Efficient Mechanisms for Level-*k* Bilateral Trading

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

Thanks to many, many people for very helpful discussions and comments on earlier versions; and to Rustu Duran for outstanding research assistance.

The research leading to these results has received 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 funding.

Revised 15 March 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.

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

I relax MS's equilibrium assumption in favor of a structural nonequilibrium "level-*k*" model of strategic thinking meant to describe people's responses to novel or complex games, adapted to handle asymmetric information.

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

Motivation

- Mechanism design often creates novel games, for which the usual learning justification for equilibrium is weak; yet the design may still 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, for which the usual learning justification for equilibrium is weak; yet the design may still 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 of the 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.

Estimates vary with the setting and population, but the estimated frequency of *L0* is normally 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.

- 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).

- 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).

- A level-*k* analysis can identify settings where equilibrium-based conclusions 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).

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

- A level-*k* analysis can 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).

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

- A level-*k* analysis can 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 may 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 fraction of truthful traders.

- Kneeland's 2013 analysis of level-*k* implementation, with illustrations including bilateral trading.

Outline

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

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

- Defining level-*k* models with asymmetric information.

- Level-*k* analysis of the double auction: *L1* aggressivness and incentive-*in*effiency; *L2* meekness and incentive-*super*effiency.

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

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

- Generalizations: Relaxing level-*k*-incentive-compatibility; relaxing the population’s being concentrated 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.

Traders' vN-M utility functions are quasilinear in money: hence they are 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 with no 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-based 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* mechanism asks traders to report their values, and makes 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.

A *direct* mechanism asks traders to report their values, and makes 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 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 respective values.

The revelation principle shows that 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”, subject to which incentive-efficient outcome functions p(·,·) and x(·,·) 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 these results discussed below will explain MS’s results in more detail.)

In CS's example with uniform value densities, MS's Theorem 2 yields a closed-form solution for the incentive-compatible form of the 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 Theorem 2 yields a closed-form solution for the incentive-compatible form of the 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 the 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 effect 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 the 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 effect 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. Thus MS's result for this example can be viewed as coincidental.)

Level-*k* models with asymmetric information

- Recall that in a level-*k* model, people anchor beliefs in a naïve model of others' reactions to the game, *L0*, and adjust them 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.

Level-*k* models with asymmetric information

- Recall that in a level-*k* model, people anchor beliefs in a naïve model of others' reactions to the game, *L0*, and adjust them 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 assume *L0*'s bids or asks are uniform over the possible values and *independent of own value*.

- This *L0* yields a hierarchy of rules via iterated best responses (called a "random level-*k* model", though only *L0* is random).

- 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 players reason contingent on others' possible values, but behaviorally far-fetched.

- And if others' private information varies over time, no amount of learning is likely to transcend such cognitive limits.

- 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 players reason contingent on others' possible values, but behaviorally far-fetched.

- And if others' private information varies over time, no amount of learning is likely to transcend such cognitive limits.

- A random level-*k* model captures people's aversion to fixed-point and contingent reasoning in a tractable way, which gives a realistic account of people's strategic thinking and their naiveté regarding others' uses of private information.

Random level-*k* models have a long history:

> "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)

Dad is worried that Son will follow a random *L1* rule, rational but sticking with his prior in the face of an offer "too good to be true".

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—giving them new information, fundamentals unchanged, cannot lead to new trades.

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

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—giving them new information, fundamentals unchanged, cannot lead to new trades.

Any such new 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, because it rests on equilibrium-like inferences such as:

> "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 Friar's Club

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—giving them new information, fundamentals unchanged, cannot lead to new trades.

Any such new 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, because it rests on equilibrium-like inferences such as:

> "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 Friar's Club

Milgrom and Stokey contrast the correct, equilibrium-based inference with a rule they call Naïve Behavior, which sticks with its prior and otherwise behaves rationally, just as random *L1* does; and another rule called First-Order Sophistication, which best responds to Naïve Behavior, just as random *L2* does.

Inspired by the No-Trade Theorem, Brocas, Carillo, Camerer, and Wang 2014 *REStud* report powerful experimental evidence from three-state betting games (close enough to zero-sum):

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

Inspired by the No-Trade Theorem, Brocas, Carillo, Camerer, and Wang 2014 *REStud* report powerful experimental evidence from three-state betting games (close enough to 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.

Inspired by the No-Trade Theorem, Brocas, Carillo, Camerer, and Wang 2014 *REStud* report powerful experimental evidence from three-state betting games (close enough to 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 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 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 (Bet, Pass):

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

The 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 (Bet, Pass):

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

Round 3 (Bet, Pass):

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, in Brocas et al.'s and several similar previous experiments, half of the subjects Bet, in patterns that varied systematically with player role and state.

Random *L1* respects simple dominance (Bet, Pass):

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

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

Random *L1* respects simple dominance (Bet, Pass):

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

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

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

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

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

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

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

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

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

And random *L3* respects three rounds of iterated weak dominance (= trembling-hand perfect Bayesian 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.

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

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

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

Assume that a player's level is independent of its value.

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

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

Assume that a player's level is independent of its value.

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.

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

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.

L1 aggressivness and incentive-*in*effiency

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 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).

Comparing the equilibrium bidding function $b_*(V) = 2V/3 + 1/12$ and the *L1* bidding function $b_1(V) = 2V/3$, an *L1* buyer bids 1/12 more aggressively (that is, bids less) than an equilibrium buyer.

An *L1* buyer's naïve model of the seller leads it to underestimate the seller's upward-shaded ask relative to equilibrium, which makes it underbid (compare Crawford and Iriberri's 2007 *ECMA* analysis of random *L1* bidding in first-price auctions).

Similarly, an *L1* seller asks 1/12 more aggressively (that is, asks more) than an equilibrium seller.

L1 buyers' and sellers' strategies have the same slope, 2/3, as equilibrium buyers' and sellers' strategies, but are 1/12 more aggressive.

If an *L1* buyer meets an *L1* seller, trade takes place iff $V \geq C + \frac{1}{2}$ (as opposed to $V \geq C + \frac{1}{4}$ for an equilibrium buyer and seller).

The value gap needed for trade is ¼ larger for *L1* buyers and sellers than for equilibrium buyers and sellers; thus ex post efficiency is lost for more value combinations than in equilibrium.

The ex ante probability of trade for *L1*s is only 1/8 = 12.5%, in comparison to the equilibrium probability of 9/32 ≈ 28%, and the largest possible interim individually rational probability of 50%.

If an *L1* buyer meets an *L1* seller, trade takes place iff $V \geq C + ½$ (as opposed to $V \geq C + ¼$ for an equilibrium buyer and seller).

The value gap needed for trade is ¼ larger for *L1* buyers and sellers than for equilibrium buyers and sellers; thus ex post efficiency is lost for more value combinations than in equilibrium.

The ex ante probability of trade for *L1*s is only 1/8 = 12.5%, in comparison to the equilibrium probability of 9/32 ≈ 28%, and the largest possible interim individually rational probability of 50%.

These highly inefficient outcomes raise the question of whether a designer who knows that all traders are *L1*s can design a mechanism that enhances their efficiency by counteracting their aggressiveness better than CS's double auction, which is incentive-efficient for equilibrium bargainers, does.

I will show that the answer is Yes.

L2 meekness and incentive-*super*effiency

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 ≤ 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 ≤ 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 ϵ [1/3, 1]

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

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 ϵ [1/3, 1].

In the interior, an *L2* buyer bids 1/36 less aggressively (higher) than an equilibrium buyer and 1/9 less aggressively than an *L1* buyer.

This happens because an *L2* buyer's *L1* model of the seller leads it to overestimate the seller's upward-shaded ask relative to equilibrium, which induces an *L2* buyer to overbid (as in Crawford and Iriberri's analysis of random *L2* bidding in first-price auctions).

Similarly, an *L2* seller's ask a_2(C) maximizes over a in [0, 1]

$$\int_a^1 \frac{a+b}{2} f\left(b_1^{-1}(b)\right) db + \int_0^a C\, f\left(b_1^{-1}(b)\right) db,$$

where $f(b_1^{-1}(b))$ is the density of an *L1* buyer's bid b_1(V) induced by the value density f(V).

For instance, if f(V) is uniform, an *L2* seller believes that the buyer's bid b_1(V) = 2V/3 is uniformly distributed on [0, 2/3], with density 3/2 there and 0 elsewhere.

An *L2* seller who believes the buyer's bid is distributed on [0, 2/3] believes that trade requires a < 2/3.

For C ≥ 2/3 it is therefore optimal for an *L2* seller to bid anything it thinks yields zero probability of trade. In the absence of dominance among such strategies, I set $a_2(C) = C$ for C in (2/3, 1].

For C < 2/3, an *L2* seller's ask $a_2(C)$ maximizes over a in [0, 2/3]

$$\int_a^{\frac{2}{3}} \frac{a+b}{2} (3/2) db + \int_0^a C\,(3/2) db.$$

The second-order condition is satisfied when f(V) is uniform.

Solving the first-order condition (3/2)(a-C) + (3/2)(2/3 - C)/2 = 0, which yields $a_2(C) = 2C/3 + 2/9$ for C in [0, 2/3].

In the interior an *L2* seller asks 1/36 less aggressively (lower) than an equilibrium seller, and 1/9 less aggressively than an *L1* seller.

This happens because an *L2* seller's *L1* model of the buyer leads it to underestimate the buyer's downward-shaded bid relative to equilibrium, which induces an *L2* seller to underask.

With uniform value densities, *L2* buyers' and sellers' strategies again have the same slope, 2/3, as equilibrium buyers' and sellers' strategies, but are 1/36 less aggressive.

When an *L2* buyer meets an *L2* seller, trade takes place if V ≥ C + 1/6, versus V ≥ C + ¼ for an equilibrium buyer and seller.

The value gap needed for trade is 1/12 smaller than for equilibrium buyers and sellers and 1/3 smaller than for *L1* buyers and sellers, so ex post efficiency is lost for fewer values than with equilibrium or *L1* bargainers.

The ex ante probability of trade is 25/72 ≈ 35%, higher than the equilibrium probability of 28% or the *L1* probability of 12.5%, but still less than the maximum individually rational probability of 50%.

These much more efficient outcomes raise the question of whether a designer who knows (say) that all traders are *L2*s can design a mechanism that further enhances their efficiency by exploiting their meekness even better than CS’s double auction.

These much more efficient outcomes raise the question of whether a designer who knows (say) that all traders are *L2*s can design a mechanism that further enhances their efficiency by exploiting their meekness even better than CS's double auction.

I will show that with uniform value densities, if we require incentive-compatibility, the answer is No.

But if we allow non-*L2*-incentive-compatible mechanisms such as the double auction, the answer may change to Yes.

Failures of the revelation principle and "level-*k* menu effects"

- How can CS's double auction yield less efficient outcomes for *L1* traders than the equilibrium-incentive-compatible direct version of CS's double auction, which is MS's equilibrium-incentive-efficient direct mechanism?

- How can CS's double auction yield more efficient outcomes for *L2* traders than MS's equilibrium-incentive-efficient mechanism?

Going from a double auction to MS's equilibrium-incentive-efficient direct mechanism (which is equilibrium-equivalent to it with uniform value densities) creates level-*k* menu effects that change the relation between a level-*k* trader's beliefs and correct beliefs:

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

- The double auction improves upon MS's equilibrium-incentive-efficient mechanism for *L2*s by *not* rectifying their beliefs and thereby preserving their beneficial meekness.

Going from a double auction to MS's equilibrium-incentive-efficient direct mechanism (which is equilibrium-equivalent to it with uniform value densities) creates level-*k* menu effects that change the relation between a level-*k* trader's beliefs and correct beliefs:

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

- The double auction improves upon MS's equilibrium-incentive-efficient mechanism for *L2*s by *not* rectifying their beliefs and thereby preserving their beneficial meekness.

In each case menu effects make the revelation principle fail for level-*k* bargainers, even for direct mechanisms and with *L0* fixed.

Such effects are a residue of *Lk*'s anchoring on *L0*, which would be eliminated by equilibrium thinking.

The way the double auction improves upon MS's equilibrium-incentive-efficient direct mechanism for *L2*s by *not* rectifying their beliefs highlights an important feature of level-*k* mechanism design that cannot arise in an equilibrium-based analysis:

Tacit Exploitation of Predictably Incorrect Beliefs ("TEPIB")

The way the double auction improves upon MS's equilibrium-incentive-efficient direct mechanism for *L2*s by *not* rectifying their beliefs highlights an important feature of level-*k* mechanism design that cannot arise in an equilibrium-based analysis:

Tacit Exploitation of Predictably Incorrect Beliefs ("TEPIB")

- "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.

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

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

- 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.

- 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) favors 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).

The failure of the Revelation Principle for level-*k* models poses a hard choice: Should we require *Lk*-incentive-compatibility or not?

- Requiring *Lk*-incentive-compatibility may make implementation simpler, more transparent, and more reliable.

- Not requiring it may unfairly favor those who are willing to lie.

The failure of the revelation principle for level-*k* models poses a hard choice: Should we require *Lk*-incentive-compatibility or not?

- Requiring *Lk*-incentive-compatibility may make implementation simpler, more transparent, and more reliable.
- Not requiring it may unfairly favor those who are willing to lie.
- Yet many analysts find direct but non-incentive-compatible mechanisms such as the first-price auction acceptable.
- And if *Lk*-incentive-compatibility is not truly needed to make a mechanism workable, relaxing it may allow efficiency gains.

The failure of the revelation principle for level-*k* models poses a hard choice: Require *Lk*-incentive-compatibility, or not?

- Requiring *Lk*-incentive-compatibility may make implementation simpler, more transparent, and more reliable.

- Not requiring it may unfairly favor those who are willing to lie.

- Yet many analysts find direct but non-incentive-compatible mechanisms such as the first-price auction acceptable.

- And if *Lk*-incentive-compatibility is not truly needed to make a mechanism workable, relaxing it may allow efficiency gains.

It's partly an empirical question, which I do not try to resolve here.

Most of my analysis requires *Lk*-incentive-compatibility, which allows a more complete analysis; but I briefly consider relaxing it.

Efficient mechanisms for level-*k* bilateral trading

In analyzing efficient mechanisms for level-k bilateral trading, I restrict attention throughout to direct mechanisms—including the double auction, in which bids are conformable to value reports—for which there is some evidence to guide the specification of *L0*.

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

(Screening traders' levels along with their values is very difficult.)

I ignore the noisiness of people's decisions, as in MS's and almost all other analyses of design.

- I first require *Lk*-incentive-compatibility: "mechanisms that are efficient in the set of *Lk*-incentive-compatible mechanisms".

- I then consider relaxing it: "*Lk*-incentive-efficient mechanisms".

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

Here my analysis closely parallels MS's analysis.

Recall that MS's Theorems 1 and 2 use conditions for equilibrium-incentive-compatibility to derive an "incentive budget constraint", subject to which 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.

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

Here my analysis closely parallels MS's analysis.

Recall that MS's Theorems 1 and 2 use conditions for equilibrium-incentive-compatibility to derive an "incentive budget constraint", subject to which 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, Theorems 1-2 yield a closed-form solution for the incentive-compatible direct form of the incentive-efficient mechanism, which then yields the same outcomes as CS's linear equilibrium of the double auction (with traders' equilibrium bidding strategies shading to mimic the effect of truthful reporting in MS's incentive-compatible mechanism).

My first result shows that MS's equilibrium-based result for uniform value densities is completely robust to level-*k* thinking when *Lk*-incentive-compatibility is required.

*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.*

My first result shows that MS’s equilibrium-based result for uniform value densities is completely robust to level-*k* thinking when *Lk*-incentive-compatibility is required.

*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. With uniform value densities, in the truthful equilibrium of MS’s equilibrium-incentive-efficient direct mechanism, each trader faces a uniform distribution of the other’s reports. *L1* traders best respond to *L0*s that also imply uniform distributions. *L1* traders’ conditions for individual rationality and incentive-compatibility therefore coincide with the analogous conditions for equilibrium traders; and because they make *L1*s report truthfully, so on for *Lk*s ad infinitum. The objective function, which reflects correct beliefs, is also the same as in MS’s equilibrium-based analysis. Q.E.D.

But to yield good outcomes for *Lk* traders, unlike for equilibrium traders, MS's equilibrium-incentive-efficient mechanism must be implemented in its direct, *Lk*-incentive-compatible form.

The examples of failures of the revelation principle above show that using the raw double auction for *L1* traders is *not* efficient in the set of *L1*-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.

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*.

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

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

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

For a given *k*, the mechanism *p*(·, ·), *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) $U_B^k(V) \geq V P_B^k(v) - X_B^k(v)$ and $U_S^k(C) \geq X_S^k(c) - C P_S^k(c)$.

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) $U_B^k(V) \geq V P_B^k(v) - X_B^k(v)$ and $U_S^k(C) \geq X_S^k(c) - C P_S^k(c)$.

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 $U_S^k(C) \geq 0$.

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

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

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

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

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

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 nonequilibrium 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,

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

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

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

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

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

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

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

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

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

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

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

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

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. Some of the "prices" are negative; but by the logic of the incentive constraints, there is no free disposal.

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

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 objective function and the constraint are linear in the $p(V,C)$, so the solution will be "bang-bang", with $p(V,C)$ = 0 or 1 a.e.

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

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

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

Given that $p(V,C) = 1$ if and only if (6.12) is satisfied, the optimal λ is set so that (6.5) holds with equality.

If the expression in

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

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

Then by Theorem B, the problem's solution is associated with a mechanism that maximizes expected total surplus among all level-*k*-incentive-compatible and level-*k*-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.

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 λ = 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*.

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 λ = 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 MS's equilibrium analysis.

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.

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.

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 $V < C$: consistent with level-*k*-interim-individually-rationality but "perverse" ex post.

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 $V < C$: consistent with level-*k*-interim-individually-rationality but "perverse" ex post.

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 $V < C$.

Such perverse trade can loosen (6.5) enough to compensate for the local loss in surplus by enabling trade for other value combinations, as illustrated in some of the 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 $V < C$: consistent with level-*k*-interim-individually-rationality but “perverse” ex post.

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 $V < C$.

Such perverse trade can loosen (6.5) enough to compensate for the local loss in surplus by enabling trade for other value combinations, as illustrated in some of the 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

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.

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

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, Figure 1 reports such mechanisms' trading regions for *L1*s' and representative linear density combinations.

(The combinations are a comprehensive coarse subset of all possible linear density combinations, with combinations excluded only for the few extreme combinations that violate Theorems B-C's monotonicity conditions for the mechanism to be truly optimal.)

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

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, Figure 1 reports such mechanisms' trading regions for *L1*s' and representative linear density combinations.

(The combinations are a comprehensive coarse subset of all possible linear density combinations, with combinations excluded only for the few 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.)

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

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

When *L1* sellers' beliefs are pessimistic relative to buyers' true densities, *L1* mechanisms exploit TEPIB to implement trading regions that are supersets of those for equilibrium-incentive-efficient mechanisms (except in Figure 1's case "0.75, 1.75", in which they overlap); *L1* mechanisms then yield higher surplus.

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

When *L1* sellers' beliefs are pessimistic relative to buyers' true densities, *L1* mechanisms exploit TEPIB to implement trading regions that are supersets of those for equilibrium-incentive-efficient mechanisms (except in Figure 1's case "0.75, 1.75", in which they overlap); *L1* mechanisms then yield higher surplus.

In two such combinations, "0.25, 1.5" and "0.25, 1.75", *L1* mechanisms have ex post perverse trade for very high *V* and *C*.

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

When *L1* sellers' beliefs are pessimistic relative to buyers' true densities, *L1* mechanisms exploit TEPIB to implement trading regions that are supersets of those for equilibrium-incentive-efficient mechanisms (except in Figure 1's case "0.75, 1.75", in which they overlap); *L1* mechanisms then yield higher surplus.

In two such combinations, "0.25, 1.5" and "0.25, 1.75", *L1* mechanisms have ex post perverse trade for very high *V* and *C*.

When *L1* sellers' beliefs are optimistic, *L1* mechanisms still exploit TEPIB, but equilibrium-incentive-efficient trading regions are supersets of the *L1* trading regions (except in cases "1.0, 0.25", "1.5, 0.25", and "1.5, 0.50", in which they overlap).

The seller's beliefs are more important than the buyer's because trade does not always occur when V > C, so initial ownership of the object breaks the symmetry between them.

The seller's beliefs are more important than the buyer's because trade does not always occur when $V > C$, so initial ownership of the object breaks the symmetry between them.

The terms in the buyer's density in the incentive budget constraints (3.5) and (6.5) are "virtual utilities", negative for low values of V but positive for high values; while the terms in the seller's density are not virtual utilities, and are always positive.

Even Figure 1's equilibrium trading regions are not symmetric across cases with buyer's and seller's densities interchanged.

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

(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.)

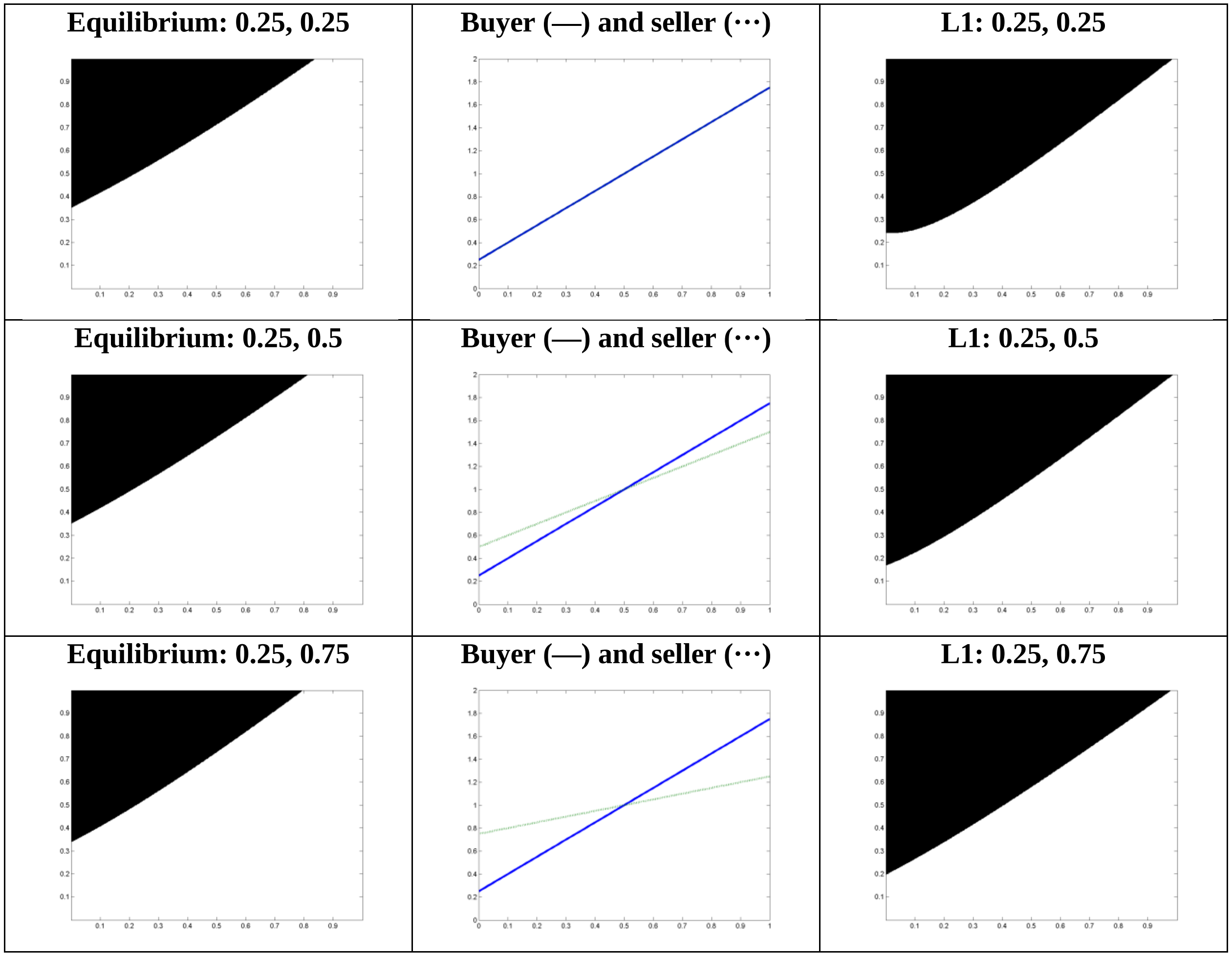


Figure 1 (cont.). Trading regions (in black) for equilibrium-incentive-efficient mechanisms and mechanisms that are efficient in the set of *L1*-incentive-compatible mechanisms

(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.)

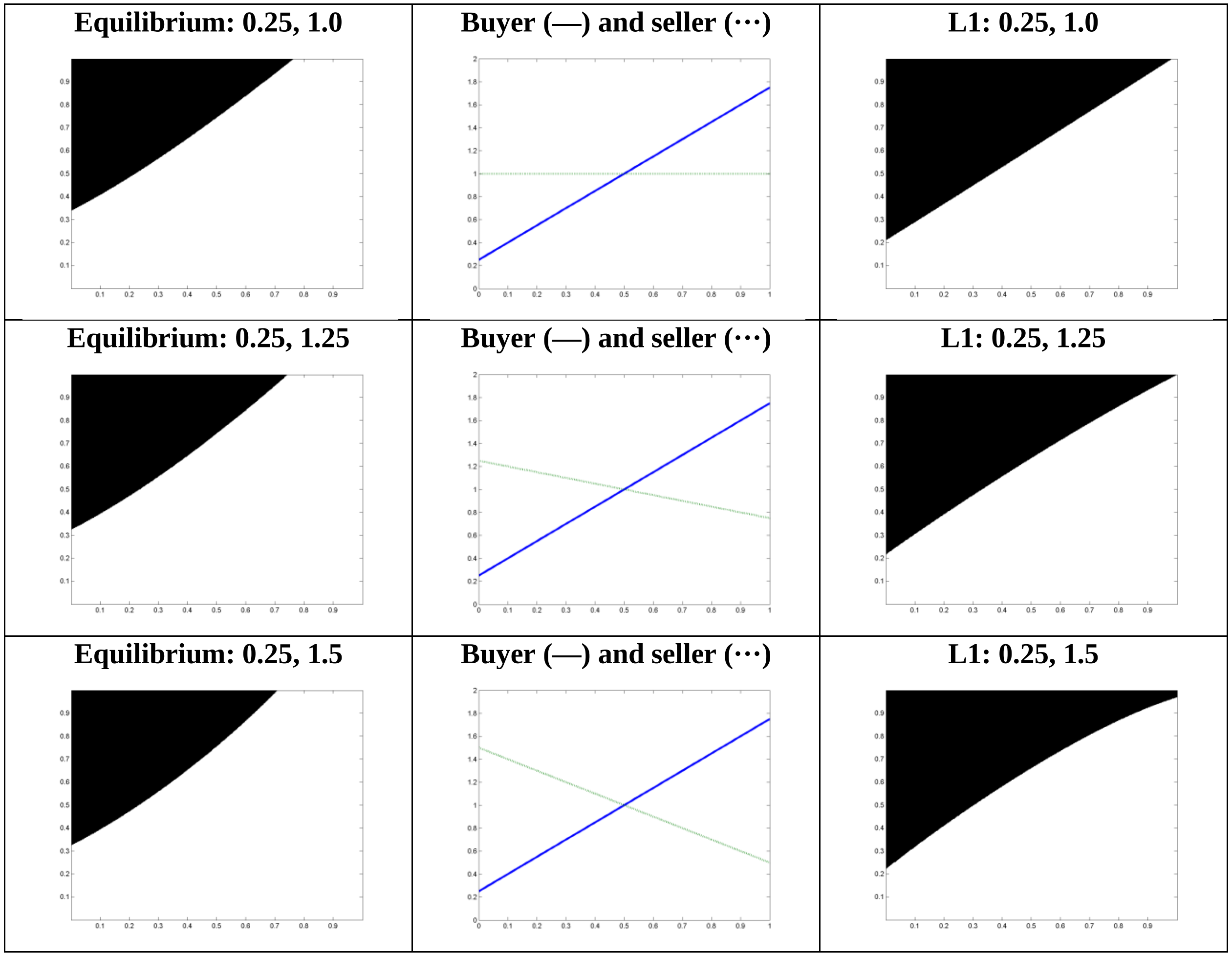


Figure 1 (cont.). Trading regions (in black) for equilibrium-incentive-efficient mechanisms and mechanisms that are efficient in the set of *L1*-incentive-compatible mechanisms

(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.)

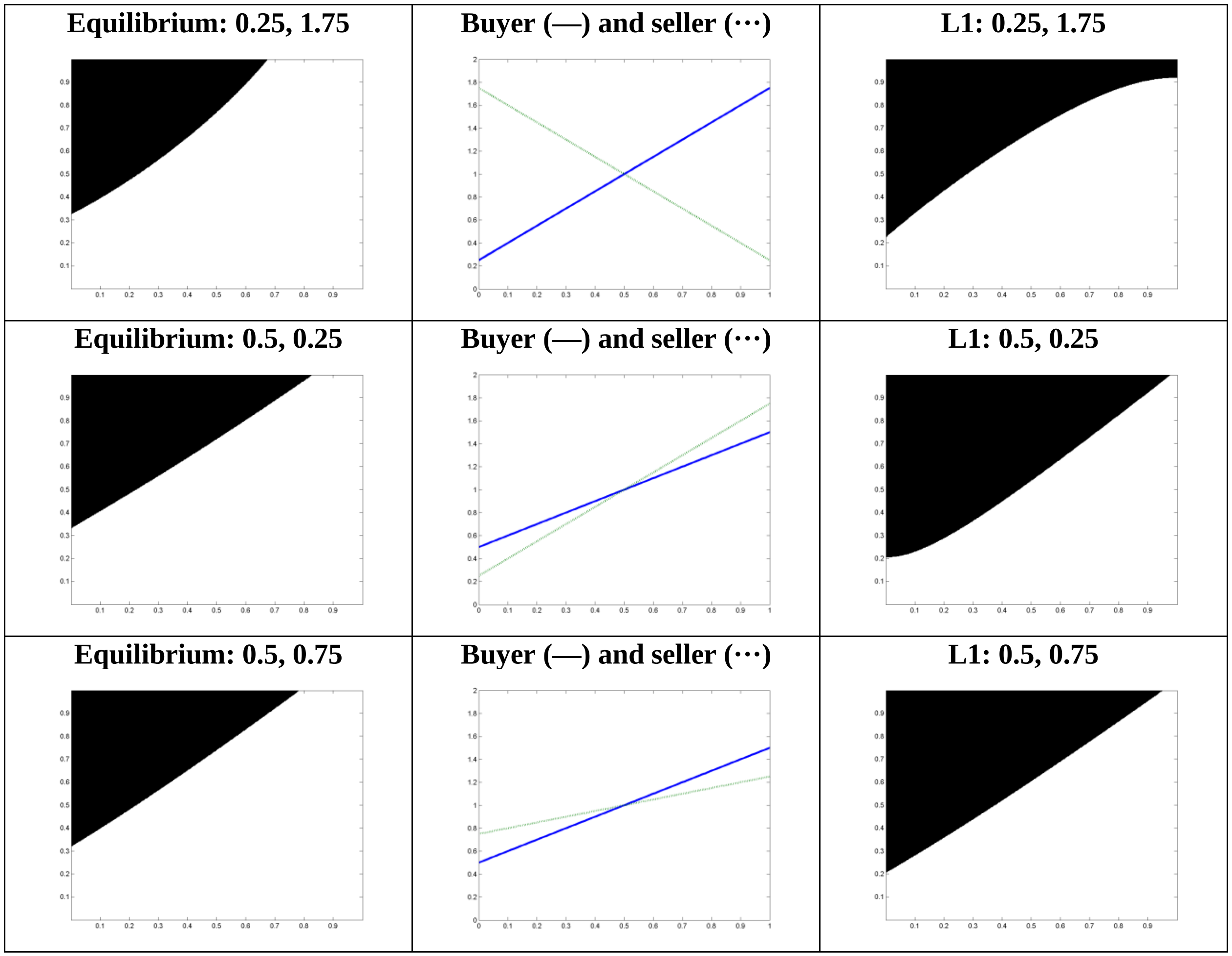


Figure 1 (cont.). Trading regions (in black) for equilibrium-incentive-efficient mechanisms and mechanisms that are efficient in the set of *L1*-incentive-compatible mechanisms

(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.)

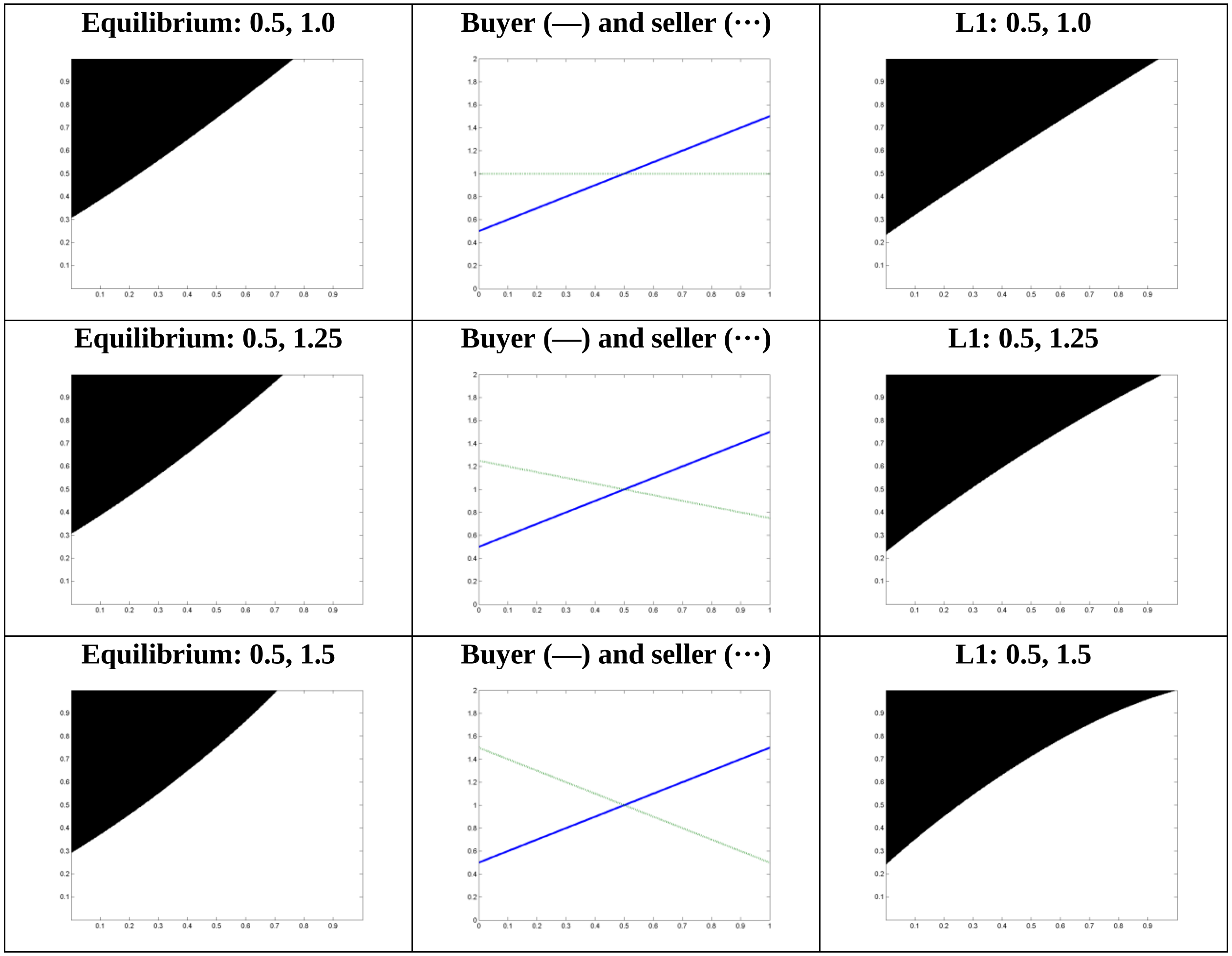


Figure 1 (cont.). Trading regions (in black) for equilibrium-incentive-efficient mechanisms and mechanisms that are efficient in the set of *L1*-incentive-compatible mechanisms

(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.)

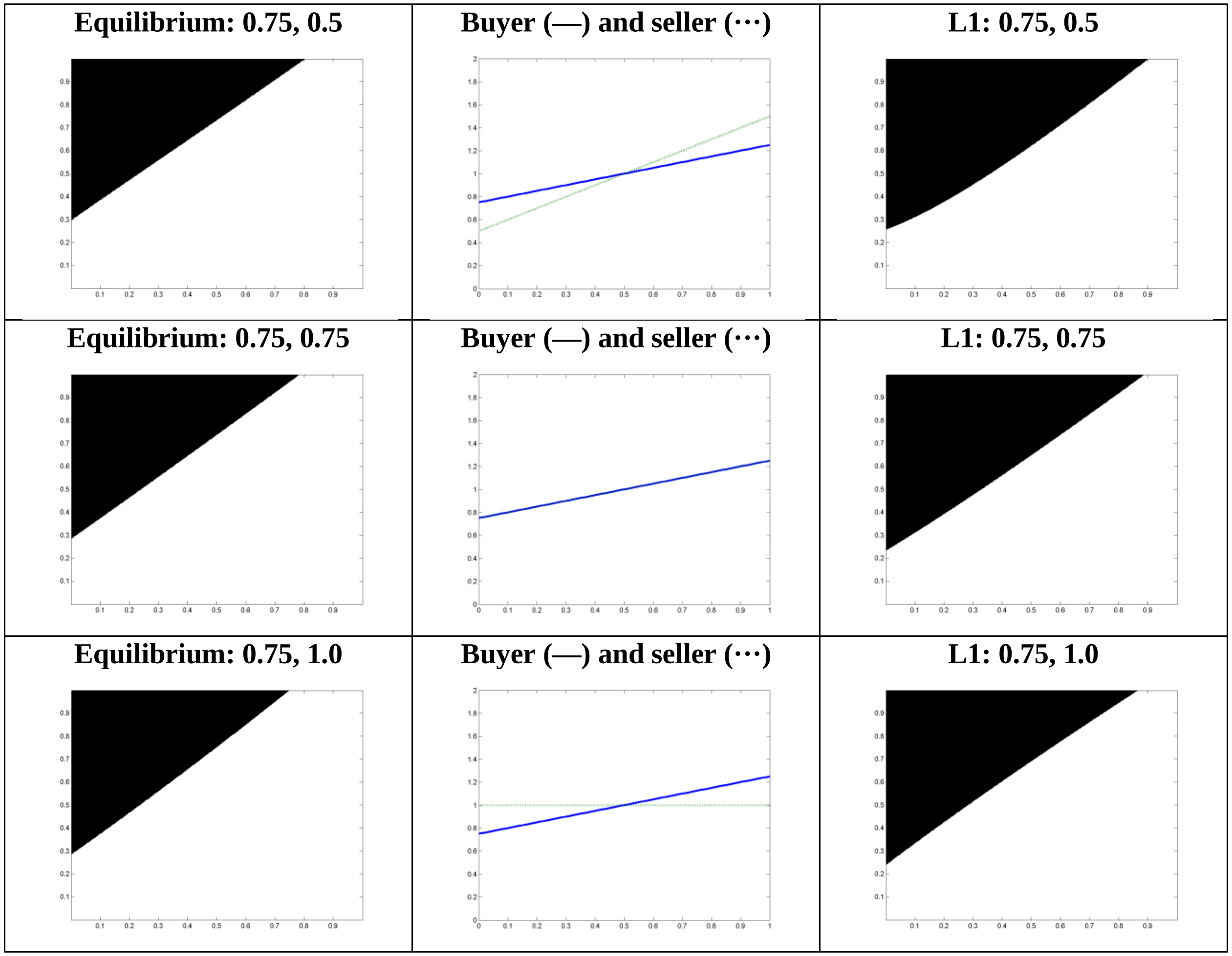


Figure 1 (cont.). Trading regions (in black) for equilibrium-incentive-efficient mechanisms and mechanisms that are efficient in the set of *L1*-incentive-compatible mechanisms

(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.)

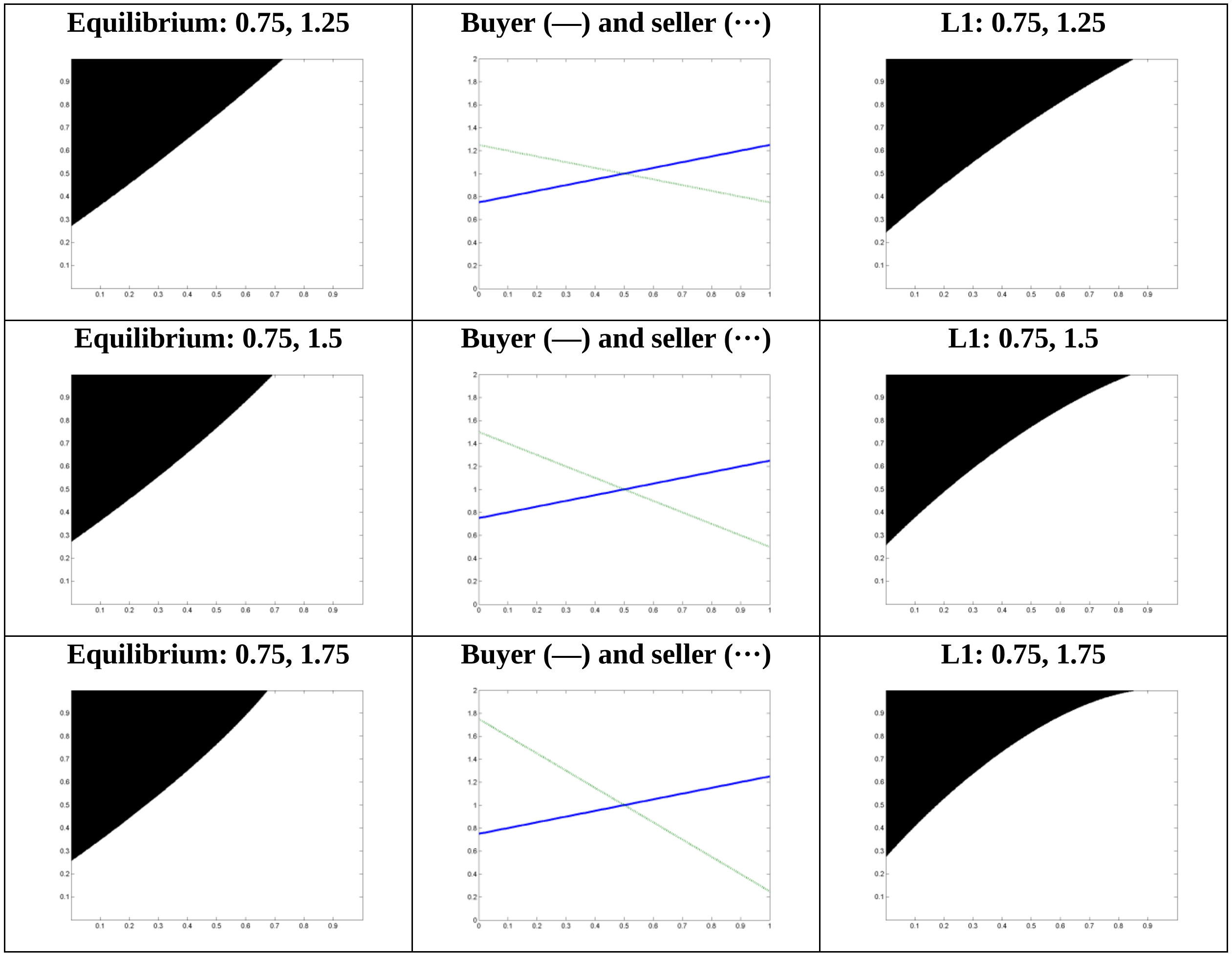


Figure 1 (cont.). Trading regions (in black) for equilibrium-incentive-efficient mechanisms and mechanisms that are efficient in the set of *L1*-incentive-compatible mechanisms

(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.)

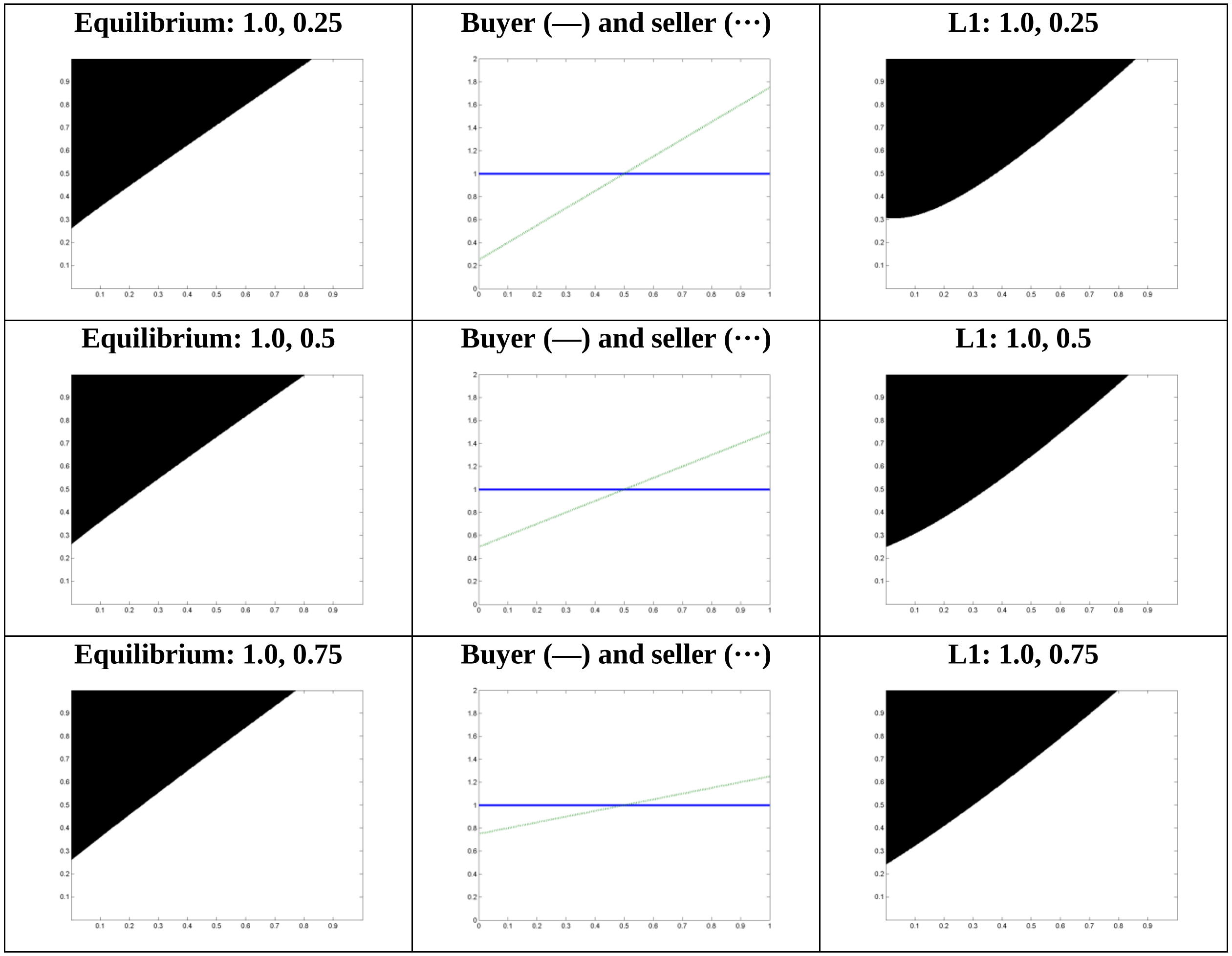


Figure 1 (cont.). Trading regions (in black) for equilibrium-incentive-efficient mechanisms and mechanisms that are efficient in the set of *L1*-incentive-compatible mechanisms

(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.)

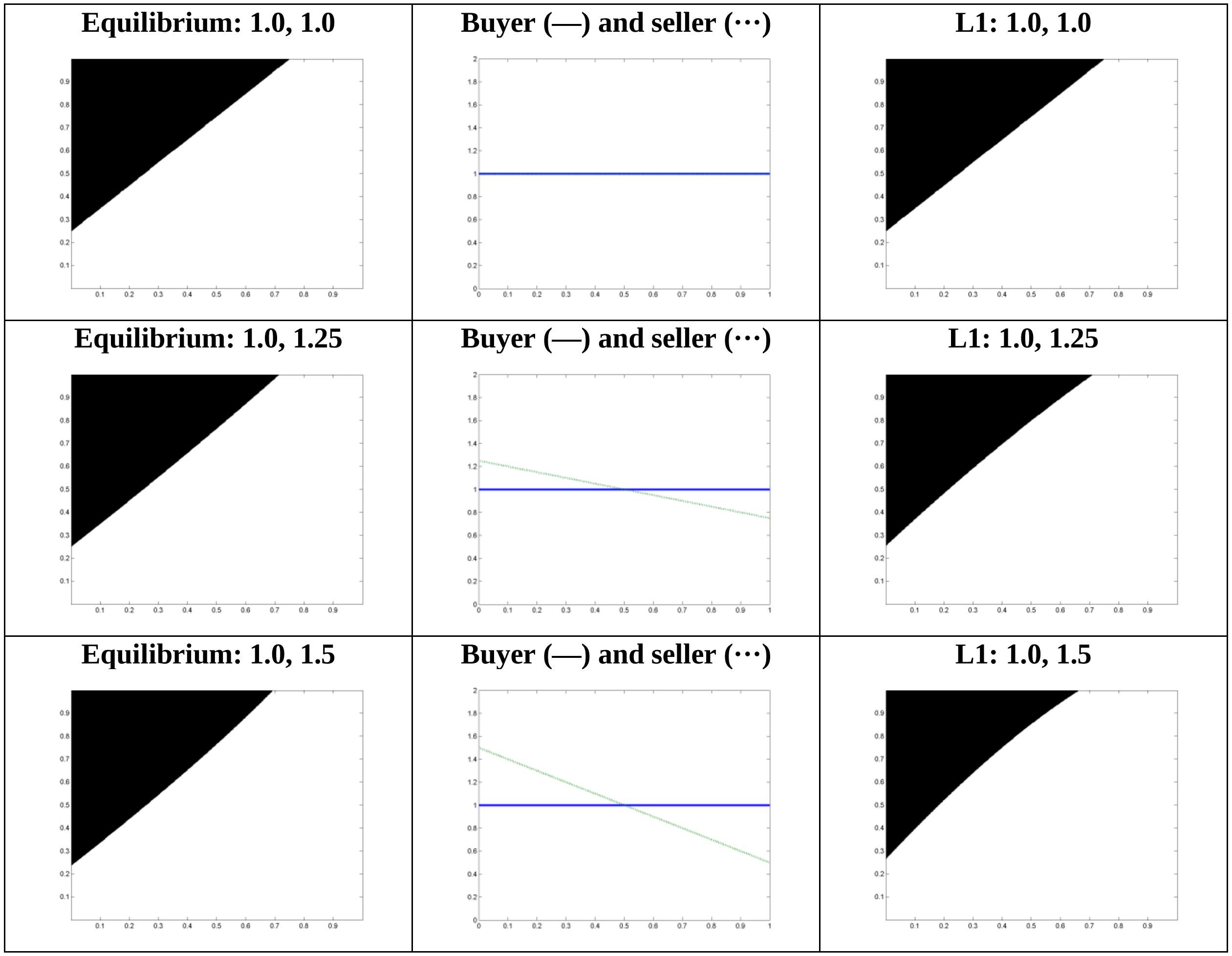


Figure 1 (cont.). Trading regions (in black) for equilibrium-incentive-efficient mechanisms and mechanisms that are efficient in the set of *L1*-incentive-compatible mechanisms

(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.)

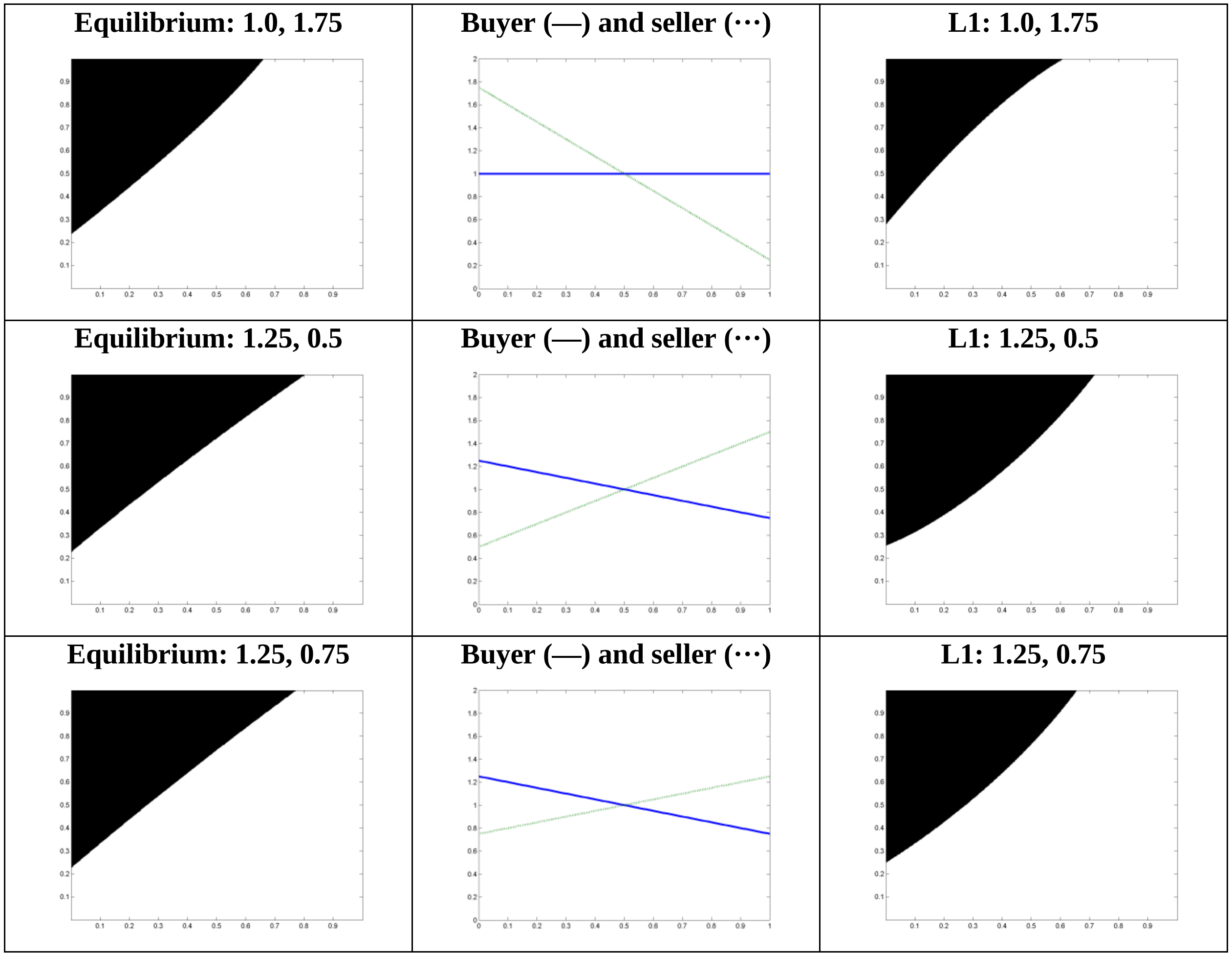


Figure 1 (cont.). Trading regions (in black) for equilibrium-incentive-efficient mechanisms and mechanisms that are efficient in the set of *L1*-incentive-compatible mechanisms

(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.)

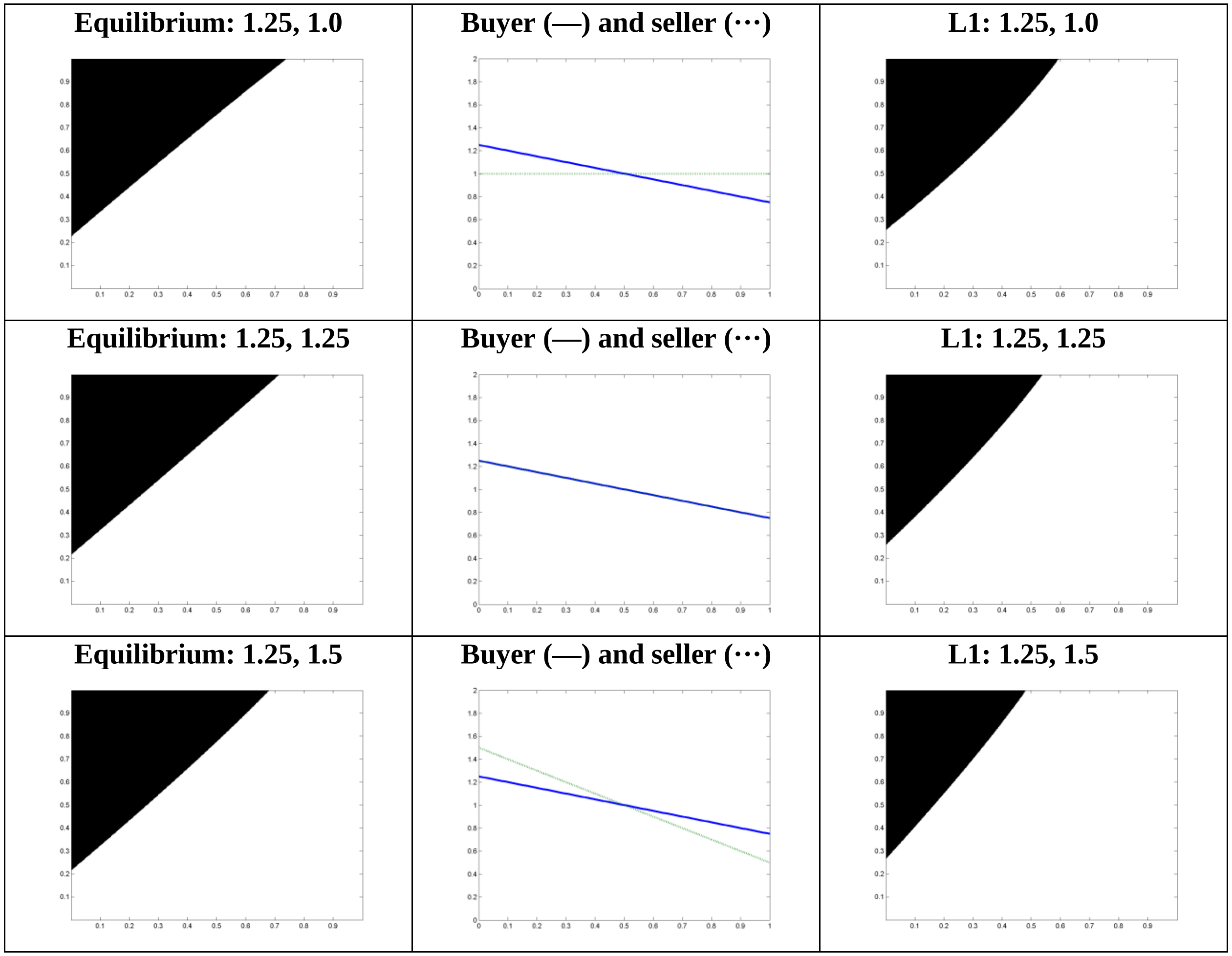


Figure 1 (cont.). Trading regions (in black) for equilibrium-incentive-efficient mechanisms and mechanisms that are efficient in the set of *L1*-incentive-compatible mechanisms

(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.)

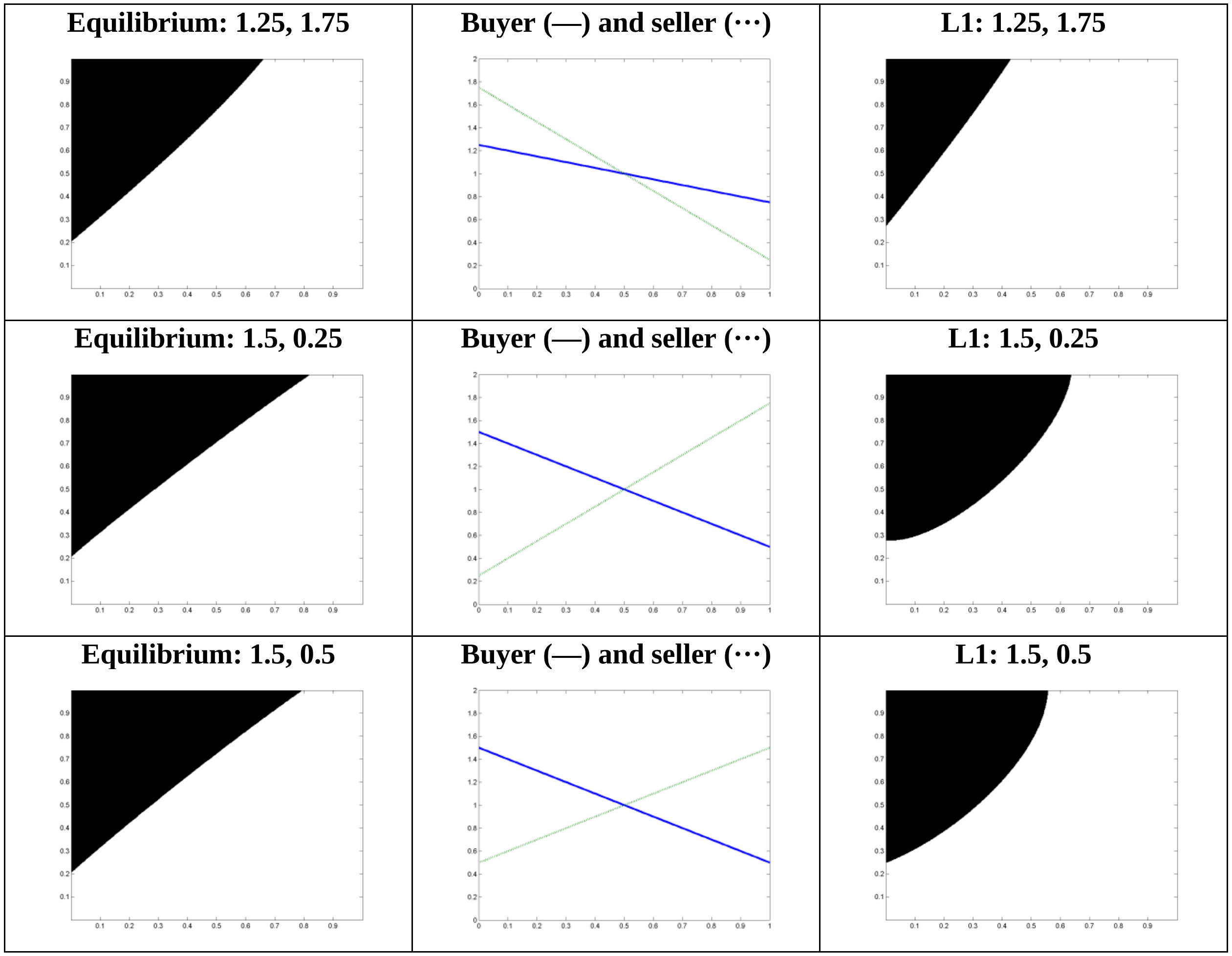
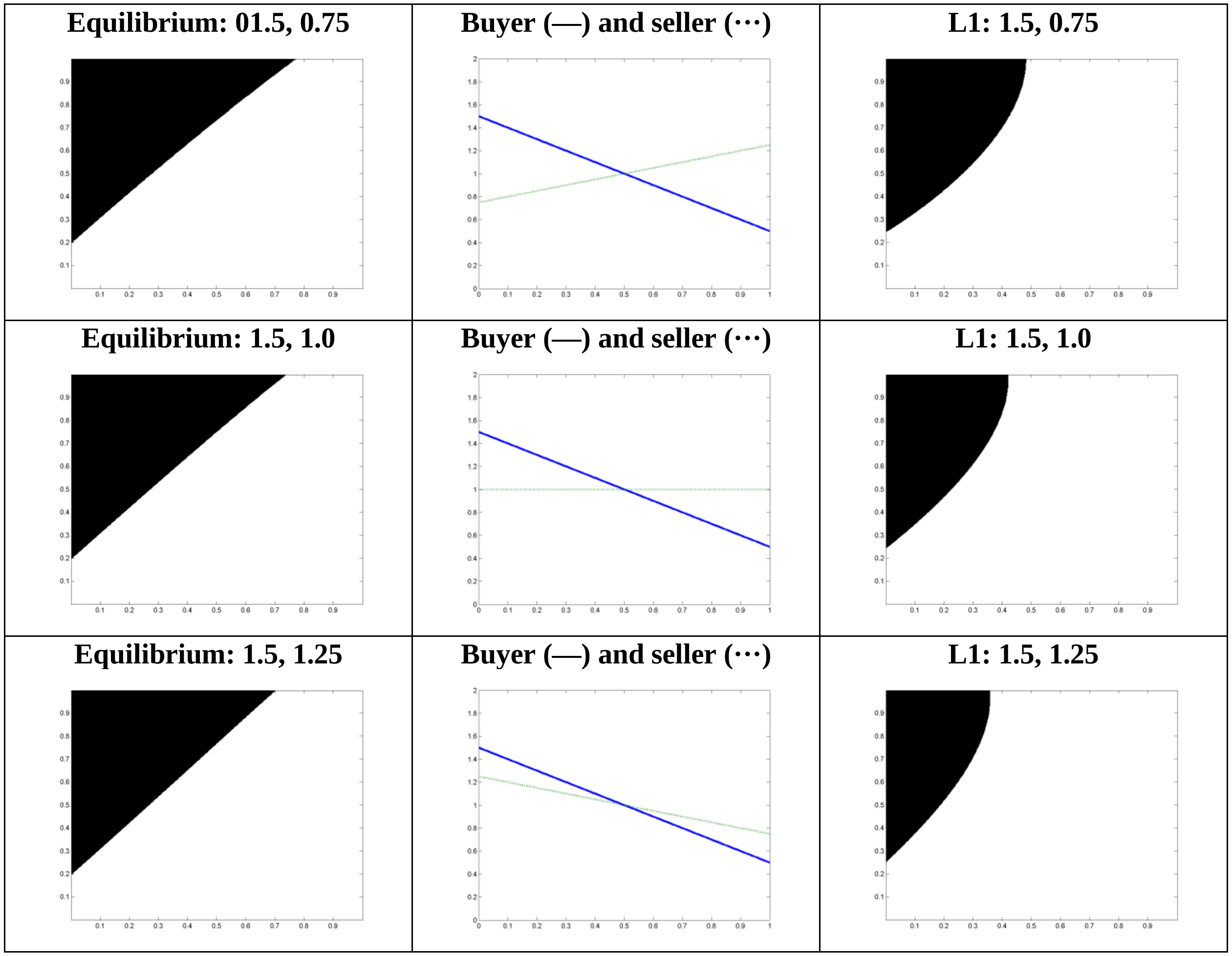


Figure 1 (cont.). Trading regions (in black) for equilibrium-incentive-efficient mechanisms and mechanisms that are efficient in the set of *L1*-incentive-compatible mechanisms

(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.)



Mechhanisms 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 the revelation principle fails for level-*k* traders, so that if *Lk*-incentive-compatibility is not required, it might be beneficial to allow direct mechanisms that are not *Lk*-incentive-compatible.

"*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 directly 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 directly 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* analysis adds to the usefulness of equilibrium-based design theory in several ways:

- A level-k model adds enough specificity to allow an analysis with power comparable to that of an equilibrium analysis, yielding characterization results that clarify the role of the equilibrium assumption in MS's analysis in several ways.

Summary

A level-*k* analysis adds to the usefulness of equilibrium-based design theory in several ways:

- A level-k model adds enough specificity to allow an analysis with power comparable to that of an equilibrium analysis, yielding characterization results that clarify the role of the equilibrium assumption in MS's analysis in several ways.

- The revelation principle fails, due to menu effects whereby the mechanism influences the correctness of level-*k* beliefs, which can favor a level-*k*-incentive-compatible mechanism over its direct *non*-level-*k*-incentive-compatible counterpart or vice versa.

- Such menu effects compel a choice about whether to require level-*k*-incentive-compatibility.

If level-*k*-incentive-compatibility is required, MS's analysis is surprisingly robust to level-*k* thinking:

- MS's result that with uniform value densities, the equilibrium-incentive-efficient direct mechanism mimics CS's linear double-auction equilibrium, is completely robust to level-*k* thinking.

If level-*k*-incentive-compatibility is required, MS's analysis is surprisingly robust to level-*k* thinking:

- MS's result that with uniform value densities, the equilibrium-incentive-efficient direct mechanism mimics CS's linear double-auction equilibrium, is completely robust to level-*k* thinking.
- MS's characterization of incentive-efficient mechanisms for general value densities is fully robust to level-*k* thinking for populations concentrated on one level.

If level-*k*-incentive-compatibility is required, MS's analysis is surprisingly robust to level-*k* thinking:

- MS's result that with uniform value densities, the equilibrium-incentive-efficient direct mechanism mimics CS's linear double-auction equilibrium, is completely robust to level-*k* thinking.
- MS's characterization of incentive-efficient mechanisms for general value densities is fully 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 analysis is surprisingly robust to level-*k* thinking:

- MS's result that with uniform value densities, the equilibrium-incentive-efficient direct mechanism mimics CS's linear double-auction equilibrium, is completely robust to level-*k* thinking.
- MS's characterization of incentive-efficient mechanisms for general value densities is fully 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 sellers' uniform beliefs are pessimistic relative to buyers' (upward-sloping) true densities.
- They then exploit TEPIB to implement trading regions that are supersets of ones for equilibrium-incentive-efficient mechanisms and thereby obtain higher expected total surplus.

- Mechanisms that are efficient among *L1*-incentive-compatible mechanisms perform better when sellers' uniform beliefs are pessimistic relative to buyers' (upward-sloping) true densities.
- They then exploit TEPIB to implement trading regions that are supersets of ones for equilibrium-incentive-efficient mechanisms and thereby obtain higher expected total surplus.
- Sometimes such mechanisms require ex post perverse trade for a small number of value combinations.

- Mechanisms that are efficient among *L1*-incentive-compatible mechanisms perform better when sellers' uniform beliefs are pessimistic relative to buyers' (upward-sloping) true densities.
- They then exploit TEPIB to implement trading regions that are supersets of ones for equilibrium-incentive-efficient mechanisms and thereby obtain higher expected total surplus.
- Sometimes such mechanisms require ex post perverse trade for a small number of 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.

- Mechanisms that are efficient among *L1*-incentive-compatible mechanisms perform better when sellers' uniform beliefs are pessimistic relative to buyers' (upward-sloping) true densities.
- They then exploit TEPIB to implement trading regions that are supersets of ones for equilibrium-incentive-efficient mechanisms and thereby obtain higher expected total surplus.
- Sometimes such mechanisms require ex post perverse trade for a small number of 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.

- 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.

The level-*k* analysis also yields several examples of nonrobustness, beyond the failure of the revelation principle and the possibility of ex post perverse trade noted above:

- 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.

The level-*k* analysis also yields several examples of nonrobustness, beyond the failure of the revelation principle and the possibility of ex post perverse trade noted above:

- 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.

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

The level-*k* analysis also yields several examples of nonrobustness, beyond the failure of the revelation principle and the possibility of ex post perverse trade noted above:

- 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.

- 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.