MATH 342: The Mathematics of Social Choice

Lecture Notes, Fall 2017

Kevin Woods, Oberlin College

Preface

These are lecture notes for an upper-level undergraduate mathematics course at Oberlin College, called MATH 342: The Mathematics of Social Choice. I created this course and have taught it four times, most recently in Fall 2017. Because there is no good textbook for the material, I have created these notes, which I provide to the students each day. It seems fairly unusual to teach this at the upper-level (the politics topics tend to be covered in "math for the liberal arts" type courses, and the economics topics are more likely to be taught in a graduate microeconomics course). But the students enjoy the mix of math, economics, politics, and computer science.

The main prerequisite is the mathematical maturity to tackle an upper-level math class. Lectures 20–29 also require the basics of integration at the level of a Calculus I course. Each lecture is designed to be worked through in a 50 minute class.

The subject matter is a mix of politics and economics. On the politics side:

- Lectures 2–5 cover voting methods, culminating in Arrow's Impossibility Theorem and the Gibbard–Satterthwaite Theorem (on strategizing in elections);
- Lectures 30–32 cover Apportionment (like in Congressional seats); and
- Lectures 33–36 cover various solutions to the fair division problem (in particular, examining what "fairness" even means).

These units are all fairly independent of each other and the rest of the text. On the economics side, we mostly cover auctions. In particular:

- Lectures 6–11 cover the theory of strategy-proof auctions, such as the second price auction and Vickrey–Clarke–Groves mechanism;
- Lectures 12–16 cover the special case of auctions in a matching market (there are multiple objects for sale, but each buyer only wants one of them);
- Lectures 17–19 cover matching markets, if money is not allowed to change hands;
- Lectures 20–23 cover background in utility theory, game theory, and continuous probability so that we can talk about strategizing (in particular, Nash Equilibria and Bayesian Nash Equilibria); and
- Lectures 24–29 cover strategizing in various types of auctions, culminating in the Revenue Equivalence Theorem.

After finishing Lectures 6–11, matching markets (Lectures 12–19) and strategizing (Lectures 20–29) are fairly independent topics.

In handouts given to the students before class, the notes would have had blank boxes that were meant to be filled in interactively. The course website has both versions, depending on whether you want to try to fill in the blanks yourself! The website also has the problem sets and syllabus used in Fall 2017.

If you're interested in using this material to teach a class, if you're interested in using it to learn something about the subject, or you simply want to make me feel good, then please do contact me. Enjoy!

Contents

1 Overview

- Voting example: Bush v Gore v Nader in Florida, 2000 election.
  - Bush: 48%, Gore: 47%, Nader: 5% (ish).
  - In plurality vote (everyone votes for 1 candidate), Bush wins.
  - Who "should" win? What does "should" mean?
  - Problem: If Nader doesn't run, Gore probably wins.
    - Can we create a different *mechanism* (a different way of voting), so that Gore wins in the three way race?
  - Problem: In plurality vote, Nader voters should "lie" and vote for Gore.
    - Can we fix this problem?
  - While both problems are fixable in this example, we will see that they can't be fixed in general.
- Auction example. I have a really good cookie.
  - Alice values it at $100.
    - Alice would pay $95 for it, and feel like she "gained" $5.
    - Alice would not pay $105 for it.
    - If offered the choice between the cookie or $100, Alice would be indifferent.
  - Bob values it at $70. Chelsea values it at $60. (Though I probably don't know any of these numbers.)
  - Who "should" get the cookie? How much "should" they pay for it?
- 1st price, sealed bid auction.
  - Each makes a bid (without seeing the other bids). Highest bid wins the cookie, and they pay their bid.
  - Suppose they each bid what they think it's worth. What happens? What's wrong?

- 2nd price, sealed bid auction.
  - Highest bid wins the cookie, and they pay the second highest bid.
  - What happens if everyone bids what they think it's worth?
  - Sounds weird, but this is the same outcome as the typical, ascending price "English" auction.
    - Bob will stop bidding once it hits $70, and Alice will win it.
  - Can anyone do better by strategizing?
- 2nd price auctions are *strategy-proof* (loosely, always a good idea to tell the truth).
  - The Holy Grail.
  - We will show works for auctions, but not for voting.
- Key ideas:
  - Mechanisms: algorithms for doing these things.
  - Game theory: how to act based on how you believe others will act.
  - Social Choice: who "should" win.
- Key topics:
  - Voting, and the impossibility of perfect mechanisms.
  - Auctions, and other places where money can be transferred, and the possibility of strategy-proof mechanisms (like 2nd price).
  - Matching markets (multiple things available, but nobody gets more than one) from many perspectives (with and without money).
  - Auctions like 1st price, where you must act based on what you think others will do, probabilistically.
  - Bargaining and distribution of wealth/utility. Who "should" get it?
  - Apportionment of congressional seats to the states, by population.
- I hope I've made this sound fun. But you need to realize that this is a 300 level math class.

- ◦ We will be proving things.
- ◦ Sometimes things will get pretty abstract.

2 Introduction to Voting

- Setup of a voting system:
  - Each voter ranks all candidates.
  - Combine these to determine either
    - Society's top choice or
    - Society's full ranking.
- Example: Plurality.
  - Each voter gives one vote to their top choice (lower rankings ignored).
  - Candidates ranked by number of votes.
- Example: Borda Count.
  - If k candidates,
    - Each voter gives k points to top choice, $k-1$ to second, ..., 1 point to last choice.
    - Candidates ranked by total number of points.
  - Example: 5 voters rank $y > z > x$ and 6 voters rank $x > y > z$:
    - Does the right candidate win?
- Some notation:
  - Voter i has preference relation $>_i$, completely ranking candidates.
  - $x >_i y$ means Voter i prefers x to y.
  - Satisfies
    - Completeness: for all $x \neq y$, either $x >_i y$ or $y >_i x$, but not both.
    - Transitivity: if $x >_i y$ and $y >_i z$, then $x >_i z$.
  - Social Choice: a function, $C(>_1, \ldots, >_n)$, yielding the elected candidate.
  - Social Ranking: a function, $R(>_1, \ldots, >_n)$, yielding a full ranking, $>$.
- What properties do we want a Social Ranking (or Choice) to have?
  - Pareto property: If $\forall i\ x >_i y$, then $x > y$.
    - In words:

- Does Plurality have this?

- Does Borda?

◦ Independence of Irrelevant Alternatives (IIA): If $\{>_i\}$ and $\{>'_i\}$ are sets of rankings (same voters and candidates, different votes) with

$$\{i: \ x >_i y\} = \{i: \ x >'_i y\},$$

then $x > y$ iff $x >' y$ (in social rankings).

- In words:

- Example: if x is winning, and everyone moves z to the bottom of their rankings, then y can't suddenly win.

- Does Borda have this?

- Does Plurality?

- "I'll have the chicken. Oh, you're out of the fish? Then I'll have the steak."

• Arrow's Theorem: Suppose you have a Social Ranking function such that:

0) It only uses information from voters' rank orderings (technically, this a part of the definition of Social Ranking Function),
1) There are at least 3 candidates,
2) Any preference ranking is possible as a vote,
3) Pareto Property is satisfied,
4) IIA is satisfied.

Then it is a dictatorship: for some i, the social ranking is always $>_i$.

◦ Proof: next class.

• An example where you can start seeing the problem.

◦ 3 voters, $x >_1 y >_1 z$, $z >_2 x >_2 y$, and $y >_3 z >_3 x$.

- If IIA true, what "ought" to happen?

- You may decide you don't like IIA anymore.
  - We'll see some weaker properties.
  - But IIA is closely tied to manipulability.

3 Arrow's Theorem

- Pareto: If everyone prefers x to y, then society must.
- IIA: Moving a third candidate, z, up or down with individual voters can't change whether $x > y$ or $y < x$ in social ranking. The only way to change the relative ranking of x and y is to change their relative ranking in some of the votes.
- Arrow's Theorem: If a social ranking function is such that:
  0) it only uses from voters' rank orderings,
  1) ≥ 3 candidates,
  2) All rankings possible as votes,
  3) Pareto
  4) IIA
  Then it is a dictatorship.
- Proof:
  - Take any three distinct candidates x, y, z.
  - Begin with something x clearly wins:

Voter	1	2	$\cdots$	n	Social
	x	x	$\cdots$	x	x
	y	y	$\cdots$	y	y

(1)

  - Now keep making the votes worse for x, while keeping x the winner, see Diagram (9).
  - For Voter 1, move y to 2nd place. x still wins.

1	2	$\cdots$	n	Social
x	x	$\cdots$	x	x
y				
	y	$\cdots$	y	

(2)

  - For Voter 1, move y to 1st. Either x or y wins.

1	2	$\cdots$	n	Social
y	x	$\cdots$	x	x or y
x				
	y	$\cdots$	y	

(3)

  - If x still wins, do the same for Voter 2 (move y to 2nd, then 1st), then Voter 3, etc.

- At some point, y must become the winner.
- Let Voter k be where that switch happens (Diagrams (4) and (5)).

1	$\cdots$	$(k-1)$	k	$(k+1)$	$\cdots$	n	Social
y	$\cdots$	y	x	x	$\cdots$	x	x
x	$\cdots$	x	y				
				y	$\cdots$	y	

(4)

1	$\cdots$	$(k-1)$	k	$(k+1)$	$\cdots$	n	Social
y	$\cdots$	y	y	x	$\cdots$	x	y
x	$\cdots$	x	x				x
				y	$\cdots$	y	

(5)

- x is still 2nd in (5).

- Move x to the bottom for Voters 1 to $k-1$, 2nd from bottom for $k+1$ to n. y still wins.

1	$\cdots$	$(k-1)$	k	$(k+1)$	$\cdots$	n	Social
y	$\cdots$	y	y				y
			x				
				x	$\cdots$	x	
x	$\cdots$	x		y	$\cdots$	y	

(6)

- Switch x and y for Voter k. y is 1st or y is 2nd to x.

1	$\cdots$	$(k-1)$	k	$(k+1)$	$\cdots$	n	Social
y	$\cdots$	y	x				$y?$
			y				$y?$
				x	$\cdots$	x	
x	$\cdots$	x		y	$\cdots$	y	

(7)

- Moving from (4) straight to (7), what happens? What can we say about (7)?
- Move y 2nd to last in 1 to $k-1$.
- Now look at z. Suppose 3rd to last in 1 to $k-1$, third to last in $k+1$ to n.
- Move z to 2nd in k.

- x still winning.

1	$\cdots$	$(k-1)$	k	$(k+1)$	$\cdots$	n	Social
			x				x
			z				
			y				
z	$\cdots$	z		z	$\cdots$	z	
y	$\cdots$	y		x	$\cdots$	x	
x	$\cdots$	x		y	$\cdots$	y	

(8)

- Move x to last in $k+1$ to n.

1	$\cdots$	$(k-1)$	k	$(k+1)$	$\cdots$	n	Social
			x				$?$
			z				
			y				
z	$\cdots$	z		z	$\cdots$	z	
y	$\cdots$	y		y	$\cdots$	y	
x	$\cdots$	x		x	$\cdots$	x	

(9)

- What can you tell me?

- Status: (9) looks unbelievably bad for x, and yet Voter k propels x to 1st. Looking like a dictator.

- Claim: Given any set of votes and any candidate $w \neq x$, if $x >_k w$, then $x > w$ in the social ranking.
  - That is, "k is a dictator for x."
  - Proof: Case 1: $w \neq z$.
    - Suppose we have a set of votes such that $x >_k w$, and assume $w > x$ in the social ranking (seeking a contradiction)
    - Move z to the top for all voters except k. For voter k, move z just below x.

1	$\cdots$	$(k-1)$	k	$(k+1)$	$\cdots$	n	Social
z	$\cdots$	z		z	$\cdots$	z	$?$
			x				
			z				
			w				

(10)

    - What do we know?

- Contradiction!

- Case 2: $w = z$. Do the same thing, but use y instead of z.

- We can finally show that Voter k is a dictator!
  - The above proof shows that k is a dictator for x.
  - But the same proof shows that *someone* is a dictator for any other w.
  - Suppose ℓ is a dictator for some w, with $\ell \neq k$.
    - Create a set of votes with $x >_k w$ and $w >_\ell x$.
    - Then $x > w$ (k a dictator for x), but $w > x$ (ℓ a dictator for w).
    - Contradiction!

4 Discussion of Arrow's Theorem

- Arrow's Theorem: If a social ranking function is such that:
  0) it only uses info from voters' rank orderings,
  1) ≥ 3 candidates,
  2) All rankings possible as votes,
  3) Pareto
  4) IIA
  Then it is a dictatorship.
- To fix this problem, we could change one of the conditions.
- 0) What if we used more / different information from the voter rankings?
  - Approval voting: you can give 1 vote to however many candidates you "approve" of.
    - IIA?
    - Thoughts?
  - Utilitarian: Rate each candidate (e.g., on a scale of 0 to 10), and sum each candidate's ratings.
    - IIA?
    - Thoughts?
- 1) Allow only 2 candidates. A 2 party system!
- 2) In particular settings, might be something underlying the preferences so that some are never expressed.
  - 1-dimensional preferences.
    - Maybe all of politics is a Conservative–Liberal spectrum, and you prefer the candidate nearest you in the spectrum.
    - Can't vote Bush $>$ Nader $>$ Gore.
    - We will see that nice mechanisms exist here.
  - Introduce money.
    - If the possible outcomes of the auction are that I get the object or not, and I pay some amount of money, what can you say about my preferences?

- 2nd price auction (or standard English ascending auction) work well.
- We'll start studying this in general, soon.

- 3) Remove the Pareto restriction.
  - Are you crazy?
  - Get even weirder possibilities, if you try to keep IIA but lose Pareto:
    Elect the candidate that a "dictator" likes least.
- 4) Weaken IIA.
  - Monotonicity:
    - If Candidate x moves up in some votes, but nothing else changes, x can't go down in the Social ranking.
    - Many systems have this (but not all!).
  - Condorcet:
    - If x would beat all other candidates in a head-to-head race (in which case x is called the Condorcet winner), then x should win.
    - Many systems don't have this.
      - Does plurality?
    - Not always a Condorcet winner, so what then?
      - 3 voters: $x > y > z$, $y > z > x$, and $z > x > y$.
  - Intensity IIA:
    - Relative rank of x and y only depends on how far apart x and y are in each individual ranking (i.e., how many candidates ranked in between them for each voter).
    - Borda Count.
  - Strategy-proof:
    - No voter has the incentive to lie about their true preferences.
    - Alas, we'll see next time that this is basically equivalent to IIA, so no good systems will have this property.
- 5) Learn to love a dictatorship.
- Process:
  - Try examples, think about moral intuitions.

- Systematize what the system will be and what properties you want the system to have.
- Prove theorems stating what systems satisfy the properties.
- Adjust properties to expand or contract the set of options.

5 Strategizing and Voting

- Arrow's theorem is about ranking candidates. What about simply choosing a winner?
- Notation:
  - $>_i$ is ith vote (complete ranking)
  - $f(>_1, >_2, \ldots, >_n)$ is choice function, picking one candidate to elect.
  - $>_{-i} = (>_1, >_2, \ldots, >_{i-1}, >_{i+1}, \ldots, >_n)$.
    - Everyone but i's votes.
    - So $f(>_{-i}, >_i)$ is candidate elected with everyone's votes.
- Pareto: if x is the top choice of all voters, then $f(>_1, >_2, \ldots, >_n) = x$.
- IIA: If $f(>_{-i}, >_i) = x$ but $f(>_{-i}, >'_i) = y$ with $x \neq y$, then
$$x >_i y \quad \text{and} \quad y >'_i x.$$
  - In words:
- Arrow–Muller–Satterwaite Theorem: If at least 3 candidates and all possible votes allowed, then any Pareto and IIA social choice function is a dictatorship.
  - Proof: almost identical.
- Same possible fixes as before.
  - For example, IIA seems strong, so might look look for different properties. (Monotone, Condorcet, etc.)
  - Maybe some property about strategizing?
- Strategy-proof: For all $>_{-i}, >_i, >'_i$,
$$f(>_{-i}, >_i) \geq_i f(>_{-i}, >'_i).$$
  - $\geq_i$ includes possibility that they are the same outcome.
  - In words?

- Can we hope to have nice voting systems that are strategy-proof?
  - Not really, because it turns out the strategy-proof is the same as IIA!
  - Pretty much every election system is manipulable.
- Gibbard–Satterthwaite Theorem: A social choice function f is IIA if and only if f is strategy-proof.
  - Proof, part 1. not IIA $\Rightarrow$ not strategy-proof:
    - Suppose it is not the case that:
      For all $>_{-i}, >_i, >_i'$,
      If $f(>_{-i}, >_i) = x$, $f(>_{-i}, >_i') = y$, $x \neq y$,
      then $x >_i y$ and $y >_i' x$.
  - Part 2, not strategy-proof $\Rightarrow$ not IIA:
    - Exists $>_{-i}, >_i, >_i'$ such that
      $$f(>_{-i}, >_i) <_i f(>_{-i}, >_i').$$
    - Let $f(>_{-i}, >_i) = x$ and $f(>_{-i}, >_i') = y$.
    - So,
  - Done!

6 Strategy-proof Auctions

- We've seen that no general voting system can be strategy-proof.
  - Fix: Let people pay money. We'll use the language of auctions.
  - This removes the "all rankings possible" hypothesis of Arrow's Theorem. Everyone prefers paying $5 to paying $10.
- We want auction mechanisms that are strategy-proof:
  - Never any regrets from having bid true valuation, no matter what everyone else bids.
  - So no need to "strategize", just tell the truth!
- Example: 2nd price auction.
- Non-example: 1st price auction.
  - If 2nd highest bid is $50 and I bid true valuation $70, in retrospect would have preferred to bid $51.
- Example: Suppose 2 cookies for sale, and everybody wants one (and only one).
  - Dexter values at $100, Elara at $70, Flora at $40.
  - What if top two bids get it, and both pay 2nd highest bid?
    - Can anyone strategize? (For concreteness, let's assume the other two tell the truth.)
  - What if top two bids get it, 1st bid pays 2nd bid, 2nd bid pays 3rd bid?
  - So what should we do?
- Example: 2 cookies for sale. Gillian wants to eat both of them!
  - She'd pay $100 for the first, and $40 for the second (decreasing marginal utility).
  - Hugh would only eat one, and values it at $70.
  - Assume Gillian will submit two bids $b_1 > b_2$ for her 1st and 2nd cookie. Hugh will submit one bid, b_3 (and a second bid of $0, let's say).
  - What if top two bids get it, both pay 3rd highest bid?

  - Hard to imagine a strategy-proof mechanism where Gillian doesn't get a free cookie. Hugh can't stop her from getting a cookie, so she could just pretend to not be that interested.
  - A possible mechanism:
    - Give Gillian one cookie free, then do a second price auction on the other one.
      - Gillian will bid b_2 and Hugh b_3.
- We will show that any auction setup has an essentially unique strategy-proof mechanism.
- Intuition:
  - Give the cookies to the people who want them most.
  - Make Agent i pay how much extra utility the other players would have gotten if i had not entered auction.
  - Example: 1 cookie.
    - Winner deprives 2nd place bidder the cookie, so pays 2nd highest bid.
  - Example: 2 cookies, with Dexter (\$100), Elara (\$70), and Flora (\$40).
    - Dexter and Elara deprive Flora of a cookie worth \$40 to her.
  - Example: 2 cookies, with Gillian (\$100 and \$40) and Hugh (\$70).
    - Hugh deprives Gillian of a cookie worth \$40 for her.
    - Gillian deprives Hugh of nothing.
  - Example: 2 cookies, with Gillian (\$100 and \$70) and Hugh (\$40).
    - Gillian deprives Hugh of a cookie worth \$40 to him.
- Intuition for why this is strategy-proof:
  - Your exact bid doesn't matter for what you pay, other than indirectly by changing outcome.
    - So your precise bid is not important.
  - If your bid changes the outcome, you did it because
    how much it helps you $-$ how much you pay > 0. And since
    how much you pay $=$ how much it hurts others,
    how much it helps you $-$ how much it hurts others > 0.
  - So your bid is an honest signal that you "deserve" it.
  - Internalizing the externality.
- Not technically the unique strategy-proof mechanism:

- Alternative mechanism: Give all of the cookies to Isabelle for free.
- Alternative mechanism: Maximum bid, b_{max}, gets cookie for free. All others get paid b_{max}.
- There is a unique strategy-proof mechanism such that:
  - it is efficient (people who want the cookies the most get them) and
  - People who don't get anything don't pay/receive money.

7 VCG mechanism

- Definitions and assumptions:
  - Agent i has a type, θ_i (probably secret).
    - In 1 cookie auction, type is their valuation of the cookie.
  - There is a set D of possible decisions (outcomes).
    - In auction, decisions are to give the cookie to somebody (or possibly nobody).
  - For any decision $d \in D$, Agent i has a utility, $v_i(d, \theta_i)$, which depends on their type.
    - Assumption: the function v_i is known to all (but value of θ_i probably secret).
    - In auction:
  - A decision function maps a vector of types to a decision.
    - Interpretation: everyone announces type, and a decision, $d(\Theta)$, is made ($\Theta = (\theta_1, \ldots, \theta_n)$).
    - In auction, announcing type is placing a bid.
      - Presumably, $d(\Theta)$ is giving cookie to highest bidder.
      - Might sometimes be reserve prices; no one gets it if the bids aren't high enough.
  - Definition: the decision function, $d(\cdot)$ is *efficient* if, for any given Θ, $d(\Theta)$ maximizes
    $$\sum_i v_i(d', \theta_i)$$
    across all $d' \in D$.
    - In auction:
  - Agent i is involved in a money transfer, $t_i(\Theta)$.
    - $t_i > 0$ means receives, $t_i < 0$ means pays.
    - In 2nd price auction:
  - A mechanism is a decision function $d(\cdot)$ together with transfer rules $t_i(\cdot)$.

- Assumption: Agent i's total utility is
$$v_i\big(d(\Theta), \theta_i\big) + t_i(\Theta).$$
  - Assumption is called *quasi-linearity* (Constant "marginal utility of income").
- A mechanism is *strategy-proof* if no agent ever has an incentive to lie:
  - For all i, for all $\Theta_{-i} = (\theta_1, \ldots, \theta_{i-1}, \theta_{i+1}, \ldots, \theta_n)$,
    for all θ_i (i's true type),
    θ_i maximizes
$$v_i\big(d(\Theta_{-i}, \theta_i'), \theta_i\big) + t_i(\Theta_{-i}, \theta_i')$$
    across all θ_i'.
  - θ_i' is a possible lie, and $v_i\big(d(\Theta_{-i}, \theta_i'), \theta_i\big)$ is true feelings about outcome brought about by the lie.
  - No regrets. Even knowing all other agent's bids, no incentive to lie. Irrelevant to you whether Θ_{-i} is truth or lie.

- For a given agent i,
  - $$\sum_{j \neq i} v_j(d(\Theta), \theta_j)$$
    measures how good the outcome is for everyone else, and
  - $$\max_{d'} \sum_{j \neq i} v_j(d', \theta_j)$$
    measures the best that *any* outcome could have been for everyone else.
  - Definition: The *Clarke Pivot Mechanism* has transfers
$$t_i(\Theta) = \sum_{j \neq i} v_j\big(d(\Theta), \theta_j\big) - \max_{d'} \sum_{j \neq i} v_j(d', \theta_j)$$
    - How much agent i "helps/hurts" other agents, before money is taken into account.
    - For auction:
    - See examples from last class.

- Theorem: Suppose $d(\cdot)$ is an *efficient* decision function. Then any of the following transfer rules yield a strategy-proof mechanism:
  1. $$t_i(\Theta) = \sum_{j \neq i} v_j\big(d(\Theta), \theta_j\big)$$
  (in 1 cookie auction, pay the max bid to everyone except the winner),

2.
$$t_i(\Theta) = h_i(\Theta_{-i}) + \sum_{j \neq i} v_j\big(d(\Theta), \theta_j\big),$$
where $h_i(\Theta_{-i})$ is any function that is independent of θ_i. This is called the Vickrey-Clarke-Groves (VCG) mechanism.

3. The Clarke Pivot mechanism:
$$t_i(\Theta) = \sum_{j \neq i} v_j\big(d(\Theta), \theta_j\big) - \max_{d'} \sum_{j \neq i} v_j(d', \theta_j).$$

- Proof:
  - 1. Let $d' = d(\Theta_{-i}, \theta_i')$, a decision agent i might bring about with (possibly fake) bid θ_i'.
    - Net utility is
$$\begin{aligned} v_i(d', \theta_i) &+ t_i(\Theta_{-i}, \theta_i') \\ &= v_i(d', \theta_i) + \sum_{j \neq i} v_j(d', \theta_j) \\ &= \sum_j v_j(d', \theta_j) \end{aligned}$$
    - Wants to bring about the d' maximizing this.
    - By definition of efficiency,
$$d(\Theta) = d(\Theta_{-i}, \theta_i)$$
yields the max, so should tell the truth.
    - Transfer rule forces Agent i to "internalize the externality".
  - 2. Agent i can't affect $h_i(\Theta_{-i})$, so still strategy-proof.
    - Agent i only cares about maximizing net utility of outcome, across all possible "lies" θ_i', and $h_i(\Theta_{-i})$ is constant wrt θ_i.
  - 3. Take
$$h_i(\Theta_{-i}) = -\max_{d'} \sum_{j \neq i} v_j(d', \theta_j).$$

8 Examples of Strategy-proof Auctions

- Notation
  - d a decision,
  - θ_i Agent i's type,
  - $v_i(d, \theta_i)$ Agent i's base utility for outcome d,
  - A mechanism includes:
    - decision $d(\Theta)$
      - the efficient $d(\Theta)$ maximizes
        $$\sum_i v_i(d', \theta_i)$$
        across all d'.
    - transfers $t_i(\Theta)$
      - negative means pays.
    - Assume i's net utility is
      $$v_i\big(d(\Theta), \theta_i\big) + t_i(\Theta)$$
- General VCG mechanism for efficient $d(\Theta)$:
  $$t_i(\Theta) = h_i(\Theta_{-i}) + \sum_{j \neq i} v_j\big(d(\Theta), \theta_j\big),$$
  where $h_i(\Theta_{-i})$ is any function that is independent of θ_i.

- Clarke Pivot mechanism for efficient $d(\Theta)$:
  $$t_i(\Theta) = \sum_{j \neq i} v_j\big(d(\Theta), \theta_j\big) - \max_{d'} \sum_{j \neq i} v_j(d', \theta_j).$$
- Some properties of the Pivot Mechanism:
  - Since $d(\Theta)$ is one possible d' in the maximization,
    $$t_i(\Theta) \leq 0.$$
    - No agent receives money.
  - If $d(\Theta)$ is the maximum d', then $t_i(\Theta) = 0$.
    - If agent doesn't affect decision, doesn't have to pay.
  - Since each $t_i(\Theta) \leq 0$,
    $$\sum_i t_i(\Theta) \leq 0.$$
    - We say this mechanism is *feasible*: no money need magically appear.

- If v_j are all nonnegative (like in an auction),
  net utility $= v_i\big(d(\Theta), \theta_i\big) + t_i(\Theta)$

$$= v_i\big(d(\Theta), \theta_i\big) + \sum_{j \neq i} v_j\big(d(\Theta), \theta_j\big) - \max_{d'} \sum_{j \neq i} v_j(d', \theta_j)$$

$$= \sum_{j} v_j\big(d(\Theta), \theta_j\big) - \max_{d'} \sum_{j \neq i} v_j(d', \theta_j)$$

$$\geq \sum_{j} v_j\big(d(\Theta), \theta_j\big) - \max_{d'} \sum_{j} v_j(d', \theta_j)$$

$$=$$

  - We say this mechanism is *individually rational*: no one regrets participating.

- An auction example:
  - Isaac, Jenny, and Karla want my 4 cookies.

$$\theta_I = (10, 8, 6, 1)$$
$$\theta_J = (9, 7, 2, 0)$$
$$\theta_K = (5, 4, 3, 0)$$

    - Type is vector of *marginal* utiilities of 1st, 2nd, 3rd, and 4th cookies.
    - If Isaac gets 1 cookie, 10 utility.
      If Isaac gets 2 cookies, $10 + 8 = 18$ utility.
    - Etc. Decreasing marginal utility.
  - A few example outcomes:

d	v_I	v_J	v_K
$d_1 =$ "I gets 2, J gets 2"	$10 + 8 = 18$	$9 + 7 = 16$	0
$d_2 =$ "J gets 2, K gets 2"	0	$9 + 7 = 16$	$5 + 4 = 9$
$d_3 =$ "I gets 3, K gets 1"	$10 + 8 + 6 = 24$	0	5

    - Among all possible decisions, which is efficient?
    - Which decision maximizes

$$\sum_{j \neq I} v_j(d', \theta_j) = v_J(d', \theta_J) + v_K(d', \theta_K)$$

    - Which decision maximizes $v_I(d', \theta_I) + v_K(d', \theta_K)$?
    - Which maximizes $v_I(d', \theta_I) + v_J(d', \theta_J)$?
  - So Isaac and Jenny should each get 2 cookies (the efficient d_1), and pivot mechanism

transfers are:

$$
\begin{aligned}
t_I(\Theta) &= \big(v_J(d_1,\theta_J) + v_K(d_1,\theta_K)\big) - \big(v_J(d_2,\theta_J) + v_K(d_2,\theta_K)\big) \\
&= (16+0) - (16+9) = -9, \\
t_J(\Theta) &= \big(v_I(d_1,\theta_I) + v_K(d_1,\theta_K)\big) - \big(v_I(d_3,\theta_I) + v_K(d_3,\theta_K)\big) \\
&= (18+0) - (24+5) = -11, \\
t_K(\Theta) &= \big(v_I(d_1,\theta_I) + v_J(d_1,\theta_J)\big) - \big(v_I(d_1,\theta_I) + v_J(d_1,\theta_J)\big) \\
&= (18+16) - (18+16) = 0.
\end{aligned}
$$

- ◦ Isaac pays 9 (deprives Karla of 5+4 utility),
  Jenny pays 11 (deprives Isaac of 6, Karla of 5),
  Karla pays 0 (hurts nobody).

- Another possible auction mechanism:
  - ◦ “Ascending Price” auction for identical goods.
  - ◦ Start the price, p, at 0, and slowly raise.
  - ◦ Agent i holds up k_i fingers if would buy k_i additional cookies at price p, but not $k_i + 1$ additional cookies.
    - Would buy each additional cookie if marginal utility is $> p$.
    - Ex: for price 7, Isaac would prefer
      2 cookies:
      net utility $(10-7) + (8-7) = 4$
      to 3 cookies:
      net utility $(10-7) + (8-7) + (6-7) = 3$.
  - ◦ Suppose n cookies left. If
    $$n - \sum_{j \neq i} k_j > 0,$$
    then Agent i immediately buys that many cookies, each at price p.
  - ◦ Continue with remainder of cookies. (k_i now represents *additional* cookies desired)
  - ◦ Example: 4 cookies,
    $\theta_I = (10, 8, 6, 1)$, $\theta_J = (9, 7, 2, 0)$, $\theta_K = (5, 4, 3, 0)$.

p	k_I	k_J	k_K	
\$0.50				
\$1				
\$2				
\$3				
\$4				
\$4.50				
\$5				
\$5.50				
\$6				

- So Isaac gets 2 cookies for $4 + 5 = 9$,
  Jenny gets 2 cookies for $5 + 6 = 11$.
- This is the pivot mechanism!
- Reasonable: get each cookie at lowest price that someone else can't prevent you.
  - That is, have to pay exactly what someone else wanted it for, which is pivot rule.
  - Lower price couldn't be strategy-proof: If Jenny could get her 2nd cookie for under \$6, Isaac would lie to try to get that cookie (his 3rd) for under \$6.
  - Higher price couldn't be strategy-proof: If Jenny were forced to pay \$6.50 for 2nd cookie, she would pretend she only wanted it for \$6.01.
- This is called an *indirect* mechanism: agents don't directly reveal type (their 4-tuple of valuations, here).
  - Instead, some more complicated procedure happens.
  - We see that it is equivalent to a *direct* mechanism: the Clarke Pivot rule.

9 VCG in Buyer-Seller Example

- Previous examples assumed the seller was not themselves an agent in the mechanism.
- Let's try 1 buyer and 1 seller.
  - Suppose value the object $\theta_b \geq 0$ and $\theta_s \geq 0$, respectively.
  - Two outcomes Trade and No Trade.
  - $v_i(d, \theta_i)$:

	b	s
T	θ_b	$-\theta_s$
NT	0	0

    - Note: This table completely determines everything else that follows.
  - Efficient $d(\Theta) =$

  - General VCG transfers:

$$t_i(\Theta) = h_i(\Theta_{-i}) + \sum_{j \neq i} v_j\big(d(\Theta), \theta_j\big)$$

	t_b	t_s
T	$h_b(\theta_s) - \theta_s$	$h_s(\theta_b) + \theta_b$
NT	$h_b(\theta_s) + 0$	$h_s(\theta_b) + 0$

  - For Clarke pivot mechanism,

$$h_b(\theta_s) = -\max_{d'} \sum_{j \neq b} v_j =$$

$$h_s(\theta_b) = -\max_{d'} \sum_{j \neq s} v_j =$$

	t_b	t_s
T	$0 - \theta_s = -\theta_s$	$-\theta_b + \theta_b = 0$
NT	$0 + 0 = 0$	$-\theta_b + 0 = -\theta_b$

  - Net utility $= v_i(d, \theta_i) + t_i(\Theta) =$

	b	s
T	$\theta_b - \theta_s$	$-\theta_s + 0 = -\theta_s$
NT	$0 + 0 = 0$	$0 - \theta_b = -\theta_b$

    - This is horrible. Not *individually rational* (net utility may be less than 0).

▫For Clarke pivot, $t_i \leq 0$, always.

▫Pivot only works well for nonnegative valuations.

∘ We can control h_i.

- If NT, probably want transfers:
- $h_b(\theta_s) =$
- $h_s(\theta_b) =$
- Yields:

	t_b	t_s
T	$-\theta_s$	θ_b
NT	0	0

- Net utility:

	b	s
T	$\theta_b - \theta_s > 0$	$\theta_b - \theta_s > 0$
NT	0	0

- Sound good? Example: $\theta_b = 80$, $\theta_s = 30$.

∘ A mechanism is *feasible* if, for all Θ,

$$\sum_i t_i(\Theta) \leq 0.$$

No money needs to come from outside system.

∘ Individually rational $\Leftrightarrow$ $\forall \Theta, i: \; v_i + t_i \geq 0$

$\Rightarrow$ (NT case for buyer): $0 + \big(0 + h_b\big) \geq 0$

$\Rightarrow$ $h_b \geq 0$. Similarly, $h_s \geq 0$.

$\Rightarrow$ If $\theta_b > \theta_s$ and so $d(\Theta) =$T,

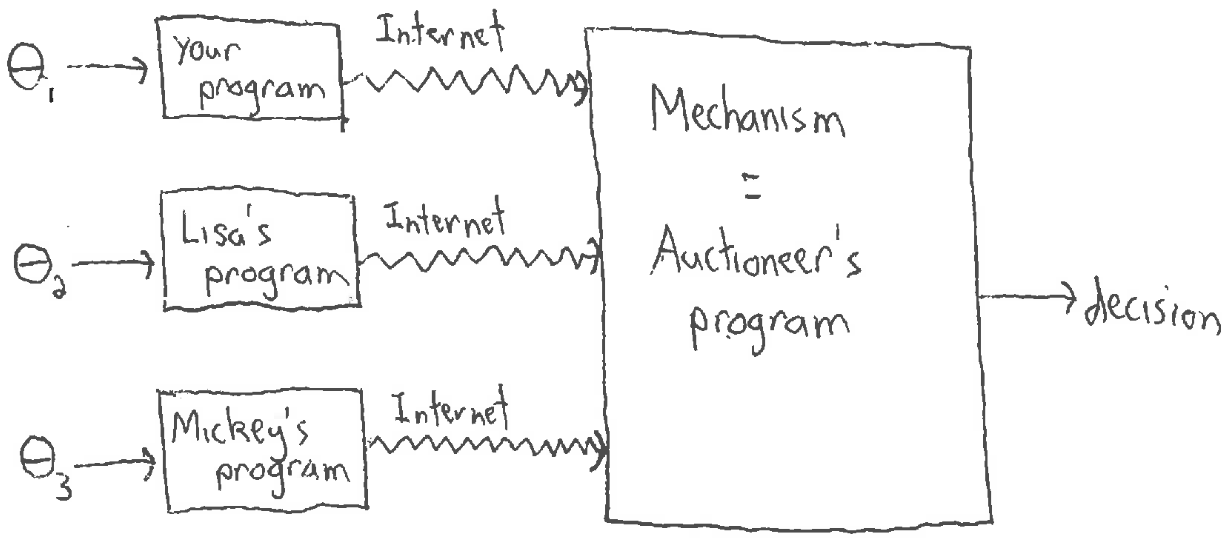
$$\begin{aligned} t_b + t_s &= (h_b - \theta_s) + (h_s + \theta_b) \\ &\geq 0 - \theta_s + 0 + \theta_b \\ &= \theta_b - \theta_s \\ &> 0 \end{aligned}$$

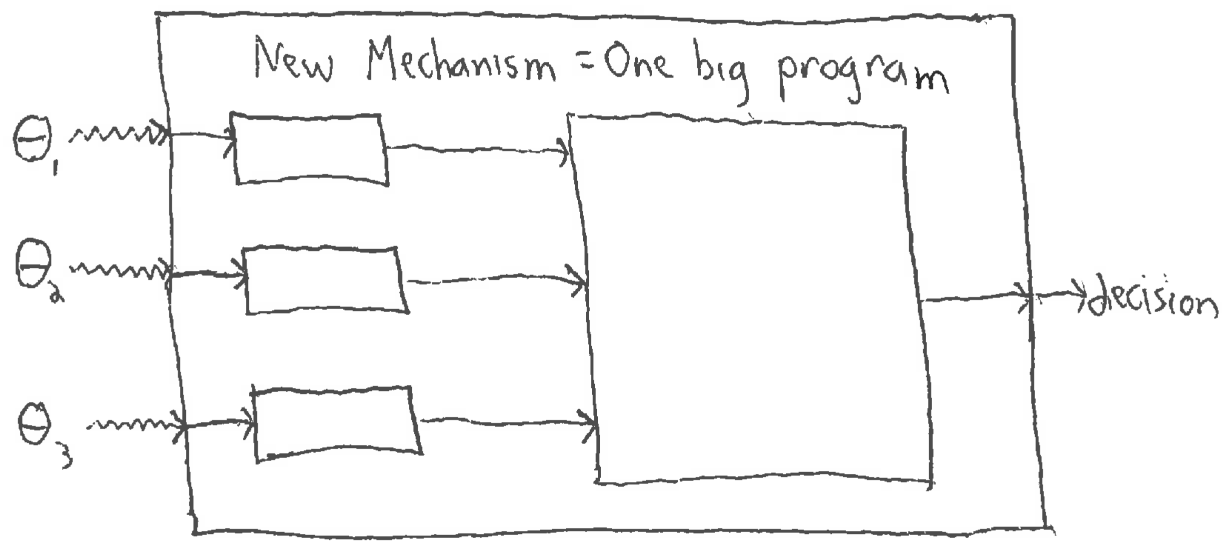
$\Rightarrow$ not feasible.

- In the buyer-seller example, no VCG mechanism is both individually rational and feasible.

10 Revelation Theorem and Preview of Green–Laffont

- In the buyer-seller example, no VCG mechanism is both individually rational and feasible.
  - Maybe some other strategy-proof mechanism is both?
  - VCG mechanism is a *direct mechanism*:
    - agents reveal type (or lie),
    - decision, transfers, etc. made based on that.
  - What about an *indirect mechanism*, like an Ascending Price auction?
    - Keep type secret, perform various actions (like raise a number of fingers) based on your type, decision based on sequence of actions.
    - Our (indirect) Ascending Price mechanism was equivalent to a direct mechanism (the Clarke Pivot mechanism).
- Theorem (Gibbard's Revelation Principle): Every indirect mechanism is equivalent to a direct mechanism.
- "Proof":
  - Imagine the indirect mechanism (e.g., ascending auction for 1 cookie) is run on the internet.
    - You type into a terminal when/how to act (e.g., bid).
    - Screen informs you of how others are acting (e.g., who's bidding how much).
  - Imagine you need to be away from the computer when the mechanism is run, but you're a good programmer.
    - Write a program to participate for you (e.g., bid in small increments until price reaches $37, then stop).
  - In fact, you could write a program that takes as input your type and then participates (e.g., if input θ_i, bid in small increments up to θ_i).
    - This is an assumption hidden in the "definition" of an indirect mechanism: you can decide how you will act, for any θ_i.
  - Why not just give your program to the auctioneer, who you trust?
    - Then you could just tell the auctioneer your type, and they could run the program on their own servers.
    - If everybody does this, this is a direct mechanism!
      - Your individual programs are now part of the mechanism.





- So indirect mechanisms can't help us get around these problems, in the Buyer-Seller example.
  - WLOG, our mechanism is direct.
  - But maybe there's some other direct mechanism that is strategy-proof, one we haven't thought of (one that isn't a VCG mechanism)?
    - Nope.
- Theorem (Green–Laffont): The only efficient, strategy-proof, direct mechanisms are VCG mechanisms. (Proof next class.)
- Corollary: In the Buyer-Seller example, no mechanism (direct or indirect) can be all of
  1) Efficient,
  2) Strategy-proof,
  3) Individually rational ($v_i + t_i \geq 0$ all i), and

4) Feasible ($\sum_i t_i \leq 0$).

- Proof of Corollary:

- Saw examples from last class of dropping either individually rational or feasible.
  - Neither option seemed great.
- Could drop efficient. Example mechanism:
  - If $\theta_s < 50 < \theta_b$, Trade. Else, No Trade.
  - If Trade, buyer pays 50, seller receives 50.
    Else, no transfers of money.
  - Strategy-proof, Individually rational, feasible.
  - Not efficient:
    - If $\theta_s < \theta_b < 50$, No Trade, but should.
    - If $50 < \theta_s < \theta_b$, No Trade, but should.
  - This is how a theoretical competitive market works.
    - Neither buyer nor seller can affect price.
    - But supply and demand "magically" set price so that ends up efficient.
      - That is, no individual can affect price, but somehow the price moves to the right level anyway.
- Could drop strategy-proof.
  - Turns out it doesn't help: For Buyer-Seller example, there is no mechanism that is efficient, individually rational, and feasible "in equilibrium", i.e., after people strategize.
  - We'll return to questions like this later, when we define the appropriate framework to think about strategizing (Nash Equilibria and Bayesian Nash Equilibria).
- For now, continue analyzing cases, like auctions, where we can have all four properties.

11 Green–Laffont Theorem

- Looking for a solution to the Buyer-Seller problem.
  - VCG mechanisms had failings (either not individually rational or not feasible)
  - Is there another mechanism we might try?
    - By Revelation Principle, might as well assume it's a direct mechanism (agent's reveal type).
    - Today: There are no other strategy-proof mechanisms, unless we give up on efficiency.
- Theorem (Green–Laffont): The only efficient, strategy-proof, direct mechanisms are VCG mechanisms.
  - Mild assumptions on Θ-space and on $v_i(\cdot)$ functions, to be discussed in proof.
- Proof:
  - We will use language of 1 cookie auction, and indicate where it might be misleading.
  - We have a direct mechanism, with type-vector Θ, valuations $v_i(d, \theta_i)$, efficient decision $d(\Theta)$, and transfers $t_i(\Theta)$.
  - Since efficient, can assume it makes same decisions as the VCG mechanisms (ties are a technicality).
  - Need to prove that the transfers correspond to a VCG mechanism, for some $h_i(\Theta_{-i})$ functions.
  - Let other bids Θ_{-i} be given.
  - Start by assuming $\theta_i = 0$.
    - Decision$=d(\Theta_{-i}, 0)$, transfer$=t(\Theta_{-i}, 0)$.
    - Define
$$h_i(\Theta_{-i}) = t_i(\Theta_{-i}, 0) - \sum_{j \neq i} v_j\big(d(\Theta_{-i}, 0), \theta_j\big).$$
      - This establishes which VCG mechanism it must be. That is, we will show that, if it agrees with a VCG mechanism for a particular θ_i ($\theta_i = 0$ here), it must agree with that same mechanism for all θ_i.
  - Starting at $\theta_i = 0$, slowly increase θ_i by small ε each time.
    - We will show that the $t_i(\Theta_{-i}, \theta_i + \varepsilon)$ is uniquely determined by the strategy-proof condition.
    - Therefore it must be the VCG payments, which are already known to be strategy-proof.
    - Assumption: set of all possible Θ is "path-connected."

- Loosely, you can get from a starting point (like $(0,\dots,0)$) to any other type vector with little ε jumps.

- Case 1: Decision doesn't change between θ_i and $\theta_i+\varepsilon$.
  - Claim $t_i(\Theta_{-i},\theta_i+\varepsilon)=t_i(\Theta_{-i},\theta_i)$.
  - Proof:
    - Suppose $t_i(\Theta_{-i},\theta_i)<t_i(\Theta_{-i},\theta_i+\varepsilon)$.
      - Who should lie?
    - Suppose $t_i(\Theta_{-i},\theta_i)>t_i(\Theta_{-i},\theta_i+\varepsilon)$.
      - Who should lie?
- Case 2: Decision changes from d to d' as change from θ_i to $\theta_i+\varepsilon$.
  - Claim:

$$t_i(\Theta_{-i},\theta_i+\varepsilon)=t_i(\Theta_{-i},\theta_i)+v_i(d,\theta_i)-v_i(d',\theta_i).$$

    - $v_i(d',\theta_i)-v_i(d,\theta_i)$ is how much an agent of type θ_i is better off in d' vs. d, so payment counteracts that.
  - Proof:
    - Since an agent of true type θ_i doesn't want to lie, we have the inequality:

      So

      $t_i(\Theta_{-i},\theta_i+\varepsilon)-t_i(\Theta_{-i},\theta_i)\le$ (⋆)
    - Since agent of true type $\theta_i+\varepsilon$ doesn't want to lie, have the inequality:

      So

      $t_i(\Theta_{-i},\theta_i+\varepsilon)-t_i(\Theta_{-i},\theta_i)\ge$ (⋆⋆)
    - Assumption: For a fixed d, $v_i(d,\theta_i)$ is continuous in θ_i.

- ⋄ So $v_i(d, \theta_i + \varepsilon) \to v_i(d, \theta_i)$ and $v_i(d', \theta_i + \varepsilon) \to v_i(d', \theta_i)$, as $\varepsilon \to 0$.

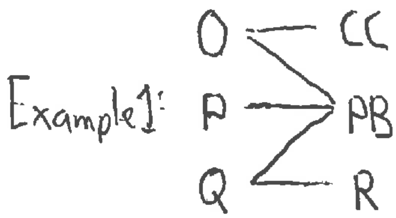
- ▫ Then ($\star$) and ($\star\star$) force

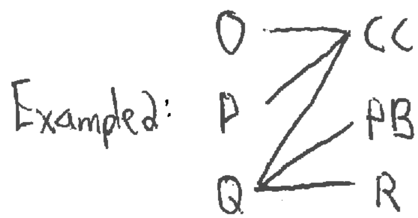
- ◦ Done!

- Ex: 1 cookie auction, highest bid among others is \$40.
  - ◦ If I bid \$0, I pay \$0 (this fixes h_i and tells us which VCG mechanism we must have).
  - ◦ As I raise my prospective bid from \$0 to \$40, I always pay \$0.
    - ▪ If I were required to pay more than \$0 (and not get the cookie!), I would lie and say \$0.
    - ▪ If I would receive money, then people who don't want the cookie at all would bid a positive amount!
  - ◦ When my bid hits \$40.01 and I win cookie, I pay \$40.
    - ▪ If I only had to pay \$39, then a bidder of true type \$39.99 would lie and say \$40.01.
    - ▪ If I had to pay \$41, then a bidder of true type \$40.01 would want to lie and say \$39.99, to avoid having to buy the cookie.
  - ◦ Any higher bid and I still pay \$40.
    - ▪ If I had to pay more than \$40, I would lie and say \$40.01.
- That is, the only strategy-proof, efficient 1 cookie auction, where a player who bids \$0 pays \$0, is the 2nd price mechanism.

12 Matchings

- Clarke Pivot mechanism seems to do well for auctions.
  - If all objects identical (but may want more than 1), we saw a nice Ascending Auction version:
    - While demand too high, increase price.
    - Equivalent to pivot mechanism.
- Feeling bold? Let's try non-identical auction example:
  - Selling jar of Peanut Butter and jar of Jelly.
    - Laura wants PB only. Would pay $5.
    - Matthew want Jelly only. Would pay $5.
    - Nellie wants both. Would pay $6 for both combined, but wouldn't pay anything for just one.
  - Efficient solution:
  - Pivot Payment:
- Okay, back up. Let's try non-identical objects, but each person wants at most one of them, since part of the problem was Nellie's desire for both.
  - Turns out to be rich, interesting topic, and we'll spend several weeks on it.
- Ophelia, Patrick, and Quentin want Chocolate Chip, Peanut Butter, or Raisin cookie. I have one of each. No one will get more than one.
  - For today, assume they find each cookie either acceptable or unacceptable (no monetary utilities).
  - Bipartite Graph:
    - Two types of vertices: people and cookies.
      - Let P be set of people vertices, and C be set of cookie vertices.
    - Bipartite: all edges go between P and C.
    - Edge between person and cookie if cookie is acceptable to the person.





- In Example 1, can everyone get an acceptable cookie?
- Matching: set of edges, at most one edge incident to each vertex.
- Perfect matching (from point of view of people): everyone gets a cookie.
  - If $|P| < |C|$ can still be a perfect matching.
    - Some cookies left over.
- How about in Example 2?

• Amazing: if there is no perfect matching, there is always a simple example showing demand exceeds supply.

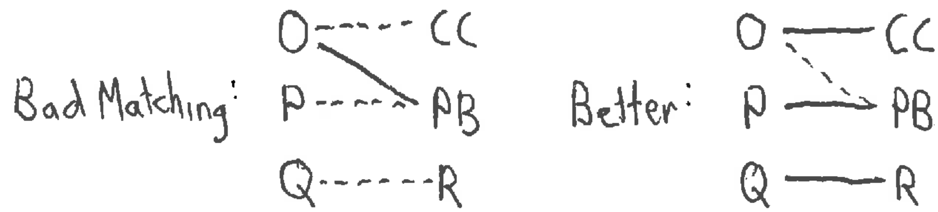
- Given $P' \subseteq P$, let
$$N(P') = \{c \in C : \ c \text{ adjacent to some } p \in P'\}.$$
  - The "neighborhood" of P'.
  - The set of cookies demanded by P'.

• Hall's (Marriage) Theorem: If there is not a perfect matching, then there exists $P' \subseteq P$ such that
$$|P'| > |N(P')| \,.$$

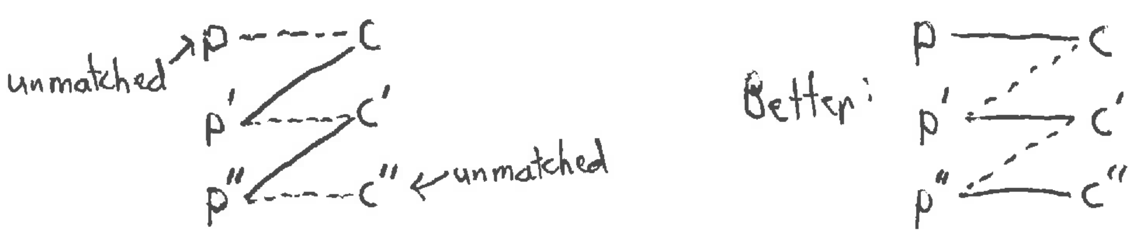
- In example, $P' = \{O, P\}$, $N(P') = \{CC\}$, $2 > 1$.
- If there is a perfect matching, you can show me one.
- If there is not, you can give me a concise reason why there isn't one: an over-demanded set of cookies.

• Warmup before proof:

- Say we have any matching that is not perfect. How might we try to make it better?
  - Example: (sold lines are in the matching, dotted lines are the other possible edges)

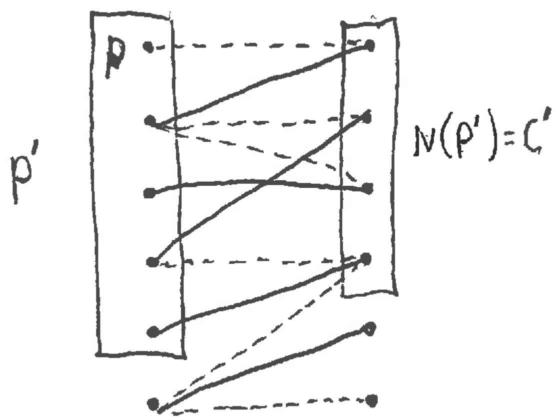


- Case 1. If (p, c) is an edge and neither are matched. Match them.
  - Q—R edge above.
- Case 2: Perform a switch like:
  - O—CC and P—PB, instead of O—PB.
  - In general, look for an *augmenting path*:
    - A) Start with an unmatched $p \in P$.
    - B) Pick a dotted/unmatched edge from p to some $c \in C$. Is c matched?
      - If not, add edge (p, c) (Case 1).
      - If so, continue to Step C:
    - C) Follow the solid/matched edge from c to some $p' \in P$. Are there more (necessarily dotted/unmatched) new edges out of p' to follow?
      - If not, failed to find an augmenting path.
      - If so, continue to Step D:
    - D) Follow dotted/unmatched edge to c'. Is c' matched?
      - If so, loop back to Step C.
      - If not, you have an augmenting path:
        - dotted-solid-dotted-$\cdots$-dotted such that:
        - First vertex, p, is unmatched.
        - Last vertex (in C) is unmatched.
        - k solid lines, $k + 1$ dotted, some k.
    - E) Switch the matching along this path. One more person (p) will be matched.



- Proof of Hall's Theorem essentially says that, if there is any way to make a matching better, this approach will work.

- Proof of Hall's Theorem:
  - Suppose there is no perfect matching.
  - Take a matching, μ, that is "maximum" (matches as many as possible).
    - μ has an unmatched $p \in P$.
    - There are no augmenting paths in μ.
  - Look at all alternating dotted-solid-dotted-$\cdots$ paths starting at p.
    - Every time they follow a dotted line into C, they follow a solid line back to P (or else would be an improving augmenting path).



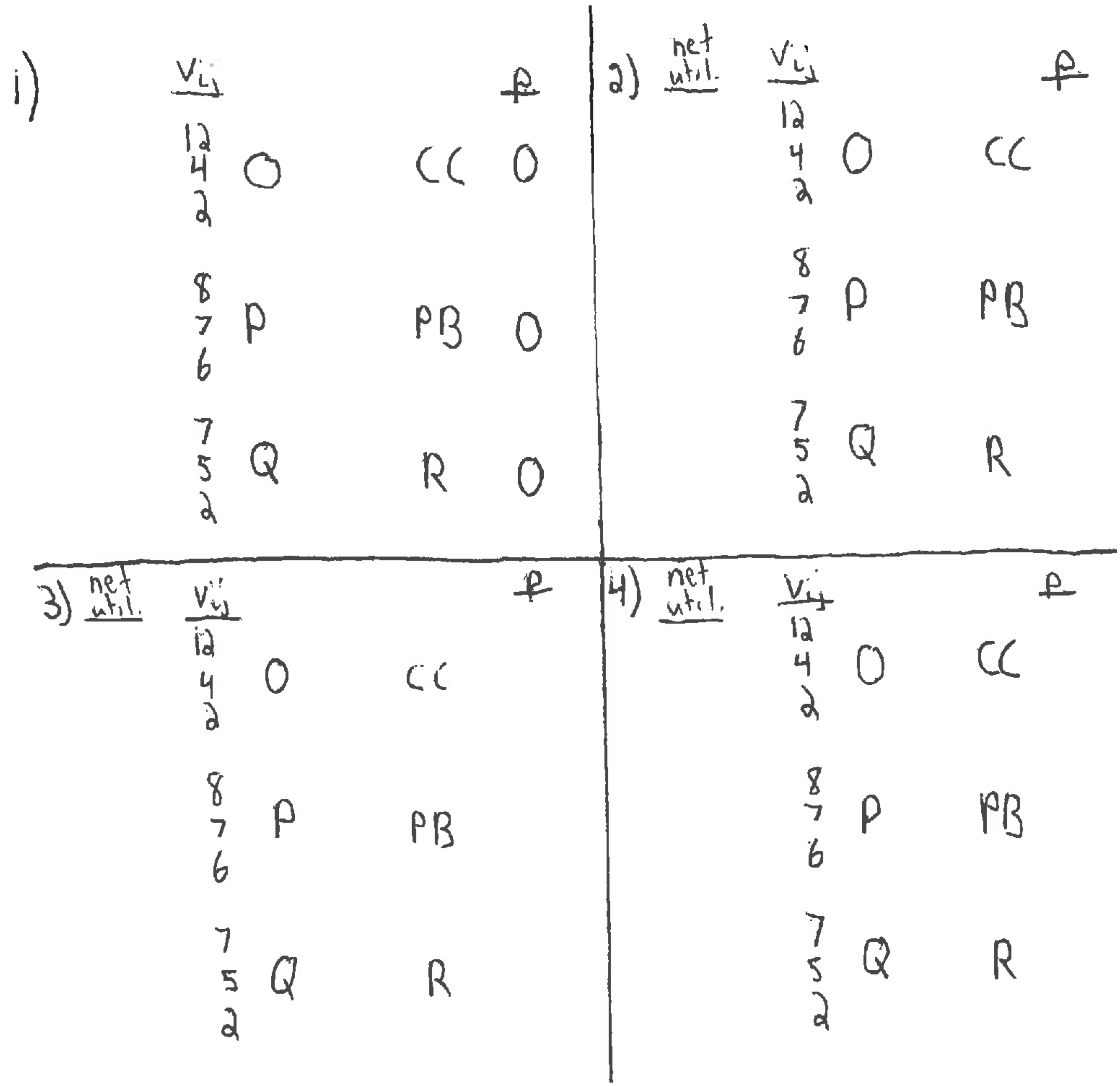
    - Let P' be all $p' \in P$ reachable from p by *some* alternating path. ($p \in P'$, as well).
    - Let C' be all $c \in C$ reachable from p by some alternating path.
    - Claim 1: $C' = N(P')$:
      - $\subseteq$: If $c \in C'$, got to it by following a dotted line from P', so $c \in N(P')$.
      - $\supseteq$: If $c \in N(P')$, connected to some $p' \in P'$ by either
        - a dotted line, in which case c follows p' in an alternating path, so $c \in C'$, or
        - a solid line, in which case how did p' get into P'?
          - New vertices only added to P' by following solid line back from C', and only solid line to p' is from c.
          - So c was already in C' before p' got put into P'.
        - Either way, $c \in C'$.
    - Claim 2: $|P'| > |C'|$:
      - Everything in C' is matched (to a unique element of P'), because no augmenting paths means always a solid line back to P'.
      - The matching gives a bijection between C' and a subset of $P' \setminus \{p\}$ (p not matched).
    - So $|P'| > |C'| = |N(P')|$.

13 Matching Markets

- Setup:
  - We re-introduce money now. Non-identical cookies are for sale.
  - Buyers have nonnegative valuations for each cookie.
  - Buyers want at most 1 cookie.
    - This is essential! PB&J example.
  - Quasi-linear utility (valuation − price paid).
    - Results hold without this, but messy.
  - Same number of buyers as cookies.
    - If too many buyers, add dummy cookies that everyone values at 0.
    - If too many cookies, add dummy buyers that value every cookie at 0.
  - Seller will sell the cookies no matter what.
    - Every buyer will get a cookie.
    - Most theorems still true without this (as long as seller doesn't strategize), but messier.
  - Every buyer simply pays for a cookie.
    - Can prove that any "stable" solution looks like this (no side payments).
- Definitions:
  - Buyers: $1 \leq i \leq n$, generally indexed with i.
  - Cookies: $1 \leq j \leq n$, generally indexed with j.
  - $v_{ij} \geq 0$ is value of Cookie j to Buyer i.
  - p_j is price of Cookie j.
  - $p = (p_1, \dots, p_n)$ is vector of prices.
  - If Buyer i gets Cookie j, net utility is
$$v_{ij} - p_j.$$
  - Let $m_i = \max_j(v_{ij} - p_j)$.
    - Best possible net utility for Buyer i.
  - Let $D_i(p) = \{j : \ v_{ij} - p_j = m_j\}$.
    - Set of objects that i would be happiest with, at current prices.
    - The "demand set" for Buyer i.
  - Let $G(p)$ be the demand graph:

  - Bipartite with vertices = buyers and cookies,
  - edge i—j iff $j \in D_i(p)$.
- ○ Prices are *quasi-stable* if $G(p)$ has a perfect matching.
  - Everyone gets one of their "best deals".
- ○ Prices are *stable* if, furthermore, $m_i \geq 0$, for all i.
  - Individually rational.

- Ascending Price Mechanism:
  - ○ 0) Prices start at 0.
  - ○ 1) If $G(p)$ has a perfect matching, you're done!
    - Sell according to that matching, at prices p.
  - ○ 2) If no perfect matching, there is some set, B, of buyers such that
$$|B| > |N(B)|\,.$$
    - Hall's Theorem! Demand exceeds supply!
  - ○ 3) Let $N(B)$ be a *minimal* such set, and raise prices evenly among $N(B)$.
    - E.g, increase each of them by 1 dollar, then by another dollar, etc.
    - Leave other prices the same.
    - If, e.g., O and P demand CC, and Q demands PB, don't raise prices on both CC and PB.
  - ○ 4) When $G(p)$ changes, return to Step 1.


1)
v_{ij}
p
12 4 2 O CC 0
8 7 6 P PB 0
7 5 2 Q R 0
2) net util.
v_{ij}
p
12 4 2 O CC
8 7 6 P PB
7 5 2 Q R
3) net util.
v_{ij}
p
12 4 2 O CC
8 7 6 P PB
7 5 2 Q R
4) net util.
v_{ij}
p
12 4 2 O CC
8 7 6 P PB
7 5 2 Q R

- Theorem:
  - 1) This auction eventually terminates (with a perfect matching).
  - 2) The final prices, p, are quasi-stable.
    - Immediate, from definitions.
  - 3) For any other quasi-stable prices, q, $p_j \leq q_j$ for all j.
    - p is optimal for each buyer, at least among quasi-stable prices.
  - 4) One of the p_j is 0.
    - If more real buyers than cookies, this will be a dummy cookie.
  - 5) Final prices are stable.
  - 6) Final matching is efficient:
    - Maximizes $\sum_i v_{i,\mu(i)}$ among all matchings, μ.
  - 7) Final prices are the Clarke Pivot prices!
    - Revelation principle!
  - 8) It is strategy-proof for buyers to act using their true valuations.
    - Immediate from 7.

- Proof of 1):
  - Assume all valuations are multiples of some ε. Increase prices by ε each time (messier without this assumption).
  - Define "potential function",
    $\Phi(p) = \sum_j p_j + \sum_i m_i$.
    - Seller revenue, plus utilities of buyers' current "dream cookies".
    - Favorite cookies probably over-demanded, and as prices go up, buyers will "settle".
  - We will:
    - a) Show $\Phi(p)$ decreases at every step by at least ε, and
    - b) give a lower bound for $\Phi(p)$.
    - Therefore, the mechanism must stop.
  - a) We have buyers B such that, in $G(p)$, $|B| > |N(B)|$.
    - Increase prices in $N(B)$ by ε.
      - $\sum_j p_j$ goes up by:
      - For each $i \in B$, m_i goes down by:
      - For $i \notin B$, m_i:
      - So $\sum_i m_i$ goes down by at least
    - Conclusion:
  - b) Take the matching i—i (may make some people unhappy).

$$m_i = \max_j (v_{ij} - p_j) \geq v_{ii} - p_i.$$

- So

$$\begin{aligned}\Phi(p) &= \sum_j p_j + \sum_i m_i \\ &\geq \sum_j p_j + \sum_i (v_{ii} - p_i) \\ &= \sum_j p_j + \sum_i v_{ii} - \sum_i p_i \\ &= \sum_i v_{ii}.\end{aligned}$$

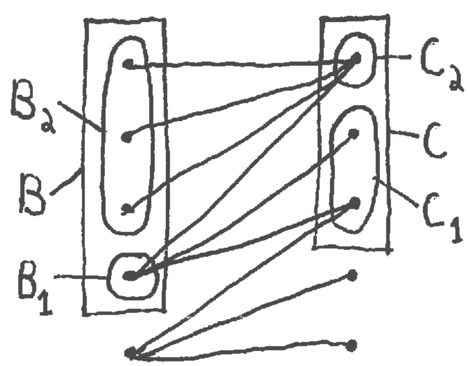
- This is a lower bound, independent of p.

○ Done!

• Proof of 2): immediate from definitions.

14 Matching Markets, part 2

- Proof of 3): For any other quasi-stable prices, q, $p_j \leq q_j$ for all j.
  - Suppose not, seeking a contradiction. Let q be quasi-stable, and look at the first point in the auction where some $p_j > q_j$.
  - In particular, let
    - B be an over-demanding set of buyers.
    - $C = N(B)$ be the set of items whose prices are being raised.
      - $|C| < |B|$, and this is a minimal set of over-demanded cookies.
      - We will eventually contradict minimality: If you raised prices above q_j, it's because you screwed up the mechanism and raised the wrong prices.
    - $p_j \leq q_j$, for all j, currently (about to change)
    - $C_1 = \{j \in C : \; p_j = q_j\}$ currently.
      - These will shortly have $p_j > q_j$.
      - C_1 is nonempty (or else there is no problem at the moment).
    - $B_1 = \{i \in B : \; D_i(p) \cap C_1 \neq \emptyset\}$, the buyers in B who are happy with something in C_1.
    - $C_2 = C \setminus C_1$ and $B_2 = B \setminus B_1$.



  - We will prove that
    a) $N(B_2) \subseteq C_2$ and
    b) $|C_2| < |B_2|$.
    - This contradicts that $C = N(B)$ was a minimal over-demanded set.
      - Shouldn't have been raising prices on C_1.
    - a)

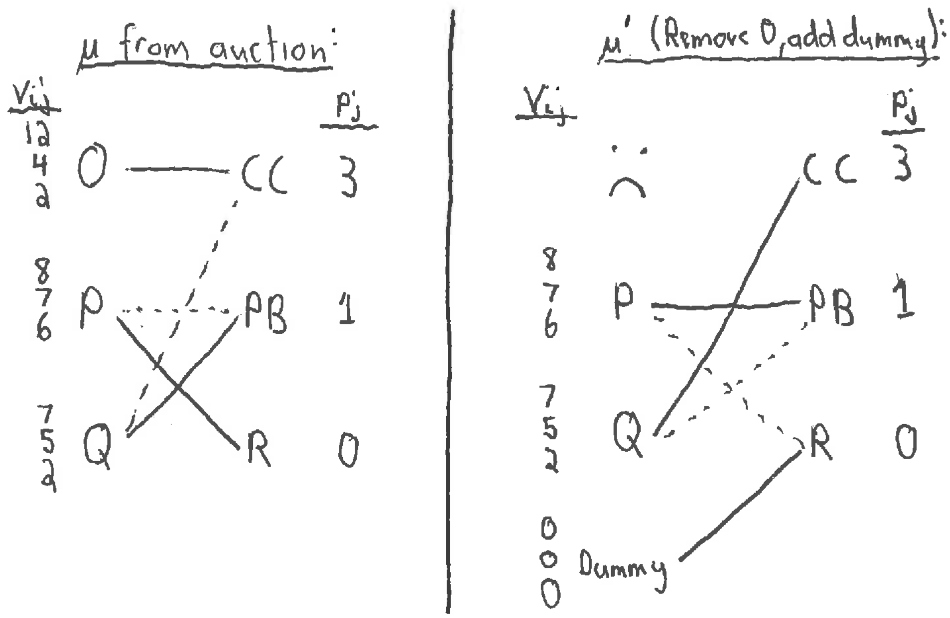
- b) Imagine raising all prices from p_j to q_j, in one go.
  - Prices in C_1 stay the same.
    Prices in C_2 go up.
    Prices not in C stay same or go up.
  - Claim: for $i \in B_1$, $D_i(q) \subseteq C_1$.
  - So $N(B_1) \subseteq C_1$ for prices q.
  - Since q is quasi-stable, everyone in B_1 can be matched to something in $N(B_1) \subseteq C_1$.
  - Let's compare some cardinalities:

- Proof of 4): One of the p_j is 0.
  - Suppose all of the p_j are positive, seeking a contradiction.
  - Look at new prices $p_j - \varepsilon$:
- Proof of 5): Final prices are stable.
  - By 2) (final prices quasi-stable), need to show $m_i \geq 0$ for all i.
  - By 4), some $p_{j_0} = 0$.

$$m_i = \max_j (v_{ij} - p_j) \geq v_{ij_0} - p_{j_0} = v_{ij_0} \geq 0.$$

15 Matching Markets, part 3

- Proof of 6): Final matching is efficient
  (Maximizes $\sum_i v_{i,\mu(i)}$ among all matchings, μ,
  where $\mu(i)$ is cookie that Buyer i gets in μ.).
  - Will show stronger result:
    quasi-stability $\Rightarrow$ efficient matching.
  - Fix quasi-stable prices, p, and let μ be the resulting matching.
  - For each Buyer i, $v_{i,\mu(i)} - p_{\mu(i)}$ maximizes
    $v_{ij} - p_j$ over all j (definition of quasi-stable).
  - So $\sum_i \left(v_{i,\mu(i)} - p_{\mu(i)}\right)$ is maximized across all matchings (it's actually maximized across all functions μ : Buyers $\to$ Cookies, because each term in the sum is maximized).
    $$\sum_i \left(v_{i,\mu(i)} - p_{\mu(i)}\right) = \sum_i v_{i,\mu(i)} - \sum_j p_j.$$
  - $-\sum_j p_j$ is fixed across all matchings.
  - Therefore $\sum_i v_{i,\mu(i)}$ is maximized across all matchings.
    - This is what efficient means!
  - Done! Amazing!
    - Feels like I'm pulling a fast one, but I promise I'm not.
    - Only needed quasi-stability.
    - cf. "First Welfare Theorem" (microeconomics): Equilibria / stable solutions are efficient.
    - Like VCG mechanism, appropriately setting prices aligned individual interests with global interest.
    - Used supply and demand to find appropriate prices.
- Proof of 7): Final prices are the Clark Pivot prices!
  - Let's see an example first:



- μ is the matching the auction comes up with.
- Remove Ophelia from auction, and add dummy buyer who will take the free cookie.
  - μ' is stable, so it is optimal matching for P and Q (by 6).
- Ophelia's pivot price:

- Why it worked:
  - P and Q indifferent between μ and μ' after paying.
  - In μ', Q and P pay an additional
$$(3-1)+(1-0)=3-0=3.$$
    "Telescoping sum".
  - Therefore raw valuation for P and Q combined must be 3 more in μ' than in μ.
  - So pivot mechanism must have $p_{CC}=3$.
  - Notice:
    - μ has alternating solid-dotted-· · · -solid path from Ophelia to the free Raisin cookie.
    - μ' switches dotted and solid, and gives Dummy the leftover free Raisin cookie.
- In general, let μ be the matching the auction ends on.
  - By 6), maximizes $\sum_i v_{i,\mu(i)}$, so same efficient decision as pivot mechanism.
- Let's figure out pivot payment for Buyer 1 (need to show it is $p_{\mu(1)}$).

- Look at paths starting at Buyer 1, alternating solid-dotted-solid-$\cdots$.
- Claim: For prices from Ascending price mechanism, one of these paths eventually hits a Cookie j_0 with $p_{j_0} = 0$.
  - Proof: Suppose not, seeking contradiction.
    - Let B be all buyers reached along such paths, and C be all cookies reached.
    - Assuming $p_j > 0$, for all $j \in C$.
    - $N(C) = B$, and everyone in B matched to something in C:
      - Same as Hall's Theorem proof.
      - Here we start solid-dotted-$\cdots$ instead of dotted-solid-$\cdots$.
    - Nobody outside of B is demanding things from C.
    - Since everything in C has price > 0, can lower prices in C by ε.
    - μ is still stable:
    - This contradicts 3), which said p_j were buyer-optimal, among quasi-stable prices.
- By claim, have an alternating path starting at Buyer 1, solid-dotted-$\cdots$-solid, ending at Cookie j_0, with $p_{j_0} = 0$.
- Let μ' be matching if we take μ, remove Buyer 1, and switch solid/dotted in the above path.
  - The free Cookie j_0 is unmatched. Give it to a new Dummy Buyer, who will demand it.
- Then:
  - μ' is stable for other buyers, so it maximizes
    $$\sum_{i=2}^{n} v_{i,\mu'(i)}$$
    over all matchings, by 6).
  - $v_{i,\mu(i)} - p_{\mu(i)} = v_{i,\mu'(i)} - p_{\mu'(i)}$, for $i \neq 1$, since both edges in $G(p)$.

◦ So Clarke Pivot prices are:

$$\begin{aligned}
\sum_{i=2}^{n} v_{i,\mu'(i)} - \sum_{i=2}^{n} v_{i,\mu(i)} &= \sum_{i=2}^{n} v_{i,\mu'(i)} - \sum_{j=1}^{n} p_j + \sum_{j=1}^{n} p_j - \sum_{i=2}^{n} v_{i,\mu(i)} \\
&= \left(\sum_{i=2}^{n} v_{i,\mu'(i)} - p_{\mu'(i)} \right) - p_{j_0} + \sum_{j=1}^{n} p_j - \sum_{i=2}^{n} v_{i,\mu(i)} \\
&= \left(\sum_{i=2}^{n} v_{i,\mu(i)} - p_{\mu(i)} \right) - 0 + \sum_{j=1}^{n} p_j - \sum_{i=2}^{n} v_{i,\mu(i)} \\
&= \sum_{i=2}^{n} v_{i,\mu(i)} - \sum_{i=2}^{n} p_{\mu(i)} + \sum_{j=1}^{n} p_j - \sum_{i=2}^{n} v_{i,\mu(i)} \\
&= p_{\mu(1)}.
\end{aligned}$$

Done!

- Proof of 8): It is strategy-proof for buyers to act using their true valuations.
  - Immediate because Clarke Pivot mechanism is strategy-proof.

16 Stability and Strategy-proofness

- No precise definitions today.
- If you don't like the strictness of the auction mechanism we've been looking at (a very particular set of prices has to be raised at each step), consider the following:
  - Opening bids for all objects are $0.
  - A buyer not currently winning any objects may bid on any object.
    - Assume they bid exactly ε over the current bid, for some fixed ε, and that they bid on an object maximizing their net utility (at current prices).
  - Bidding continues until no one wants to bid anymore.
  - In Problem Set 5: you will show that the final prices will be within $2k\varepsilon$ of the pivot / Ascending Price mechanism, where k is the number of non-dummy objects.
    - Example: 1 cookie (and 1 dummy), Roger values it at 5, Sara values it at 3, $\varepsilon = \$0.01$:
- For matchings with money, the Ascending Price mechanism is strategy-proof and gives each buyer best price among all stable (quasi-stable) prices.
- In more general settings, loosely:
  - "Stable" means
    - Individually rational, and
    - There is no way to break this deal and make a new one in such a way that both seller and the "swooping" buyers are better off.
  - Example: 1 cookie, Roger values it at 5, Sara values it at 3. Stable solutions:
    - In the matching with money context, the new definition of stable agrees with the old one (no one is jealous), but this is not obvious (and not true in other settings).
  - Nice thing: stable implies efficient!
  - Strategy-proof mechanisms *hopefully* yield stable solutions (e.g., if Roger gets cookie for $1, Sara can probably strategize to get it for $2).
  - Knowing other player's bids, a player can often lie to bring about the stable solution that is best for themselves.
    - In example, Roger can pretend to want it for $3.01, to get it for his best stable price.

  - One can prove that, indeed, a strategy-proof mechanism must give an agent their best outcome, among all stable outcomes.

- Example: 2 identical cookies.
  Twila values one at 5, second for 3 more.
  Ursula values one at 6, second for 1 more.
  Stable solutions:
      Twila pays between 1 and 5 for one.
      Ursula pays between 3 and 6 for one.
      Efficient!
      \$1 and \$3 are Pivot payments.

  - Note: The "no jealousy" version of stability (that Ursula doesn't want Twila's cookie *instead of* her own) doesn't coincide here. Only that Ursula can't swoop in to *also* take Twila's.

- Example: Laura values PB (only) at 5.
  Matthew values Jelly (only) at 5.
  Nellie wants (only) both, for \$6 combined.

  - Pivot outcome (\$1 each) is not itself stable!

    - Best individual outcomes for Laura and Matthew are not simultaneously stable, unlike in previous examples.

    - To make strategy-proof, auctioneer must commit to selling at prices that feel stupid.

- Loose Definition: Items are substitutes if decreasing price of one does not increase demand for the other.

  - Matching? Yes! Only want one.

  - Identical items? Yes! Decreasing marginal utility.

  - PB&J? No! Nellie would pay more for Jelly if PB were cheaper. They are *complements*.

- Loose Theorem: If items are substitutes, the Pivot mechanism is stable.

- Stability is still a powerful idea for non-auctions.

- Example: Buyer-Seller.

  - Vera has cookie and would sell for 3.
    Walter would buy for 5.
    Stable solutions:

- Example: Public Good.
  - Once bought, everyone can use and get utility. $100 Coffee maker!
  - Xavier values it at 40.
    Yancy values it at 80.
    Stable solutions:
- Thinking about stability shows why finding a strategy-proof mechanism didn't work well with these.
  - Optimal stable solutions for the agents don't "align" with each other.
  - Amazingly for matching (and also multiple identical item) auctions, buyers stable solutions do align, and strategy-proof is sensible!

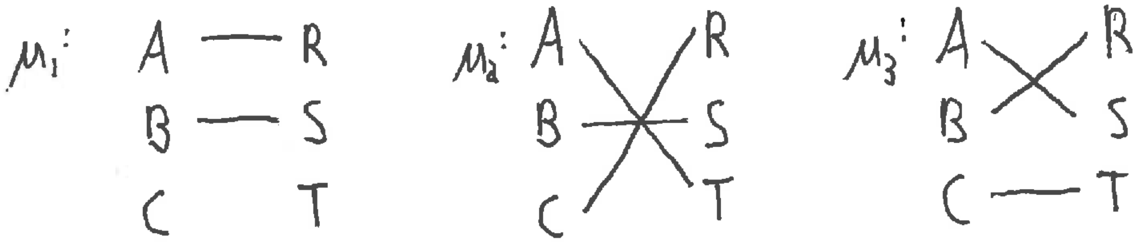
17 Matching without Money

- We're seeing limitations of "strategy-proof". We'll talk about strategizing soon.
  - But first, one more example: matching without money (but with strict preferences on outcomes).
- Setup:
  - Two groups, M (men) and W (women).
  - Matching = straight marriage.
    - Will see an example of how gay marriage has very different properties.
  - No money changes hands.
    - Students choosing public schools.
      Residency programs choosing doctors.
      Organ donation (Al Roth's Nobel prize!)
      Undergrads choosing houses.
  - If m and w matched in μ, write $\mu(m) = w$ and $\mu(w) = m$.
  - We allow someone to stay single, and write $\mu(m) = m$. (not a perfect matching).
  - A person has strict preferences (no ties) among the opposite group.
  - Some may be "unacceptable" (prefer to be single).

- Example:

Al	Bob	Chris	Rachel	Susie	Theresa
R	S	R	A	A	A
S	R	T	B	C	B
	T	S	C	B	C

  - A prefers R to S to "single" to T.



  - Is μ_1 "stable"? (whatever that means)

  - Is μ_2 "stable"?

  - Is μ_3 "stable"?

- Definition: μ is *unstable* if either
  1) Some $k \in M \cup W$ finds $\mu(k)$ unacceptable (cf, individually rational) or
  2) Exists $m \in M, w \in W$ such that w prefers m to $\mu(w)$ and m prefers w to $\mu(m)$. (cf, loose definition from last class).
  - μ is *stable* otherwise.
- Is there always a stable matching? Yes:
- Women Propose Algorithm.
  - Everyone starts unattached.
  - Each woman proposes to her favorite (acceptable) man.
  - A particular man rejects all of his proposals, except for his favorite (acceptable) option among the proposals.
  - If his absolute favorite woman proposes, he accepts. Otherwise, he stalls his favorite among the proposals. ("I'll think about it").
  - Repeat. Every woman who is neither married nor stalled proposes to her next favorite *unmarried*, acceptable man. Men accept their absolute favorite *unmarried* woman, if she proposes; else stalls their favorite among the proposals.
  - When there are no new proposals, each man "settles": accepts his current proposal.

   - Call the answer μ_W.
- In example:

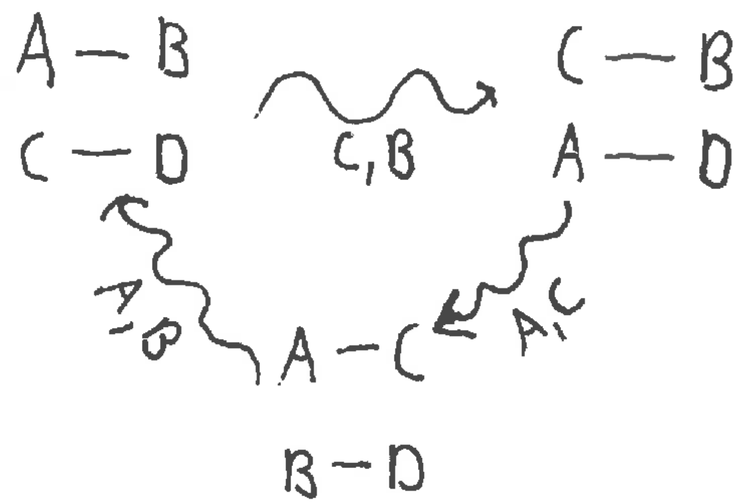
- This always yields a stable matching:
   - Can anyone be matched to an unacceptable person?
   - Suppose m and w prefer each other to $w' = \mu_W(m)$ and $m' = \mu_W(w)$.
      - At some point, w' proposes to m, and m eventually accepts. How long will m stall?
      - Must w eventually propose to m?
      - Since w proposes to m, he will reject w' (now, if w' already proposed; later, if not). Contradiction.
- Similarly, there is a Men Propose Algorithm, yielding a stable matching. Call it μ_M.
   - In example:

- Battle within genders feels most salient:
  A and C prefer R, but only A matched with R.
  R, S, and T prefer A.
- Battle between genders is more subtle: If people are realistic and only look at *stable* matchings:
   - In this example (but not always), μ_W and μ_M are the only stable matchings.
   - All women prefer μ_W to μ_M,
     and all men prefer μ_M to μ_W.

- Will show this is a general phenomenon!

- Gay marriage example:

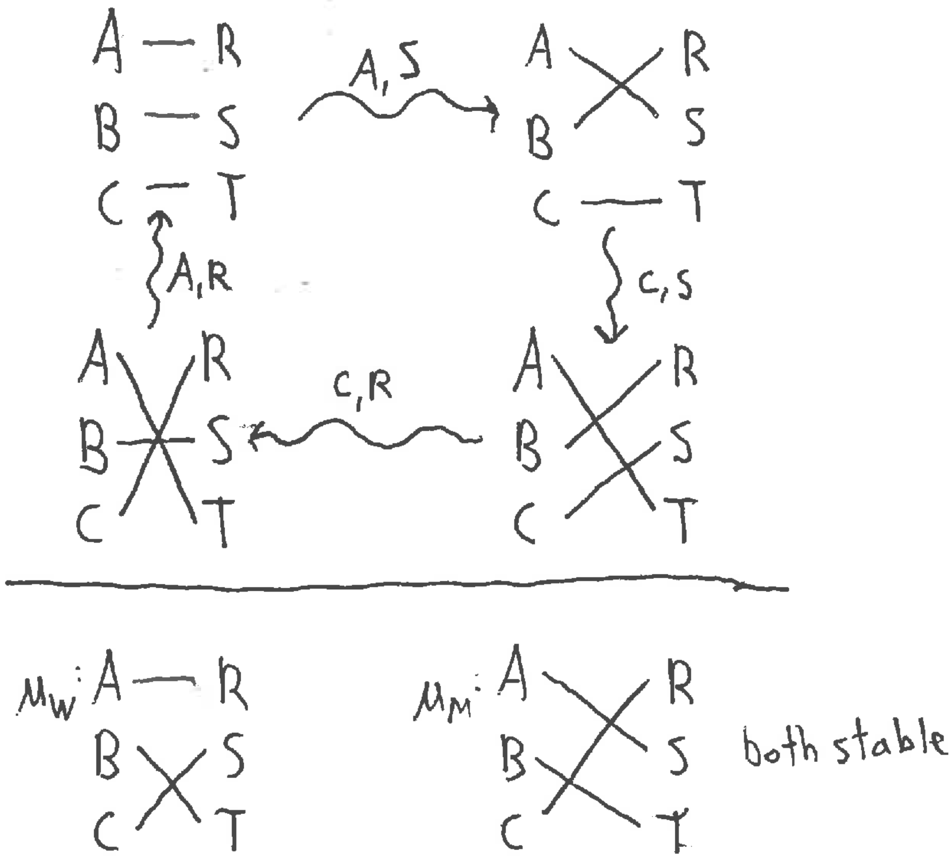
A	B	C	D
B	C	A	A
C	A	B	B
D	D	D	C



  - ◦ Bipartite nature is essential.

- Example: Unstable matchings may not stabilize naturally:

A	B	C	R	S	T
S	R	R	A	C	A
R	T	S	C	A	C
T	S	T	B	B	B

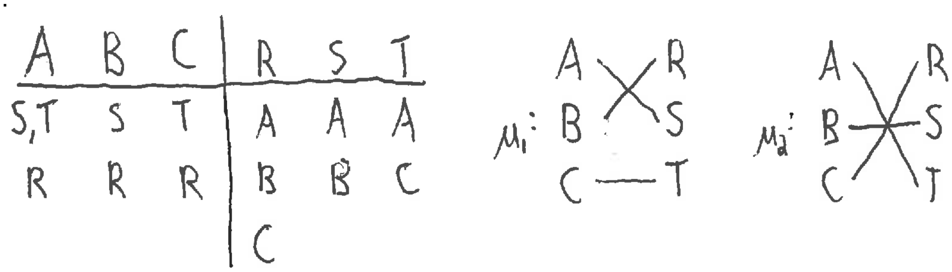


18 Battle of the Genders

- Theorem (Battle of the Genders):
  - 1) Among all stable matchings, no woman does better than μ_W.
  - 2) If μ and μ' are stable matchings such that all men like μ as well as they like μ', then all women like μ' as well as they like μ.
  - 3) Among all stable matchings, no woman does worse than μ_M.
  - 4) μ_W is strategy-proof for women.
    - cf 1) and 4) with earlier comments on stability and strategy-proof.
- Proof of 1) Among all stable matches, no woman does better than μ_W:
  - Say that m is *achievable* for w if exists a stable matching μ with $\mu(w) = m$.
  - Claim: in Woman Propose matching, no woman is ever rejected by one of her achievable men (let "rejection" include the case where he is already matched when she first considers proposing).
    - This implies 1), because a woman may propose to some unachievable men, but they must eventually reject her, and eventually she will propose to her favorite achievable man. Since he won't reject her, he must eventually say yes.
    - Proof of claim: Seeking contradiction, suppose some woman is rejected by an achievable man.
      - Let (w, m) be the *first* such pair, chronologically.
      - m rejects w, but $\mu(w) = m$ for some stable matching μ.
      - Case 1: w is unacceptable to m. Contradiction, because:
      - Case 2: w is rejected by m in favor of some w' (who is accepted or stalled).
        - w' has not been rejected by an achievable man (since w is the first). Contradiction, because:

- Proof of 2) If μ and μ' are stable matchings such that all men like μ as well as they like μ', then all women like μ' as well as they like μ:
  - Suppose not. Suppose w prefers $m = \mu(w)$ to $m' = \mu'(w)$, $m \neq m'$. Contradiction. because:

- Proof of 3) Among all stable matchings, no woman does worse than μ_M:
  - Given a stable μ, analog of 1) implies all men do as well in μ_M as in μ. Therefore 2) implies that no women does better in μ_M than μ.
- "Proof" of 4) μ_W is strategy-proof for women:
  - If you like someone, just ask him! The worst he can do is say no! Some unachievable men may say no, but you never stood a shot anyway.
  - By proof of part 1, your best achievable man will wait for you (no achievable man will reject you, including saying "yes" to another woman).
- Note: Strict preferences are necessary for this theorem.
  - Example:



  - μ_1 and μ_2 are the only stable matchings.

    Prefers μ_1:

    Prefers μ_2:

- So Women Propose Algorithm seems extra nice if we don't care about the preferences of the men!
- Many natural examples where one side has no real preferences.
  - Example: People looking for housing W=people. M=houses/dorm rooms.
    - If people have a priority order, each dorm could have that preference list.
    - If people already living in a house have "squatter's rights", they could be at the top of their house's list... can't be rejected.
- Example: Hospitals and med students looking for residencies.

- Many-to-one matching: hospitals may get more than one resident.
- Still natural "Hospital Proposes" and "Resident Proposes" algorithms.
- Can prove that Hospital Propose is best for hospitals among stable matchings. But it is not strategy-proof for hospitals!
  - Example: A gets 2 residents; B and C get 1.

A	B	C	R	S	T	U
R	R	T	C	B	A	A
S	S	R	A	A	C	B
T	T	S	B	C	B	C
U	U	U				

- Hospital Proposes (and only stable matching):

A	R
B	S
C	T
	U

.

- If A lies and gives preference R>U (S and T unacceptable; demand reduction):

A	R
B	S
C	T
	U

.

  - Not stable under true preferences:
  - All three hospitals do as well or better when A lies!

- History of Residency Match program:
  - Through the 40's, a free-for-all.
    - Hospitals would choose residents two years in advance to beat out other hospitals. Unstable.
    - Law clerking is still somewhat like this. Voluntary common deadlines that many place ignore. "Exploding" offers.
  - 50's to 90's: Hospital Propose Mechanism.
    - Stable; Good for hospitals, who were creating the mechanism.
  - Late 90's to present.
    - Residents lose faith in system. Resentment builds (it matters a lot more to them than to the hospitals).
    - Residents start trying to strategize.
      - Though empirical evidence says that's very hard to do. Very few stable matchings, so strategizing can't help much. More likely to hurt yourself because of your lack of information.
    - Switched to residents propose system.
      - Hospitals a little worse off (but not much). Gained "buy-in" from residents, for whom it is strategy-proof.
      - Also helped with a growing problem: Couples could jointly "propose" to same hospital.
        - Getting matched at same hospital is a "complementary good" which does cause problems.
        - No guarantee of stable solution, and couples can strategize.
        - But seems to do a really good job, in practice.

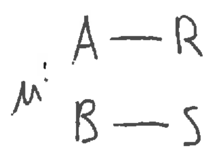
19 Optimality (Optional)

- What is a "good" outcome for a set of agents? What is the "best outcome"?
  - Agents are often at odds with each other.
- Matching with Money:
  - For the *buyers*, the Ascending Auction result is *dominant* among all *stable* solutions. It is the best for each buyer individually.
    - Buyers are not "at odds" with each other, at least if you only consider stable/realistic solutions.
    - Loosely, this what is needed for a strategy-proof solution.
  - For buyers and sellers as a group, this is no longer true (it is the *worst* stable outcome for the sellers).
    - But any stable result is *utilitarian* among all possible solutions, meaning it maximizes the sum of all net utilities:
      - Stable ⇒ efficient, maximizes
        $$\sum_{i \text{ a buyer}} v_i(d)$$
        (consumer surplus).
      - Payments don't affect sum of net utilities: If I pay you \$5, I'm \$5 sadder and you're \$5 happier.
      - We're just deciding how to divide up the consumer surplus.
- Matching without money:
  - For the women, Women-propose is *dominant* among all stable solutions. Rejoice (women)! Dominance is a rare thing.
    - Again, this is loosely the reason that this is strategy-proof for the women.
  - But can't talk about utilitarian, because no utilities!
- Definitions: In general, let a group, A, of agents, and a set, D, of outcomes be specified.
  - $d \in D$ is *dominant* (for A, among D) if $\forall i \in A$, $\forall d' \in D$, Agent i likes d as well as likes d'.
  - $d \in D$ is *utilitarian* if maximizes
    $$\sum_{i \in A} \text{utility of } d',$$
    among all $d' \in D$.
    - Requires a utility function and interpersonal comparison of utility.
      - \$5 to me is the same as \$5 to you, e.g.
      - Sketchy, and we don't always have a clear utility function.

- $d \in D$ is *strongly Pareto optimal* (SPO) if there is no $d' \in D$ making *some* agents $i \in A$ *strictly* better off than they were in d and all agents *at least as* well off (can't help some without hurting others).
- $d \in D$ is *weakly Pareto optimal* (WPO) if there is no $d' \in D$ making *all* agents $i \in A$ *strictly* better off than they were in d.
- Dominant ⇒ SPO ⇒ WPO:

• For A = all men and all women together, many matchings are SPO (among D = all matchings):

A	B	R	S
S			B
R			A

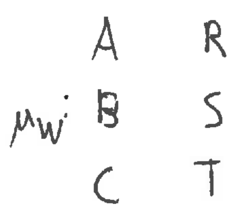


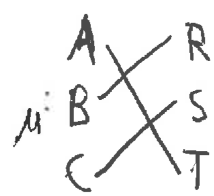
- Without quantifying utility, "helping three while hurting one" isn't clearly good or bad.

• How about A = just women?

- μ_W is dominant for women, among D = all *stable* matchings.
- How about for a broader class of matchings?

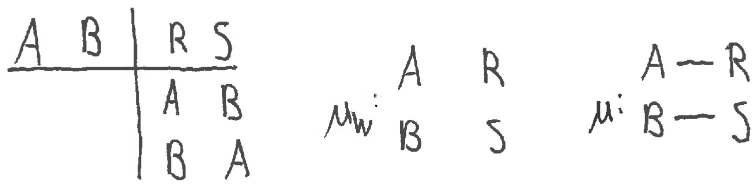
men			women		
A	B	C	R	S	T
R	T	R	B	A	A
S	R	S	A	B	B
T	S	T	C	C	C





  - μ_W not SPO for women:
  - μ not stable:
  - Women could collude to make μ happen in Woman-Propose mechanism:
    - Some better off in this collusion, others no worse off (μ_W not SPO).

- Another example:



  - μ_W not WPO for women, because μ better for both.
  - μ is not individually rational.
  - μ_W is WPO for women, among all individually rational matchings.
    - Women can't collude in Women-Propose to make *all* women strictly better off.

- Theorem: μ_W is WPO for women, among all individually rational matchings.
- Theorem: No subset of women can lie about their preferences and make all members of that subset strictly better off in Woman-Propose.
  - "Group-strategy-proof". Though can collude if some are ok with being no worse off.
  - Stronger than the previous theorem: a subset can't make themselves strictly better off, even by hurting other women.
  - Proofs: kinda long and intricate.
- Matching with money.
  - Collusion very possible: 1 cookie; if Zhao values it at \$20 and Aimee at \$10:
  - Not WPO for buyers among *all* possibilities. But dominant for buyers among all stable possibilities.
- In any situation (not just matching) where money can change hands arbitrarily, utilitarian, SPO, and WPO are equivalent:
  - not SPO $\Rightarrow$ not utilitarian: if d' makes some people better off and no one worse off, then sum of utilities is higher for d'.
  - not WPO $\Rightarrow$ not SPO: done!
  - not utilitarian $\Rightarrow$ not SPO:
    - If d not utilitarian, exists some d' with
    $$\sum_i u_i(d') > \sum_i u_i(d)$$
    (u_i are net utilities).
    - Make d' happen. Then:
      - $\forall i$, Agent i puts $u_i(d')$ dollars into a pot.
      Net utility = 0.

- $\forall i$, Agent i takes $u_i(d)$ out of pot.
  Net utility $= u_i(d)$.
- $\varepsilon = (\sum_i u_i(d') - \sum_i u_i(d))$ left in pot.
- Everyone takes ε/n out of pot.
  Net utility $= u_i(d) + \varepsilon/n$.

- Everyone is ε/n happier than they were in d, so all strictly better off!

20 Utility

- Efficiency is awesome. In any situation (not just matching) where money can change hands arbitrarily, *not efficient* means there is a way to make *everybody* strictly better off.
  - And if there's a way to make everyone better off, surely we're doing something wrong!
  - Proof: If d not efficient, exists some d' with
$$\sum_i v_i(d', \theta_i) > \sum_i v_i(d, \theta_i)$$
  - Let current transfers t_i be given, so that net utilities are currently $v_i(d) + t_i$.
  - Let $\varepsilon = \left(\sum_i v_i(d') - \sum_i v_i(d)\right)/n > 0$.
  - Make d' happen instead, with new transfers
$$t'_i = t_i - v_i(d') + v_i(d) + \varepsilon.$$
    - New net utility is
$$v_i(d') + t'_i = v_i(d) + t_i + \varepsilon.$$
    - Everyone is ε happier than they were before!
    - Also
$$\begin{aligned}\sum_i t'_i &= \sum_i t_i - \sum_i v_i(d') + \sum_i v_i(d) + n\varepsilon \\ &= \sum_i t_i,\end{aligned}$$
      so the change is budget neutral.
- Welfare Theorems of Market Economics (roughly):
  1) Market equilibria / stable prices are efficient.
  2) Any efficient outcome can be achieved as a market equilibrium by changing the "initial endowment" (redistribute wealth).
  - Translation: Market equilibria are awesome, because efficiency is awesome. If you think someone "deserves" the cookie, even though they do not have the highest valuation for it, then give them more money and let them decide whether to buy it.
- A major problem:
  - There are many efficient outcomes. Which one is "best"?
  - Taking $100 from me and giving it to Bart will lead to a different equilibrium that makes Bart happier and me less happy.
  - Interpersonal comparison of utility is hard and fraught. Does that $100 make Bart or me happier? And do we care?
- A start is to define utility $U(\cdot)$ more carefully, for an individual:
  - We've been assuming that $U(\cdot)$ is linear in money transfer, but not always true.

- Assumption: people can always choose which of two things they prefer, and those choices are "consistent."
  - I might give you some money! Arbitrarily decide $U(\$\ 0) = 0$ and $U(\$1\text{ million}) = 100$.
  - Which do you prefer:
    1) \$1 million, guaranteed, or
    2) 50-50 shot at \$2 million (\$0 otherwise).
  - $U(\text{option 1}) > U(\text{option 2})$.
  - For what p are you indifferent between p chance at \$2 million versus \$1 million guaranteed?
    - We define $U(\cdot)$ as if you are maximizing your *expected utility*. So:
      $$100 = U(\$1\text{ mill}) = U(\text{this lottery}) =$$
    - So $U(\$2\text{ mill}) =$
  - For what p are you indifferent between a p chance at \$1 mill versus a sure thing of \$100,000?
    - So:
  - If one assumes that an individual person chooses among these lotteries "consistently", you can prove that this utility function is well-defined.
- In any situation, if you fix 2 reference points (like \$0 and \$1 mill) arbitrarily, then this gives a utility function. So you can say things like:
  - "I like \$100,000 p times as much as \$1 mill, in relation to \$0.
  - Or in a voting context, for example:
    - "If $U(\text{Bush}) = 0$ and $U(\text{Nader}) = 1$, then $U(\text{Gore}) = 0.5$."
    - "I am indifferent between Gore as president and a 50-50 coin flip between Bush and Nader."
    - "I like Nader twice as much as Gore, in relation to Bush."
  - This communicates more than a preference relation, while still avoiding interpersonal comparisons of utility.

  - Even though $U(\text{Nader}) = 1$ for you, and $U(\text{Nader}) = 1$ for me, we aren't saying that choosing Nader over Bush makes us both equally happy. We fixed two outcomes' utilities arbitrarily.

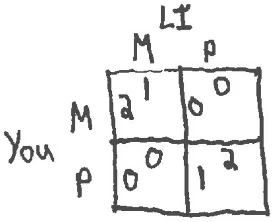
- We are going to switch to thinking about strategizing, where the outcome will be uncertain. The assumption will be that people act so as to maximize expected utility.

21 Game Theory

- You have a Love Interest (LI).
  - You're hoping to run into them tonight, either at a Math Lecture (M) or a Play (P).
- Scenario 1: You get +2 utils of happiness for M,
  and +1 utils if LI goes to same event as you.
  Your personal utilities:

You \ LI	M	P
M	3	2
P	0	1

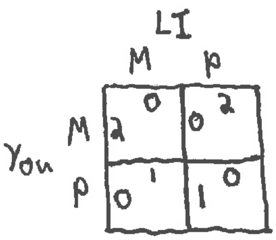
  - If you knew LI going to M, what would you do?
    Always maximize utility.
  - If you knew LI going to P, what would you do?
  - M is a *dominant strategy* for you:
    - No matter what opponent does, you cannot do better than M.
    - M maximizes your utility, regardless of what opponent does.
  - This is what strategy-proof means.
    - Strengths:
      Don't have to strategize.
      Don't even have to know LI's utilities.
    - Weaknesses:
      Doesn't always exist (Buyer-Seller, Public Good).
      Seems that mechanism designer may give up a lot (why not run a first price auction and try for larger revenue?)
- Scenario 2: Want to be with LI above all else, but hope it is at M.
  LI also wants to meet, but prefers P.
  Both players' utilities:



- Is there a dominant strategy?

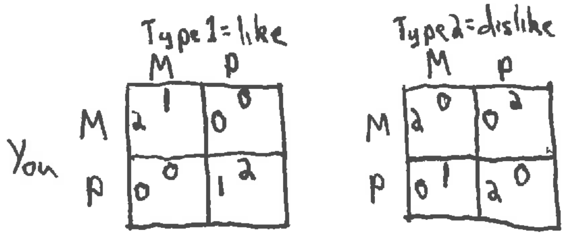
- Must strategize. Your optimal choice depends on what they do, and vice versa. Negotiation through friends ends at either (M,M) or (P,P).
- (M,M) is a *Nash Equilibrium*:
  - Given that LI does M, your utility is maximized by doing M, and vice versa.
  - No one has any incentive to deviate when everyone else follows Nash strategy.
- (P,P) also a Nash equilibrium.

- Scenario 3: LI wants to avoid you.



  - At (M,M), does anyone want to deviate?
  - At (M,P), does anyone want to deviate?
  - At (P,P), does anyone want to deviate?
  - At (P,M), does anyone want to deviate?
  - No Nash Equilibrium.

- Scenario 4: 50/50 whether LI likes you back.
  - LI will be one of two different types, each with probability 0.5.
    LI knows own type and can act based on it.
    You do not know LI's type.
    You have only one possible type, in this example (so everyone knows it).



(oops. bottom right 2 should be a 1)

- A *strategy* is a list of what an individual will do, conditional on their own type. (You will commit to M or P before knowing LI's type.)
  - Ex: Perhaps your strategy is M,
    and LI's strategy is (M,P):
    M if type 1, P if type 2.
  - If you know LI's strategy is (M,P), and you follow M, your expected utility is:
  - Suppose you consider following strategy P instead. Expected utility:
- If you know LI's strategy is (M,P), then your best option in M.
- If LI knows your strategy is M, what should they do?
- (M, (M,P)) is a *Bayesian Nash Equilibrium*:
  - Given that LI does (M,P) (depending on their type), your expected utility is maximized by doing M. Given that you are doing M, LI should do (M,P).
  - No one has any incentive to deviate when everyone else follows Bayesian Nash strategy. When you are considering deviating, you know:
    - Your type.
    - Probabilities that someone else has a given type, but not which type they actually are.
    - How they will act, based on every possible type they might be.
- Is you do P, LI does (P,M) a BNE?

- ◦ Only one BNE: do what you want and hope they show up! (They will if they like you.)
- ◦ BNE is what we will need to analyze first price auctions: you know how much you value the object, but don't know how much others do.

- Three common time points to consider:
  - ◦ *ex ante*: before any types are known. This is the point of view of the auctioneer when setting up the auction; don't yet know the types of any bidders.
  - ◦ in the *interim*: you know your type but no one else's. This is the point of view when you are deciding what to bid.
  - ◦ *ex post*: everything is known. Auction is over.
  - ◦ Dominant strategies have no regrets, ex post, no matter what other players do.
  - ◦ Nash equilibria have no regrets, ex post, assuming other players follow Nash strategy.
  - ◦ Bayesian Nash equilibria have no qualms, in the interim, assuming other players follow Bayesian Nash strategy. May have ex post regret.
    - ▪ Seeing that LI went to P, you regret (ex post) going to M.
    - ▪ In 1st price auction, seeing that next highest bid was $30, you regret (ex post) bidding $50.
    - ▪ But at the time you were making the decision, you couldn't have done better (in expected utility).

22 Probability

- Strategy-proof has gotten us this far.
  - The Good: VCG mechanism, Matching Market mechanism, Women-propose mechanism for the women.
  - The Bad: Gibbard-Satterthwaite Theorem, Buyer-Seller example, Public Goods Example, Women-propose mechanism for the men.
- Now we will talk about strategizing. How should I react to what others will do, e.g., in a 1st price auction?
  - You can only guess about what others will do, because don't know their valuation.
  - Need probability.
- $P(A)$ is the "probability that event A happens".
  - Ex: Rolling die, $P(\text{at least a 4}) = \frac{3}{6} = \frac{1}{2}$.
- If A and B are *disjoint* events (they can't both happen, i.e., $P(A \text{ and } B) = 0$), then
$$P(A \text{ or } B) = P(A) + P(B).$$
  - Ex: $P(1 \text{ or } 2) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$
  - Ex: Does $P(\geq 4 \text{ or odd}) = \frac{3}{6} + \frac{3}{6}$?

- In general,
$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B).$$
  - $P(\geq 4 \text{ or odd}) = \frac{3}{6} + \frac{3}{6} - \frac{1}{6} = \frac{5}{6}$.
- If A and B are *independent* events, then
$$P(A \text{ and } B) = P(A) \cdot P(B).$$
  - Ex: Two dice rolls, $P(\text{both 1's}) = \frac{1}{6} \cdot \frac{1}{6}$.
  - Ex: Does $P(\text{top two cards in a shuffled deck both red})$
$$= \frac{26}{52} \cdot \frac{26}{52}?$$

- In general,
$$P(A \text{ and } B) = P(A) \cdot P(B \big| A),$$
where $P(B \big| A)$ is the *conditional probability* that B occurs, given that A occurs (pronounce $\big|$ as "given" to remember this).
  - $P(\text{2nd red} \big| \text{1st red}) = \frac{25}{51}$, so
$$P(\text{both red}) = \frac{26}{52} \cdot \frac{25}{51}.$$

- If we *fix* A, $P(\cdot \big| A)$ is a probability distribution.
  - $P(\text{2nd red} \mid \text{1st red}) = \frac{25}{51}$ and
    $P(\text{2nd black} \mid \text{1st red}) = \frac{26}{51}$.
- Often useful rearrangement:

$$P(B \big| A) = \frac{P(A \text{ and } B)}{P(A)}.$$

- Independent means $P(B \big| A) = P(B)$.

• Continuous probability:

- Pick a random real number, X, from $[0,1]$ *uniformly* (all choices equally likely).

  $P\left(X = \frac{\sqrt{\pi}}{e}\right) =$

  $P\left(X = \frac{1}{2}\right) =$

  $P\left(X \leq \frac{1}{2}\right) =$

  $P\left(X \leq \frac{\sqrt{\pi}}{e}\right) =$

  For $0 \leq x \leq 1$, $P\left(X \leq x\right) =$

• Notation:

- X is a random variable (takes values in $\mathbb{R}$).
  - Convention: Lower case x is a possible value of X (mentally plug in an actual number for x to help understand).
    So $\forall x, P(X = x) = 0$.
- $F(x) = P(X \leq x)$ is the *cumulative distribution function* (cdf).
  - In general, for F to be a cdf on some interval $[a, b]$ (a and b may be $\mp\infty$), need

    $F(a) =$

    $F(b) =$

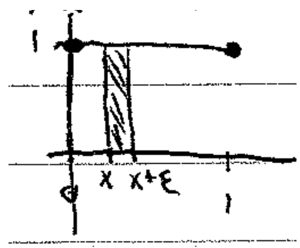
    and, in between, F is
- Let $f(x) = F'(x)$; called the *density function*.
  - Convention: using lower case letter of cdf.

• $P(c \leq X \leq d) = F(d) - F(c) = \int_c^d f(x)\, dx$.

- Probabilities are areas under density curve.

- $$\begin{aligned} P(x \le X \le x+\varepsilon) &= F(x+\varepsilon) - F(x) \\ &= \frac{F(x+\varepsilon) - F(x)}{\varepsilon}\varepsilon \\ &\approx \end{aligned}$$
  - $f(x)$ is not a probability, but it does give relative likelihood of ending up "near" x.
- Ex: $F(x) = P(X \le x) = x$ on $[0,1]$ (uniform)
  $f(x) = 1$, for $0 \le x \le 1$.

  Graph of $f(x)$:



- To be a density function on $[a,b]$, need $f(x) \ge 0$ and $\int_a^b f(x)\,dx = 1$.
- Density functions are not probabilities, so be careful! Generally, start with cdf, because that is actually a probability.
- Example: Let Θ_1 and Θ_2 be i.i.d. (independent and identically distributed), uniformly on $[0,1]$.
  - For example, two independent valuations of a cookie.
  - Let $X = \max\{\Theta_1, \Theta_2\}$.
    Compute $F(x)$ and $f(x)$, cdf and density functions for X:
    - $P(\frac{1}{2} \le X \le 1) =$

23 Expected Value

- If we do something (like roll a die) a bunch of times, what will the outcome be, on average?
  - Ex: Average roll of a die =
$$\frac{1}{6}\cdot 1+\frac{1}{6}\cdot 2+\frac{1}{6}\cdot 3+\frac{1}{6}\cdot 4+\frac{1}{6}\cdot 5+\frac{1}{6}\cdot 6=3.5.$$
- Definition: If X is a discrete random variable, then $E[X]$, the *expected value* of X, is
$$\sum_{\text{outcomes } x} P(X=x)\cdot x.$$
  - Sum of probabilities times values.
  - When you see the word "expected", think "average".
- If h is a function,
$$E\big[h(X)\big]=\sum_{x} P(X=x)\cdot h(x).$$
  Sum of probabilities times values.
  - Example: E[die roll, doubled] is
$$\frac{1}{6}\cdot 2+\frac{1}{6}\cdot 4+\frac{1}{6}\cdot 6+\frac{1}{6}\cdot 8+\frac{1}{6}\cdot 10+\frac{1}{6}\cdot 12=7.$$
  - Example: E[die roll, squared] is
$$\frac{1}{6}\cdot 1+\frac{1}{6}\cdot 4+\frac{1}{6}\cdot 9+\frac{1}{6}\cdot 16+\frac{1}{6}\cdot 25+\frac{1}{6}\cdot 36=15.1\overline{6}.$$
- Note: $E[kX]=kE[X]$, but,
  in general, $E\big[h(X)\big]\neq h\big(E[X]\big)$.
  - $3.5^2\neq 15.1\overline{6}$.

- Cool Fact (take MATH 335 to prove it!):
  $E[X+Y] = E[X] + E[Y]$, even if X and Y are not independent.
  - $E[\text{sum of 2 die rolls}] = 3.5 + 3.5 = 7$. Much easier than
  $$\frac{1}{36} \cdot 2 + \frac{2}{36} \cdot 3 + \cdots + \frac{1}{36} \cdot 12 = 7.$$
- Continuous Probability:
  X a random variable on $[a,b]$.
  $F(x) = P(X \leq x)$ is cdf.
  $f(x) = F'(x)$ is density function.
  $P(c \leq X \leq d) = F(c) - F(d) = \int_c^d f(x)\,dx$.
  $P(x \leq X \leq x+\varepsilon) \approx f(x) \cdot \varepsilon$.
- Divide up all of $[a,b]$ into little intervals $x_0 = a$, $x_1 = a+\Delta x$, $x_2 = a+2\Delta x$, $\ldots x_n = b$. Then

  $$E[X] \approx \sum_i P(x_i \leq X \leq x_{i+1}) \cdot x_i$$

  $\approx$ [ ]

  [ ]

  [ ]

  [ ]
- Definition: $E[X] = \int_a^b f(x) \cdot x\,dx$.
  - "summing" the "probability" times the value, where "summing" is integrating, "probability" is $f(x)\,dx$, and value is x.
  - $E\big[h(X)\big] = \int_a^b f(x)h(x)\,dx$.
- Example: X uniform $[0,1]$, $F(x) = x$, $f(x) = 1$.

  $E[X] =$ [ ]
- Example: Θ_1, Θ_2 i.i.d. uniform $[0,1]$.
  $X = \max\{\Theta_1, \Theta_2\}$. $F(x) = x^2$, $f(x) = 2x$.

  $E[X] =$ [ ]
- Example: Suppose your valuation for a cookie is \$0.80 and you bid \$0.60 in a first price auction.
  Suppose their is one other bidder, whose bid B_2 you know will be uniformly distributed

$[0, 1]$.

- Your net utility, $u(B_2)$, is a function of B_2:

$$u(B_2) = \begin{cases} 0.8 - 0.6 = 0.2 & \text{if } B_2 < 0.6, \\ 0 & \text{if } B_2 > 0.6. \end{cases}$$

$$\begin{aligned} E[\text{util}] &= \int_0^1 1 \cdot u(b_2)\, db_2 \\ &= \int_0^{0.6} u(b_2)\, db_2 + \int_{0.6}^1 u(b_2)\, db_2 \\ &= \int_0^{0.6} 0.2\, db_2 + \int_{0.6}^1 0\, db_2 \\ &= 0.6 \cdot 0.2 + 0.4 \cdot 0 = 0.12. \end{aligned}$$

- Or, your net utility is a function of the discrete variable whose values are win and $lose$.

$$\begin{aligned} E[\text{util}] &= P(\text{win})u(\text{win}) + P(\text{lose})u(\text{lose}) \\ &= P(B_2 < 0.6) \cdot 0.2 + P(B_2 > 0.6) \cdot 0 \\ &= 0.6 \cdot 0.2 + 0.4 \cdot 0 = 0.12. \end{aligned}$$

- Second way is quicker, but it only worked because net utility depended *only* on whether you won or lost.
- Wouldn't work in 2nd price auction, for example, because net utility depends directly on B_2:

$$u(B_2) = \begin{cases} 0.8 - B_2 & \text{if } B_2 < 0.6, \\ 0 & \text{if } B_2 > 0.6. \end{cases}$$

$$\begin{aligned} E[\text{util}] &= \int_0^1 1 \cdot u(b_2)\, db_2 \\ &= \int_0^{0.6} u(b_2)\, db_2 + \int_{0.6}^1 u(b_2)\, db_2 \\ &= \int_0^{0.6} 0.8 - b_2\, db_2 + \int_{0.6}^1 0\, db_2 \\ &= 0.8b_2 - 0.5b_2^2\Big|_0^{0.6} + 0 \\ &= 0.3. \end{aligned}$$

24 First Price Auction, $n = 2$, uniform distribution

- Assumption: People only care about expected value of utility, as measured in dollars.
  - Indifferetn between \$10 and 50-50 shot at \$20.
  - People are risk neutral. Should be true for small dollar amounts.
- 2 agents bidding on a cookie.
  - Let Θ_1 be your type, Θ_2 other agent's type, i.i.d. (independent and identically distributed) uniform $[0, 1]$.
  - You will know Θ_1 when you bid, but not Θ_2 (in the interim).
  - First price auction. Highest bid wins and pays their bid.
- Suppose you know Agent 2 will bid exactly Θ_2, whatever that is.
  - Knowing $\Theta_1 = \theta_1$, what should you bid?
  - Say you bid b.
    - $0 \leq b \leq 1$ (stupid to bid $>$ 1, because Agent 2 will bid at most 1).
    - Let's calculate $U(b) = E[\text{net utility when bid } b]$ (the long way, which always works:)
    - Maximize $U(b)$ over $0 \leq b \leq 1$:
- But Agent 2 won't bid Θ_2. What if they bid $\Theta_2/2$?
  - Say you bid b. What is the optimal b?
    - Stupid to bid $b > \frac{1}{2}$:

- P(you win) =

  $U(b) = E$[net utility when bid b]

  =

- Maximize $U(b)$ over $0 \leq b \leq \frac{1}{2}$:

- Bidding $B_1 = \Theta_1/2$, $B_2 = \Theta_2/2$ is a Bayesian Nash Equilibrium!
  - Given that you know $\Theta_1 = \theta_1$ and that agent 2 will follow strategy $B_2 = \Theta_2/2$, bidding $b_1 = \theta_1/2$ is (interim) optimal, in expectation.
    - Of course, you will probably have ex post regrets... maybe could have won with a lower bid, or maybe could have increased bid to change from losing to winning.
  - By symmetry, Agent 2 is also following optimal strategy, given that you are bidding $B_1 = \Theta_1/2$.
- The seller creates the mechanism. Wants good E[revenue] = E[max bid].
  - This is ex ante: before either Θ_1 or Θ_2 known to seller.
  - Assuming y'all follow the equilibrium strategy:
    - $E[\max\{\Theta_1, \Theta_2\}] = \frac{2}{3}$, from p:111.
    - E[max bid] =
- Is this better for the seller than a second price auction?
  - Assume y'all follow the dominant strategy: $b_1 = \Theta_1$, $b_2 = \Theta_2$, in second price auction.
  - Let X be the revenue, the second highest of Θ_1 and Θ_2.

- c.d.f for X:

$$\begin{aligned} F(x) &= P(X \leq x) \\ &= 1 - P(X > x) \\ &= 1 - P(\Theta_1 > x, \Theta_2 > x) \\ &= 1 - P(\Theta_1 > x)P(\Theta_2 > x) \\ &= 1 - (1-x)(1-x) = 2x - x^2. \end{aligned}$$

- Density function:

$$f(x) = F'(x) = 2 - 2x.$$

- Expected revenue:

$E[X] =$

25 Risk Aversion

- Setup from last time: Let Θ_1 be your type, Θ_2 other agent's type, i.i.d. uniform $[0,1]$.
- First price auction: Bayesian Nash equilibrium is that you both bid according to $b(\Theta_i) = \Theta_i/2$. We showed auctioneer's (ex ante) expected revenue is $1/3$. Another way to see that:
  - In the interim, you know $\Theta_1 = \theta_1$. Assuming that you both follow the BNE strategy:
    - $P(\text{you win}) = P(\Theta_2/2 < \theta_1/2) = \theta_1$.
    - Your interim expected payment, $m(\theta_1)$, is
  - Then your ex ante expected payment is

    $E[m(\Theta_1)] =$

  - Agent 2's ex ante expected payment is same. So auctioneer's ex ante expected revenue is

$$E[\text{agent 1's payment}] + E[\text{agent 2's payment}] = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}.$$

- Now let's assume both agents are *risk averse*.
  - A common way to simulate that:
    use net utility $= \sqrt{\text{surplus}}$, where
    surplus = valuation - price paid.
    - A \$40 surplus (e.g., value cookie at \$90 but only pay \$50) is only worth twice as much to you as a \$10 surplus.
    - You would bet at most \$10 for a 50/50 shot at \$40.
    - Other concave down functions, like utility=log(surplus), are also commonly used.
- For second price auction, risk aversion doesn't change anything. Dominant strategy is always the best option, period, no regrets. Seller's (ex ante) expected revenue is still $1/3$.
  - You are always maximizing surplus (ex post), so you are always maximizing $\sqrt{\text{surplus}}$.
- What if seller tries a first price auction with these risk averse agents (with rest of setup the same)?
  - Might get a different answer than last time, because maximizing $E[\text{surplus}]$ is not the

same as maximizing $E[\sqrt{\text{surplus}}] \neq \sqrt{E[\text{surplus}]}$.

- Assume we've found a *symmetric* Bayesian Nash Equilibrium (both use same bidding function).
  - There is a $b(\cdot)$ so that Agent i bids $b(\Theta_i)$.
  - You know $\Theta_1 = \theta_1$. You know Agent 2 will bid $b(\Theta_2)$, whatever that is.
  - Since it is a BNE, the bid that maximizes your net utility must be $b(\theta_1)$.
- Can we figure out what $b(\cdot)$ is, from this information? Our goal will be to find a differential equation for $b(\cdot)$, and then solve it.
  - Suppose you are thinking of bidding b^*, for some b^*.
    - $b(0) \leq b^* \leq b(1)$ (bidding over $b(1)$ is stupid).
    - A neat trick: this means $b^* = b(\psi)$, for some ψ. You are thinking of bidding *as if* your type were ψ (but otherwise following BNE).
    - Since we have a BNE, the optimal ψ must be θ_1, so that you bid $b^* = b(\theta_1)$.
  - The probability that you win is:

  - Expected net utility is:

  - Maximized at the critical point (also check $\psi = 0$ and $\psi = 1$):

$$0 = \frac{d}{d\psi}\left(\psi\sqrt{\theta_1 - b(\psi)}\right)$$

$$0 = 1 \cdot \sqrt{\theta_1 - b(\psi)} \ + \ \psi \cdot \frac{1}{2\sqrt{\theta_1 - b(\psi)}} \cdot \big(-b'(\psi)\big)$$

$$0 = 2\big(\theta_1 - b(\psi)\big) - \psi \cdot b'(\psi)$$

  - Since we know we are at a BNE, the solution to this must be $\psi = \theta_1$, so

$$0 = 2\big(\theta_1 - b(\theta_1)\big) - \theta_1 \cdot b'(\theta_1) \qquad \text{(A differential equation!)}$$

$$2\theta_1 = 2b(\theta_1) + \theta_1 b'(\theta_1)$$

$$2\theta_1^2 = 2\theta_1 b(\theta_1) + \theta_1^2 b'(\theta_1)$$

  - Why did I do that?

- Change variable to t. Integrate $2t^2$ from 0 to θ_1:

- Magic? Risk aversion causes you to fear missing out on the tasty cookie! So you bid 33% more than before. And gives seller 33% more profits!

26 Buyer-Seller Example

- One buyer, one seller, with respective valuations Θ_1 and Θ_2.
  - If $\Theta_1 > \Theta_2$, there are gains to be made by selling.
    - If price is p, then:
      Buyer Surplus $= \Theta_1 - p$,
      Seller Surplus $= p - \Theta_2$, and
      Total Surplus $= \Theta_1 - \Theta_2$.
      If no sale, total surplus $= 0$.
- Suppose Θ_1 and Θ_2 are i.i.d., uniform $[0, 1]$.
- Suppose we could get buyer and seller to reveal true type, so sale will always happen if $\Theta_1 > \Theta_2$ (this would be an efficient mechanism). What is the expected (ex ante) total surplus?
- First compute, for fixed $\Theta_1 = \theta_1$, interim expected total surplus:
  $\int_0^1 f(\theta_2)\cdot$ surplus $d\theta_2 =$

  (In WolframAlpha.com, you can type "Integrate t1 - t2 with respect to t2 from 0 to t2", and feel free to do this on the next problem set.)
- Then the ex ante expected surplus is

$$\int_0^1 f(\theta_1)(\text{Exp. surplus, given } \theta_1)\, d\theta_1 =$$

- Alas, no strategy-proof mechanism is efficient (and also individually rational and feasible).
  - Can show that there is also no mechanism with a BNE that is efficient and feasible, while being either ex post or interim individually rational. (Latter means, if know your type but not others, expected net utility is nonnegative).

- There are solutions that are ex ante individually rational (before you know your own type), but these don't seem that useful.
- What if we give up on efficiency? Do we lose much of the possible gains of trade?

- Mechanism: Buyer and seller reveal types θ_1 and θ_2 (or lie). If $\theta_1 > 0.5$ and $\theta_2 < 0.5$, they trade for price $p = 0.5$.
  - Strategy-proof: can't affect price, and definitely worth buying/selling at that price.
  - It is as if the market has established a "take it or leave it" price of $0.50, for both buyer and seller. Not efficient though.
  - Given that $\Theta_1 = \theta_1$.
    If $\theta_1 \leq 0.5$, interim expected surplus is 0.
    If $\theta_1 > 0.5$, interim expected surplus is:
  - So ex ante expected surplus is:
- Less than the ideal of $1/6$, but not too shabby.
- What other mechanisms might you try?

27 Buyer-Seller Example, Part 2

- Buyer and seller have respective valuations Θ_1 and Θ_2, i.i.d. uniform $[0, 1]$.
  - Ideal (both reveal true type without strategizing) has ex ante expected total surplus $1/6$.
  - Strategy-proof mechanism where only price allowed is .5 has ex ante expected total surplus $1/8$. By symmetry, buyer and seller have ex ante individual surplus $1/16$.
- What if the seller sets a price p?
  - Buyer has no incentive to lie: either buy it at p (if $\Theta_1 > p$) or don't.
  - Seller has an incentive to "lie": set $p > \Theta_2$.
  - Maximize seller's interim expected surplus / net utility (i.e., we're assuming risk neutral):
    - Seller knows $\Theta_2 = \theta_2$.
$$\begin{aligned} P(\text{sell}) &= P(\Theta_1 > p) = 1 - P(\Theta_1 \leq p) = 1 - p. \\ E[\text{utility}] &= P(\text{sell})u(\text{sell}) + P(\text{no sell})u(\text{sell}) \\ &= (1-p)(p-\theta_2) + p \cdot 0 \\ &= -p^2 + (1+\theta_2)p - \theta_2. \end{aligned}$$
    - Downward facing parabola with maximum at the critical point
$$-2p + (1+\theta_2) = 0, \text{ i.e., } \quad p = \frac{1}{2}\theta_2 + \frac{1}{2}.$$
    - Seller scales valuation halfway up to 1.
  - Seller's interim expected surplus (substitute $p = \frac{1}{2}\theta_2 + \frac{1}{2}$):
$$-\left(\frac{1}{2}\theta_2 + \frac{1}{2}\right)^2 + (1+\theta_2)\left(\frac{1}{2}\theta_2 + \frac{1}{2}\right) - \theta_2 = \frac{1}{4}(1-\theta_2)^2.$$
  - Seller's ex ante expected surplus:
$$\int_0^1 1 \cdot \frac{1}{4}(1-\theta_2)^2 \, d\theta_2 = 1/12.$$
    - This is also seller's "fair share" of the ideal total surplus of $1/6$.
  - Buyer's interim expected surplus, knowing $\Theta_1 = \theta_1$ and $p = \frac{1}{2}\Theta_2 + \frac{1}{2}$, but not Θ_2:
    - Will buy if
$$\theta_1 > \frac{1}{2}\Theta_2 + \frac{1}{2}, \text{ i.e., } \Theta_2 < 2\theta_1 - 1.$$
    - **Careful!** If $\theta_1 \leq 1/2$, never buys, and expected surplus is 0.

- If $\theta_1 > 1/2$, interim expected buyer's surplus:

$$\int_0^1 f(\theta_2) \cdot \text{surplus}\, d\theta_2$$
$$= \int_0^{2\theta_1 - 1} 1 \cdot \left(\theta_1 - \left(\frac{1}{2}\theta_2 + \frac{1}{2}\right)\right) d\theta_2 + \int_{2\theta_1 - 1}^1 1 \cdot 0 \, d\theta_2$$
$$= \theta_1^2 - \theta_1 + \frac{1}{4}.$$

- Buyer's ex ante expected surplus:

$$\int_0^{1/2} 1 \cdot 0 \, d\theta_1 + \int_{1/2}^1 1 \cdot \theta_1^2 - \theta_1 + \frac{1}{4} \, d\theta_1 = 1/24.$$

- Bad for buyer. Ex ante expected total surplus is $\frac{1}{12} + \frac{1}{24} = \frac{1}{8}$, no better than $p = 0.5$ mechanism.

- How about the "split the difference" mechanism: buyer and seller name respective prices b_1 and b_2; if $b_1 > b_2$, then sale happens for price $(b_1 + b_2)/2$.
  - One can show that the following is a Bayesian Nash Equilibrium:

$$b_1(\Theta_1) = \frac{2}{3}\Theta_1 + \frac{1}{12},$$
$$b_2(\Theta_2) = \frac{2}{3}\Theta_2 + \frac{1}{4}.$$

  - Buyer shades bid $1/3$ of the way down to 0.25 (note seller always reports $b_2 \geq 0.25$, so no sale will happen if $\Theta_1 < 0.25$). Seller shades bid $1/3$ of the way up to 0.75.
  - Sale happens when $b_1(\Theta_1) > b_2(\Theta_2)$, i.e.,

$$\Theta_1 - \Theta_2 > 1/4.$$

(cool way to see how this can be inefficient).

  - Can show ex ante expected surplus is $9/128$ for the buyer and for the seller. So full table:

Mechanism	Buyer	Seller	Total surp.
Ideal	?	?	0.167
$p = 0.5$	0.063	0.063	0.125
Seller sets p	0.041	0.083	0.125
Split the diff.	0.070	0.070	0.141

  - Can show that "spit the difference" maximizes expected total surplus, across all mechanisms with Bayesian Nash Equilibria that are feasible and individually rational. And "seller sets p" maximizes expected seller surplus.
- Let's check, in split the difference: if seller bids $\frac{2}{3}\Theta_2 + \frac{1}{4}$, then buyer's optimal bid is $\frac{2}{3}\Theta_1 + \frac{1}{12}$.
  - Suppose buyer is considering bidding b^*.
  - There will be a sale iff

$$b^* > b_2(\Theta_2) = \frac{2}{3}\Theta_2 + \frac{1}{4}, \text{ i.e., } \frac{3}{2}\left(b^* - \frac{1}{4}\right) > \Theta_2.$$

- In which case buyer's utility is

$$\theta_1 - \frac{1}{2}\big(b^* + b_2(\Theta_2)\big) = \theta_1 - \frac{1}{2}\left(b^* + \frac{2}{3}\Theta_2 + \frac{1}{4}\right).$$

- Therefore the buyer's expected utility across all possible Θ_2 is

$$\begin{aligned}&\int_0^1 f(\theta_2)u(\theta_2)\,d\theta_2\\ &= \int_0^{\frac{3}{2}\left(b^*-\frac{1}{4}\right)} 1\cdot\left(\theta_1 - \frac{1}{2}\left(b^* + \frac{2}{3}\theta_2 + \frac{1}{4}\right)\right)\,d\theta_2 + \int_{\frac{3}{2}\left(b^*-\frac{1}{4}\right)}^1 1\cdot 0\,d\theta_2\\ &= -\frac{9\left(b^*\right)^2}{8} + \frac{3b^*\theta_1}{2} + \frac{3b^*}{16} - \frac{3\theta_1}{8} + \frac{3}{128}.\end{aligned}$$

- This is a concave down parabola (as a function of b^*), whose maximum is achieved at the critical point

$$-\frac{9}{4}b^* + \frac{3}{2}\theta_1 + \frac{3}{16} = 0, \text{ i.e., } b^* = \frac{2}{3}\theta_1 + \frac{1}{12}.$$

- That's what we wanted!

28 1st Price Auctions, general distributions

- 1st price auction, n risk neutral bidders (seek to maximize their expected surplus).
  - Assume Θ_i chosen i.i.d., cdf $F(\theta_i)$ and density function $f = F'$, distributed on some $[0, c]$.
- Let's look for a symmetric BNE:
  - Some function b such that agent i bids $b(\Theta_i)$.
  - If agents $2, \ldots, n$ follow this strategy of bidding $b(\Theta_i)$, agent 1 maximizes interim expected utility by bidding $b(\theta_1)$, where $\Theta_1 = \theta_1$.
  - Agent 1 knows $\Theta_1 = \theta_1$, know others will bid $b(\Theta_i)$, but doesn't know $\Theta_2, \ldots, \Theta_n$.
  - We are going to assume that we've found the symmetric BNE, $b(\cdot)$, use properties of BNE to find a differential equation for $b(\cdot)$, and then solve the differential equation.
- We expect $b(\cdot)$ to have the following properties (can prove these must be the case for "nice" distributions):
  - $b(0) = 0$.
  - $b(\theta_i)$ increases with θ_i.
- Given $\Theta_1 = \theta_1$, suppose thinking of bidding b^*. What b^* maximizes interim expected utility?
  - Certainly $0 \leq b^* \leq b(c)$.
  - $b^* = b(\psi)$ for some ψ. Bidding "as if" an agent of type ψ who is plugging into b.
- Let $Y = \max\{\Theta_2, \ldots, \Theta_n\}$, a random variable.
  - Let $G(y)$ be cdf for Y. $G(y) = [F(y)]^{n-1}$.
  - Let $g(y) = G'(y)$ be the density function.
  - You will win if $b(\Theta_i) < b(\psi)$ for $2 \leq i \leq n$.
  - Since b increasing, you will win if $\Theta_i < \psi$ for $2 \leq i \leq n$.
$$P(\text{win}) = P(Y < \psi) = G(\psi).$$
- Interim expected utility when bidding $b(\psi)$ is

  - What ψ maximizes this?

- Since we're assuming we found a BNE, must have $\psi = \theta_1$ be the optimal ψ.

  - Switching to t:

$$g(t)t = \frac{d}{dt}\big[G(t)b(t)\big].$$

  - Integrating between 0 and θ_1 (i.e., we're using that this optimality equation is true for all t up to θ_1):

- We're done!

$$b(\theta_1) = \frac{1}{G(\theta_1)} \int_0^{\theta_1} g(t)t\,dt.$$

- If Θ_i is uniform $[0,1]$, then

$$\begin{aligned} G(y) &= \big(F(y)\big)^{n-1} = y^{n-1}, \\ g(y) &= G'(y) = (n-1)y^{n-2}, \\ b(\theta_1) &= \frac{1}{G(\theta_1)} \int_0^{\theta_1} g(t)t\,dt \\ &= \frac{1}{\theta_1^{n-1}} \int_0^{\theta_1} (n-1)t^{n-2} \cdot t\,dt \\ &= \frac{1}{\theta_1^{n-1}} \left(\frac{n-1}{n} t^n \Big|_0^{\theta_1} \right) \\ &= \frac{n-1}{n}\theta_1. \quad \text{Familiar?} \end{aligned}$$

- Let's play. For fixed θ_1, $P(Y \le t \mid Y \le \theta_1)$ is a (conditional) probability distribution.

- It has some cdf $\hat{G}(t)$, valid on $0 \le t \le \theta_1$:

$$\begin{aligned}\hat{G}(t) &= P(Y \le t \bigm| Y \le \theta_1) \\ &= \frac{P(Y \le t \text{ and } Y \le \theta_1)}{P(Y \le \theta_1)} \\ &= \frac{P(Y \le t)}{P(Y \le \theta_1)} \\ &= \frac{G(t)}{G(\theta_1)}.\end{aligned}$$

- Density function: $\hat{g}(t) = \hat{G}'(t) = \frac{g(t)}{G(\theta_1)}$.
- So,

$$\begin{aligned}b(\theta_1) &= \frac{1}{G(\theta_1)} \int_0^{\theta_1} g(t)t\,dt \\ &= \int_0^{\theta_1} \frac{g(t)}{G(\theta_1)} t\,dt \\ &= \int_0^{\theta_1} \hat{g}(t)t\,dt \\ &= E[Y \bigm| Y \le \theta_1] \\ &= E[\text{second highest valuation} \mid \text{I'm the highest}]\end{aligned}$$

- You bid exactly what you would expect to have to pay if:
  This were a second price auction instead,
  Everybody bid their dominant strategy, Θ_i, and
  You win the second price auction.
- Sounds like revenue from 1st and 2nd price auctions end up being the same!

29 The Revenue Equivalence Theorem

- The Revenue Equivalence Theorem: Suppose we have any auction system where
  1. the *highest bid wins*,
  2. payment can be any function of the set of bids, except that a bidder who *bids nothing will pay nothing*,
  3. the buyers are risk neutral, and
  4. Θ_i are i.i.d. with cdf F on $[0, c]$.

  Let $b(\cdot)$ be a symmetric Bayesian Nash Equilibrium. Then the ex ante expected revenue is the same as in the first price auction.
  - Examples: 2nd price auction. Or "all pay" auction, where everyone pays their bid, but only winner gets object (penny auctions online, or – loosely – advertising in a senate campaign).
  - Takeaway: Why do a complicated auction when 2nd price auction gives the same revenue... and buyers will have no need to strategize!
- Non-examples
  - Reserve price (highest bid may not win). For fixed reserve price, there is still revenue equivalence (see last problem set).
  - Bidders pay fixed cost to be able to enter (so bidders who bid nothing might pay something). Turns out to be equivalent to a reserve price.
  - Buyers risk averse. 1st price may generate more revenue.
  - If object has "true" value (e.g., resellable work of art) and buyers are uncertain about it (violates i.i.d.), 2nd price auction may generate more revenue.
    - If you bid highest, you probably overestimated value (WInner's Curse). Knowing this, buyers will scale back bids, especially in 1st price.
- Proof:
  - Assume others bid $b(\Theta_i)$, and you are considering bidding $b(\psi)$.
    - By the definition of BNE, your optimal ψ will be $\psi = \theta_1$.
  - $Y = \max\{\Theta_2, \ldots, \Theta_n\}$ is random variable.
    $G(y) = [F(y)]^{n-1}$ is cdf for Y.
    $g(y) = G'(y)$ is density function.
    One can prove $b(\cdot)$ must be increasing.
    We compute

$$\begin{aligned} P(\text{win}) &= P\big(b(\Theta_i) < b(\psi),\ \text{for } 2 \le i \le n\big) && \text{(since highest bid wins)} \\ &= P(\Theta_i < \psi,\ \text{for } 2 \le i \le n) && (b(\cdot)\text{ increasing}) \\ &= P(Y < \psi) \\ &= G(\psi). \end{aligned}$$

    (doesn't depend on auction type!)

- Define $m(\psi)$ to be your interim expected payment. We will derive a formula for $m(\cdot)$.
  - Note: for first price auction, we saw

$$\begin{aligned} m(\psi) &= P(\text{win}) \cdot b(\psi) + P(\text{lose}) \cdot 0 \\ &= G(\psi) \cdot \frac{1}{G(\psi)} \int_0^{\psi} g(t)t\,dt \\ &= \int_0^{\psi} g(t)t\,dt. \end{aligned}$$

- Interim expected utility is

  - $m(\psi)$ summarizes everything you need to know about a possibly complicated auction payment system, when you're deciding whether to bid $b(\psi)$.

- Maximized at

- Since BNE, maximized at $\psi = \theta_1$:

$$\begin{aligned} 0 &= g(\theta_1)\theta_1 - m'(\theta_1) \\ \Rightarrow\ m'(t) &= g(t)t \text{ for all } t. \end{aligned}$$

- So

$$\begin{aligned} m(\theta_1) &= m(\theta_1) - m(0) && \text{(bidding 0 pays 0)} \\ &= \int_0^{\theta_1} m'(t)\,dt = \int_0^{\theta_1} g(t)t\,dt. \end{aligned}$$

- Your (interim) expected payment does not depend on the auction! (Note that it agrees with what we derived for the 1st price auction).
- Your (agent 1's) ex ante expected payment:

- Seller's ex ante expected revenue:

- This is a formula that doesn't depend on the auction type!

- Example: uniform $[0,1]$, so $F(\theta) = \theta$,
  $G(t) = [F(t)]^{n-1} = t^{n-1}$, and
  $g(t) = G'(t) = (n-1)t^{n-2}$.

30 Apportionment

- Article 1, Section 2 of US Constitution: "Representatives ... shall be apportioned among the several States ... according to their respective Numbers [revisited every 10 years after a census]. Each State shall have at least one representative."
- Sounds easy.
- Example: Total population: $p = 10$ million.
  Number of seats in House: $h = 10$.
  Fair district size: $d_s = {}^{p}/_{h} = 1$ million
  (called the "standard divisor").
  State populations: $p_1, \ldots, p_n$, with $p_1 + \cdots + p_n = p$.
  Fair quota: $q_i = {}^{p_i}/_{d_s}$.
  Actual allocation: nonnegative integers, a_i;
  (Constitution requires *positive* integers,
  but other situations might not).
  Want $a_i \approx q_i$.

State	Population	Fair Quota	Allocation
1	$p_1 = 1.45$ mill	$q_1 = 1.45$	$a_1 =$
2	$p_2 = 3.40$ mill	$q_2 = 3.40$	$a_2 =$
3	$p_3 = 5.15$ mill	$q_3 = 5.15$	$a_3 =$
	$p = 10.00$ mill	$h = 10.00$	$h = 10$

- What should we do?
- Hamilton Method (proposed but not implemented for 1st apportionment in 1791):
  - Order states in decreasing order of the fractional parts of q_i.
  - Find k such that $a_i = \lceil q_i \rceil$ for $1 \le i \le k$ and $a_i = \lfloor q_i \rfloor$ for $k < i \le n$ gives exactly $a_1 + \cdots + a_n = h$.
  - In above example:
- Looks good. Suppose we add a seat, so $h = 11$.
  Fair district size $= d_s = {}^{10 \text{ million}}/_{11} = 0.909$ million.

State	Population	Fair Quota	Allocation
1	1.45 mill	1.60	
2	3.40 mill	3.74	
3	5.15 mill	5.67	
	10.00 mill	11.00	11

- Uh-oh. Alabama paradox (would have hurt Alabama in 1880's, if used).
- Notation: h is the old House size, p_i the old populations, a_i the old result of the method, and h', p'_i, a'_i are the new numbers/results.

- Definition: An allocation method is *house monotone* if whenever $h' > h$ and all $p'_i = p_i$, then

$$a'_i \geq a_i, \text{ for all } i.$$

  - Hamilton method is not house monotone.

- What if we increase seats to accommodate a new state?

State	Population	Fair Quota	Allocation
1	1.45 mill	1.50	
2	3.40 mill	3.51	
3	5.15 mill	5.31	
4	2.60 mill	2.68	
	12.60 mill	13.00	13

- Uh-oh. Oklahoma Paradox (Oklahoma's becoming a state in 1907 would have hurt New York, if used).

- Definition: An allocation method is *state monotone* if, whenever a new state $n+1$ is added such that $h' > h$ and $a'_{n+1} = h' - h$, then

$$a'_i = a_i \text{ for } 1 \leq i \leq n.$$

  - Technicality: we want an allocation method to be defined for a fixed set of states. So we model "adding a new state" as assuming $p_{n+1} = 0$ yields $a_{n+1} = 0$. Then raising p_{n+1} above zero corresponds to adding a state.

- What if populations simply change?

State	Population	Fair Quota	Allocation
1	1.47 mill	1.55	
2	3.38 mill	3.56	
3	4.65 mill	4.89	
	9.50 mill	10.00	10

- Uh-oh.

- Definition: An allocation is *population monotone* if it is always true that:

$$\text{if } a'_i < a_i \text{ and } a'_j > a_j, \text{ then}$$
$$\text{either } p'_i < p_i \text{ or } p'_j > p_j.$$

  - In words?

- This is a fairly strong property.

- Theorem: If an allocation method is population monotone, then it is house monotone and state monotone.
  - Prove the contrapositives:
    - Not house monotone $\Rightarrow$
    - Not state monotone $\Rightarrow$.
- Definition: An allocation method is *order preserving* if

$$p_i < p_j \Rightarrow a_i \leq a_j.$$

- Definition: An allocation method satisfies the *quota property* if it always has either

$$a_i = \lfloor q_i \rfloor \text{ or } \lceil q_i \rceil .$$

  - Hamilton rule is order preserving and satisfies the quota property.
- Theorem: No order preserving allocation method satisfies both quota property and population monotone.
  - This is an impossibility theorem, like Arrow's Theorem for voting, or the lack of good mechanisms for Buyer-Seller and for Public Good.
  - Proof: This example will fail:

State	Population	Fair Quota	Allocation
1	69,900	6.99	
2	5,200	0.52	
3	5,000	0.50	
4	19,900	1.99	
	100,000	10.00	10

    - By quota property, a_2 is either 0 or 1.
      But it can't be 0:

State	Population	Fair Quota	Allocation
1	68,000	8.02	
2	5,500	0.65	
3	5,600	0.66	
4	5,700	0.67	
	84,800	10.00	10

- By quota property, a'_2 is either 0 or 1.
  But it can't be 1:

- Contradicting population monotone:

31 Divisor methods for apportionment

- Hamilton's method starts with a fixed house size, h, and computes $d_s = p/h$, a fair district size called the *standard divisor*.
  - The non-integral fair quotas are $q_i = p_i/d_s$. Call these the *standard* quotas.
  - Hamilton's method rounds the q_i's with the highest fractional parts up (and the rest down) so as to get a House of size *exactly* h.
- Back in the day, the size of the House was of less concern. (It has been $h = 435$ since 1930.)
- Jefferson proposed starting with a divisor d (and not with h, which would define d_s), then computes $q_i = p_i/d$, allocations by $a_i = \lfloor q_i \rfloor$, and finally lets $h = a_1 + \cdots + a_n$.
- Example:

State	Population	$d = 1.1$ mill q_i	a_i	$d = 1$ mill q_i	a_i	$d = 0.91$ mill q_i	a_i
1	1.8 mill	1.65	1	1.8	1	1.98	1
2	8.2 mill	7.45	7	8.2	8	9.02	9
	$p = 10$ mill		$h = 8$		$h = 9$		$h = 10$

- *Jefferson's method*, as now defined, isn't quite this. It does use an h set in advance.
  - For example, if $h = 10$:
    - First try using $d = d_s = {}^{10\text{ mill}}/10 = 1$ mill and $a_i = \lfloor q_i \rfloor$. But this only creates 9 seats.
    - Slowly decrease d (increasing q_i) until a new seat is added (around $d = 0.91$ mill), yielding the desired $h = 10$.
  - Note that the more "natural" Hamilton method would have used standard divisor $d_s = 1$ mill, $a_1 = 2$, $a_2 = 8$. So there is a difference.
  - Jeferson's method favors large states.
    - Coincidentally, Jefferson was from Virginia, which was large in the 1790's.
    - Large states have a lot of power and (his original form) was used 1790–1830.
- The *Adams method* (after John Quincy): Do the same thing, but round *up*, $a_i = \lceil q_i \rceil$.
  - Example: $h = 10$, $p = 10$ mill, standard divisor $= 1$ mill.

State	Population	$d = 1$ mill q_i	a_i	$d = 1.11$ mill q_i	a_i
1	1.2 mill	1.2	2	1.08	2
2	8.8 mill	8.8	9	7.93	8
	$p = 10$ mill		$h = 11$		$h = 10$

- So $a_1 = 2$, $a_2 = 8$.
- Favors small states. Coincidentally, Massachusetts was small back then. Never used.

- It became clear that Jefferson's method favored large states by too much (could easily get several seats more than their standard quota).
  - Daniel Webster, a mastermind of the Compromise of 1850 (delineating where slavery would be allowed in new territories, in an effort to prevent secession), had another compromise.
  - *Webster's method*: don't round up or down, just round:
    - If the fractional part of q_i is bigger than 0.5, round up, else round down. Adjust d to get the right h.
- Webster's and Hamilton's method were both popular 1840–1900, with no clear favorite. Several times, Congress specifically looked for an h that would made these two methods agree, and went with that.
  - Then it became clear how often Hamilton's method suffers from Alabama/Oklahoma/population change paradoxes (it's not population/state/house monotone).
  - Webster's method was the clear favorite from 1900 to 1930. It does not have these problems.
- Theorem: These divisor methods (Jefferson, Adams, Webster, plus two more we will see soon) are population monotone (and therefore state and house monotone).
- Proof: p_i, a_i, d old populations/allocations/divisor. p_i', a_i', d' new.
  - Suppose $a_i' < a_i$ and $a_j' > a_j$ (hypothesis of population monotone).
  - Need to show $p_i' < p_i$ or $p_j' > p_j$.
  - Since $a_i' < a_i$,
$$\frac{p_i'}{d'} < \frac{p_i}{d}, \quad \text{i.e.,} \quad p_i' < \frac{d'}{d} p_i$$
    (all of these methods round p_i/d in some consistent way, to get the allocation).
  - Since $a_j' > a_j$,
$$\frac{p_j'}{d'} > \frac{p_j}{d}, \quad \text{i.e.,} \quad p_j' > \frac{d'}{d} p_j.$$
  - If $\frac{d'}{d} \leq 1,$ then $p_i' < \frac{d'}{d} p_i \leq p_i,$ and
    If $\frac{d'}{d} > 1,$ then $p_j' > \frac{d'}{d} p_j > p_j$
  - This is what we wanted. So the method is population monotone.
- Of course, since they are population monotone and order-preserving, they must not satisfy the quota property.
- We would be using Webster's method today if it weren't for
  a Census department bureaucrat,
  a Harvard mathematician,

the National Academy of Sciences,
And the Great State of Arkansas.

- Joseph Hill worked for the census department in the 1910's.
  - He noticed that, if $q_i \in [m, m+1)$, then Webster's method uses the cutoff $m + 1/2$ to decide whether to round up or down.
  - $m + 1/2$ is the arithmetic mean of m and $m+1$.
  - *Hill's method*: use the geometric mean instead. cutoff $= \sqrt{m(m+1)}$.
- For completeness, *Dean's method* (1830's, U Vermont mathematician), use the harmonic mean instead; cutoff:

$$\frac{2}{1/m + 1/m+1} = \frac{2m(m+1)}{2m+1}.$$

- Cutoffs:

Range	Adams	Dean	Hill	Webster	Jefferson
$[0,1)$	0	0	0	0.5	1
$[1,2)$	1	1.33	1.41	1.5	2
$[2,3)$	2	2.40	2.45	2.5	3
$[3,4)$	3	3.43	3.46	3.5	4
$[10,11)$	10	10.48	10.49	10.5	11
$[50,51)$	50	50.495	50.498	50.5	51

  - Dean and Hill cutoffs increase with m, approaching $m + 1/2$.
  - min $\leq$ harmonic mean $\leq$ geometric mean $\leq$ arithmetic mean $\leq$ max.
  - So listed in order of favoring small states to favoring large states.
  - Adams/Dean/Hill all have $[0,1)$ cutoff 0, ensuring all states get at least 1 seat. For other methods, this is a separate step.
- 1941:
  - Harvard mathematician, Edward Huntington, revived Hill's method and advocated for it.
  - Congress was tired of fighting about it, and they asked the National Academy of Sciences what it thought.
  - The Academy favored Hill, based on the symmetry of this set of options (plus another symmetry we will see next time). They thought it was the best balance between large state and small state concerns.
  - Though it must sometimes fail the quota property, it is very rare, and it didn't fail on any historical data they looked at. (Same is true for Webster's, but Adams and Jefferson often have major quota violations).
  - Congress listened to the mathematicians!

  - It was surely a coincidence that the Hill method gave the Democratic-leaning state of Arkansas an extra seat, and Congress was controlled by the Democrats.

- Hill's method has been used ever since.

  - Montana once sued Congress when they thought they deserved 2 seats instead of the 1 that Hill's method gave them. The Supreme Court upheld the status quo.

- Some examples from 2010 census:

State	Pop (mill)	Std q_i	Adams	Dean	Hamilt.	Hill	Webst.	Jeff.
CA	37.3	52.42	**50**	52	52	53	53	**55**
TX	25.1	35.32	**34**	35	35	**35**	**36**	**37**
MN	5.30	7.46	8	7	**8**	7	7	7
RI	1.06	1.49	2	2	2	**2**	**1**	1
MT	0.98	1.26	2	**2**	1	**1**	1	1

  - Notice the extreme quota violations of Adams and Jefferson methods.
  - Notice the small state / large state continuum. (Hamilton isn't technically on the continuum).
  - Notice Hamilton randomly gives Minnesota (a medium state) an extra seat, above what Hill's method would do. Hamilton's method is slightly erratic, as evidenced by the paradoxes.
  - In Hill vs. Webster (the two considered most reasonable), Hill gives Rhode Island a seat at the expense of Texas.
  - Montana sued... they want the Dean method.

32 Comparing Divisor Methods

- Divisor methods: Fix h. Try a divisor d.
  - Compute $q_i = p_i/d$, the quotas for this given divisor, d (not necessarily the *standard* quota, which requires that d is the standard divisor, $d_s = p/h$).
  - If $q_i \in [m, m+1)$, m an integer:
    let a_i be q_i rounded up to $m+1$ or down to m, depending on whether q_i is $\geq$ or $<$ the cutoff:

Adams	Dean	Hill	Webster	Jefferson
m	$\frac{2m(m+1)}{2m+1}$	$\sqrt{m(m+1)}$	$m+\frac{1}{2}$	$m+1$

  - If $a_1 + \cdots + a_n \neq h$, adjust d and repeat, as necessary.
- Analyzing Webster's Method:
  - If q_i got rounded to a_i, that means:

$$\begin{aligned} & a_i - \frac{1}{2} \leq q_i < a_i + \frac{1}{2} \\ \Rightarrow \quad & a_i - \frac{1}{2} \leq \frac{p_i}{d} < a_i + \frac{1}{2} \\ \Rightarrow \quad & d > \frac{p_i}{a_i + 1/2} \quad \text{and} \quad d \leq \frac{p_i}{a_i - 1/2}. \end{aligned}$$

  - This is true for all i. So for a pair (i, j):

$$\begin{aligned} & \frac{p_j}{a_j + 1/2} < d \leq \frac{p_i}{a_i - 1/2} \\ \Rightarrow \quad & \frac{p_j}{a_j + 1/2} < \frac{p_i}{a_i - 1/2} \\ \Rightarrow \quad & \frac{2a_j + 1}{p_j} > \frac{2a_i - 1}{p_i} \\ \Rightarrow \quad & \frac{a_j}{p_j} + \frac{a_j + 1}{p_j} > \frac{a_i}{p_i} + \frac{a_i - 1}{p_i} \\ \Rightarrow \quad & \frac{a_i}{p_i} - \frac{a_j}{p_j} < \frac{a_j + 1}{p_j} - \frac{a_i - 1}{p_i}. \end{aligned}$$

  - What does $\frac{a_i}{p_i}$ signify?

  - Suppose i is better represented than j, per capita (LHS positive, measures the difference for these two).
    - Then giving a seat from i to j makes j better represented per capita than i (RHS positive).
    - And, in fact, it increases

$$\left| \frac{a_i}{p_i} - \frac{a_j}{p_j} \right|,$$

the (absolute) difference in representation per capita.

- Webster's method: The unique (all of these steps are reversible) method that *locally* minimizes the discrepancy in representation per capita, i.e., there is no way to make any of the
$$\left|\frac{a_i}{p_i} - \frac{a_j}{p_j}\right|,$$
smaller by shifting a seat from i to j.

- How about Dean's method?
  - Similarly, start with
$$\frac{2(a_i-1)a_i}{2(a_i-1)+1} \leq q_i < \frac{2a_i(a_i+1)}{2a_i+1}.$$
Same manipulations yield
$$\frac{p_j}{a_j} - \frac{p_i}{a_i} < \frac{p_i}{a_i-1} - \frac{p_j}{a_j+1}.$$
  - What does $\frac{p_i}{a_i}$ represent?

  - Dean's method is the unique method that *locally* minimizes the discrepancy in district size, i.e., there is no way to make any of the
$$\left|\frac{p_i}{a_i} - \frac{p_j}{a_j}\right|$$
smaller by shifting a seat from i to j.

- What do you think? Is one better?
- What about Hill's method?

$$\begin{aligned}
&\sqrt{a_i(a_i-1)} \leq \frac{p_i}{d} < \sqrt{a_i(a_i+1)} \\
\Rightarrow\quad &\forall i,j \quad \frac{p_j}{\sqrt{a_j(a_j+1)}} < d \leq \frac{p_i}{\sqrt{a_i(a_i-1)}} \\
\Rightarrow\quad &\forall i,j \quad \frac{p_j \cdot p_j}{a_j(a_j+1)} < \frac{p_i \cdot p_i}{a_i(a_i-1)} \\
\Rightarrow\quad &\forall i,j \quad \frac{p_j/a_j}{p_i/a_i} < \frac{p_i/(a_i-1)}{p_j/(a_j+1)} \quad \text{and} \quad \frac{a_i/p_i}{a_j/p_j} < \frac{(a_j+1)/p_j}{(a_i-1)/p_i}.
\end{aligned}$$

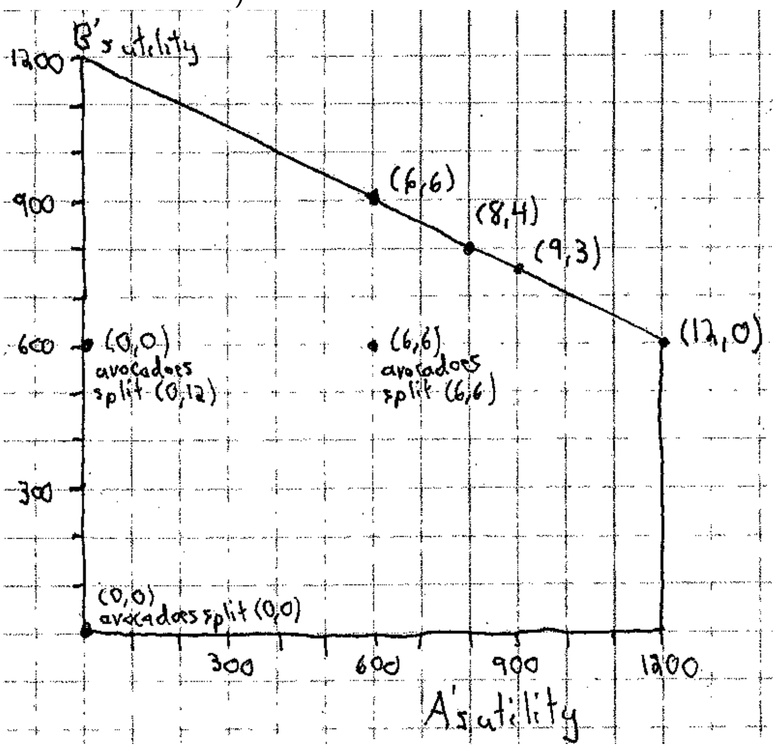
  - No way to improve the *relative* sizes of p_i/a_i (district size) or the *relative* sizes of a_i/p_i (per capita representation), by switching a seat.
    - Dean is about *absolute* difference in district size, and Webster is about *absolute* difference in representation per capita.
  - Symmetry! Hill is "in the middle". Another symmetry that helped National Academy of Sciences suggest Hill.
- What do you think?
- How about *global* optimality (these were all local optimality)?

- Can show that Webster is the global minimum of
$$\frac{1}{p}\sum_i p_i\left(\frac{a_i}{p_i}-\frac{h}{p}\right)^2,$$
which is the variance in per capita representation (a_i/p_i) across the whole population.
- Can show that Hill is global minimum of
$$\frac{1}{h}\sum_i a_i\left(\frac{p_i}{a_i}-\frac{p}{h}\right)^2,$$
which is the variance in district size (p_i/a_i) across all districts.

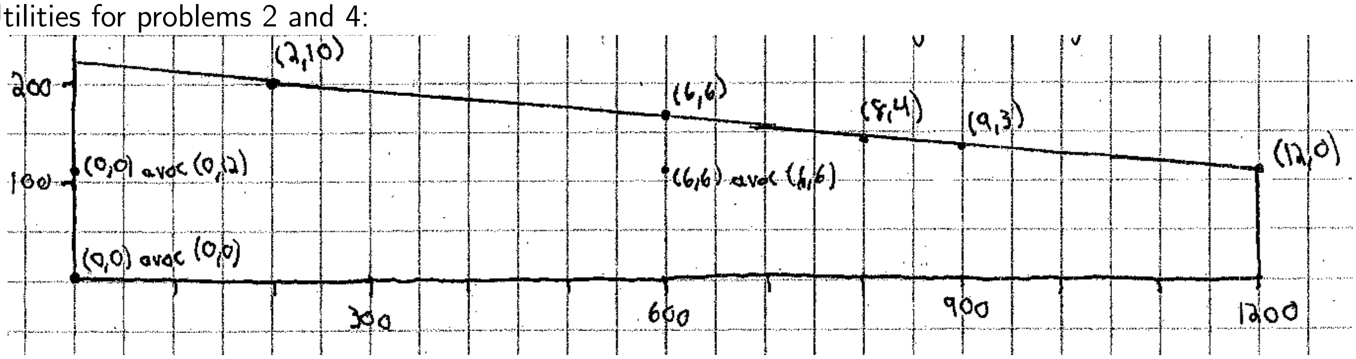
- What do you think?
  - The math can provide some measurements by which to judge (minimizing district size, minimizing per capita representation, avoiding population paradoxes, avoiding quota problems).
  - The math can say which measurements are consistent with each other (you can't have them all).
  - You have to decide which measurements are more important.
  - But the math is good for cutting to the essence of the problem of doing this "fairly".

33 Dividing the Spoils

- For problems 1, 3, 5 from PS9, below is the set of possible ordered pairs of utilities from splitting the fruit up. Listed points are the grapefruit split (B gets all of the avocados, unless otherwise indicated).



- Utilities for problems 2 and 4:

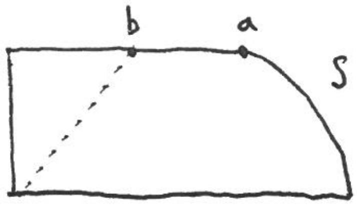


- What properties are you looking for in a split?
- How does your intuition change across the different problems?

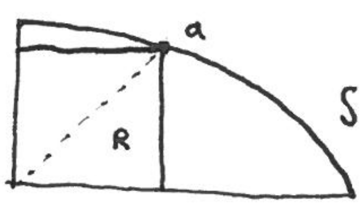
34 Egalitarian and Utilitarian

- Different outcomes affect different people differently. What is a fair way of thinking about what to do, based on the different utilities of the outcomes? What principles can we use?
- Assumptions:
  - An outcome can be associated to a point $u \in \mathbb{R}^n$, where u_i is utility of i^{th} person for that outcome.
    - If two outcomes correspond to the same point, everyone (and therefore society) is indifferent. Called the *Welfarism* assumption.
      - Doesn't depend on who deserves something, who has a right to something, peculiarities of the situation, etc.
    - Let $S \subseteq \mathbb{R}^n$ be the set of possible outcomes/utilities.
  - S is convex.
    - A possible explanation: People only care about expected utilities. If $a, b \in S$ are utilities of two outcomes, then the lottery with utility
    $$pa + (1-p)b,$$
    some p with $0 \leq p \leq 1$, should be a possible outcome, so the line between a and b should be in S.
  - S is closed and bounded. This lets us maximize functions over S.
  - Free disposal of utility. If $a \in S$ and $0 \leq b \leq a$ (meaning $0 \leq b_i \leq a_i$, for all i), then $b \in S$.
  - Let $n = 2$. Only for sake of pictures and concreteness. All results hold for larger n.
  - $(0, 0) \in S$ is the *default point* / *disagreement point*.
    - This is a reference point for comparison, which may be useful in determining what outcome should happen.
    - E.g, would happen if no agreement can be reached, or the "status quo", or some obvious default (like each get 6 grapefruits/avocadoes).
    - This point is $(0, 0)$ without loss of generality (if it is some other point, d, replace S with $S - d$).
  - See pictures from last class.
- Definition: Let $\mathcal{C}$ be the set of all subsets of $\mathbb{R}^n$ that meet the above assumptions. A *social choice function* is a function
$$f : \mathcal{C} \to \mathbb{R}^n \quad \text{with} \quad f(S) \in S.$$
That is, from among the set S of outcomes, it chooses a single one.
  - How do you do this in the "right" way?

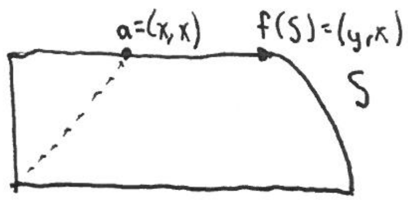
- Possible Properties of f:
  - Individual Rationality (IR): $f(S) \geq (0,0)$.
  - Symmetry: If ρ is the reflection over the line $y = x$, then
$$f\big(\rho(S)\big) = \rho\big(f(S)\big).$$
  - Strong Pareto: If $a, b \in S$ with $b \leq a$ and $b \neq a$, then $f(S) \neq b$. (not SP means there is a way for someone to do better without others doing worse; we shouldn't choose that point).
  - Weak Pareto: If $a, b \in S$ with $b < a$, then $f(S) \neq b$. (not WP means there is a way for everyone to do better).
  - Monotone: if $S \subseteq T$, then $f(S) \leq f(T)$. More options can't hurt anyone.
- Definition: The *egalitarian* social choice function is given by
$$f(S) = (x, x),$$
where x is the largest number with $(x, x) \in S$.
- Theorem: No social choice function satisfies Symmetry, SP, and Monotone. Egalitarianism is the unique social choice function satisfying Symmetry, WP, and Monotone. (Also satisfies IR.)
- Proof:
  - Egalitarian satisfies Symmetry, WP, Monotone, and IR:
  - Egalitarian does not satisfy SP:

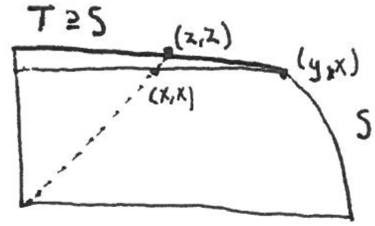


$b \leq a$, $b \neq a$, but $f(S) = b$.
  - If f satisfies Symmetry, WP, and Monotone, then f is the egalitarian function:
    - Given S, let $a = (x, x) \in S$ be the egalitarian solution. Let $R = \{b : \ b \leq a\} \subseteq S$.



- The only way $f(S)$ is not the egalitarian (x, x) is if it is (y, x), $y > x$, and S is as pictured left:





(or the symmetric version).

- Then look at T as pictured on right:

- So What? I think monotonicity captures the essence of egalitarianism: more options should help everyone.
  - If we relax monotonicity, might be able to add in other interesting axioms.
- For example, (loosely) we might say that we should be able to decide between options a and b, based solely on the difference $a - b$. That is, decide between two options solely based on how much better/worse off people are. In particular, the disagreement/default point should be unimportant in deciding between a and b.
  - This, together with Symmetry and WP, leads to *utilitarianism*: $f(S) = a$, where a maximizes $\sum_i a_i$ over S.
  - Note: utilitarianism is not individually rational. Doesn't even consider the default point.
- What kind of monotonicity does utilitarianism have?
  - Weak Monotone: Suppose $S \subseteq T$, $a = f(T)$, and $a \in S$. Then $f(S) = a$.
    - If currently choosing an option a from T, and *other* options disappear, then should still choose the same option, a.
    - Feels a lot like Independence of Irrelevant Alternatives.
    - Any social choice function which maximizes some function on S has weak monotonicity:
      if a maximizes a function over T, then it still maximizes the function over $S \subseteq T$.

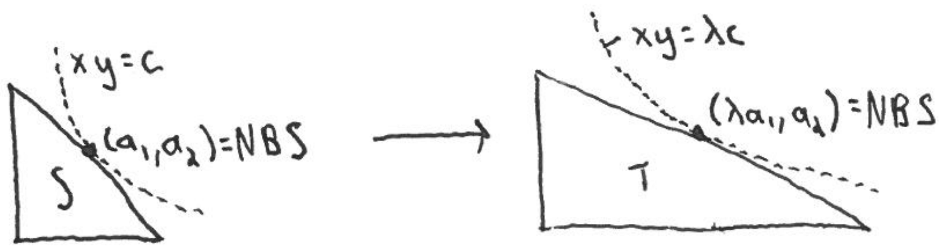
▫Utilitarian: maximizes $\sum_i a_i$.

▫Egalitarian: maximizes $\min_i a_i$.

▫What else? Next time, we will try maximizing $\prod_i a_i$.

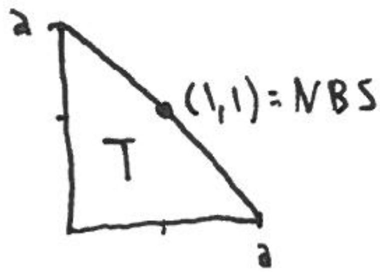
35 Nash Bargaining Solution

- Definition: The *Nash Bargaining Solution* is $f(S) = a$, where a maximizes $\prod_i b_i$, over all $b \in S$ (with b nonnegative).
  - It satisfies Symmetry, Strong Pareto, Weak Monotonicity.
  - Also satisfies *Scale Invariance*: if $f(S) = a$ and
    $$T = \{(b_1, \ldots, b_{i-1}, \lambda b_i, b_{i+1}, \ldots, b_n) : \ b \in S\}$$
    for some $\lambda > 0$ (that is, player i scales their utility by λ), then
    $$f(T) = (a_1, \ldots, a_{i-1}, \lambda a_i, a_{i+1}, \ldots, a_n).$$
    - Going from S to T multiplies $\prod_i b_i$ by λ, for every $b \in S$, so NBS satisfies Scale Invariance.



- So what?
  - Recall that, for deciding on my personal utility function, I need to specify the utility of 2 points: e.g., $u(0) = 0$ and $u(\$1\text{ mill}) = 100$.
    - Every other utility can be calculated by thinking about lotteries.
  - In our current setting, we're forcing
    $$u(\text{default point}) = 0,$$
    but we are not forcing a scale.
    - What does 100 utils mean to me? Is it the same that it means to you?
  - In fruit example, we saw two possible arbitrary decisions, that the scale should be measured in mg vitamin F or in willingness to pay.
    - But it's unclear that we can/should always do something like that.
    - Just because I'm richer and am willing to pay more, am I getting more "happiness"?
    - And for large dollar amounts, "willingness to pay" isn't an actual utility scale: $u(\$2\text{ mill}) \neq 2 \cdot u(\$1\text{ mill})$.
  - Scale invariance gets around this, because the scale doesn't matter (if you multiply all utilities by 2, will get same solution).
  - Scale invariance avoids Interpersonal Comparison of Utility. You can set up whatever scale you like. 100 utils for me doesn't have to be the same thing as 100 utils for you.
- Theorem: The only social choice function, f, satisfying Symmetry, Weak Pareto, Weak Monotone, and Scale Invariance is the Nash Bargaining Solution.

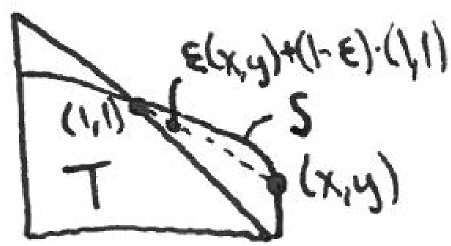
- Proof:
  - Given S, scale S so that NBS is $(1, 1, \ldots, 1)$. Need to show that $f(S) = (1, \ldots, 1)$.
  - Let $T = \{x \in \mathbb{R}^n : \; x_i \geq 0, \; \sum_i x_i \leq n\}$. We assume $n = 2$, for convenience.



  - $f(T) = (1, 1)$:
  - $S \subseteq T$: Suppose $(x, y) \in S$ with $x + y > 2$. Let's examine

$$\varepsilon \cdot (x, y) + (1 - \varepsilon) \cdot (1, 1),$$

which is in S by convexity:



  - $f(S) = (1, 1)$, as desired:

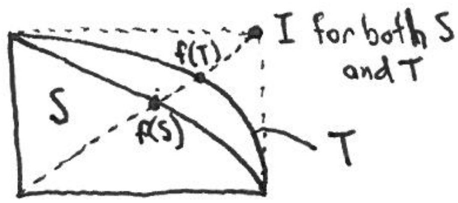
- Definition: To find the *Kalai-Smorodinsky Bargaining solution*, let $I \in \mathbb{R}^n$ be defined by

$$I_i = \max_{b \in S} b_i$$

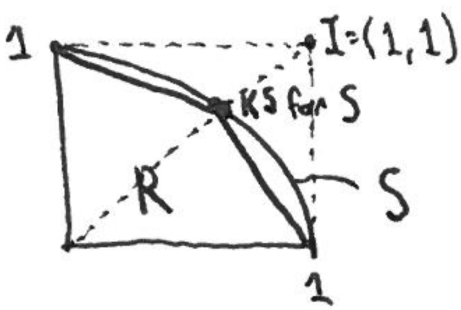
(player i's ideal outcome).
Define $f(S)$ to be the Pareto Optimal solution along the line segment from $(0, \ldots, 0)$ to I.

- ◦ Each player gets the same proportion of their ideal.

- Property: *Idealist Monotonicity*: If $S \subseteq T$ and the ideal points are the same for S and T, then $f(S) \leq f(T)$.
  - ◦ If there are more options, but the ideal outcomes stay the same, no one is worse off.



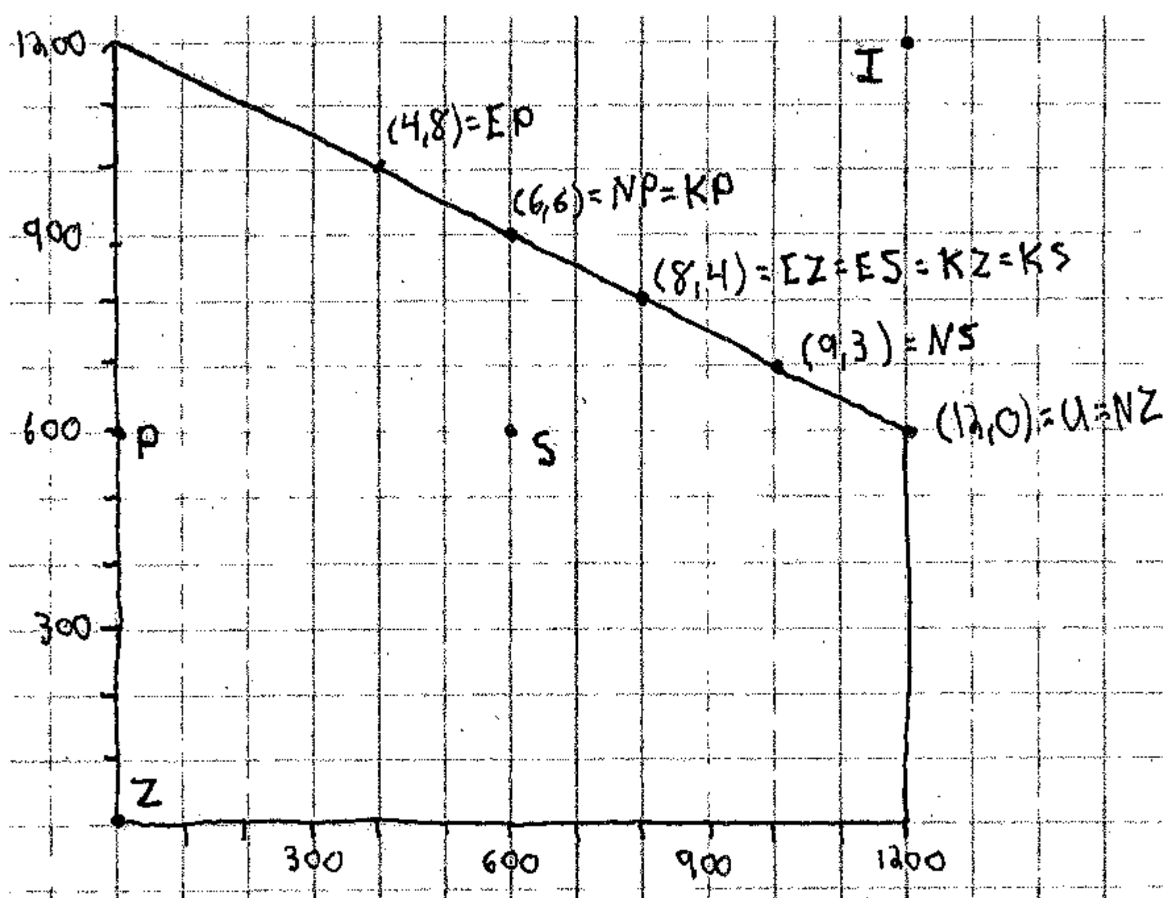
- Theorem: The only social choice function satisfying Symmetry, Strong Pareto (even Weak Pareto), Scale Invariance, and Idealist Monotonicity is the KS solution.
- Proof:
  - ◦ Scale S so that $I = (1, 1 \ldots, 1)$.
  - ◦ Let a be the KS solution for S and let R be as shown in the picture.



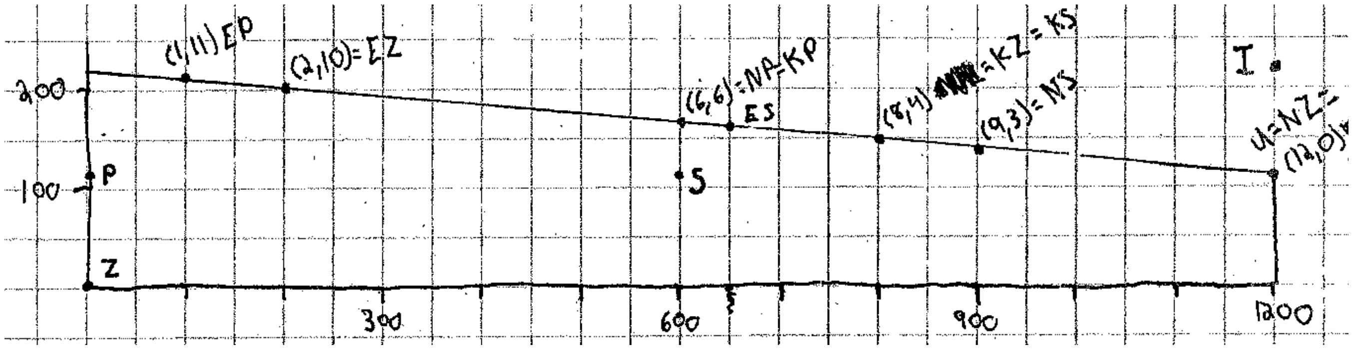
  - ◦ $f(R) = a$, using Symmetry and WP, as before.
  - ◦ $R \subseteq S$, by convexity.
  - ◦ $f(S) = a$, by Idealist Monotonicity.
- Corollary: No social choice function satisfies Symmetry, Weak Pareto, Scale Invariance, Weak Monotonicity, and Idealist Monotonicity.
  - ◦ Proof: NBS and KS are the unique functions satisfying all but one of these.
- Thoughts?
  - ◦ Symmetry and Weak Pareto seem essential.
  - ◦ Scale Invariance is essential if you don't believe in Interpersonal Comparison of Utility.
  - ◦ Which version of Monotonicity do you like better?

36 Dividing the Spoils Conclusion

- Returning to grapefruits and avocados.
  - 12 grapefruit, 12 avocado to split.
    Alexis gets 100 utils per grapefruit and 0 utils per avocado. Brenda gets 50 utils per fruit.
  - Possible default points:
    Z = all goes to waste,
    S = Split evenly (each gets 6 of each fruit)
    P = Pareto default: B gets the avocado, and the grapefruit spoil.
  - I = Ideal point = most each could get individually from all fruit.
  - Possible allocations:
    E = Egalitarian = maximal, equal utility gains
    U = Utilitarian = maximal sum of utilities (doesn't depend on default point)
    N = Nash Bargaining = maximal product of utilities
    K = Kalai-Smorodinsky Bargaining = proportional to ideal.



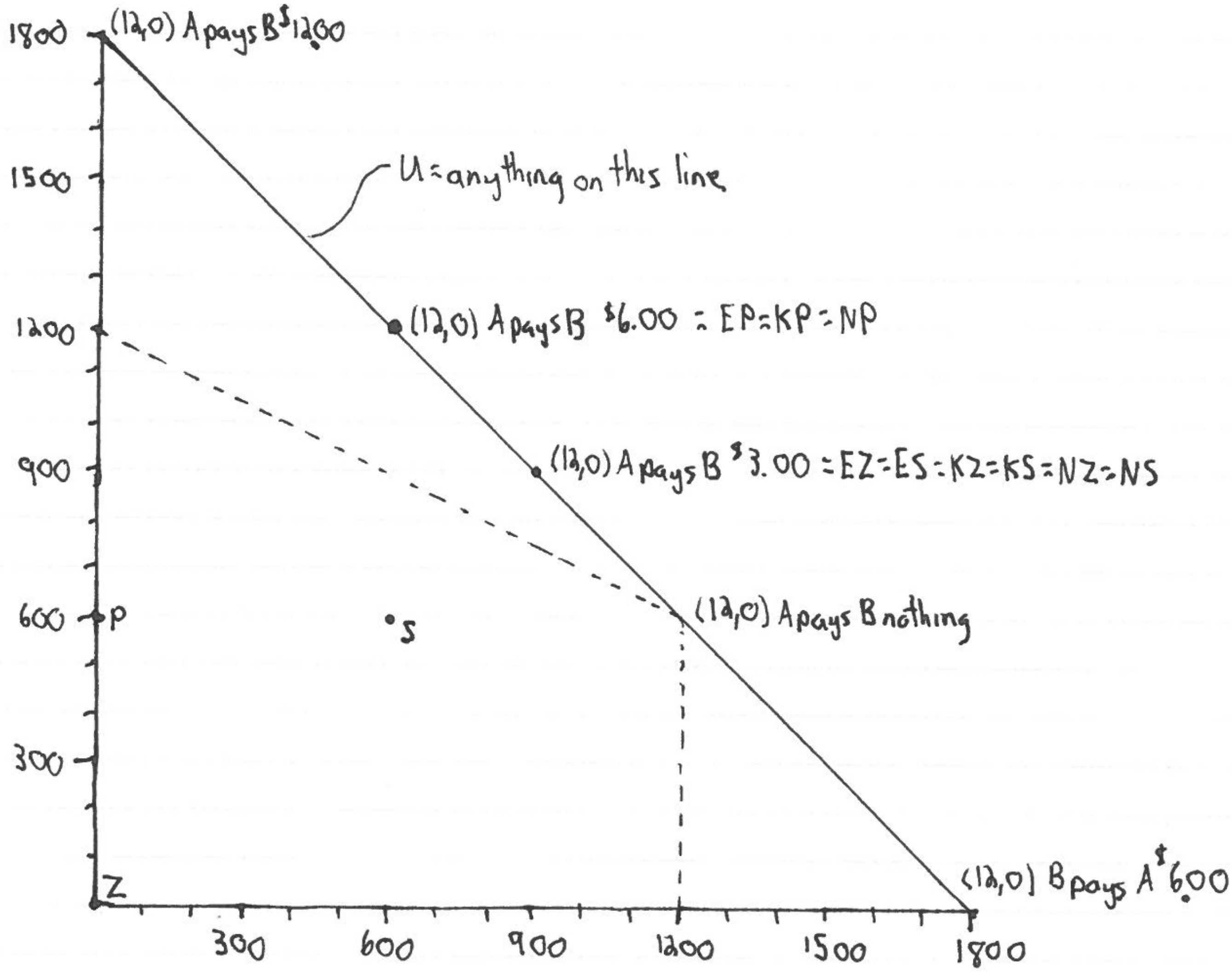
- When Brenda only gets 9.09 utils per fruit:



- Thoughts?

- Yaari and Bar-Hillel (1984) asked Israeli students for the "just" solution.
  - Needs-based (vitamin F):
    In First picture, strongly for (8,4).
    In Second picture, more spread out between (2,10), (6,6), and (12,0).
  - Taste-based: tended to be more spread out between (8,4) and (12,0) in either scenario.

- Look at First picture again, but allow money to change hands.
  - Expands the set of possibilities.
  - Assumption: Utility = Money.



  - Many utilitarian solutions.
  - Egalitarian, Nash, and Kalai-Smorodinksy all agree.
  - Breaks problem into 2 pieces:
    1. Find the *efficient* solution.

2. Decide how to fairly divide the generated utility.

    - Allows everyone to do better. It's a nice way to think about the power of markets.

- Two strains of analysis:
    - What is "good" / "reasonable"?
        - From earlier topics: IIA, Pareto, Efficient, Condorcet.
        - From this topic: Symmetry, Monotonicity, Pareto, Scale Invariance.
    - What is likely to happen, assuming agents act rationally/selfishly?
        - From earlier topics: Strategy-proof, Stability, Bayesian Nash, Revenue maximization, Individual Rationality.
        - How about in this setting?
- Imagine the two players are bargaining.
    - Player 1 has proposed the outcome with utilities (a_1, a_2). Player 2 has proposed (b_1, b_2).
    - By conceding, Player 1 would give up $a_1 - b_1$.
      By pushing hard, negotiations might fail, and Player 1 would give up $a_1 - 0 = a_1$.
    - A measure of Player 1's bargaining power is ${}^{(a_1 - b_1)}/_{a_1}$.
    - Player 1 has greater bargaining power than Player 2 means:
    - Now you might try to design some bargaining game; e.g., players with less bargaining power are more likely to make concessions; or if a player pushes too hard there is a probability of negotiations breaking down.
        - Many of these games lead to the outcome of the Nash Bargaining Solution.
        - If the NBS is proposed, there is no counter-proposal a player can make that puts them in a stronger bargaining position.

37 Conclusion

- This class has been concerned with creating mechanisms / algorithms that somehow choose an outcome.
  - The Mathematics of *Social* Choice: multiple people with their own ideas, desires, etc. How do you balance these?
- Axiomatization:
  - Define properties that we would like our mechanism to have.
    - This helps us focus on what is important/essential.
  - Describe mechanisms that have these properties.
    - The sad part: Impossibility Theorems. Can't always have everything you want.
    - The happy part: Possibility Theorems. Look for particular settings where we can have desirable properties.
- Two viewpoints / types of properties:
  - Normative: what *should* happen?
  - Descriptive: what *will* happen, assuming that people act selfishly, may lie, etc.?
- Started off mostly normative.
  - Voting Systems
    - Normative properties: Pareto Property, Independence of Irrelevant Alternatives, Condorcet.
      - Impossibility Theorem: Arrow. Can't have reasonable properties plus IIA.
      - Possibility Theorems:
        - If have a linear political spectrum (Bush – Gore – Nader), many nice properties (e.g., always a Condorcet winner).
        - Checked what properties different systems have.
      - Common tension: do you prefer simplicity, or better properties?
        - The simplest, plurality, is pretty horrible.
        - More complicated systems had especially nice properties like always electing the Condorcet winner, if it exists (Instant Runoff with Borda count, or minimizing Kemeny-Young distance).
    - Descriptive property: Strategy-Proof.
      - Impossibility Theorem: Gibbard-Satterthwaite (Strategy-proof is basically IIA).
      - Possibility Theorems:

        ⋄ Works in linear political spectrum.
        ⋄ Works if money is allowed to change hands!

- A shift towards thinking more descriptively.
  - Buying things (auctions, buyer-seller, public goods).
    - Normative: Feasible, Efficient / Maximizing total surplus.
    - Descriptive: Individually rational, Strategy-proof.
    - Possibility Theorem: Efficient, Strategy proof mechanisms are exactly the VCG mechanisms.
      - Clarke's Pivot mechanism often makes sense in auction-like settings.
      - Ascending price auction for multiple identical items (holding up fingers to indicate demand).
    - Impossibility Theorems:
      - Can't reconcile Strategy-proof with Feasible, Efficient, and Individually Rational in Buyer-Seller and Public Good examples.
  - Matching markets with and without money.
    - Possibility Theorems: Several mechanisms are reasonable (Ascending-price mechanism, Women-propose mechanism).
    - Im/Possibility Theorem: They are strategy-proof, if you only consider one side (Buyers, Women).
    - Im/Possibility "Theorem": Strategy proof mechanisms work out well when, among all "stable" solutions, all of the agents agree about which one is best.
      - That is, conflict among agents is only apparent; if they restrict themselves to realistic solutions, the conflict disappears.
      - Good examples: Women in Women-propose, Buyers in matching auction, Auctions of one or more identical items.
      - Bad Examples: Men and Women together, Seller as agent, Public Goods, Auctions with complementary items (PB&J).
  - Perhaps Strategy-proof is too restrictive.
    - Game Theory, Probability, Expected Value, Expected Utility.
    - Bayesian Nash Equilibria is our new descriptive property.
    - Possibility Theorems: Can find BNE for first-price auctions, auctions with reserves, etc.
    - Impossibility Theorem: Still can't create feasible, efficient, individually rational mechanisms if seller is an agent or if public good (we didn't prove this).

- Possibility Theorem: If allow not to be efficient, can come up with solutions:
  - Seller sets price, Auction with reserve (seller sets min price), Fixed price, Go halvesies, Meet in the middle.
  - And can quantify degree of efficiency and revenue. "Meet in middle" did better than simpler, strategy-proof options.
- Impossibility Theorem: BNE can't help increase revenue over the strategy-proof version.
  - Again, a tension between simplicity and other properties.

- Finally, a return to normative concerns.
  - Apportionment
    - Properties: House Monotone, State Monotone, Population Monotone, Quota Property.
    - Impossibility Theorem: Can't have all of these.
    - Possibility Theorem: Divisor methods (particularly Hill and Webster) are population monotone and can do a good job, e.g., they minimize various discrepancies, both pairwise and globally.
  - Fair Distribution
    - Properties: Symmetry, SP, WP, Scale Invariance, Monotone, Weak Monotone, Idealist Monotone.
    - Impossibility Theorem: Can't have all of these.
    - Possibility Theorems: Can have some of these, for various mechanisms:
      - Egalitarian, Utilitarian, Nash Bargaining, Kalai-Smorodinsky.
  - Apportionment and Fair Distribution are both about dividing up something.
    - In apportionment, the objects were simple: h identical (indivisible) seats.
    - In fair distribution, the objects were more complicated, but we simplified the rest:
      - Continuous case (e.g., by flipping a coin, or assuming divisible objects).
      - Symmetry of "equal claims".
    - You can imagine many other different ways to try to divide things. But we're out of time!