

Bringing Real Market Participant's Real Preferences into the Lab: An Experiment that Changed the Course Allocation Mechanism at Wharton

Eric Budish

University of Chicago Booth School of Business

Judd B. Kessler

The Wharton School, University of Pennsylvania

August 2, 2017

Applied Economics Workshop

Market Design: Theory → Practice

- Promise of Market Design: abstract economic theory can be used to solve real-world resource allocation problems
 - Auctions: wireless spectrum, online advertising
 - Matching: medical labor markets, public schools, kidney exchange
- Often, new academic work is needed to help bring market design from theory → practice
- This talk: a novel kind of laboratory experiment bringing real agents' real preferences into the lab
 - Key feature is ability to test the common “agents report their type” assumption, a major issue for the design we study (Budish 2011)

Problem: Combinatorial Assignment

- Combinatorial assignment is well known to be a hard market design problem
 - Indivisible objects to be allocated to agents
 - Agents' preferences are over bundles of the objects
 - Monetary transfers are prohibited
- Examples: course allocation, shift scheduling, allocating shared computational resources, the food bank problem
- Theory: mostly impossibility theorems (Papai 2001; Ehlers and Klaus 2003; Hatfield 2009)
 - Only mechanisms that are SP + ex-post efficient are dictatorships
 - Unfair ex-post (and inefficient ex-ante)
- Practice: mechanisms in practice have important flaws (Sonmez and Unver 2003, 2010; Budish and Cantillon 2012)

Course Allocation at Wharton

- In Fall 2011, as part of a curriculum review, Wharton convened a committee to reevaluate its mechanism for course allocation
- Its fake-money auction mechanism had been having the kinds of problems you'd expect given theoretical criticisms of the mechanism (Sonmez and Unver 2003, 2010)
 - Similar mechanism used widely at many other schools: Booth, Columbia, Haas, Kellogg, Ross, Sloan, Stern, Yale SOM...

Top Ten Reasons that Wharton Students Hate the Auction

ACADEMICS

NATALIE BOLCH (WG '12)

The sentiment behind the Auction is simple: with limited supply of fantastic classes and professors, why not set up a system where students can spend more on classes they really want? Theoretically, this process should ensure that all students get to take at least a few classes at Wharton that they personally deem desirable. Practically, however, that's just not the case. This fact, combined with conversations with many disgruntled Wharton students, has led me to compile the following list:

The Top 10 Reasons Wharton Students Hate the Auction:

1. Even with historical prices, it's hard to know what to bid

Notionally, the prices for classes are supposed to be set by students based on each course's desirability. Yeah, maybe in 1997. Today, students' bids have little to do with how much they value particular classes and, instead, are based solely on history. Bidding amounts typically form as follows: 1) Decide what classes to take; 2) Search past auctions for bidding history; 3) Be super strategic and bid the historical price plus 10%; 4) Reevaluate if you don't have enough points. The results? A system with price inflated bids based far less on a person's eagerness to take a particular course and far more on a conglomeration of past clearing prices.

2. It feeds greed - the easiest way to profit is to cheat other students out of

classes by hedging your bets

How many students do you know with absurd amounts of points who choose to bid on a plethora of classes, simply to win early and profit later? By hedging for profit, these students are, in essence, taking advantage of their position of power to penalize their fellow classmates, preventing many of them from taking classes that could benefit their future careers. Isn't there something to be said for having honor amongst thieves? Come on Wharton, even honey badgers don't attack each other!

3. The only real round is Round 1

Let's get real.

The only true round of the auction is Round 1. First years get completely screwed without even having the ability to enter

before the cheapest round of the auction is over and any second year who underbid for a great class by even 10 points, may get penalized to the point of being forbidden to take the class altogether. A big shout out to Sigglekow here. His must-take Strategy and Competitive Advantage class cleared for 7635 points in Round 5 last year, more points than I could have acquired if I waived the entire spring semester. I suppose though, that it works out in the end. Professors with so much pull apparently assign 75-page papers. With only one week left of first semester, I'm Sigglekow-free and thanking my lucky stars!



4. I could've bid all of my points and I still wouldn't be going to China

Enough said.

5. I started 2nd year with half the points of my friend, simply because I actually took the core

Am I the only person who thinks all students should start on an even playing field at the beginning of their second year? I've wracked my brain and still can't figure out why first year students get to profit off of classes they choose and

make nothing off of core classes in which they are forced to enroll. Last I checked, no one was excited about OPIM Crystal Ball simulations. I vote for a second year point reset, where students each have 5000 points to start and can divvy them up as desired. Perhaps then,

those of us without business undergrad degrees would be able to compete for fall classes.

6. I spent twice as much on a class as my friend because I bid in a different round

The point here is simple. I just don't get it. Why are prices so cheap in Rounds 1 and 8 and so expensive in the middle? It seems unless you're a perfect guesser, you could wind up paying exponentially for a class on which another student only spent 200.

7. Figuring out what to bid takes forever. I'd rather go to pub

Successfully navigating the auction is a class in itself. Is it just me or do you sometimes feel like you need a PhD to bid on classes for your MBA?

8. I can't even take the best classes in my own major

Why is it that students don't get priority for classes in their own major? Professor Nakahara's Real Estate Development course, a requirement for real estate majors, went for 4003 points in Round 1 last spring. As a result, many real estate majors have spent their entire tenure at Wharton saving up auction points, in fear that they wouldn't be able to fulfill their major due to a lack of currency. One hundred thousand dollars later and with plans to pursue a real estate career, you'd think you'd get priority over students who basically want to audit the class for fun.

9. I bid on 10 classes, got all of them, and may take 4

What's with the 1% taking over the Wharton world? I don't know about you but I've talked to at least a few students who are about ready to start an Occupy Wharton movement. This 1% has somehow managed to get so rich that they have the ability to bid on and buy more classes than they need in early rounds "just in case." Their action not only occupies desired spots, but also it drives up prices. It's about time the government intervened.

10. There must be better models out there

One of these days maybe we'll take a page out of Stanford's book, offer our professors coconuts, and call it a day. 🥥

CEEI

- Budish (2011) proposes a new mechanism for combinatorial assignment, called “approximate competitive equilibrium from equal incomes” (CEEI):
 - Students report their type: preferences over schedules (u_i 's)
 - Students are assigned approx. equal budgets (b_i 's)
 - Computer finds approx. market clearing prices (p 's)
 - Each student i receives a schedule that maximizes her utility subject to budget
- Properties (for formal statements, see Budish 2011)
 - *Efficiency*: approx. ex-post Pareto efficient
 - *Fairness*: approx. envy free
 - *Incentives*: strategy-proof in the large (Azevedo and Budish, 2013)

CEEI

- Attractive in theory...but suitable for practice?
- Key issue is preference reporting: in the model, “agents report their type,” which is clearly crazy in practice
 - Hundreds of millions of schedules in a given semester, for each agent we need a rank-order preference over all of them
 - Can agents do this “accurately enough” to reap the benefits of CEEI?
- We needed a new type of experiment to test the assumption
 - The endowed preferences methodology is not going to work here
 - We want to know if they can communicate preferences represented in their own mind into something the mechanism can use
 - Testing translation from English → Latin

Key Features of the Design

1. Realistic market participants' real preferences
 - Real market participants: Wharton MBA students
 - Report real preferences over real schedules of Wharton classes
 - Realistic preference reporting language
 - Professionally designed user interface
2. “Binary comparison” tool
 - While generating a rank order list over schedules is hard (likely inaccurate), comparing two schedules is easy (likely accurate)
 - Allows us to obtain isolated data on preference reporting language and data on overall performance

Roadmap

- Introduction
- **Experimental Design**
- Results on Efficiency and Fairness
- Preference Reporting
- Discussion

Experimental Procedure

- 8 sessions, 14 to 19 subjects per session (N=132), Nov-Dec 2011
 - Only Wharton MBA students were allowed to participate
- Overall flow of each session with parts yielding data underlined:
 - Read general instructions
 - Look over subset of Spring 2012 course offerings for 5 minutes
 - Read instructions about and use first mechanism (CEEI or BPA)
 - Answer survey questions about first schedule and mechanism
 - Read instructions about and use second mechanism (BPA or CEEI)
 - Answer survey questions about second schedule and mechanism
 - Make binary comparisons of pairs of schedules
 - Answer additional survey questions about both mechanisms
 - Write free-response comments after using both mechanisms

Bidding Points Auction (BPA)

- Implementation very similar to practice (e.g., same user interface)
- All students have same budget of 5,000 points
- Round 1: Auction
 - Students bid points for courses (5 minutes to select bids)
 - For a course with k seats, the k highest bidders get it and pay the $k+1^{\text{th}}$ highest bid
- Rounds 2-4: Double Auction
 - Both buying and selling
 - Can set an ask price for a course you currently have (2:30 per round)
 - Trade occurs at minimum market-clearing price
- In practice and in the lab, most of the action is in Round 1

CEEI

- Report preferences (10 minutes)
 - Told that only responsibility is to report preferences
 - Told that system would find market clearing prices and they would get best schedule they could afford at those prices
 - Explicitly advised to report their preferences truthfully: “You get the best schedule possible simply by telling the computer your true values for courses.”
- Each student assigned a budget of 5,000 points plus a small random amount extra (to break ties)
- The computer finds approximate market clearing prices
- Each student is allocated the bundle that maximizes her utility given reported preferences, budgets, and prices

Preference Reporting Language for CEEI

- Two ingredients: cardinal item values and pairwise adjustments
- Cardinal item values
 - Students instructed to submit a value of 100 for favorite section
 - Students instructed to submit a relative value (between 1 and 100) for any other section they have interest in taking
 - Cardinal item values over individual courses induce ordinal preferences over bundles of courses
 - If values for $\{A,B,C,D\}$ are $\{100,80,50,10\}$, and need at most 2, ordinal preferences are $\{A,B\} > \{A,C\} > \{B,C\} > \{A,D\} > \{A\} > \dots$
- Pairwise adjustments
 - Can enter positive or negative adjustments for any pair of sections
 - Simple way to express certain kinds of complements / substitutes
- “Top-ten” widget



COURSE REGISTRATION COURSE MATCHING SYSTEM

Assign Values

My Top 10 Schedules

First Survey

Second Survey

Third Survey

Fourth Survey

MY ADJUSTMENTS

Courses	Individual Value	Combined Adjustment	Combined Value	
ACCT742003, ACCT897402	64 + 27 = 91	-91	0	delete

MY VALUES

Search:

Course ▲	Title ◇	Instructor ◇	Meeting ◇	Credit ◇	Open ◇	Value ◇	Apply Adjustment
ACCT742003	PROBLEMS IN FIN REPORTIN	LAMBERT R	MW 1:30 PM-3:00 PM	1.00	5	64	<input type="checkbox"/>
ACCT897402	TAXES AND BUS STRATEGY	BLOUIN J	MW 12:00 PM-1:30 PM	1.00	4	27	<input type="checkbox"/>
FNCE726003	ADVANCED CORP FINANCE	VAN WESEPE	TR 12:00 PM-1:30 PM	1.00	5	39	<input type="checkbox"/>
FNCE728003	CORPORATE VALUATION	CICHELO M	MW 3:00 PM-4:30 PM	1.00	4	27	<input type="checkbox"/>
FNCE750001	VENT CAP & FNCE INNOVAT	WESSELS D	MW 1:30 PM-3:00 PM	1.00	4	55	<input type="checkbox"/>
FNCE750002	VENT CAP & FNCE INNOVAT	WESSELS D	MW 3:00 PM-4:30 PM	1.00	4	32	<input type="checkbox"/>
FNCE891001	Corporate Restructuring	JENKINS M	TR 1:30 PM-3:00 PM	1.00	4	45	<input type="checkbox"/>
LGST806407	NEGOTIATIONS	BRANDT A	W 3:00 PM-6:00 PM	1.00	3	34	<input type="checkbox"/>
LGST806409	NEGOTIATIONS	DIAMOND S	R 3:00 PM-6:00 PM	1.00	3	0	<input type="checkbox"/>

Screenshot: Top-10 Schedules Widget

MY TOP 10 SCHEDULES

Given the values you reported, your agent thinks these are your 10 favorite schedules. Your agent will try to buy you these schedules, in this order. Note that depending on the market clearing prices, the schedule you get may not appear on this list, but your agent will buy you the best schedule that you can afford.

Schedule Value: 200							
	Sun	Mon	Tue	Wed	Thu	Fri	Sat
8AM							
9AM							
10AM							
11AM							
12PM		TAXES AND BUS STRATEGY	ADVANCED CORP FINANCE	TAXES AND BUS STRATEGY	ADVANCED CORP FINANCE		
1PM		VENT CAP & FNCE INNOVAT	Corporate Restructuring	VENT CAP & FNCE INNOVAT	Corporate Restructuring		
2PM							
3PM				NEGOTIATION			
4PM							
5PM							
6PM							

Schedule Value: 193							
	Sun	Mon	Tue	Wed	Thu	Fri	Sat
8AM							
9AM							
10AM							
11AM							
12PM		TAXES AND BUS STRATEGY	ADVANCED CORP FINANCE	TAXES AND BUS STRATEGY	ADVANCED CORP FINANCE		
1PM		VENT CAP & FNCE INNOVAT	Corporate Restructuring	VENT CAP & FNCE INNOVAT	Corporate Restructuring		
2PM							
3PM		CORPORATE VALUATION		CORPORATE VALUATION			
4PM							
5PM							
6PM							

Binary Comparisons

- After playing both mechanisms, students asked their preference between specifically chosen pairs of schedules (“slightly prefer,” “prefer” or “strongly prefer”)
- Premise: preferences over schedules are complex, binary comparisons are simple
- Our binary comparisons were selected to let us generate tests for:
 - *Efficiency*: CEEI schedule vs. BPA schedule (asked first and asked last with schedule order reversed)
 - Allows us to construct a social welfare measure
 - *Fairness* (i.e., Envy): CEEI schedule vs. others’ CEEI schedules; BPA schedule vs. others’ BPA schedules
 - Allows us to assess differences in envy across mechanisms
 - *Preference Reporting Language*: All comparisons are tests of reporting, but added a few additional “close” comparisons.

Screenshot: Binary Comparisons



Wharton

UNIVERSITY of PENNSYLVANIA




COURSE REGISTRATION COURSE MATCHING SYSTEM

[Assign Values](#)
[My Top 10 Schedules](#)
[First Survey](#)
[Second Survey](#)
[Third Survey](#)
[Fourth Survey](#)

What is your preference between the schedules below?

	Sun	Mon	Tue	Wed	Thu	Fri	Sat
8AM							
9AM							
10AM			Corp Dev: Merg & Acquis		Corp Dev: Merg & Acquis		
11AM							
12PM		Taxes and Bus Strategy		Taxes and Bus Strategy			
1PM		Problems in Fin Reportin	Corporate Restructuring	Problems in Fin Reportin	Corporate Restructuring		
2PM							
3PM		Leg Asp Entrepren					
4PM							
5PM							
6PM							
7PM							
8PM							

	Sun	Mon	Tue	Wed	Thu	Fri	Sat
8AM							
9AM							
10AM							
11AM			CORP DEV: MERG & ACQUIS		CORP DEV: MERG & ACQUIS		
12PM		TAXES AND BUS STRATEGY	ADVANCED CORP FINANCE	TAXES AND BUS STRATEGY	ADVANCED CORP FINANCE		
1PM		PROBLEMS IN FIN REPORTIN	Corporate Restructuring	PROBLEMS IN FIN REPORTIN	Corporate Restructuring		
2PM							
3PM							
4PM							
5PM							
6PM							
7PM							
8PM							

Strongly Prefer A	Prefer A	Slightly Prefer A	Slightly Prefer B	Prefer B	Strongly Prefer B
					

Incentives

- Downside is that behavior in experiment is not incentivized
 - Not about binary comparisons in general (they *could be* incentivized)
 - We designed study such that lack of incentives works against us finding a difference between CEEI and BPA
- There are two concerns that arise due to lack of incentives
 1. Do not exert the same effort as if it were real
 - Think of “triers” and “non-triers”
 2. Subjects want to deliberately make one mechanism look good
 - Recruiting did not explicitly mention course allocation
 - Subjects representative of all Wharton MBAs on demographics as well as behavior in and attitude towards the Wharton Auction

Experimental Subjects

Representativeness of Experimental Subjects

	Subjects 132	Wharton MBAs 1660	
Demographics			p-value
First Year Student	51.7%	50.8%	0.83
Female	42.0%	47.0%	0.27
From United States	37.1%	34.3%	0.52
Finance Major	23.5%	25.7%	0.57
Total Registered Credits	17.1	17.0	0.96
Wharton Credits	11.5	11.3	0.56
No race reported	14.4%	21.1%	0.07*
GPA	Subjects directionally higher		0.14
Auction Behavior			
Points at Start of Spring Auction	6899.6	6966.4	0.79
Points in 4th Round of Spring Auction	4992.3	4960.7	0.92
	Subjects 62	Wharton MBAs 731	
Auction Beliefs (Second years only)			
Reported Auction effectiveness	4.69	4.68	0.96

Roadmap

- Introduction
- Experimental Design
- **Results on Efficiency and Fairness**
- Preference Reporting
- Discussion

Efficiency and Fairness

- Fundamentally want to test whether agents can report their preferences accurately enough to reap the benefits of CEEI
 - Use binary comparison data to test whether CEEI outperforms BPA
- Interested in how imperfect preference reporting harms mechanism performance
 - Reported preference data to assess how CEEI would have done relative to BPA without reporting mistakes
 - Gives a sense of the upper bound on CEEI's performance
- One-sided tests given predictions of Budish (2011)
 - Only care whether agents report preferences accurately enough
 - Report individual-subject level and market-session level results

Efficiency and Fairness

	Aggregation Level	
Data	<i>Individual-Subject</i>	<i>Market-Session</i>
<i>Binary Comparison</i>		
<i>Reported Preference</i>		

Efficiency

- Binary comparison directly asked students which they preferred between their CEEI and BPA schedules
 - Given importance of this question we asked each student twice
 - First and last binary comparison (presentation order switched)
 - “Prefer” CEEI or BPA if prefer it both times, otherwise “Indifferent”
- Simple measure of ex-ante social welfare
 - Identifies what a social planner would prefer
 - Important to the Wharton Committee
- Look at subject-level and session-level results
 - Individual-Subject: Count each subject as an individual observation
 - Market-Session: Classify each session based on majority rule

Efficiency

Session	Students in the Session	Prefer CEEI	Prefer BPA	Identical Outcome	Indeterminate Preference	Voting Result
1	18	8	5	1	4	CEEI
2	15	5	5	3	2	tie
3	19	10	4	2	3	CEEI
4	16	5	4	4	3	CEEI
5	18	8	8	0	2	tie
6	14	6	5	2	1	CEEI
7	18	8	6	2	2	CEEI
8	14	6	5	3	0	CEEI
All	132	56	42	17	17	6-0-2

- Individual level: 98 students demonstrated a preference; of those, 56-42 in favor of CEEI (binomial probability test $p=0.094$)
- Session level: majority prefers CEEI 6-0 with 2 ties ($p=0.016$)

Efficiency

	Aggregation Level	
Data	<i>Individual-Subject</i>	<i>Market-Session</i>
<i>Binary Comparison</i>		
<i>Reported Preference</i>		

Efficiency

	Aggregation Level	
Data	<i>Individual-Subject</i>	<i>Market-Session</i>
<i>Binary Comparison</i>	(A) 56 - Prefer CEEI 42 - Prefer BPA 17 - Identical outcome 17 - Indeterminate preference p=0.094	(B) 6 - Prefer CEEI 0 - Prefer BPA 2 - Tie p=0.016
<i>Reported Preference</i>		

Efficiency

- Can conduct the same analysis using reported preferences
 - “Prefer” CEEI or BPA if reported preferences suggest higher utility from one schedule, otherwise “Indifferent”
- Again look at subject-level and session-level results
- 114 students favored either CEEI or BPA schedule
 - Subject: 69% prefer CEEI schedule to BPA schedule
 - Session: CEEI wins 7-0 with one tie

Efficiency

	Aggregation Level	
Data	<i>Individual-Subject</i>	<i>Market-Session</i>
<i>Binary Comparison</i>	<p>(A)</p> <p>56 - Prefer CEEI 42 - Prefer BPA 17 - Identical outcome 17 - Indeterminate preference</p> <p>$p=0.094$</p>	<p>(B)</p> <p>6 - Prefer CEEI 0 - Prefer BPA 2 - Tie</p> <p>$p=0.016$</p>
<i>Reported Preference</i>	<p>(C)</p> <p>79 - Prefer CEEI 35 - Prefer BPA 17 - Identical outcome 1 - Indeterminate preference</p> <p>$p<0.001$</p>	<p>(D)</p> <p>7 - Prefer CEEI 0 - Prefer BPA 1 - Tie</p> <p>$p=0.008$</p>

Efficiency

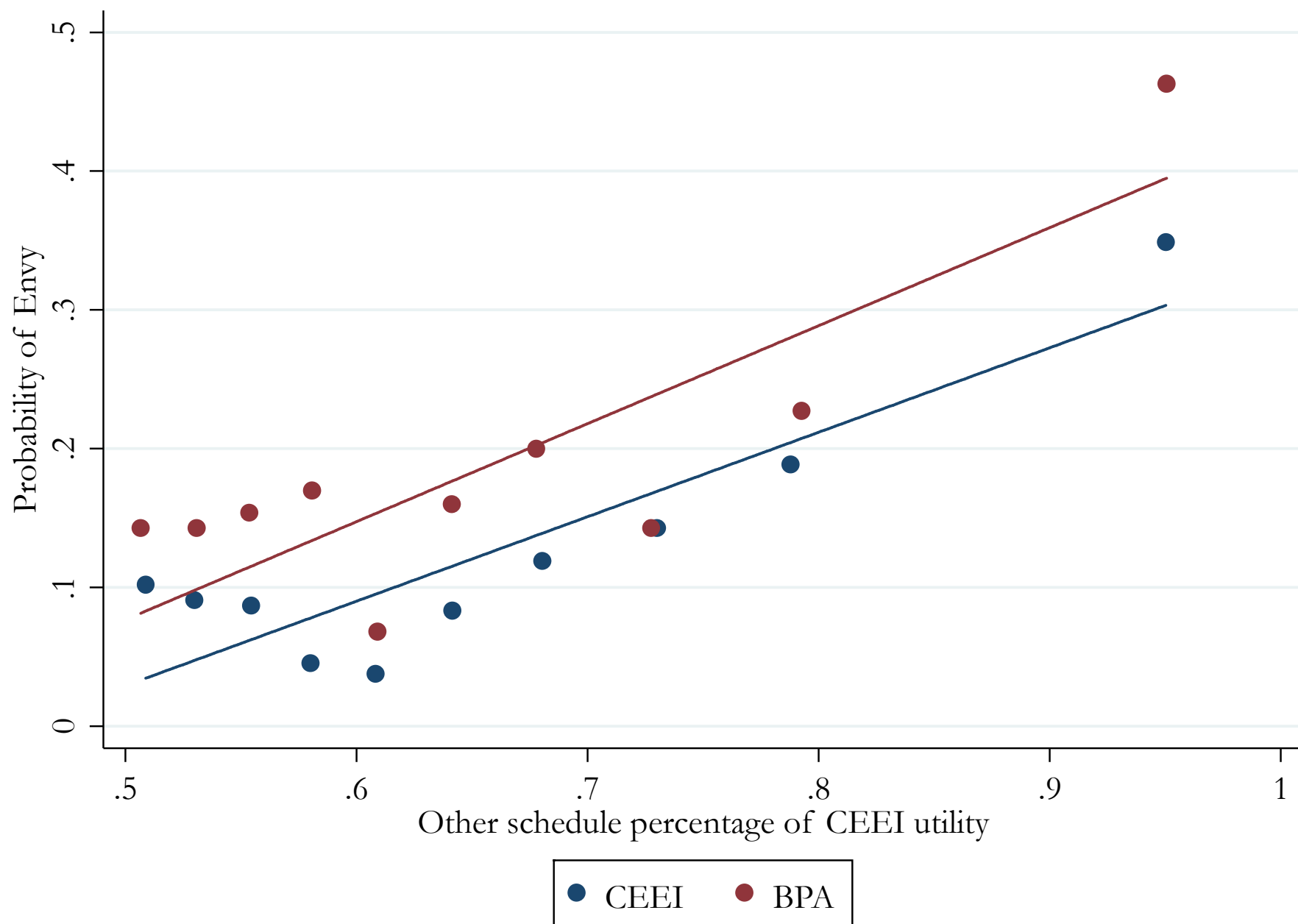
- Can compare binary comparison to reported preference results
 - 57% (56/98) prefer CEEI based on binary comparisons
 - 69% (79/114) prefer CEEI based on reported preferences
 - One-sided test of proportions yields $p=0.033$
- Suggests preference reporting mistakes measurably harmed mechanism performance
 - Large potential gains from having more training on use of the reporting language

Efficiency: Robustness (Binary Comparisons)

	Aggregation Level	
Data	<i>Individual-Subject</i>	<i>Market-Session</i>
<i>“Prefer” or “Strongly Prefer”</i>	47 - Prefer CEEI 32 - Prefer BPA 17 - Identical outcomes 36 - Indeterminate preference $p=0.057$	5 - Prefer CEEI 1 - Prefer BPA 2 - Tie $p=0.109$
<i>Average Intensity</i>	59 - Prefer CEEI 47 - Prefer BPA 17 - Identical outcomes 9 - Indeterminate preference $p=0.143$	5 - Prefer CEEI 1 - Prefer BPA 2 - Tie $p=0.109$

Fairness

- Fairness results are based on envy (Foley 1967; Varian 1974)
 - Student A *envies* student B if A prefers B's schedule to her own
- CEEI approximately eliminates envy; BPA has no such guarantee
 - Envy in CEEI because of small randomness in budgets (Budish 2011)
- We asked up to 6 binary comparisons for each mechanism
 - Ask student A whether they prefer A's schedule or B's schedule
 - Random draws from set of schedules with utility $\geq 0.5 \times \text{CEEI utility}$
 - Assume anything with utility $< 0.5 \times \text{CEEI utility}$ is not envied
 - Subjects saw average of 4 utility comparisons for each mechanism
 - Count whether subject experiences less envy in CEEI or BPA



Fairness

	Aggregation Level	
Data	<i>Individual-Subject</i>	<i>Market-Session</i>
<i>Binary Comparison</i>		
<i>Reported Preference</i>		

Fairness

	Aggregation Level	
Data	<i>Individual-Subject</i>	<i>Market-Session</i>
<i>Binary Comparison</i>	(E) 40 - Less Envy CEEI 23 - Less Envy BPA 65 - No Envy either 4 - Same Envy both $p=0.021$	(F) 5 - Less Envy CEEI 1 - Less Envy BPA 2 - Tie $p=0.109$
<i>Reported Preference</i>		

Fairness

- Again we report the same tests using reported preferences
 - We expect the results to be stark because CEEI approximately eliminates envy

Fairness

	Aggregation Level	
Data	<i>Individual-Subject</i>	<i>Market-Session</i>
<i>Binary Comparison</i>	(E) 40 - Less Envy CEEI 23 - Less Envy BPA 65 - No Envy either 4 - Same Envy both $p=0.021$	(F) 5 - Less Envy CEEI 1 - Less Envy BPA 2 - Tie $p=0.109$
<i>Reported Preference</i>		

Fairness

	Aggregation Level	
Data	<i>Individual-Subject</i>	<i>Market-Session</i>
<i>Binary Comparison</i>	<p>(E)</p> <p>40 - Less Envy CEEI 23 - Less Envy BPA 65 - No Envy either 4 - Same Envy both $p=0.021$</p>	<p>(F)</p> <p>5 - Less Envy CEEI 1 - Less Envy BPA 2 - Tie $p=0.109$</p>
<i>Reported Preference</i>	<p>(G)</p> <p>35 - Less Envy CEEI 4 - Less Envy BPA 93 - No Envy either 0 - Same Envy both $p<0.001$</p>	<p>(H)</p> <p>8 - Less Envy CEEI 0 - Less Envy BPA 0 - Tie $p=0.004$</p>

Fairness

- Can compare binary comparison to reported preference results
 - 63% (40/69) experience less envy under CEEI based on binary comparisons
 - 90% (35/39) experience less envy under CEEI based on reported preferences
 - One-sided test of proportions yields $p=0.002$
- Again suggests preference reporting mistakes measurably harmed mechanism performance
 - Large potential gains from having more training on use of the reporting language

Fairness: Robustness (Binary Comparisons)

	Aggregation Level	
Data	<i>Individual-Subject</i>	<i>Market-Session</i>
<i>“Prefer” or “Strongly Prefer”</i>	(E) 36 - Less Envy CEEI 14 - Less Envy BPA 80 - No Envy either 2 - Same Envy both $p=0.021$	(F) 6 - Less Envy CEEI 0 - Less Envy BPA 2 – Tie $p=0.016$
	(G) 31 - Less Envy CEEI 17 - Less Envy BPA 65 - No Envy either 19- Same Envy both $p=0.030$	(H) 5 - Less Envy CEEI 2 - Less Envy BPA 1 – Tie $p=0.227$

Efficiency and Fairness Discussion

- Subjects are able to report preferences accurately enough for CEEI to outperform BPA with respect to efficiency and fairness
 - Most comparisons significant at conventional levels
- But magnitudes are modest overall on binary comparisons, large difference when compared to reported preference results
 - Prefer CEEI: goes from 57% to 69%
 - Less Envy CEEI: goes from 63% to 90%
- Suggests preference-reporting mistakes played an important role

Roadmap

- Introduction
- Experimental Design
- Results on Efficiency and Fairness
- **Preference Reporting**
- Discussion

Preference Reporting

- Every binary comparison tests preference-reporting language
 - Reported preferences generates a prediction about what subject will choose in the binary comparison
 - Say comparison is “Accurate” if binary comparison choice is consistent with reported preferences; otherwise a “Contradiction”
- 84.4% accurate, 15.6% contradictions (N=1,662 comparisons)
 - But 76.4% of students have at least one contradiction
- Conceptually, two reasons preference reports \neq true preferences:
 1. Limitations of the preference-reporting language itself
 2. Difficulty using the preference-reporting language for things that the language allows them to report

Limitations of Preference-Reporting Language

- Maybe there are some kinds of package preferences that students care about and cannot report with the reporting language
- Set of non-expressible preferences is vast. Wharton committee suggested we look at scheduling considerations.
- We look at two types of “elegant schedules”

Limitations of Preference-Reporting Language

- Maybe there are some kinds of package preferences that students care about and cannot report with the reporting language
- Set of non-expressible preferences is vast. Wharton committee suggested we look at scheduling considerations.
- We look at two types of “elegant schedules”
 - **Balanced schedule** (class on Mon-Thurs)

Type of Comparison	Comparisons of this Type	Accurate	Contradictions		
			Weak Preference	Preference	Strong Preference
Neither has it	23	73.91%	13.04%	13.04%	0%
Only higher rated has it	66	86.40%	3.03%	6.06%	4.55%
Only lower rated has it	51	70.59%	7.84%	15.69%	5.88%
Both have it	1,522	84.95%	5.06%	6.57%	3.42%

(p=0.036)

Limitations of Preference-Reporting Language

- Maybe there are some kinds of package preferences that students care about and cannot report with the reporting language
- Set of non-expressible preferences is vast. Wharton committee suggested we look at scheduling considerations.
- We look at two types of “elegant schedules”
 - **Contiguous schedule** (not more than one 1.5-hour break)

Type of Comparison	Comparisons of this Type	Accurate	Contradictions		
			Weak Preference	Preference	Strong Preference
Neither has it	52	88.46%	9.63%	0%	1.92%
Only higher rated has it	199	87.44%	4.52%	4.02%	4.02%
Only lower rated has it	192	81.77%	7.29%	7.81%	3.12%
Both have it	1,219	84.17%	4.76%	7.55%	3.53%

(p=0.120)

Difficulty with Preference-Reporting Language

- Do students have difficulty reporting cardinal preference info?
- We drop comparisons with adjustments and look at comparisons where there is “rank dominance,” which we call *ordinal comparisons*
 - Ex: {1,3,5,7,9} rank dominates {2,4,6,8,10}
- A comparison in which neither schedule rank dominates is called a *cardinal comparison*.

Type of Comparison	Comparisons of this Type	Accurate	Contradictions		
			Weak Preference	Preference	Strong Preference
All	1,580	84.18%	5.32%	6.96%	3.54%
Ordinal	1,207	89.06%	4.06%	4.39%	2.49%
Cardinal	373	68.36%	9.38%	15.28%	6.97%

($p < 0.01$)

(Result holds if we control non-parametrically for utility difference: cardinal comparisons are 16.1 percentage points more likely to be a contradiction versus 11% baseline.)

Lessons for Implementation

- Preference reporting is difficult for students
 - Training should clearly focus on how to better report cardinal utilities (Wharton has done this)
 - Training may also focus on how to use adjustments — those who used the tool seemed to have used them well, but many students did not use the tool (see paper)
- Preference-reporting language has limitations in that it does not allow students to easily report all their preferences
 - Enhance reporting language to report common forms of non-additive preferences

Roadmap

- Introduction
- Experimental Design
- Results on Efficiency and Fairness
- Preference Reporting
- **Discussion**

Summary

- CEEI outperforms BPA on all measures of efficiency and fairness.
 - Most comparisons significant at conventional levels
- Subjects had significant difficulty with preference reporting
 - Meaningfully harmed performance, though not enough to undermine overall case for CEEI
 - Positive spin: improving preference reporting accuracy would dramatically improve performance
- Experiment persuaded Wharton administration to adopt CEEI
- Experiment also guided implementation in modest ways (training on preference reporting, UI)

COURSE AUCTION



Allocation
Errors



Inequitable



Unpredictable



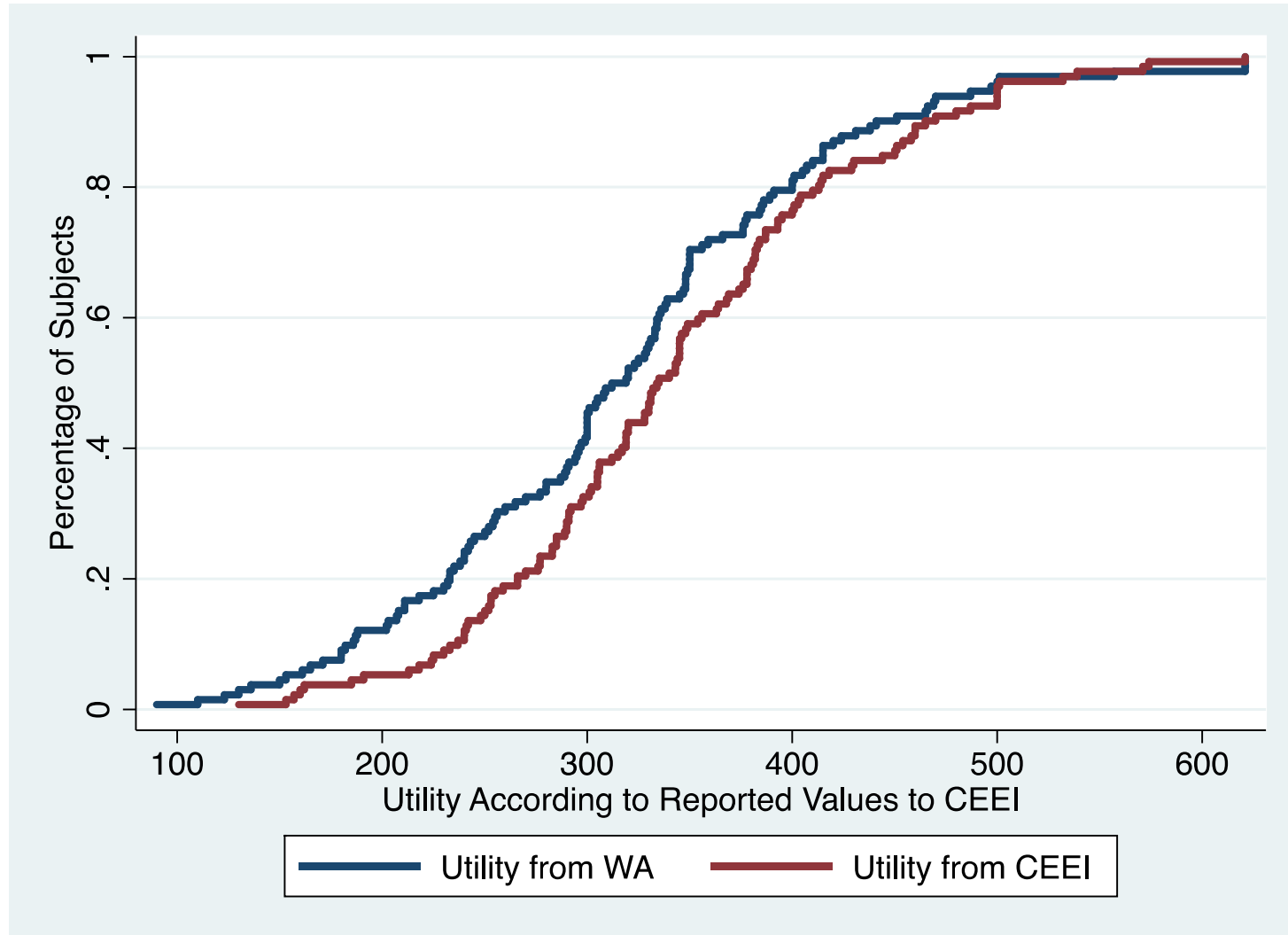
Simple, Optional, Innovative

Welcome to COURSE MATCH

Program evaluation data is starting to come in...

- There is evidence that CEEI is more fair than the BPA
 - Look at the 20 most popular courses in each year given prices
 - Under the BPA: 31% of students got 0; 6% of students got 3+
 - Under CEEI: 13% of students got 0; and 0% of students got 3+
- Annual student survey: % of students responding “Agree / Strongly Agree” to...
 - Fairness: “{The course auction, Course match} allows for a fair allocation of classes”: 28% in 2013 to 65% in 2014
 - Allocative Efficiency: “I was satisfied with my schedule from {the course auction system, course match}”: 45% to 64%
 - Overall Satisfaction: “Please rate the effectiveness of the {course auction, course match} system”: 24% to 53%

Efficiency with Reported Values



- Close to FOSD, but only SOSD since some people “win” auction

Efficiency with Reported Values

