assumed to make decisions uniformly distributed over the feasible decisions.

- Following Camerer, Ho, and Chong (2004 *QJE*), Crawford and Iriberri (2007 *ECMA*), and CKNP, I take *L0*'s decisions to be uniform over the feasible decisions, *independent of own value*.

- Specifically, in a direct mechanism for the bilateral trading setting, I take *L0*'s decisions to be uniform over the range [0, 1].

  (If I allowed bounded overlapping supports, as in CS and MS, this corresponds to assuming that *L0* is uniform on the overlap, which traders have enough information to identify.)

- This *L0* yields a hierarchy of rules via iterated best responses.

- One can imagine more refined specifications, e.g. with an *L0* buyer's bid (seller's ask) uniform below (above) its value instead of over the entire range, thus eliminating dominated strategies.

- But *L0* is not an actual player: It is a player's naïve model of other players—others whose values he does not observe.

- One can imagine more refined specifications, e.g. with an *L0* buyer's bid (seller's ask) uniform below (above) its value instead of over the entire range, thus eliminating dominated strategies.

- But *L0* is not an actual player: It is a player's naïve model of other players—others whose values he does not observe.

- It is logically possible that *Lk* players initially reason contingent on others' possible values, but behaviorally far-fetched.

- One can imagine more refined specifications, e.g. with an *L0* buyer's bid (seller's ask) uniform below (above) its value instead of over the entire range, thus eliminating dominated strategies.

- But *L0* is not an actual player: It is a player's naïve model of other players—others whose values he does not observe.

- It is logically possible that *Lk* players initially reason contingent on others' possible values, but behaviorally far-fetched.

- A level-*k* model with *L0* uniform over the feasible decisions and *independent of own value* captures people's aversion to fixed-point and complex contingent reasoning in a tractable way.

This extended level-*k* model has a long history:

Milgrom and Stokey's (1982 *JET*) "No-Trade Theorem" shows that if traders in an asset market start out in market equilibrium—Pareto-efficient, given their information—then giving them new information, fundamentals unchanged, cannot lead to new trades.

For, any such trades would make it common knowledge that both had benefited, contradicting the efficiency of the initial equilibrium.

This extended level-*k* model has a long history:

Milgrom and Stokey's (1982 *JET*) "No-Trade Theorem" shows that if traders in an asset market start out in market equilibrium—Pareto-efficient, given their information—then giving them new information, fundamentals unchanged, cannot lead to new trades.

For, any such trades would make it common knowledge that both had benefited, contradicting the efficiency of the initial equilibrium.

This result has been called the Groucho Marx Theorem:

> "I sent the club a wire stating, 'Please accept my resignation. I don't want to belong to any club that will accept people like me as a member'."
>
> —Groucho Marx, Telegram to the Beverly Hills Friars' Club

This extended level-*k* model has a long history:

Milgrom and Stokey's (1982 *JET*) "No-Trade Theorem" shows that if traders in an asset market start out in market equilibrium—Pareto-efficient, given their information—then giving them new information, fundamentals unchanged, cannot lead to new trades.

For, any such trades would make it common knowledge that both had benefited, contradicting the efficiency of the initial equilibrium.

This result has been called the Groucho Marx Theorem:

> "I sent the club a wire stating, 'Please accept my resignation. I don't want to belong to any club that will accept people like me as a member'."
> —Groucho Marx, Telegram to the Beverly Hills Friars' Club

In speculating on why zero-sum trades occur despite the theorem, Milgrom and Stokey contrast Groucho's equilibrium inference with their rules Naïve Behavior, which sticks with its prior but behaves rationally otherwise, as *L1* does; and First-Order Sophistication, which best responds to Naïve Behavior, as *L2* does.

There is a growing body of evidence that this extended level-*k* model gives a realistic account of the main patterns of people's strategic thinking and "informational naiveté", their failure to attend to how others' incentives depend on their private information.

Crawford and Iriberri (2007 *ECMA*) showed that the model gives a coherent account of subjects' overbidding and vulnerability to the winner's curse in initial responses in classic auction experiments.

Brown, Camerer, and Lovallo (2012 *AEJ Micro*) use the model to explain film-goers' failure to draw negative inferences from studios' withholding weak movies from critics before release.

Camerer et al. (2004 *QJE*) suggested that a cognitive hierarchy analogue of this model could explain zero-sum betting, an important phenomenon in real life as it is in fiction.

"Son…One of these days in your travels, a guy is going to show you a brand-new deck of cards on which the seal is not yet broken. Then this guy is going to offer to bet you that he can make the jack of spades jump out of this brand-new deck of cards and squirt cider in your ear. But, son, do not accept this bet, because as sure as you stand there, you're going to wind up with an ear full of cider."

—Obadiah ("The Sky") Masterson, quoting his father in Damon Runyon (*Guys and Dolls: The Stories of Damon Runyon*, 1932)

Camerer et al. (2004 *QJE*) suggested that a cognitive hierarchy analogue of this model could explain zero-sum betting, an important phenomenon in real life as it is in fiction.

> "Son…One of these days in your travels, a guy is going to show you a brand-new deck of cards on which the seal is not yet broken. Then this guy is going to offer to bet you that he can make the jack of spades jump out of this brand-new deck of cards and squirt cider in your ear. But, son, do not accept this bet, because as sure as you stand there, you're going to wind up with an ear full of cider."
>
> —Obadiah ("The Sky") Masterson, quoting his father in Damon Runyon (*Guys and Dolls: The Stories of Damon Runyon*, 1932)

Here, Dad is worried that Son is an *L1*: rational except for sticking with his prior in the face of an offer that is "too good to be true".

Brocas, Carillo, Camerer, and Wang (2014 *REStud*) report powerful experimental evidence for the level-k model from three-state betting games (effectively zero-sum):

player/state	**A**	**B**	**C**
1	25	5	20
2	0	30	5

Brocas, Carillo, Camerer, and Wang (2014 *REStud*) report powerful experimental evidence for the level-*k* model from three-state betting games (effectively zero-sum):

player/state	**A**	**B**	**C**
1	25	5	20
2	0	30	5

There are three ex ante equally likely states, A, B, C.

Player 1 privately learns either that the state is {A or B} or that it is C; simultaneously, player 2 privately learns either that the state is A or that it is {B or C}.

Players then simultaneously choose to Bet or Pass.

Brocas, Carillo, Camerer, and Wang (2014 *REStud*) report powerful experimental evidence for the level-*k* model from three-state betting games (effectively zero-sum):

player/state	**A**	**B**	**C**
1	25	5	20
2	0	30	5

There are three ex ante equally likely states, A, B, C.

Player 1 privately learns either that the state is {A or B} or that it is C; simultaneously, player 2 privately learns either that the state is A or that it is {B or C}.

Players then simultaneously choose to Bet or Pass.

A player who chooses Pass, or who chooses Bet while the other chooses Pass, earns 10 in any state.

If both players choose Bet, they get their respective payoffs in the table for whichever state occurs.

All this is publicly announced (to induce common knowledge).

The betting game has a unique trembling-hand perfect Bayesian equilibrium, identifiable via iterated weak dominance. (There's also an imperfect equilibrium in which both players always Pass.)

Round 1 (Bet, Pass):

player/state	A	B	C
1	25	5	20
2	0	30	5

The betting game has a unique trembling-hand perfect Bayesian equilibrium, identifiable via iterated weak dominance. (There's also an imperfect equilibrium in which both players always Pass.)

Round 1 (Bet, Pass):

player/state	A	B	C
1	25	5	20
2	0	30	5

Round 2:

player/state	A	B	C
1	25	5	20
2	0	30	5

The betting game has a unique trembling-hand perfect Bayesian equilibrium, identifiable via iterated weak dominance. (There's also an imperfect equilibrium in which both players always Pass.)

Round 1 (Bet, Pass):

player/state	A	B	C
1	25	5	20
2	0	30	5

Round 2:

player/state	A	B	C
1	25	5	20
2	0	30	5

Round 3:

player/state	A	B	C
1	25	5	20
2	0	30	5

In equilibrium, no betting takes place in any state (although player 1 is willing to bet in state C).

Despite this clear equilibrium prediction, half of Brocas et al.'s subjects Bet, in patterns that varied systematically with the player role and state (as in several similar previous experiments).

L1 respects simple dominance:

player/state	**A**	**B**	**C**
1	25	5	20
2	0	30	5

Despite this clear equilibrium prediction, half of Brocas et al.'s subjects Bet, in patterns that varied systematically with the player role and state (as in several similar previous experiments).

L1 respects simple dominance:

player/state	**A**	**B**	**C**
1	25	5	20
2	0	30	5

But if all subjects were *L1s*, 100% of player 1s and 67% of player 2s would be willing to bet, many more than in Brocas et al.'s data.

Further, 100% of subjects would be willing to bet in states B and C, which is also not true in Brocas et al.'s data.

However, *L2* respects two rounds of iterated weak dominance:

player/state	A	B	C
1	25	5	20
2	0	30	5

And *L3* respects three rounds of iterated weak dominance (= trembling-hand-perfect equilibrium in this 3-dominance-solvable game):

player/state	A	B	C
1	25	5	20
2	0	30	5

However, *L2* respects two rounds of iterated weak dominance:

player/state	A	B	C
1	25	5	20
2	0	30	5

And *L3* respects three rounds of iterated weak dominance (= trembling-hand-perfect equilibrium in this 3-dominance-solvable game):

player/state	A	B	C
1	25	5	20
2	0	30	5

Brocas et al. find clusters of subjects whose behavior corresponds to each of *L1*, *L2*, and *L3*; and also a cluster of "irrational" players.

The level-*k* model fits subjects' decisions (and information searches) better than equilibrium or any homogeneous model.

Level-*k* analysis of bilateral trading via double auction

I first apply the level-*k* model to CS's trading environment, focusing on their leading example with uniform value densities.

I assume that a trader's level is independent of its value.

I set *L0*'s frequency to zero, and focus on homogeneous populations of *L1*s or *L2*s, which allows the simplest possible illustration of the main points.

Level-*k* analysis of bilateral trading via double auction

I first apply the level-*k* model to CS's trading environment, focusing on their leading example with uniform value densities.

I assume that a trader's level is independent of its value.

I set *L0*'s frequency to zero, and focus on homogeneous populations of *L1*s or *L2*s, which allows the simplest possible illustration of the main points.

Denote the buyer's bidding strategy $b_i(V)$ and the seller's asking strategy $a_i(C)$, where the subscripts denote levels i = 1,2.

Despite multiplicity of equilibria, a level-*k* model makes generically unique predictions, conditional on population level frequencies.

L1 traders

An *L1* buyer believes that the seller's ask is uniformly distributed on [0, 1], independent of its value.

Optimization yields $b_1(V) = 2V/3$ (in the interior).

(*L1*'s optimal strategy is independent of the value densities: unlike *L2*'s, which depends on the seller's density, or an equilibrium trader's strategy, which depends on both densities.)

Similarly, an *L1* seller's ask $a_1(C) = 2C/3 + 1/3$ (in the interior).

With $b_1(V) = 2V/3$, an *L1* buyer bids 1/12 more aggressively (that is, bids less) than an equilibrium buyer with $b_*(V) = 2V/3 + 1/12$.

(Compare Crawford and Iriberri's 2007 *ECMA* analysis of *L1* bidding in first-price auctions.)

With $b_1(V) = 2V/3$, an *L1* buyer bids 1/12 more aggressively (that is, bids less) than an equilibrium buyer with $b_*(V) = 2V/3 + 1/12$.

(Compare Crawford and Iriberri's 2007 *ECMA* analysis of *L1* bidding in first-price auctions.)

An *L1* buyer's naïve model of the seller underestimates the distribution of the seller's upward-shaded ask, relative to equilibrium; and this optimism makes an *L1* buyer overbid.

With $b_1(V) = 2V/3$, an *L1* buyer bids 1/12 more aggressively (that is, bids less) than an equilibrium buyer with $b_*(V) = 2V/3 + 1/12$.

(Compare Crawford and Iriberri's 2007 *ECMA* analysis of *L1* bidding in first-price auctions.)

An *L1* buyer's naïve model of the seller underestimates the distribution of the seller's upward-shaded ask, relative to equilibrium; and this optimism makes an *L1* buyer overbid.

With $a_1(C) = 2C/3 + 1/3$, an *L1* seller asks 1/12 more aggressively (more) than an equilibrium seller with $a_*(C) = 2C/3 + ¼$.

L1 traders' strategies both have the same 2/3 shading as equilibrium traders' strategies, but are 1/12 more aggressive.

If an *L1* buyer meets an *L1* seller, trade takes place iff $V \geq C + \frac{1}{2}$.

Compared to the condition for equilibrium trade, $V \geq C + \frac{1}{4}$, ex post efficiency is lost for more value combinations.

If an *L1* buyer meets an *L1* seller, trade takes place iff $V \geq C + \frac{1}{2}$.

Compared to the condition for equilibrium trade, $V \geq C + \frac{1}{4}$, ex post efficiency is lost for more value combinations.

For *L1*s the ex ante probability of trade is 1/8 = 12.5%, versus the equilibrium probability 9/32 ≈ 28% and the largest possible interim individually rational probability 50%.

These inefficient outcomes raise the question whether a designer who knows that all traders are *L1*s can design a mechanism that enhances their efficiency by counteracting their aggressiveness.

If an *L1* buyer meets an *L1* seller, trade takes place iff $V \geq C + ½$.

Compared to the condition for equilibrium trade, $V \geq C + ¼$, ex post efficiency is lost for more value combinations.

For *L1*s the ex ante probability of trade is 1/8 = 12.5%, versus the equilibrium probability 9/32 ≈ 28% and the largest possible interim individually rational probability 50%.

These inefficient outcomes raise the question whether a designer who knows that all traders are *L1*s can design a mechanism that enhances their efficiency by counteracting their aggressiveness.

I will show that (whether or not *L1*-incentive-compatibility is required) the answer is Yes.

L2 traders

An *L2* buyer's bid $b_2(V)$ maximizes over b ϵ [0, 1]

$$\int_0^b \left[V - \frac{a+b}{2}\right] g(a_1^{-1}(a))\, da + \int_b^1 0\, da,$$

where $g(a_1^{-1}(a))$ is the density of an *L1* seller's ask $a_1(C)$ induced by the value density g(C).

For instance, if g(C) is uniform, an *L2* buyer believes that the seller's ask $a_1(C)$ = 2C/3 + 1/3 is uniformly distributed on [1/3, 1], with density 3/2 there and 0 elsewhere.

An *L2* buyer who believes the seller's ask is distributed on [1/3, 1] believes that trade requires $b > 1/3$.

For $V \leq 1/3$ it is then optimal for an *L2* buyer to bid anything it thinks yields 0 probability of trade: In the absence of dominance among such strategies, I set $b_2(V) = V$ for V in [0, 1/3].

An *L2* buyer who believes the seller's ask is distributed on [1/3, 1] believes that trade requires $b > 1/3$.

For $V \leq 1/3$ it is then optimal for an *L2* buyer to bid anything it thinks yields 0 probability of trade: In the absence of dominance among such strategies, I set $b_2(V) = V$ for V in [0, 1/3].

For $V > 1/3$, an *L2* buyer's bid $b_2(V)$ maximizes over $b \in [1/3, 1]$

$$\int_{1/3}^{\mathrm{b}} \left[V - \frac{a+b}{2}\right] (3/2)\mathrm{d}a.$$

The second-order condition is satisfied.

Solving the first-order condition $(3/2)(V - b) - (3/4)(V - 1/3) = 0$ yields $b_2(V) = 2V/3 + 1/9$ for $V \in [1/3, 1]$.

Similarly, an *L2* seller's ask $a_2(C) = 2C/3 + 2/9$ (in the interior).

With $b_2(V) = 2V/3 + 1/9$, an *L2* buyer bids 1/36 less aggressively (higher) than an equilibrium buyer with $b_*(V) = 2V/3 + 1/12$ and 1/9 less aggressively than an *L1* buyer with $b_1(V) = 2V/3$.

(Compare Crawford and Iriberri's 2007 *ECMA* analysis of *L2* bidding in first-price auctions.)

With $b_2(V) = 2V/3 + 1/9$, an *L2* buyer bids 1/36 less aggressively (higher) than an equilibrium buyer with $b_*(V) = 2V/3 + 1/12$ and 1/9 less aggressively than an *L1* buyer with $b_1(V) = 2V/3$.

(Compare Crawford and Iriberri's 2007 *ECMA* analysis of *L2* bidding in first-price auctions.)

An *L2* buyer's *L1* model of the seller overestimates the distribution of the seller's ask, relative to equilibrium; and this pessimism makes an *L2* buyer overbid.

With $b_2(V) = 2V/3 + 1/9$, an *L2* buyer bids 1/36 less aggressively (higher) than an equilibrium buyer with $b_*(V) = 2V/3 + 1/12$ and 1/9 less aggressively than an *L1* buyer with $b_1(V) = 2V/3$.

(Compare Crawford and Iriberri's 2007 *ECMA* analysis of *L2* bidding in first-price auctions.)

An *L2* buyer's *L1* model of the seller overestimates the distribution of the seller's ask, relative to equilibrium; and this pessimism makes an *L2* buyer overbid.

With $a_2(C) = 2C/3 + 2/9$, an *L2* seller asks 1/36 less aggressively (lower) than an equilibrium seller, and 1/9 less aggressively than an *L1* seller.

L2 traders' strategies both have the same 2/3 shading as *L1* and equilibrium traders' strategies, but are 1/36 less aggressive.

If an *L2* buyer meets an *L2* seller, trade takes place iff $V \geq C + 1/6$.

Compared to the condition for equilibrium trade, $V \geq C + 1/4$, ex post efficiency is lost for fewer value combinations.

If an *L2* buyer meets an *L2* seller, trade takes place iff $V \geq C + 1/6$.

Compared to the condition for equilibrium trade, $V \geq C + ¼$, ex post efficiency is lost for fewer value combinations.

For *L2*s the ex ante probability of trade is $25/72 \approx 35\%$, versus the equilibrium probability 28% or the *L1* probability 12.5%, but still less than the largest interim individually rational probability 50%.

These unexpectedly efficient outcomes raise the question of whether a designer who knows all traders are *L2*s can design a mechanism that makes bargaining more efficient than in equilibrium by exploiting their *un*agressiveness in the auction.

If an *L2* buyer meets an *L2* seller, trade takes place iff $V \geq C + 1/6$.

Compared to the condition for equilibrium trade, $V \geq C + 1/4$, ex post efficiency is lost for fewer value combinations.

For *L2*s the ex ante probability of trade is $25/72 \approx 35\%$, versus the equilibrium probability 28% or the *L1* probability 12.5%, but still less than the largest interim individually rational probability 50%.

These unexpectedly efficient outcomes raise the question of whether a designer who knows all traders are *L2*s can design a mechanism that makes bargaining more efficient than in equilibrium by exploiting their *un*agressiveness in the auction.

If *L2*-incentive-compatibility is required, the answer with uniform value densities is No; but if non-*L2*-incentive-compatible direct mechanisms are allowed, the answer is Yes.

Mechanism design for level-*k* bilateral trading

Throughout the level-*k* analysis of design, I restrict attention to direct mechanisms, with decisions conformable to value reports.

I focus on populations of *L1*s and *L2*s, known to the designer; and I (mostly) assume that the population is concentrated on one level.

As in MS's and most other analyses of design, I ignore the noisiness of people's decisions.

I define incentive-efficiency notions for correct beliefs; but I derive incentive constraints from level-*k* beliefs.

I define incentive-efficiency notions for correct beliefs; but I derive incentive constraints from level-*k* beliefs.

I use “incentive-compatible” in the narrow sense, for direct mechanisms in which it is optimal for people to report truthfully.

But when I relax it, people are still assumed to best respond, even if such responses need not be truthful, as in a first-price auction.

I define incentive-efficiency notions for correct beliefs; but I derive incentive constraints from level-*k* beliefs.

I use "incentive-compatible" in the narrow sense, for direct mechanisms in which it is optimal for people to report truthfully.

But when I relax it, people are still assumed to best respond, even if such responses need not be truthful, as in a first-price auction.

"Level-*k*-incentive-compatibility" and "level-*k*-interim-individual-rationality" parallel the standard notions, "equilibrium-incentive-compatibility" and "equilibrium-interim-individual-rationality".

Should level-*k*-incentive-compatibility be required?

In an equilibrium analysis, via the revelation principle, the choice between an equilibrium-incentive-compatible mechanism and a non-equilibrium-incentive-compatible mechanism that seeks to implement the same outcomes (e.g. as the double auction does in CS's and MS's example with uniform value densities) is neutral.

Should level-*k*-incentive-compatibility be required?

In an equilibrium analysis, via the revelation principle, the choice between an equilibrium-incentive-compatible mechanism and a non-equilibrium-incentive-compatible mechanism that seeks to implement the same outcomes (e.g. as the double auction does in CS's and MS's example with uniform value densities) is neutral.

As will be seen, the choice is *not* neutral in a level-*k* analysis.

Should level-*k*-incentive-compatibility be required?

In an equilibrium analysis, via the revelation principle, the choice between an equilibrium-incentive-compatible mechanism and a non-equilibrium-incentive-compatible mechanism that seeks to implement the same outcomes (e.g. as the double auction does in CS's and MS's example with uniform value densities) is neutral.

As will be seen, the choice is *not* neutral in a level-*k* analysis.

Some analysts of design, e.g. in school choice or combinatorial auctions, have argued that incentive-compatibility is essential, though mostly in equilibrium analyses where it is neutral.

Other analysts are willing to consider non-incentive-compatible mechanisms like the Boston Mechanism or first-price auctions.

Should level-*k*-incentive-compatibility be required?

In an equilibrium analysis, via the revelation principle, the choice between an equilibrium-incentive-compatible mechanism and a non-equilibrium-incentive-compatible mechanism that seeks to implement the same outcomes (e.g. as the double auction does in CS's and MS's example with uniform value densities) is neutral.

As will be seen, the choice is *not* neutral in a level-*k* analysis.

Some analysts of design, e.g. in school choice or combinatorial auctions, have argued that incentive-compatibility is essential, though mostly in equilibrium analyses where it is neutral.

Other analysts are willing to consider non-incentive-compatible mechanisms like the Boston Mechanism or first-price auctions.

I first require *Lk*-incentive-compatibility and then consider relaxing it (while maintaining the assumption that people best respond).

Requiring *Lk*-incentive-compatibility: Mechanisms that are efficient in the set of *Lk*-incentive-compatible mechanisms

Recall that MS's Theorems 1 and 2 use conditions for equilibrium-incentive-compatibility to derive an "incentive budget constraint", subject to which, for traders with quasilinear utility functions, an equilibrium-incentive-efficient mechanism must maximize the sum of traders' ex ante expected utilities.

MS's Corollary 1 shows that no incentive-compatible individually rational mechanism is ex post Pareto-efficient with probability one.

In CS's example with uniform value densities, MS's Theorems 1-2 yield a closed-form solution for the incentive-compatible form of their incentive-efficient mechanism, which transfers the object when the reported values satisfy $v \geq c + 1/4$, at price $(v + c + 1/2)/3$.

In CS's example with uniform value densities, MS's Theorems 1-2 yield a closed-form solution for the incentive-compatible form of their incentive-efficient mechanism, which transfers the object when the reported values satisfy $v \geq c + ¼$, at price $(v + c + ½)/3$.

When *Lk*-incentive-compatibility is required, MS's characterization of the incentive-efficient mechanism with uniform value densities is completely robust to level-*k* thinking:

*Theorem A. With uniform value densities, MS's equilibrium-incentive-efficient direct mechanism is also efficient in the set of level-*k*-incentive-compatible mechanisms for any population of levels with* k *> 0, known or concentrated on one level or not.*

In CS's example with uniform value densities, MS's Theorems 1-2 yield a closed-form solution for the incentive-compatible form of their incentive-efficient mechanism, which transfers the object when the reported values satisfy $v \geq c + \frac{1}{4}$, at price $(v + c + \frac{1}{2})/3$.

When *Lk*-incentive-compatibility is required, MS's characterization of the incentive-efficient mechanism with uniform value densities is completely robust to level-*k* thinking:

*Theorem A. With uniform value densities, MS's equilibrium-incentive-efficient direct mechanism is also efficient in the set of level-*k*-incentive-compatible mechanisms for any population of levels with* k > *0, known or concentrated on one level or not.*

Proof. Follows easily by induction from the facts that, with uniform value densities, *L1* traders have the same beliefs as in the truthful equilibrium of MS's equilibrium-incentive-efficient mechanism; and that the objective functions are both based on correct beliefs.

“Level-*k* menu effects” and the revelation principle

MS showed that with uniform value densities, their incentive-efficient mechanism yields the same outcomes as CS’s linear double-auction equilibrium, with traders shading their bids and asks to mimic the effect of truthful reporting in MS’s mechanism.

“Level-*k* menu effects” and the revelation principle

MS showed that with uniform value densities, their incentive-efficient mechanism yields the same outcomes as CS’s linear double-auction equilibrium, with traders shading their bids and asks to mimic the effect of truthful reporting in MS’s mechanism.

But my examples show that the raw double auction is *not* efficient in the set of *L1*-incentive-compatible mechanisms: To yield good outcomes for *L1* traders, MS’s equilibrium-incentive-efficient mechanism must be implemented in *L1*-incentive-compatible form.

“Level-*k* menu effects” and the revelation principle

MS showed that with uniform value densities, their incentive-efficient mechanism yields the same outcomes as CS’s linear double-auction equilibrium, with traders shading their bids and asks to mimic the effect of truthful reporting in MS’s mechanism.

But my examples show that the raw double auction is *not* efficient in the set of *L1*-incentive-compatible mechanisms: To yield good outcomes for *L1* traders, MS’s equilibrium-incentive-efficient mechanism must be implemented in *L1*-incentive-compatible form.

The examples also show that the non-*L2*-incentive-compatible double auction yields more efficient outcomes for *L2* traders than MS’s equilibrium-incentive-efficient mechanism.

In each case, the choice between an incentive-compatible and a non-incentive-compatible mechanism is *not* neutral.

Even for direct mechanisms and with *L0* fixed, the choice influences the correctness of level-*k* beliefs, via Crawford et al.'s (2009 *JEEA)* "level-*k* menu effects":

- MS's incentive-compatible mechanism neutralizes *L1*s' aggressiveness in the double auction by rectifying their beliefs.

- The non-incentive-compatible double auction, when feasible, improves upon MS's mechanism for *L2*s by not rectifying their beliefs and preserving their *un*agressiveness.

Level-*k* menu effects are residues of *Lk*'s anchoring beliefs on *L0*, which would be eliminated by equilibrium thinking.

Because these menu effects influence beliefs, not preferences, they are not irrational per se, as other framing effects are.

Requiring *Lk*-incentive-compatibility: Mechanisms that are efficient in the set of *Lk*-incentive-compatible mechanisms

Turning to general value densities, the payoff-relevant aspects of a direct mechanism are still outcome functions $p(\cdot, \cdot)$ and $x(\cdot, \cdot)$, where buyer and seller report values v and c, and $p(v, c)$ is the probability the object transfers, for expected payment $x(v, c)$.

Requiring *Lk*-incentive-compatibility: Mechanisms that are efficient in the set of *Lk*-incentive-compatible mechanisms

Turning to general value densities, the payoff-relevant aspects of a direct mechanism are still outcome functions $p(\cdot, \cdot)$ and $x(\cdot, \cdot)$, where buyer and seller report values v and c, and $p(v, c)$ is the probability the object transfers, for expected payment $x(v, c)$.

For a mechanism (p, x), $f^k(v; p, x)$ and $F^k(v; p, x)$ are the density and distribution function of an *Lk* seller's beliefs and $g^k(c; p, x)$ and $G^k(c; p, x)$ of an *Lk* buyer's.

Requiring *Lk*-incentive-compatibility: Mechanisms that are efficient in the set of *Lk*-incentive-compatible mechanisms

Turning to general value densities, the payoff-relevant aspects of a direct mechanism are still outcome functions *p*(·, ·) and *x*(·, ·), where buyer and seller report values *v* and *c*, and *p*(*v*, *c*) is the probability the object transfers, for expected payment *x*(*v*, *c*).

For a mechanism (*p*, *x*), $f^k(v; p, x)$ and $F^k(v; p, x)$ are the density and distribution function of an *Lk* seller's beliefs and $g^k(c; p, x)$ and $G^k(c; p, x)$ of an *Lk* buyer's.

With *L0* uniform on [0, 1], $f^1(v; p, x) \equiv 1$ and $g^1(c; p, x) \equiv 1$.

If $\beta_1(V; p, x)$ is an *L1* buyer's response to (*p*, *x*) with value *V* and $\alpha_1(C; p, x)$ is an *L1* seller's response to (*p*, *x*) with cost *C*, $f^2(v; p, x) \equiv f(\beta_1^{-1}(v; p, x))$ and $g^2(c; p, x) \equiv g(\alpha_1^{-1}(c; p, x))$.

As in MS's analysis, we can write the buyer's and seller's expected monetary payments, probabilities of trade, and utilities as functions of their value reports *v* and *c*.

$$X_B^k(v) = \int_0^1 x(v,\hat{c})g^k(\hat{c})d\hat{c}, \qquad X_S^k(c) = \int_0^1 x(\hat{v},c)f^k(\hat{v})d\hat{v},$$

$$(6.1)\ P_B^k(v) = \int_0^1 p(v,\hat{c})g^k(\hat{c})d\hat{c}, \qquad P_S^k(c) = \int_0^1 p(\hat{v},c)f^k(\hat{v})d\hat{v},$$

$$U_B^k(v) = vP_B^k(v) - X_B^k(v), \qquad U_S^k(c) = X_S^k(c) - cP_S^k(c).$$

For a given k, the mechanism $p(\cdot, \cdot)$, $x(\cdot, \cdot)$ is Lk-incentive-compatible iff truthful reporting is optimal given Lk beliefs.

That is, if for every V, v, C, and c in [0, 1],

(6.2) $\quad U_B^k(V) \geq VP_B^k(v) - X_B^k(v) \text{ and } U_S^k(C) \geq X_S^k(c) - CP_S^k(c)$.

For a given *k*, the mechanism *p*(·, ·), *x*(·, ·) is *Lk*-incentive-compatible iff truthful reporting is optimal given *Lk* beliefs.

That is, if for every *V*, *v*, *C*, and *c* in [0, 1],

(6.2) $U_B^k(V) \geq VP_B^k(v) - X_B^k(v)$ and $U_S^k(C) \geq X_S^k(c) - CP_S^k(c)$.

The mechanism *p*(·, ·), *x*(·, ·) is (interim) *Lk*-individually rational iff for every *V* and *C* in [0, 1],

(6.3) $U_B^k(V) \geq 0$ and $U_S^k(C) \geq 0$.

For models with known, homogeneous populations of *L1*s or *L2*s, my Theorems B and C parallel MS's (Theorems 1-2) characterization of equilibrium-incentive-efficient mechanisms.

Theorem B. *For any known population of L1 or L2 traders concentrated on one level,* k*, and any level-k-incentive-compatible mechanism,*

$$U_B(0) + U_S(1) = \min_{\mathrm{V}\in[0,1]} U_B(V) \ + \min_{\mathrm{C}\in[0,1]} U_S(C)$$

(6.4) $= \int_0^1 \int_0^1 \left(\left[V - \frac{1-F(V)}{f(V)}\right] \left[\frac{g^k(C;p,x)}{g(C)}\right] - \left[C + \frac{G(C)}{g(C)}\right] \left[\frac{f^k(V;p,x)}{f(V)}\right] \right) p(V,C) f(V) g(C) dC dV.$

Furthermore, if p(·,·) is any function mapping [0, 1]×[0, 1] into [0, 1], there exists a function x(·, ·) such that (p, x) is level-k-*incentive-compatible and level*-k-*interim-individually rational if and only if* $P_B^k(\cdot)$ *is weakly increasing for all (p, x),* $P_S^k(\cdot)$ *is weakly decreasing for all (p, x), and*

(6.5) $0 \le \int_0^1 \int_0^1 \left\{ \left[V - \frac{1-F(V)}{f(V)}\right] \left[\frac{g^k(C;p,x)}{g(C)}\right] - \left[C + \frac{G(C)}{g(C)}\right] \left[\frac{f^k(V;p,x)}{f(V)}\right] \right\} p(V,C) f(V) g(C) dC dV.$

With correct beliefs, $g^k(C; p, x) \equiv g(C)$ and $f^k(V; p, x) \equiv f(V)$, (6.5) is equivalent to MS's incentive budget constraint.

Because level-*k* beliefs happen to be correct for uniform value densities (for all *k*), that equivalence implies Theorem A.

With correct beliefs, $g^k(C;p,x) \equiv g(C)$ and $f^k(V;p,x) \equiv f(V)$, (6.5) is equivalent to MS's incentive budget constraint.

Because level-*k* beliefs happen to be correct for uniform value densities (for all *k*), that equivalence implies Theorem A.

Proof. The proof follows MS's, adjusted for level-*k* beliefs. By (6.1), $P_B^k(\cdot)$ is weakly increasing and $P_S^k(\cdot)$ is weakly decreasing for any given (*p*, *x*), which as in MS's proof yields necessary and sufficient conditions for incentive-compatibility:

(6.6) $$U_B^k(V) = U_B^k(0) + \int_0^V P_B^k(v)dv \text{ and}$$
$$U_S^k(C) = U_S^k(1) + \int_C^1 P_S^k(c)dc \text{ for all } V \text{ and } C.$$

(6.6) implies that $U_B^k(V)$ is weakly increasing and $U_S^k(C)$ is weakly decreasing, and shows that $U_B^k(0) \geq 0$ and $U_S^k(1) \geq 0$ suffice for individual rationality for all *V* and *C* as in (6.3).

To derive the incentive budget constraint (6.5), analogous to MS's (2), note that,

$$\int_0^1 \int_0^1 V\, p(V, \hat{c}) g^k(\hat{c}) d\hat{c} f(V) dV - \int_0^1 \int_0^1 C\, p(\hat{v}, C) f^k(\hat{v}) d\hat{v} g(C) dC =$$

$$\int_0^1 U_B^k(V) f(V) dV + \int_0^1 U_S^k(C) g(C) dC =$$

$$(6.7) \qquad U_B^k(0) + \int_0^1 \int_0^V P_B^k(v) dv f(V) dV + U_S^k(1) + \int_0^1 \int_C^1 P_C^k(c) dc g(C) dC =$$

$$U_B^k(0) + U_S^k(1) + \int_0^1 [1 - F(v)] P_B^k(v) dv + \int_0^1 G(c) P_S^k(c) dc =$$

$$U_B^k(0) + U_S^k(1) + \int_0^1 \int_0^1 [1 - F(v)] p(v, \hat{c}) g^k(\hat{c}) d\hat{c} dv + \int_0^1 \int_0^1 G(c) p(\hat{v}, c) f^k(\hat{v}) d\hat{v} dc\,.$$

Equating the first and last expression in (6.7) yields (6.4), which implies (6.5) whenever the mechanism is individually rational.

Finally, given (6.3) and that $P_B^k(\cdot)$ is increasing and $P_S^k(\cdot)$ is decreasing, MS's (pp. 270-271) transfer function

(6.8) $x(v,c) = \int_0^V v\, d[P_B^k(v)] - \int_0^C c\, d[-P_S^k(c)] + \int_0^1 c[1 - G(c)]\, d[-P_S^k(c)]$

makes (*p*, *x*) level-*k*-incentive-compatible and level-*k*-interim individually rational. Q.E.D.

Theorem C. *For any known population of L1 or L2 traders concentrated on one level, if there exists a mechanism (p, x) that is level-k-incentive-compatible and maximizes traders' ex ante expected total surplus*

$$\int_0^1 \int_0^1 (V - C)\, p(V,C) f(V) g(C) dC dV$$

s.t. $U_B^k(0) = U_S^k(1) = 0$ *and (6.5), then that mechanism is efficient in the set of level-k-incentive-compatible and level-k-interim-individually-rational mechanisms. Further, if*

(6.9) $$\left[V - \frac{1-F(V)}{f(V)}\right]\left[\frac{g^k(C;p,x)}{g(C)}\right] - \left[C + \frac{G(C)}{g(C)}\right]\left[\frac{f^k(V;p,x)}{f(V)}\right]$$

is increasing in V *and decreasing in* C *for any given (p, x), then that mechanism is efficient in the set of level-k-incentive-compatible and level-k-interim-individually-rational mechanisms.*

Proof. The proof adapts the proof of MS's Theorem 2.

Choose $p(\cdot, \cdot)$ to maximize ex ante expected total surplus subject to $0 \leq p(\cdot, \cdot) \leq 1$, $U_B^k(0) = U_S^k(1) = 0$, and (6.5).

The problem is like a consumer's budget problem, with a continuum of trade probabilities $p(V, C)$ analogous to goods priced linearly.

Proof. The proof adapts the proof of MS's Theorem 2.

Choose $p(\cdot, \cdot)$ to maximize ex ante expected total surplus subject to $0 \leq p(\cdot, \cdot) \leq 1$, $U_B^k(0) = U_S^k(1) = 0$, and (6.5).

The problem is like a consumer's budget problem, with a continuum of trade probabilities $p(V, C)$ analogous to goods priced linearly.

Some of the "prices" are negative; but by the logic of the incentive constraints, there is no free disposal, so the analogy is not misleading.

Proof. The proof adapts the proof of MS's Theorem 2.

Choose $p(\cdot, \cdot)$ to maximize ex ante expected total surplus subject to $0 \leq p(\cdot, \cdot) \leq 1$, $U_B^k(0) = U_S^k(1) = 0$, and (6.5).

The problem is like a consumer's budget problem, with a continuum of trade probabilities $p(V, C)$ analogous to goods priced linearly.

Some of the "prices" are negative; but by the logic of the incentive constraints, there is no free disposal, so the analogy is not misleading.

A solution exists because continuity of the value densities ensures continuity of the objective function and the constraint, and the feasible region is compact.

Proof. The proof adapts the proof of MS's Theorem 2.

Choose $p(\cdot, \cdot)$ to maximize ex ante expected total surplus subject to $0 \le p(\cdot, \cdot) \le 1$, $U_B^k(0) = U_S^k(1) = 0$, and (6.5).

The problem is like a consumer's budget problem, with a continuum of trade probabilities $p(V, C)$ analogous to goods priced linearly.

Some of the "prices" are negative; but by the logic of the incentive constraints, there is no free disposal, so the analogy is not misleading.

A solution exists because continuity of the value densities ensures continuity of the objective function and the constraint, and the feasible region is compact.

Because the $p(\cdot, \cdot)$ enter the problem linearly, the solution will be bang-bang with $p(V, C) = 0$ or 1 a.e.; but because there is a continuum of $p(\cdot, \cdot)$, (6.5) will still hold with equality at the solution.

Form the Lagrangean, but for ease of notation without separately pricing out the $p(V,C) \leq 1$ constraints:

$$\int_0^1 \int_0^1 (V - C)\, p(V,C) f(V) g(C)\, dC dV$$

$$(6.10)\quad + \lambda \int_0^1 \int_0^1 \left\{ \left[V - \frac{1-F(V)}{f(V)} \right] \left[\frac{g^k(C;p,x)}{g(C)} \right] - \left[C + \frac{G(C)}{g(C)} \right] \left[\frac{f^k(V;p,x)}{f(V)} \right] \right\} p(V,C) f(V) g(C)\, dC dV$$

$$= \int_0^1 \int_0^1 \left((V - C) + \lambda \left\{ \left[V - \frac{1-F(V)}{f(V)} \right] \left[\frac{g^k(C;p,x)}{g(C)} \right] - \left[C + \frac{G(C)}{g(C)} \right] \left[\frac{f^k(V;p,x)}{f(V)} \right] \right\} \right) p(V,C) f(V) g(C)\, dC dV.$$

Form the Lagrangean, but for ease of notation without separately pricing out the $p(V,C) \leq 1$ constraints:

$$\int_0^1 \int_0^1 (V-C)\, p(V,C) f(V) g(C)\, dC dV$$

(6.10) $$+\lambda \int_0^1 \int_0^1 \left\{\left[V - \frac{1-F(V)}{f(V)}\right]\left[\frac{g^k(C;p,x)}{g(C)}\right] - \left[C + \frac{G(C)}{g(C)}\right]\left[\frac{f^k(V;p,x)}{f(V)}\right]\right\} p(V,C) f(V) g(C)\, dC dV$$

$$= \int_0^1 \int_0^1 \left((V-C) + \lambda\left\{\left[V - \frac{1-F(V)}{f(V)}\right]\left[\frac{g^k(C;p,x)}{g(C)}\right] - \left[C + \frac{G(C)}{g(C)}\right]\left[\frac{f^k(V;p,x)}{f(V)}\right]\right\}\right) p(V,C) f(V) g(C)\, dC dV.$$

The Kuhn-Tucker conditions require $\lambda \geq 0$,

(6.11) $(V-C) + \lambda\left\{\left[V - \frac{1-F(V)}{f(V)}\right]\left[\frac{g^k(C;p,x)}{g(C)}\right] - \left[C + \frac{G(C)}{g(C)}\right]\left[\frac{f^k(V;p,x)}{f(V)}\right]\right\} \leq 0$

when $p(V,C) = 0$, and

(6.12) $(V-C) + \lambda\left\{\left[V - \frac{1-F(V)}{f(V)}\right]\left[\frac{g^k(C;p,x)}{g(C)}\right] - \left[C + \frac{G(C)}{g(C)}\right]\left[\frac{f^k(V;p,x)}{f(V)}\right]\right\} \geq 0$

when $p(V,C) = 1$.

(6.11)-(6.12) are analogous to marginal-utility-to-price ratios determining which goods to buy and which not to buy.

If the expression in the statement of the Theorem,

(6.9) $$\left[V - \frac{1-F(V)}{f(V)}\right]\left[\frac{g^k(C;p,x)}{g(C)}\right] - \left[C + \frac{G(C)}{g(C)}\right]\left[\frac{f^k(V;p,x)}{f(V)}\right]$$

is increasing in *V* and decreasing in *C* for any given (*p, x*), then $p(V,C)$ in (6.11)-(6.12) and thus $P_B^k(V)$ and $P_S^k(C)$ in (6.1) are respectively increasing and decreasing.

By Theorem B, the problem's solution is then associated with a mechanism that maximizes expected total surplus among all level-*k*-incentive-compatible and level-*k*-interim-individually-rational mechanisms. Q.E.D.

Theorem C's condition that the expression in (6.9) is increasing in *V* and decreasing in *C* for all (*p, x*) is the level-*k* analogue of MS's (Theorem 2) equilibrium-based condition for $p(V, C) = 1$.

Theorem C's condition with the true, equilibrium densities $f(V)$ and $g(C)$ replacing the level-*k* beliefs $f^k(V; p, x)$ and $g^k(C; p, x)$ reduces to MS's condition.

Theorem C's condition that the expression in (6.9) is increasing in *V* and decreasing in *C* for all (*p, x*) is the level-*k* analogue of MS's (Theorem 2) equilibrium-based condition for $p(V, C) = 1$.

Theorem C's condition with the true, equilibrium densities $f(V)$ and $g(C)$ replacing the level-*k* beliefs $f^k(V; p, x)$ and $g^k(C; p, x)$ reduces to MS's condition.

MS's condition is satisfied when the true densities fit into Myerson's (1981) "regular case", which rules out strong hazard rate variations in the wrong direction.

The level-*k* version of the condition, on (6.9), jointly restricts the true densities and level-*k* beliefs in a similar way.

Properties of mechanisms that are efficient in the set of *Lk*-incentive-compatible mechanisms

Comparing the level-*k* incentive budget constraint
(6.5)

$$0 \leq \int_0^1 \int_0^1 \left\{ \left[V - \frac{1-F(V)}{f(V)} \right] \left[\frac{g^k(C;p,x)}{g(C)} \right] - \left[C + \frac{G(C)}{g(C)} \right] \left[\frac{f^k(V;p,x)}{f(V)} \right] \right\} p(V,C) f(V) g(C) dC dV.$$

with MS's equilibrium-based incentive budget constraint, and comparing the level-*k* Kuhn-Tucker condition

(6.12) $(V-C) + \lambda \left\{ \left[V - \frac{1-F(V)}{f(V)} \right] \left[\frac{g^k(C;p,x)}{g(C)} \right] - \left[C + \frac{G(C)}{g(C)} \right] \left[\frac{f^k(V;p,x)}{f(V)} \right] \right\} \geq 0$
when $p(V,C) = 1$

with the equilibrium-based Kuhn-Tucker condition shows that the design features that foster equilibrium-incentive-efficiency also foster efficiency in the set of level-*k*-incentive-compatible mechanisms, although level-*k* beliefs give them different weights.

There are, however, important differences in the level-*k* analysis.

First, in contrast to MS's Corollary 1 it is now theoretically possible that the optimal $\lambda = 0$, so that from (6.12), $p(V, C) = 1$ if and only if $V \geq C$ (ignoring ties), (6.5) is satisfied even then, and the level-*k*-optimal mechanism is ex post efficient with probability 1.

This can be seen, tediously, by trying to adapt MS's proof of Corollary 1 for a population concentrated on one level, *k*.

Second, unless a mechanism that is efficient in the set of level-*k*-incentive-compatible mechanisms happens to induce correct beliefs (as with uniform value densities, by Theorem A), it must use tacit exploitation of predictably incorrect beliefs (“TEPIB”), a design feature with no counterpart in an equilibrium analysis:

● “Predictably” via the level-*k* model.

● “Exploitation” in the benign sense that traders’ incorrect beliefs are used only for their benefit.

● “Tacit” in that the mechanism does not actively deceive traders.

A thought-experiment clarifies the influence of TEPIB:

Suppose one could exogenously increase the pessimism of traders' level-*k* beliefs relative to the truth, in the sense of first-order stochastic dominance.

Then substituting (6.1) into (6.6) shows that, other things equal, that would loosen the incentive budget constraint (6.5).

Because traders' beliefs enter the problem only through (6.5), that would increase maximized expected total surplus.

It is not possible to exogenously change traders' beliefs, but the tradeoffs in (6.5) reflect the influence of pessimism or optimism.

TEPIB favors trade at (*V, C*) combinations for which traders' non-equilibrium beliefs make the "prices" (in curly brackets) in

(6.12) $(V - C) + \lambda\left\{\left[V - \frac{1-F(V)}{f(V)}\right]\left[\frac{g^k(C;p,x)}{g(C)}\right] - \left[C + \frac{G(C)}{g(C)}\right]\left[\frac{f^k(V;p,x)}{f(V)}\right]\right\} \geq 0$
when $p(V,C) = 1$

more favorable than for equilibrium beliefs.

In particular those for which $\frac{f^k(V;p,x)}{f(V)} > 1$ and/or $\frac{g^k(C;p,x)}{g(C)} < 1$ are favored more than in an equilibrium-incentive-efficient mechanism.

And for *k* = 2 (*L1* beliefs don't depend on the mechanism) TEPIB favors mechanisms that increase the advantages of such trades.

TEPIB also suggests that viewing “robust” mechanism design as achieving equilibrium-incentive-efficient outcomes under weaker behavioral assumptions is too narrow. For example:

- A second-price auction seems more robust than an equilibrium-revenue-equivalent first-price auction, because it yields the equilibrium outcome for any mixture of level-*k* bidders.

TEPIB also suggests that viewing “robust” mechanism design as achieving equilibrium-incentive-efficient outcomes under weaker behavioral assumptions is too narrow. For example:

- A second-price auction seems more robust than an equilibrium-revenue-equivalent first-price auction, because it yields the equilibrium outcome for any mixture of level-*k* bidders.

- But auction design for *L1*s (at least) tends to favor first-price auctions, which make *L1*s overbid, yielding revenue higher than in equilibrium or in a second-price auction, which makes *L1* bidders mimic equilibrium (Crawford and Iriberri 2007, CKNP).

Finally, Theorem C shows that a mechanism that is efficient in the set of level-k-incentive-compatible mechanisms may involve trade for some value combinations with $V < C$: “perverse” ex post (though consistent with level-k-interim-individual-rationality):

Some of the “prices” in the incentive budget constraint (6.5) are negative (with no free disposal), which makes (6.12) consistent with some trade when $V < C$.

Trade when $V < C$ can loosen (6.5) enough to compensate for the local loss in surplus by enabling trade for other value combinations, as illustrated in examples below.

Finally, Theorem C shows that a mechanism that is efficient in the set of level-k-incentive-compatible mechanisms may involve trade for some value combinations with $V < C$: “perverse” ex post (though consistent with level-k-interim-individual-rationality):

Some of the “prices” in the incentive budget constraint (6.5) are negative (with no free disposal), which makes (6.12) consistent with some trade when $V < C$.

Trade when $V < C$ can loosen (6.5) enough to compensate for the local loss in surplus by enabling trade for other value combinations, as illustrated in examples below.

(MS’s Theorem 2 shows that such perverse trade cannot occur in an equilibrium-incentive-efficient mechanism. MS note however that their transfer function may require payment by a buyer who does not get the object, which is not ex-post-individually rational.)

Examples of mechanisms that are efficient in the set of level-*k*-incentive-incentive-compatible mechanisms

As in MS's analysis, closed-form solutions are available only with uniform value densities; but the mechanism that is efficient in the set of level-*k*-incentive-compatible mechanisms then induces correct beliefs by Theorem A, so that TEPIB has no influence.

To illustrate TEPIB, I report computed trading regions for such mechanisms for *L1*s and selected combinations of linear densities.

(Figure 1 in the paper reports trading regions for a comprehensive coarse subset of linear density combinations, with combinations excluded only for extreme combinations that violate Theorems B-C's monotonicity conditions for the mechanism to be truly optimal.)

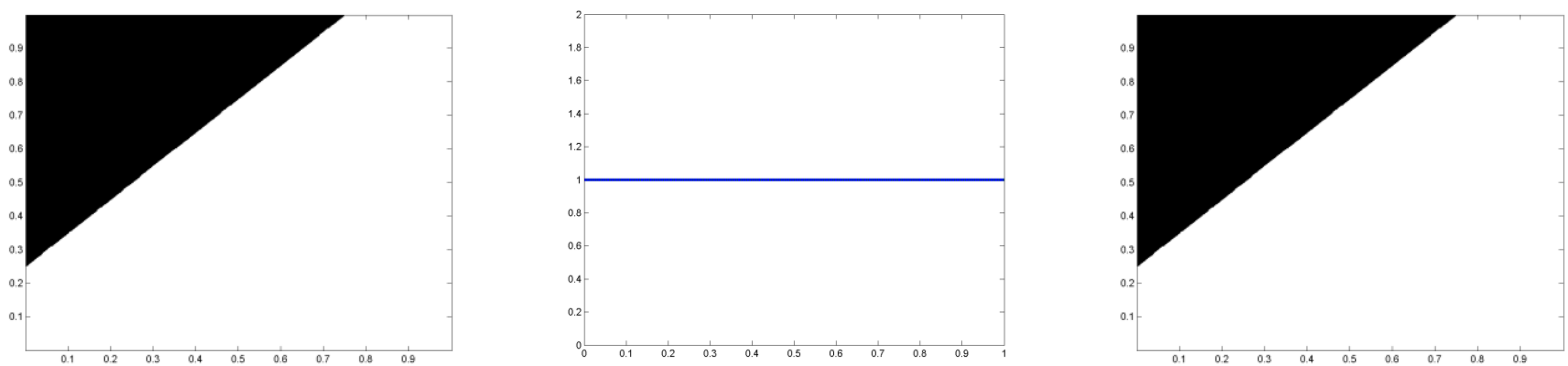
(For *L2*s, with $f^2(v) \equiv f(\beta_1^{-1}(v; p, x))$ and $g^2(c) \equiv g(\alpha_1^{-1}(c; p, x))$, (6.5) and (6.12) depend on the transfer function *x*(·, ·) as well as *p*(·, ·), making the dimensionality of search too high.)

From Figure 1. Trading regions (in black) for equilibrium-incentive-efficient mechanisms and mechanisms that are efficient in the set of *L1*-incentive-compatible mechanisms

The case of uniform value densities, for which closed-form solutions are possible

(Buyer's value V is on the vertical axis; seller's value C is on the horizontal axis. All value densities are linear; "x, y" means the buyer's density f(V) satisfies f(0) = x and f(1) = 2-x, and the seller's density g(C) satisfies g(0) = y and g(1) = 2-y.)

Equilibrium: 1.0, 1.0 **Buyer (—) and seller (···)** **L1: 1.0, 1.0**

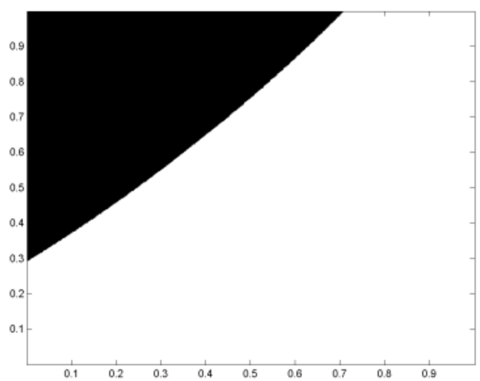


At least for populations concentrated on one level, mechanisms that are efficient in the set of *L1*-incentive-compatible mechanisms are similar to equilibrium-incentive-efficient mechanisms in most respects.

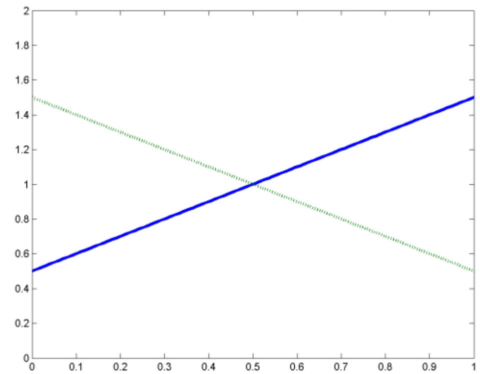
But neither equilibrium nor *L1* trading regions are symmetric with buyer's and seller's densities interchanged, because trade need not occur when V > C, breaking the symmetry

(Buyer's value V is on the vertical axis; seller's value C is on the horizontal axis. All value densities are linear; "x, y" means the buyer's density f(V) satisfies f(0) = x and f(1) = 2-x, and the seller's density g(C) satisfies g(0) = y and g(1) = 2-y.)

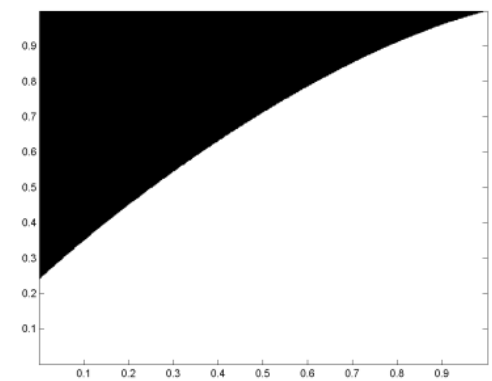
Equilibrium: 0.5, 1.5



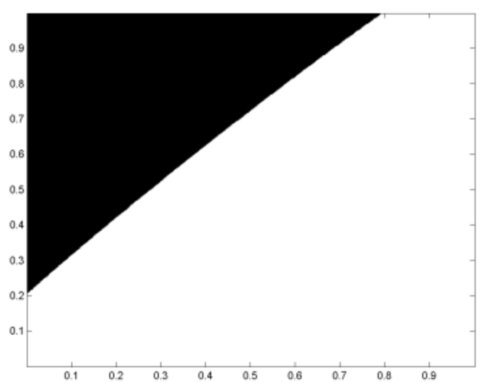
Buyer (—) and seller (···)



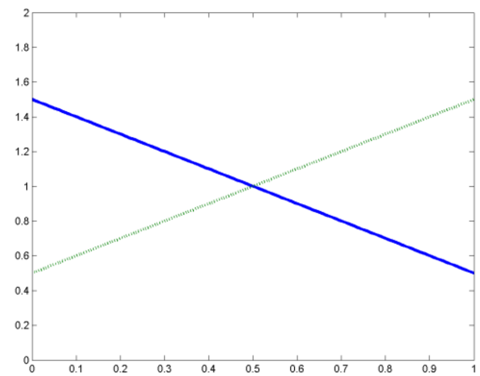
L1: 0.5, 1.5



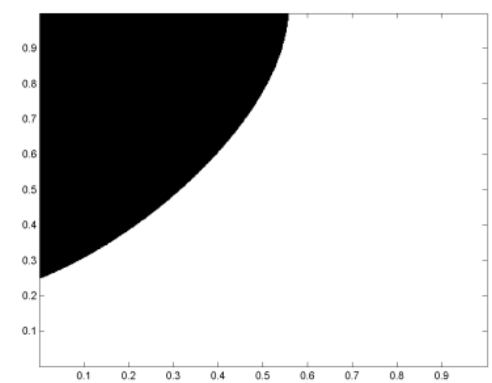
Equilibrium: 1.5, 0.5



Buyer (—) and seller (···)



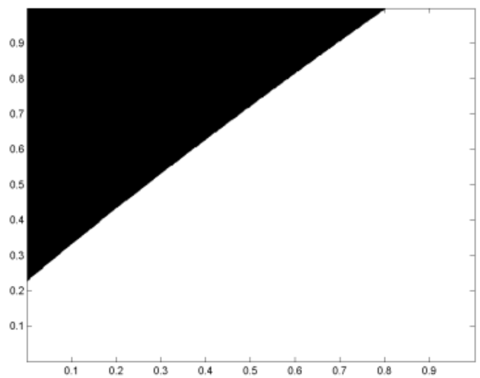
L1: 1.5, 0.5



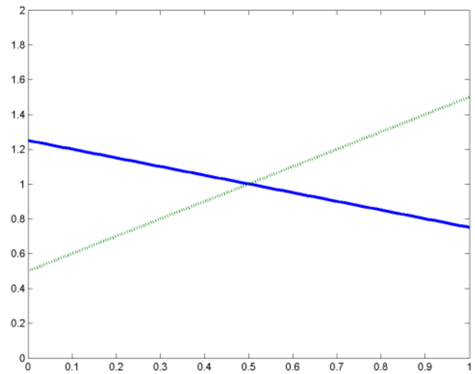
Cases where an *L1* buyer's optimism about the seller yields smaller *L1* trading regions

(Buyer's value V is on the vertical axis; seller's value C is on the horizontal axis. All value densities are linear; "x, y" means the buyer's density f(V) satisfies f(0) = x and f(1) = 2-x, and the seller's density g(C) satisfies g(0) = y and g(1) = 2-y.)

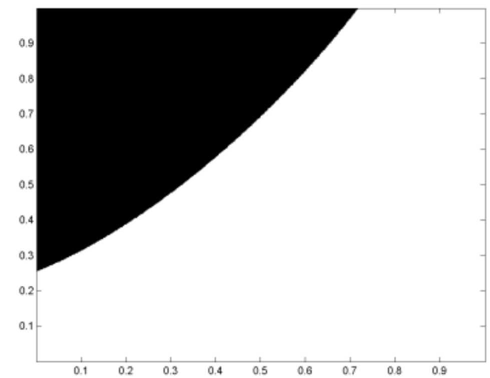
Equilibrium: 1.25, 0.5



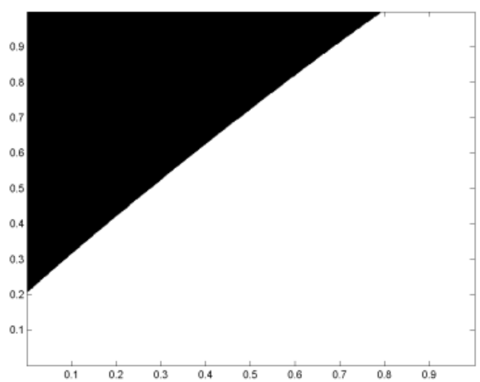
Buyer (—) and seller (···)



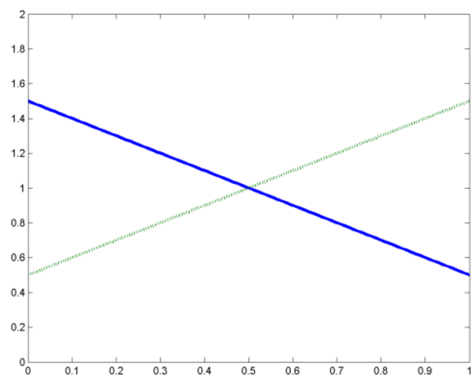
L1: 1.25, 0.5



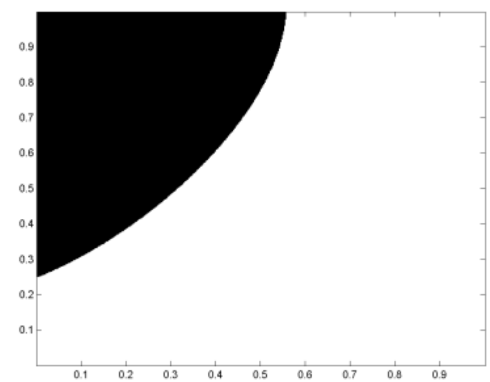
Equilibrium: 1.5, 0.5



Buyer (—) and seller (···)



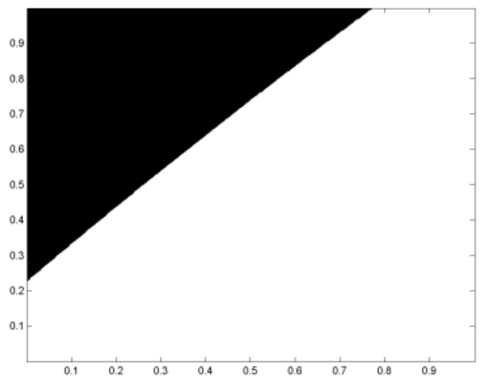
L1: 1.5, 0.5



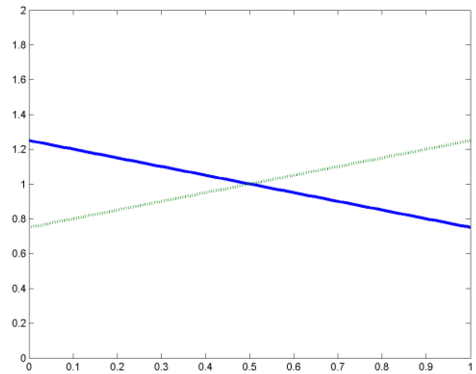
Cases where an *L1* seller's optimism about the buyer yields smaller *L1* trading regions

(Buyer's value V is on the vertical axis; seller's value C is on the horizontal axis. All value densities are linear; "x, y" means the buyer's density f(V) satisfies f(0) = x and f(1) = 2-x, and the seller's density g(C) satisfies g(0) = y and g(1) = 2-y.)

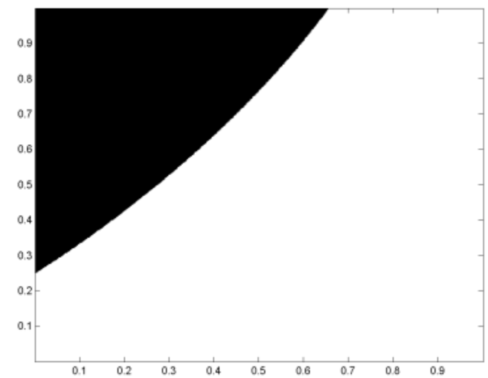
Equilibrium: 1.25, 0.75



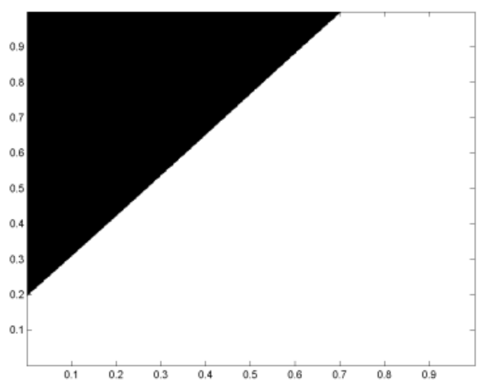
Buyer (—) and seller (···)



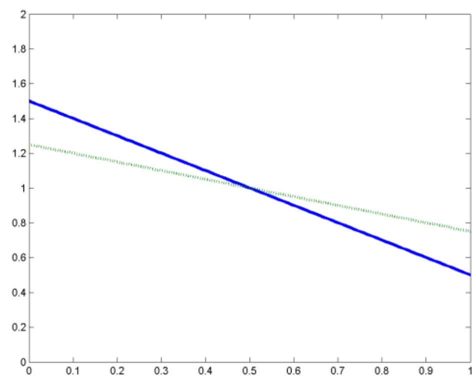
L1: 1.25, 0.75



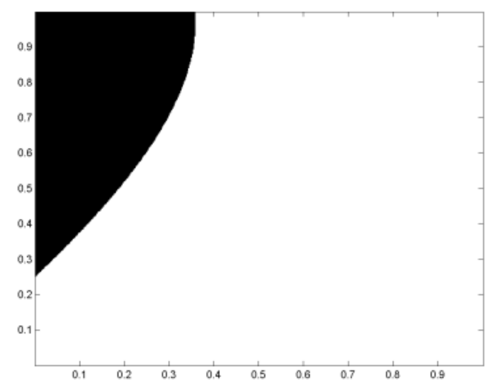
Equilibrium: 1.5, 1.25



Buyer (—) and seller (···)



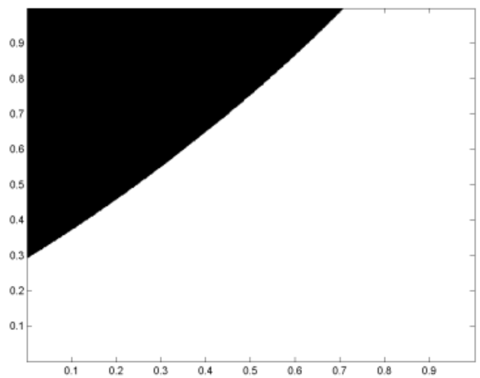
L1: 1.5, 1.25



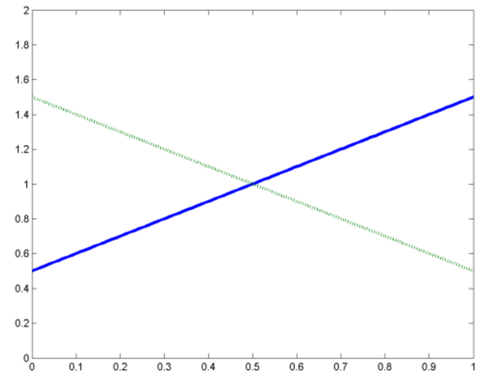
Cases where an *L1* buyer's pessimism about the seller yields larger *L1* trading regions

(Buyer's value V is on the vertical axis; seller's value C is on the horizontal axis. All value densities are linear; "x, y" means the buyer's density f(V) satisfies f(0) = x and f(1) = 2-x, and the seller's density g(C) satisfies g(0) = y and g(1) = 2-y.)

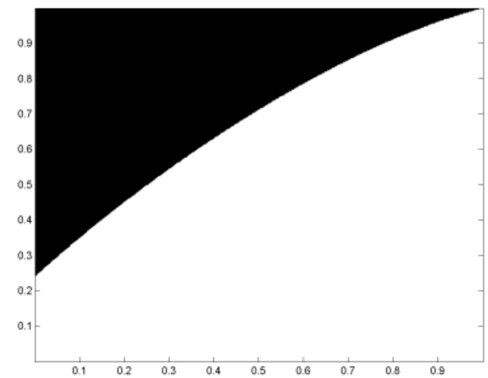
Equilibrium: 0.5, 1.5



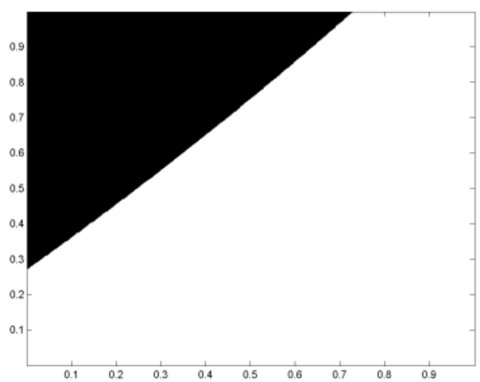
Buyer (—) and seller (···)



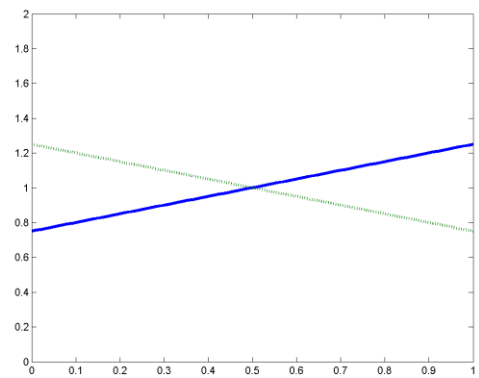
L1: 0.5, 1.5



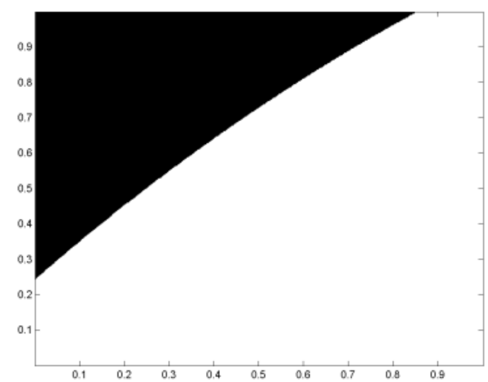
Equilibrium: 0.75, 1.25



Buyer (—) and seller (···)



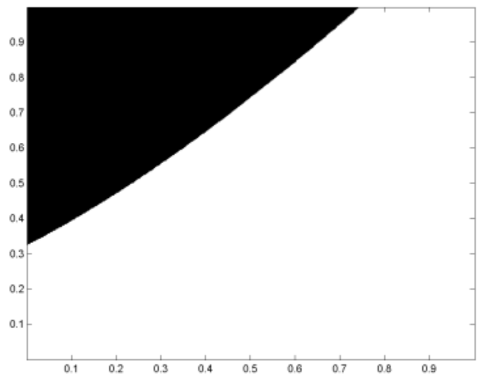
L1: 0.75, 1.25



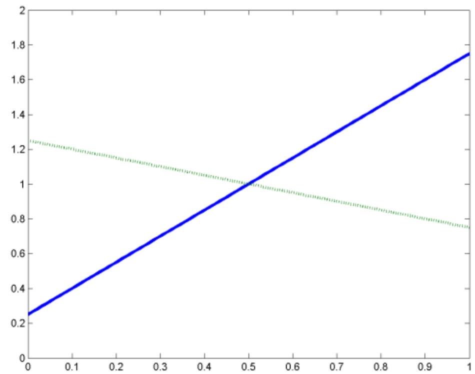
Cases where an *L1* seller's pessimism about the buyer yields larger *L1* trading regions

(Buyer's value V is on the vertical axis; seller's value C is on the horizontal axis. All value densities are linear; "x, y" means the buyer's density f(V) satisfies f(0) = x and f(1) = 2-x, and the seller's density g(C) satisfies g(0) = y and g(1) = 2-y.)

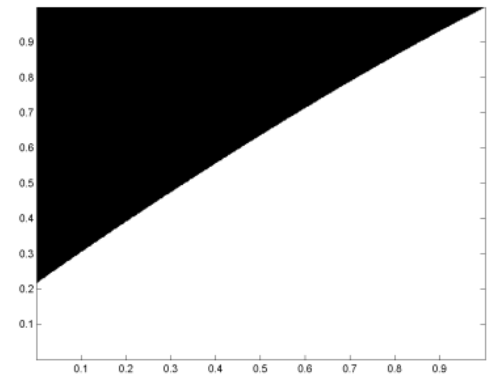
Equilibrium: 0.25, 1.25



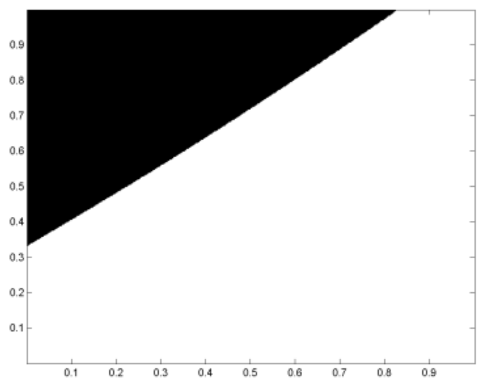
Buyer (—) and seller (⋯)



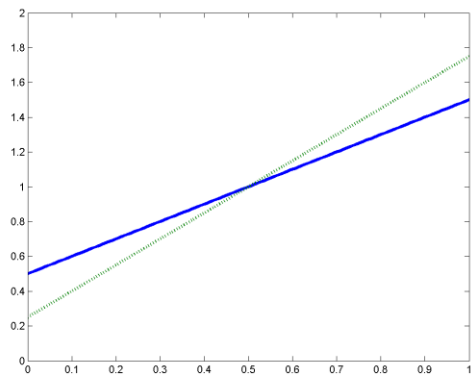
L1: 0.25, 1.25



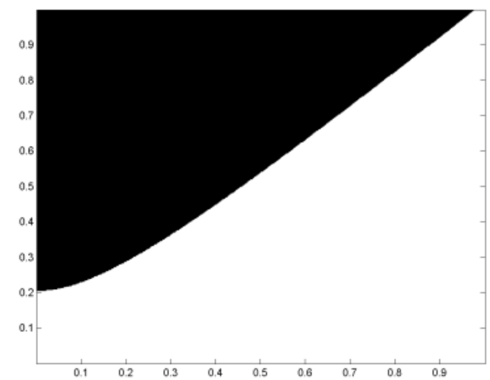
Equilibrium: 0.5, 0.25



Buyer (—) and seller (⋯)



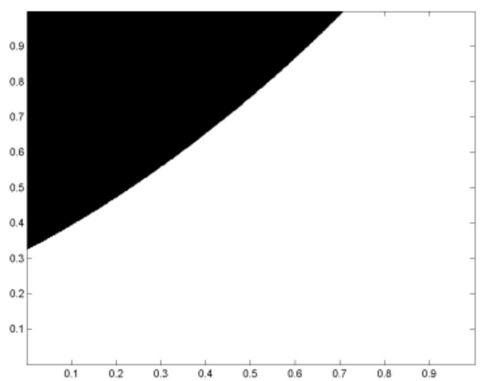
L1: 0.5, 0.25



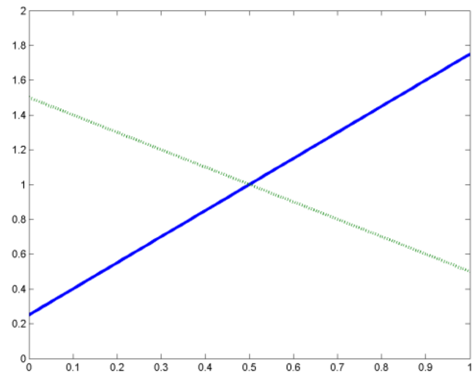
Cases where an *L1* seller's pessimism about the buyer yields larger *L1* trading regions, with some ex-post perverse trade for very high *V* and *C*

(Buyer's value V is on the vertical axis; seller's value C is on the horizontal axis. All value densities are linear; "x, y" means the buyer's density f(V) satisfies f(0) = x and f(1) = 2-x, and the seller's density g(C) satisfies g(0) = y and g(1) = 2-y.)

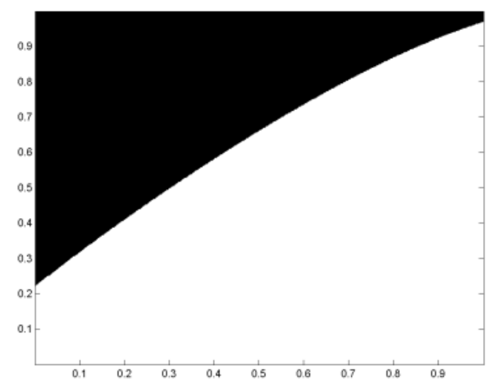
Equilibrium: 0.25, 1.5



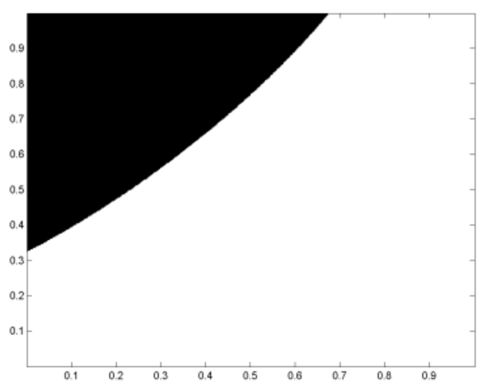
Buyer (—) and seller (···)



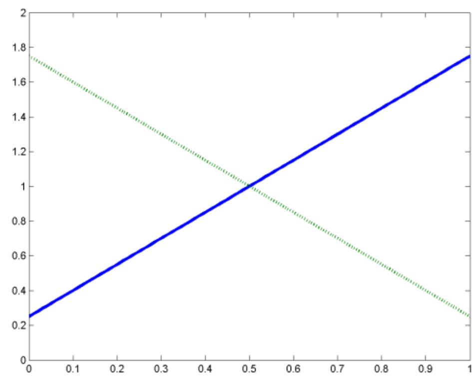
L1: 0.25, 1.5



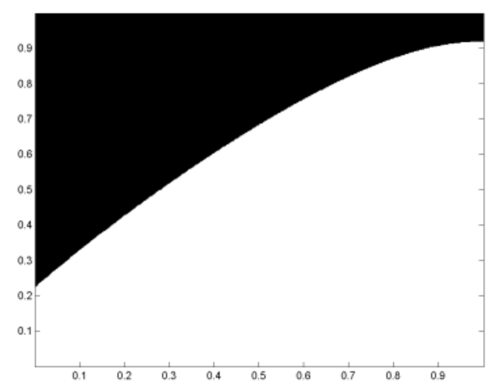
Equilibrium: 0.25, 1.75



Buyer (—) and seller (···)



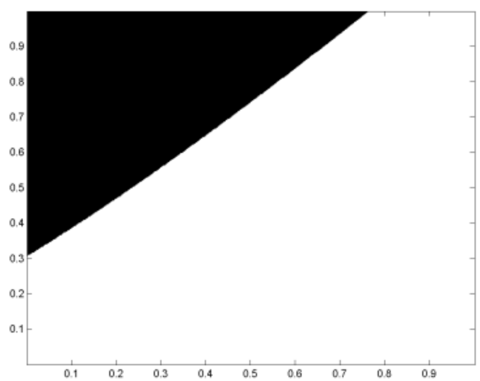
L1: 0.25, 1.75



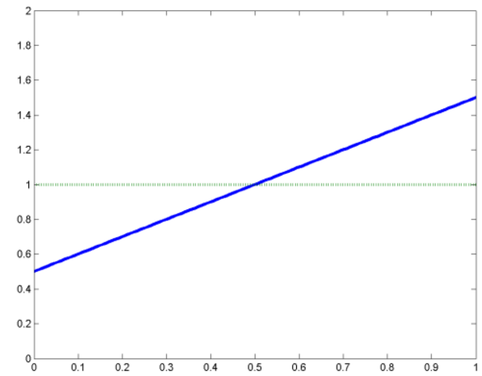
Cases where the optimism/pessimism intuition about sizes of trading regions fails, because the comparisons' changes in true densities affect the objective function as well

(Buyer's value V is on the vertical axis; seller's value C is on the horizontal axis. All value densities are linear; "x, y" means the buyer's density f(V) satisfies f(0) = x and f(1) = 2-x, and the seller's density g(C) satisfies g(0) = y and g(1) = 2-y.)

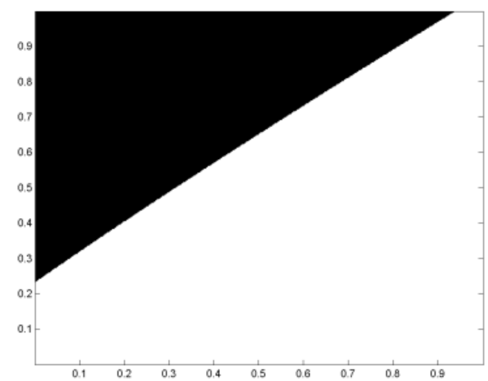
Equilibrium: 0.5, 1.0



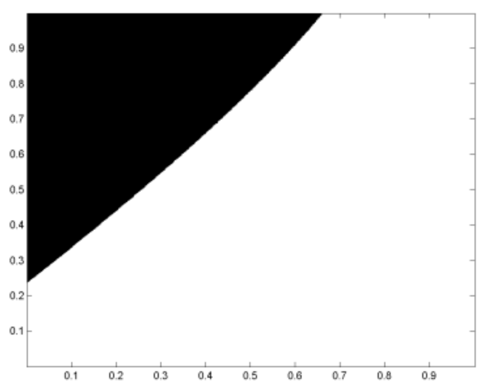
Buyer (—) and seller (···)



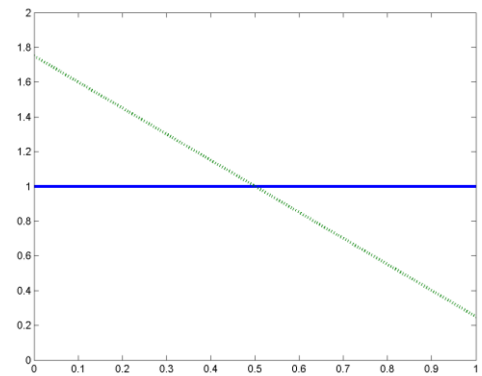
L1: 0.5, 1.0



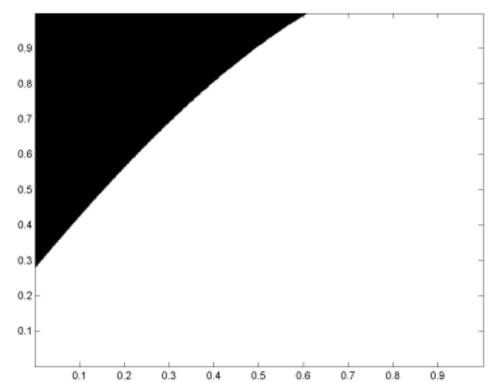
Equilibrium: 1.0, 1.75



Buyer (—) and seller (···)



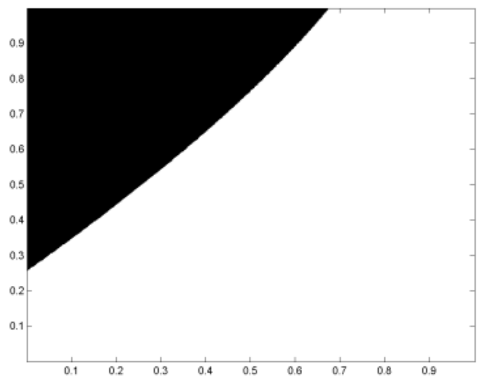
L1: 1.0, 1.75



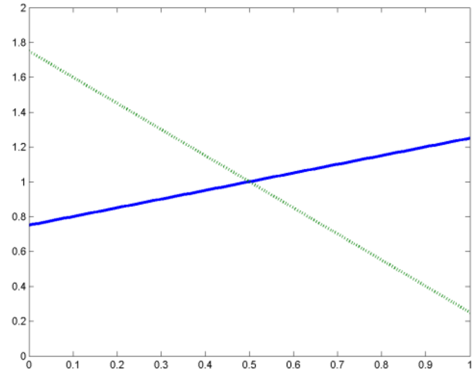
Cases where the equilibrium and *L1* trading regions overlap

(Buyer's value V is on the vertical axis; seller's value C is on the horizontal axis. All value densities are linear; "x, y" means the buyer's density f(V) satisfies f(0) = x and f(1) = 2-x, and the seller's density g(C) satisfies g(0) = y and g(1) = 2-y.)

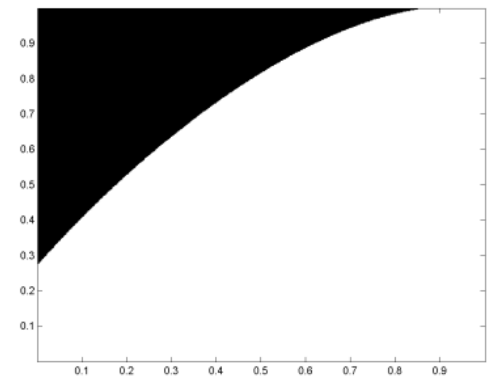
Equilibrium: 0.75, 1.75



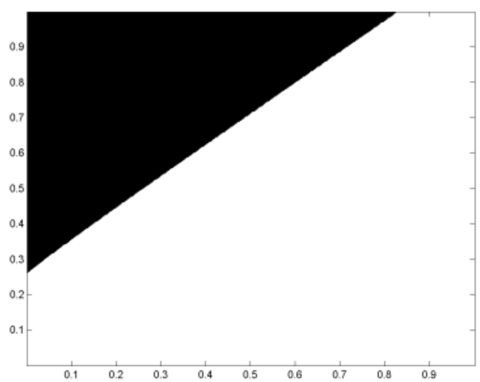
Buyer (—) and seller (···)



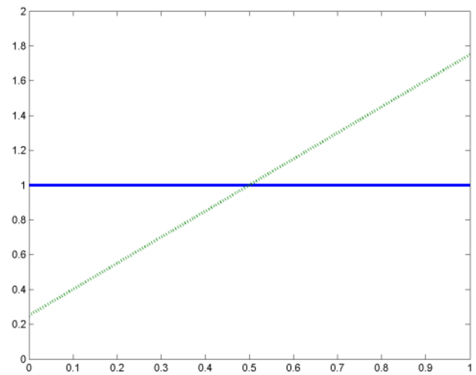
L1: 0.75, 1.75



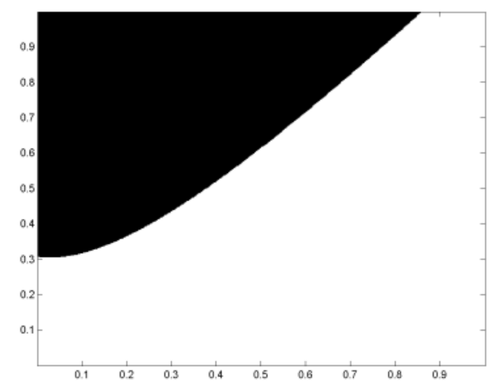
Equilibrium: 1.0, 0.25



Buyer (—) and seller (···)



L1: 1.0, 0.25



Mechanisms that are efficient in the set of level-*k*-incentive-efficient mechanisms with general value densities and heterogeneous level populations

This case has no counterpart in MS's analysis.

Continuing to assume that level-*k*-incentive-compatibility is required, and allowing general value densities, relax the assumption that the population is concentrated on one level.

Assume a known mixture of *L1* and *L2* traders, and allow only a single direct mechanism, as there is no evidence on which to base a specification of a level-k model for more complex menus.

Suppose—here it may not follow from optimization—that level-*k*-incentive-efficient mechanisms still set $U_B^k(0) = U_S^k(1) = 0$.

Then only trivial mechanisms can fully screen traders' values and levels. Screening conditions like (6.5) require different transfers for different levels, but traders would select the higher transfer.

As a result, it will normally be suboptimal to fully screen traders.

Suppose—here it may not follow from optimization—that level-*k*-incentive-efficient mechanisms still set $U_B^k(0) = U_S^k(1) = 0$.

Then only trivial mechanisms can fully screen traders' values and levels. Screening conditions like (6.5) require different transfers for different levels, but traders would select the higher transfer.

As a result, it will normally be suboptimal to fully screen traders.

If the population contains mostly *L1*s (*L2*s), the level-*k* incentive-efficient mechanism is probably optimized for *L1*s (*L2*s), ignoring the rarer level but still getting some expected surplus from it.

In general, however, the problem of screening levels interacts with the problem of screening values in complex ways, and there may be no simple structure on which values are screened.

***Lk*-incentive-efficient mechanisms relaxing *Lk*-incentive-compatibility**

This case has no counterpart in MS's analysis.

Return to a known population of traders concentrated on one level.

Recall that if *Lk*-incentive-compatibility is not required, it might be beneficial to use a non-*Lk*-incentive-compatible direct mechanism.

"*Lk*-incentive-efficient" now refers to a mechanism that cannot be improved upon by any feasible direct mechanism for *Lk* traders.

In this case one can still define a general class of feasible direct mechanisms; and the payoff-relevant aspects of a mechanism are still described by outcome functions $p(\cdot, \cdot)$ and $x(\cdot, \cdot)$.

However, even a direct mechanism's incentive effects can no longer be tractably captured via incentive constraints. Instead they must be modeled via level-k traders' responses to it.

For tractability, I focus on double auctions with reserve prices chosen by the designer, and on uniform value densities.

In this case one can still define a general class of feasible direct mechanisms; and the payoff-relevant aspects of a mechanism are still described by outcome functions $p(\cdot, \cdot)$ and $x(\cdot, \cdot)$.

However, even a direct mechanism's incentive effects can no longer be tractably captured via incentive constraints. Instead they must be modeled via level-*k* traders' responses to it.

For tractability, I focus on double auctions with reserve prices chosen by the designer, and on uniform value densities.

Reserve prices would have no effect if *L0* is uniform random on the full range of possible values [0, 1], as assumed so far.

But a restricted menu might make *Lk* players anchor beliefs instead on the restricted range of bids or asks, and that can make reserve prices useful in trading mechanisms (CKNP).

For example, in the double auction with uniform value densities, *L1* traders believe they face bids or asks uniformly distributed on [0, 1], which leads to incentive-*in*efficient outcomes.

To implement the outcome of MS's equilibrium-incentive-efficient direct mechanism via the double auction, *L1* traders have to believe that they face bids or asks uniform on [1/4, 3/4], the range of "serious" bids or asks in CS's linear double-auction equilibrium.

If *L1* traders anchor on the restricted menu, those beliefs can be induced by restricting bids to [1/4, 3/4] and asks to [1/4, 3/4]. (The upper ask limit could be raised to 1 and the lower bid limit to 0.)

Thus with uniform value densities, for *L1s* a double auction with reserve prices can take advantage of TEPIB to mimic MS's equilibrium-incentive-efficient mechanism, whose direct form is then efficient in the set of *L1*-incentive-compatible mechanisms.

(MS's general specification of feasible mechanisms implicitly allows reserve prices, and their analysis therefore shows that if equilibrium is assumed, reserve prices are not useful here.)

Computations suggest that more stringent reserve prices can further improve upon MS's equilibrium-incentive-efficient mechanism, by taking fuller advantage of TEPIB.

For *L2*s with uniform value densities, my analysis of the double auction without reserve prices shows that it can improve upon a mechanism that is efficient in the set of *L2*-incentive-efficient mechanisms, or MS's equilibrium-incentive-efficient mechanism.

Computations again suggest that reserve prices allow even more improvement, via TEPIB.

For *L2*s with uniform value densities, my analysis of the double auction without reserve prices shows that it can improve upon a mechanism that is efficient in the set of *L2*-incentive-efficient mechanisms, or MS's equilibrium-incentive-efficient mechanism.

Computations again suggest that reserve prices allow even more improvement, via TEPIB.

More generally, relaxing the restriction to level-*k*-incentive-compatible mechanisms can yield level-*k*-incentive-efficient mechanisms that differ qualitatively as well as quantitatively from equilibrium-incentive-efficient mechanisms, with substantial gains in incentive-efficiency.

Summary

- A level-*k* model adds enough specificity to allow an analysis with power comparable to that of an equilibrium analysis, clarifying the role of the equilibrium assumption in MS's analysis.

Summary

- A level-*k* model adds enough specificity to allow an analysis with power comparable to that of an equilibrium analysis, clarifying the role of the equilibrium assumption in MS's analysis.

- Via level-*k* menu effects, the choice between and *Lk*-incentive-compatible and a non-*Lk*-incentive-compatible mechanism can influences the correctness of level-*k* beliefs, and is not neutral.

- This compels a choice about whether to require *Lk*-incentive-compatibility.

- If level-*k*-incentive-compatibility is required, MS's characterization of incentive-efficient mechanisms for general value densities is robust to level-*k* thinking for populations concentrated on one level.

- As a result, the design features that foster equilibrium-incentive-efficiency in MS's analysis also foster efficiency in the set of level-*k*-incentive-compatible mechanisms, with different weights.

- If level-*k*-incentive-compatibility is required, MS's characterization of incentive-efficient mechanisms for general value densities is robust to level-*k* thinking for populations concentrated on one level.

- As a result, the design features that foster equilibrium-incentive-efficiency in MS's analysis also foster efficiency in the set of level-*k*-incentive-compatible mechanisms, with different weights.

- The level-*k* analysis reveals a novel design feature, TEPIB (tacit exploitation of predictably incorrect beliefs): Mechanisms that are efficient among level-*k*-incentive-compatible mechanisms exploit nonequilibrium beliefs, without active deception, in the benign sense of implementing outcomes that increase welfare.

- Mechanisms that are efficient among *L1*-incentive-compatible mechanisms perform better when the seller's uniform beliefs are pessimistic relative to the buyer's true density (that is, when the buyer's density first-order stochastically dominates the uniform).

- Mechanisms that are efficient among *L1*-incentive-compatible mechanisms perform better when the seller's uniform beliefs are pessimistic relative to the buyer's true density (that is, when the buyer's density first-order stochastically dominates the uniform).

- Such mechanisms then exploit TEPIB to implement trading regions that are supersets of those for equilibrium-incentive-efficient mechanisms, and obtain higher expected total surplus.

- Mechanisms that are efficient among *L1*-incentive-compatible mechanisms perform better when the seller's uniform beliefs are pessimistic relative to the buyer's true density (that is, when the buyer's density first-order stochastically dominates the uniform).

- Such mechanisms then exploit TEPIB to implement trading regions that are supersets of those for equilibrium-incentive-efficient mechanisms, and obtain higher expected total surplus.

- For some extreme densities, such mechanisms require ex post perverse trade for some extreme value combinations.

- Mechanisms that are efficient among *L1*-incentive-compatible mechanisms perform better when the seller's uniform beliefs are pessimistic relative to the buyer's true density (that is, when the buyer's density first-order stochastically dominates the uniform).

- Such mechanisms then exploit TEPIB to implement trading regions that are supersets of those for equilibrium-incentive-efficient mechanisms, and obtain higher expected total surplus.

- For some extreme densities, such mechanisms require ex post perverse trade for some extreme value combinations.

- When sellers' beliefs are optimistic, such mechanisms still exploit TEPIB, but equilibrium-incentive-efficient trading regions are usually supersets of the *L1* trading regions.

- Even if level-*k*-incentive-compatibility is required, MS's Corollary 1, that no incentive-compatible, interim individually rational mechanism can be ex post efficient with probability one, does not fully extend to level-*k* models with known populations.

- Even if level-*k*-incentive-compatibility is required, MS's Corollary 1, that no incentive-compatible, interim individually rational mechanism can be ex post efficient with probability one, does not fully extend to level-*k* models with known populations.

- Despite the possibility that anchoring beliefs on a uniform *L0* might reduce sensitivity to distributional and knowledge assumptions, *Lk* mechanisms are no less sensitive than equilibrium-incentive-efficient mechanisms.

- Even if level-*k*-incentive-compatibility is required, MS's Corollary 1, that no incentive-compatible, interim individually rational mechanism can be ex post efficient with probability one, does not fully extend to level-*k* models with known populations.

- Despite the possibility that anchoring beliefs on a uniform *L0* might reduce sensitivity to distributional and knowledge assumptions, *Lk* mechanisms are no less sensitive than equilibrium-incentive-efficient mechanisms.

- Sorting traders' levels along with their values poses formidable new analytical problems.

- And if direct but non-level-*k*-incentive-compatible mechanisms are usable, level-*k*-incentive-efficient mechanisms may differ qualitatively as well as quantitatively from equilibrium-incentive-efficient mechanisms, with possibly substantial efficiency gains.