Queueing Paradoxes

Moshe Haviv

Department of Statistics and Center for the Study of Rationality
The Hebrew University of Jerusalem

Eindhoven 2016

What is a paradox?

A statement or proposition that **seems** self-contradictory or absurd but in reality expresses a possible truth.

Jevons Paradox

In economics, the Jevons paradox occurs when technological progress increases the efficiency with which a resource is used (reducing the amount necessary for any one use), but the rate of consumption of that resource rises because of increasing demand.

Jevons Paradox

Single server queue

Jevons Paradox

Single server queue

The server increases the service rate

Jevons Paradox

Single server queue

The server increases the service rate

He believes:

Jevons Paradox

Single server queue

The server increases the service rate

He believes:

- more will join

Jevons Paradox

Single server queue

The server increases the service rate

He believes:

- more will join
- waiting time will be reduced

Jevons Paradox

Single server queue

The server increases the service rate

He believes:

- more will join
- waiting time will be reduced
- he will have more free time

Jevons Paradox

Single server queue

The server increases the service rate

He believes:

- more will join
- waiting time will be reduced
- he will have more free time

Jevons Paradox

Single server queue

The server increases the service rate

He believes:

- more will join
- waiting time will be reduced
- he will have more free time

In fact:

- more will join

Jevons Paradox

Single server queue

The server increases the service rate

He believes:

- more will join
- waiting time will be reduced
- he will have more free time

In fact:

- more will join
- no change in waiting times

Jevons Paradox

Single server queue

The server increases the service rate

He believes:

- more will join
- waiting time will be reduced
- he will have more free time

In fact:

- more will join
- no change in waiting times
- he will work more, not less

The basic queueing model (M/M/1)

- single server

The basic queueing model (M/M/1)

- single server
- first come first served (FCFS)

The basic queueing model (M/M/1)

- single server
- first come first served (FCFS)
- Poisson arrival rate λ

The basic queueing model (M/M/1)

- single server
- first come first served (FCFS)
- Poisson arrival rate λ
- exponential service rate $\mu > \lambda$ (mean of $\frac{1}{\mu}$)

The basic queueing model (M/M/1)

- single server
- first come first served (FCFS)
- Poisson arrival rate λ
- exponential service rate $\mu > \lambda$ (mean of $\frac{1}{\mu}$)
- value of service R

The basic queueing model (M/M/1)

- single server
- first come first served (FCFS)
- Poisson arrival rate λ
- exponential service rate $\mu > \lambda$ (mean of $\frac{1}{\mu}$)
- value of service R
- cost per unit of wait C

Some facts

- mean service time $1/\mu$

Some facts

- mean service time $1/\mu$
- utilization level $\rho = \lambda/\mu < 1$

Some facts

- mean service time $1/\mu$
- utilization level $\rho = \lambda/\mu < 1$
- mean time in the system

$$W = \frac{1}{\mu(1-\rho)}$$

Some facts

- mean service time $1/\mu$
- utilization level $\rho = \lambda/\mu < 1$
- mean time in the system

$$W = \frac{1}{\mu(1-\rho)}$$

- mean time from an arrival until first server idleness

$$\frac{1}{\mu(1-\rho)^2}$$

Some facts

- mean service time $1/\mu$
- utilization level $\rho = \lambda/\mu < 1$
- mean time in the system

$$W = \frac{1}{\mu(1-\rho)}$$

- mean time from an arrival until first server idleness

$$\frac{1}{\mu(1-\rho)^2}$$

  - equals the total added time to the society due to the marginal arrival

Some facts

- mean service time $1/\mu$
- utilization level $\rho = \lambda/\mu < 1$
- mean time in the system

$$W = \frac{1}{\mu(1-\rho)}$$

- mean time from an arrival until first server idleness

$$\frac{1}{\mu(1-\rho)^2}$$

  - equals the total added time to the society due to the marginal arrival
  - equals the mean time in the system for a stand-by customer

Some facts

- mean service time $1/\mu$
- utilization level $\rho = \lambda/\mu < 1$
- mean time in the system

$$W = \frac{1}{\mu(1-\rho)}$$

- mean time from an arrival until first server idleness

$$\frac{1}{\mu(1-\rho)^2}$$

  - equals the total added time to the society due to the marginal arrival
  - equals the mean time in the system for a stand-by customer

Some facts

- mean service time $1/\mu$
- utilization level $\rho = \lambda/\mu < 1$
- mean time in the system

$$W = \frac{1}{\mu(1-\rho)}$$

- mean time from an arrival until first server idleness

$$\frac{1}{\mu(1-\rho)^2}$$

  - equals the total added time to the society due to the marginal arrival
  - equals the mean time in the system for a stand-by customer

Example: assume $\lambda = 0.9$ and $1/\mu = 1$
$\Rightarrow \rho = 0.9$
$\Rightarrow$ mean time in the system 10
$\Rightarrow$ mean socially added time 100 (for 1 unit of service!)

To queue or not to queue Edleson and Hildebrand, '75

assume

$$R - \frac{C}{\mu} > 0 \;\text{ and }\; R - \frac{C}{\mu(1-\rho)} < 0$$

if nobody joins, one better join. If all join, one better do not join

To queue or not to queue Edleson and Hildebrand, '75

assume

$$R - \frac{C}{\mu} > 0 \text{ and } R - \frac{C}{\mu(1-\rho)} < 0$$

if nobody joins, one better join. If all join, one better do not join

(Nash) equilibrium: join with probability p_e where

$$R - \frac{C}{\mu(1 - p_e \rho)} = 0$$

In equilibrium, all are indifferent between joining or not
They all end up with zero

To queue or not to queue Edleson and Hildebrand, '75

assume

$$R - \frac{C}{\mu} > 0 \text{ and } R - \frac{C}{\mu(1-\rho)} < 0$$

if nobody joins, one better join. If all join, one better do not join

(Nash) equilibrium: join with probability p_e where

$$R - \frac{C}{\mu(1 - p_e\rho)} = 0$$

In equilibrium, all are indifferent between joining or not
They all end up with zero

In case μ increases

To queue or not to queue Edleson and Hildebrand, '75

assume

$$R - \frac{C}{\mu} > 0 \quad \text{and} \quad R - \frac{C}{\mu(1-\rho)} < 0$$

if nobody joins, one better join. If all join, one better do not join

(Nash) equilibrium: join with probability p_e where

$$R - \frac{C}{\mu(1 - p_e \rho)} = 0$$

In equilibrium, all are indifferent between joining or not
They all end up with zero

In case μ increases

- utility is added to nobody: all are still left with zero

To queue or not to queue Edleson and Hildebrand, '75

assume

$$R - \frac{C}{\mu} > 0 \text{ and } R - \frac{C}{\mu(1-\rho)} < 0$$

if nobody joins, one better join. If all join, one better do not join

(Nash) equilibrium: join with probability p_e where

$$R - \frac{C}{\mu(1 - p_e \rho)} = 0$$

In equilibrium, all are indifferent between joining or not
They all end up with zero

In case μ increases

- utility is added to nobody: all are still left with zero
- the server works more, not less: $p_e \rho$ goes up with μ

social optimization: join with probability p_s where

$$p_s = \arg\max_{0<p<p_e} p\lambda(R - \frac{C}{\mu(1-p\rho)})$$

social optimization: join with probability p_s where

$$p_s = \arg\max_{0<p<p_e} p\lambda(R - \frac{C}{\mu(1-p\rho)})$$

$$R - \frac{C}{\mu(1-p_s\rho)^2} = 0$$

In social optimization, the society is indifferent whether the marginal customer joins or not

social optimization: join with probability p_s where

$$p_s = \arg \max_{0<p<p_e} p\lambda(R - \frac{C}{\mu(1-p\rho)})$$

$$R - \frac{C}{\mu(1-p_s\rho)^2} = 0$$

In social optimization, the society is indifferent whether the marginal customer joins or not

- those who do not join, end with nothing

social optimization: join with probability p_s where

$$p_s = \arg \max_{0<p<p_e} p\lambda(R - \frac{C}{\mu(1-p\rho)})$$

$$R - \frac{C}{\mu(1-p_s\rho)^2} = 0$$

In social optimization, the society is indifferent whether the marginal customer joins or not

- those who do not join, end with nothing
- those who join, gain something

social optimization: join with probability p_s where

$$p_s = \arg\max_{0<p<p_e} p\lambda\left(R - \frac{C}{\mu(1-p\rho)}\right)$$

$$R - \frac{C}{\mu(1-p_s\rho)^2} = 0$$

In social optimization, the society is indifferent whether the marginal customer joins or not

- those who do not join, end with nothing
- those who join, gain something

social optimization: join with probability p_s where

$$p_s = \arg \max_{0<p<p_e} p\lambda(R - \frac{C}{\mu(1 - p\rho)})$$

$$R - \frac{C}{\mu(1 - p_s\rho)^2} = 0$$

In social optimization, the society is indifferent whether the marginal customer joins or not

- those who do not join, end with nothing
- those who join, gain something

In case μ increases:

- more join, $\lambda p_s = \mu - \sqrt{C\mu/R}$

social optimization: join with probability p_s where

$$p_s = \arg\max_{0<p<p_e} p\lambda(R - \frac{C}{\mu(1-p\rho)})$$

$$R - \frac{C}{\mu(1-p_s\rho)^2} = 0$$

In social optimization, the society is indifferent whether the marginal customer joins or not

- those who do not join, end with nothing
- those who join, gain something

In case μ increases:

- more join, $\lambda p_s = \mu - \sqrt{C\mu/R}$
- social utility goes up, $(\sqrt{R\mu} - \sqrt{C})^2$

social optimization: join with probability p_s where

$$p_s = \arg \max_{0<p<p_e} p\lambda(R - \frac{C}{\mu(1-p\rho)})$$

$$R - \frac{C}{\mu(1-p_s\rho)^2} = 0$$

In social optimization, the society is indifferent whether the marginal customer joins or not

- those who do not join, end with nothing
- those who join, gain something

In case μ increases:

- more join, $\lambda p_s = \mu - \sqrt{C\mu/R}$
- social utility goes up, $(\sqrt{R\mu} - \sqrt{C})^2$
- those who join wait less, $\sqrt{R/(C\mu)}$

social optimization: join with probability p_s where

$$p_s = \arg\max_{0<p<p_e} p\lambda(R - \frac{C}{\mu(1-p\rho)})$$

$$R - \frac{C}{\mu(1-p_s\rho)^2} = 0$$

In social optimization, the society is indifferent whether the marginal customer joins or not

- those who do not join, end with nothing
- those who join, gain something

In case μ increases:

- more join, $\lambda p_s = \mu - \sqrt{C\mu/R}$
- social utility goes up, $(\sqrt{R\mu} - \sqrt{C})^2$
- those who join wait less, $\sqrt{R/(C\mu)}$
- the server works more, not less, $p_s\rho = 1 - \sqrt{C/(\mu R)}$

The Downs-Thomson paradox

- commuters have the option to drive to work or to take the bus

The Downs-Thomson paradox

- commuters have the option to drive to work or to take the bus
- the more who opt for driving, the longer driving takes

The Downs-Thomson paradox

- commuters have the option to drive to work or to take the bus
- the more who opt for driving, the longer driving takes
- the more who opt for the bus, the shorter is the ride

The Downs-Thomson paradox

- commuters have the option to drive to work or to take the bus
- the more who opt for driving, the longer driving takes
- the more who opt for the bus, the shorter is the ride
- assume that all using one of the options (any), is not an equilibrium

The Downs-Thomson paradox

- commuters have the option to drive to work or to take the bus
- the more who opt for driving, the longer driving takes
- the more who opt for the bus, the shorter is the ride
- assume that all using one of the options (any), is not an equilibrium
- in equilibrium,

The Downs-Thomson paradox

- commuters have the option to drive to work or to take the bus
- the more who opt for driving, the longer driving takes
- the more who opt for the bus, the shorter is the ride
- assume that all using one of the options (any), is not an equilibrium
- in equilibrium,
    - some drive, the rest take the bus

The Downs-Thomson paradox

- commuters have the option to drive to work or to take the bus
- the more who opt for driving, the longer driving takes
- the more who opt for the bus, the shorter is the ride
- assume that all using one of the options (any), is not an equilibrium
- in equilibrium,
    - some drive, the rest take the bus
    - both options come with the same amount of time

The Downs-Thomson paradox

- suppose the road capacity is increased

The Downs-Thomson paradox

- suppose the road capacity is increased
- in the new equilibrium, more drive, less take the bus (still identical times)

The Downs-Thomson paradox

- suppose the road capacity is increased
- in the new equilibrium, more drive, less take the bus (still identical times)
- since fewer use the bus now, the longer it takes

The Downs-Thomson paradox

- suppose the road capacity is increased
- in the new equilibrium, more drive, less take the bus (still identical times)
- since fewer use the bus now, the longer it takes
- hence, the longer is driving as well

The Downs-Thomson paradox

- suppose the road capacity is increased
- in the new equilibrium, more drive, less take the bus (still identical times)
- since fewer use the bus now, the longer it takes
- hence, the longer is driving as well
- all are worse off

The Downs-Thomson paradox

- suppose the road capacity is increased
- in the new equilibrium, more drive, less take the bus (still identical times)
- since fewer use the bus now, the longer it takes
- hence, the longer is driving as well
- all are worse off
- **Social optimization:** no roads at all, only buses (or only roads with huge capacity)

A cab or driving a car Alimeimounga, Solomon & Ziedins, '05

- driving a car: an M/M/1 model. If λ use it, it takes $1/(\mu - \lambda)$. $\mu > \lambda$ and μ the road capacity

A cab or driving a car Alimeimounga, Solomon & Ziedins, '05

- driving a car: an M/M/1 model. If λ use it, it takes $1/(\mu - \lambda)$. $\mu > \lambda$ and μ the road capacity
- The cab leaves when the 7th arrives

A cab or driving a car Alimeimounga, Solomon & Ziedins, '05

- driving a car: an M/M/1 model. If λ use it, it takes $1/(\mu - \lambda)$.
  $\mu > \lambda$ and μ the road capacity
- The cab leaves when the 7th arrives
- Unobservable. No regrets

A cab or driving a car Alimeimounga, Solomon & Ziedins, '05

- driving a car: an M/M/1 model. If λ use it, it takes $1/(\mu - \lambda)$. $\mu > \lambda$ and μ the road capacity
- The cab leaves when the 7th arrives
- Unobservable. No regrets
- Poisson arrivals, rate Λ

A cab or driving a car Alimeimounga, Solomon & Ziedins, '05

- driving a car: an M/M/1 model. If λ use it, it takes $1/(\mu - \lambda)$. $\mu > \lambda$ and μ the road capacity
- The cab leaves when the 7th arrives
- Unobservable. No regrets
- Poisson arrivals, rate Λ

A cab or driving a car Alimeimounga, Solomon & Ziedins, '05

- driving a car: an M/M/1 model. If λ use it, it takes $1/(\mu - \lambda)$. $\mu > \lambda$ and μ the road capacity
- The cab leaves when the 7th arrives
- Unobservable. No regrets
- Poisson arrivals, rate Λ

Assume:

- $\mu > \Lambda/3$: otherwise, all using the bus is an equilibrium

A cab or driving a car Alimeimounga, Solomon & Ziedins, '05

- driving a car: an M/M/1 model. If λ use it, it takes $1/(\mu - \lambda)$.
  $\mu > \lambda$ and μ the road capacity
- The cab leaves when the 7th arrives
- Unobservable. No regrets
- Poisson arrivals, rate Λ

Assume:

- $\mu > \Lambda/3$: otherwise, all using the bus is an equilibrium
- $\mu < \Lambda$: otherwise, all driving is an equilibrium

A cab or driving a car Alimeimounga, Solomon & Ziedins, '05

- driving a car: an M/M/1 model. If λ use it, it takes $1/(\mu - \lambda)$. $\mu > \lambda$ and μ the road capacity
- The cab leaves when the 7th arrives
- Unobservable. No regrets
- Poisson arrivals, rate Λ

Assume:

- $\mu > \Lambda/3$: otherwise, all using the bus is an equilibrium
- $\mu < \Lambda$: otherwise, all driving is an equilibrium

A cab or driving a car Alimeimounga, Solomon & Ziedins, '05

- driving a car: an M/M/1 model. If λ use it, it takes $1/(\mu - \lambda)$. $\mu > \lambda$ and μ the road capacity
- The cab leaves when the 7th arrives
- Unobservable. No regrets
- Poisson arrivals, rate Λ

Assume:

- $\mu > \Lambda/3$: otherwise, all using the bus is an equilibrium
- $\mu < \Lambda$: otherwise, all driving is an equilibrium

Equilibrium: $p_e\Lambda$ use the bus where p_e obeys

$$\frac{3}{p_e\Lambda} = \frac{1}{\mu - (1 - p_e)\Lambda}$$

p_e decreases with μ but $3/(p_e\Lambda)$ increases in μ

The prisoner's dilemma

	A	B
A	(2,2)	(0,3)
B	(3,0)	(1,1)

The prisoner's dilemma

	A	B
A	(2,2)	(0,3)
B	(3,0)	(1,1)

- (A, A) is socially optimal

The prisoner's dilemma

	A	B
A	(2,2)	(0,3)
B	(3,0)	(1,1)

- (A, A) is socially optimal
- (A, A) is not an equilibrium

The prisoner's dilemma

	A	B
A	(2,2)	(0,3)
B	(3,0)	(1,1)

- (A, A) is socially optimal
- (A, A) is not an equilibrium
- (B, B) is the unique equilibrium

The prisoner's dilemma

	A	B
A	(2,2)	(0,3)
B	(3,0)	(1,1)

- (A, A) is socially optimal
- (A, A) is not an equilibrium
- (B, B) is the unique equilibrium
- action B is a dominant action (for both)

The prisoner's dilemma

	A	B
A	(2,2)	(0,3)
B	(3,0)	(1,1)

- (A, A) is socially optimal
- (A, A) is not an equilibrium
- (B, B) is the unique equilibrium
- action B is a dominant action (for both)

The prisoner's dilemma

	A	B
A	(2,2)	(0,3)
B	(3,0)	(1,1)

- (A, A) is socially optimal
- (A, A) is not an equilibrium
- (B, B) is the unique equilibrium
- action B is a dominant action (for both)

Suppose both A and B exist. Better having someone (government, dictator, regulator) who will remove option B.

The prisoner's dilemma

	A	B
A	(2,2)	(0,3)
B	(3,0)	(1,1)

- (A, A) is socially optimal
- (A, A) is not an equilibrium
- (B, B) is the unique equilibrium
- action B is a dominant action (for both)

Suppose both A and B exist. Better having someone (government, dictator, regulator) who will remove option B.

Suppose the status-quo was only A. An inventor comes with another option, B.

The prisoner's dilemma

	A	B
A	(2,2)	(0,3)
B	(3,0)	(1,1)

- (A, A) is socially optimal
- (A, A) is not an equilibrium
- (B, B) is the unique equilibrium
- action B is a dominant action (for both)

Suppose both A and B exist. Better having someone (government, dictator, regulator) who will remove option B.

Suppose the status-quo was only A. An inventor comes with another option, B.

- if I was the only one who could enjoy it, great for me (and who cares about the others)

The prisoner's dilemma

	A	B
A	(2,2)	(0,3)
B	(3,0)	(1,1)

- (A, A) is socially optimal
- (A, A) is not an equilibrium
- (B, B) is the unique equilibrium
- action B is a dominant action (for both)

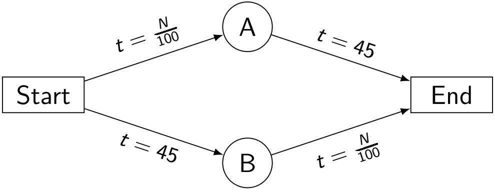
Suppose both A and B exist. Better having someone (government, dictator, regulator) who will remove option B.

Suppose the status-quo was only A. An inventor comes with another option, B.

- if I was the only one who could enjoy it, great for me (and who cares about the others)
- if all can use it, we are all worse off

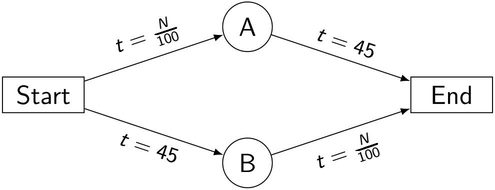
Breass Paradox

4000 commuters



Breass Paradox

4000 commuters

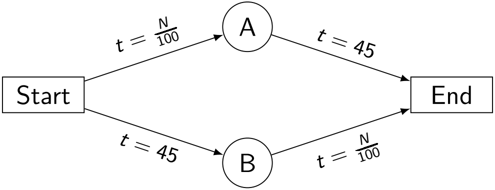


Equilibrium: 2000 go in each route
Social optimization: same behavior

Individual cost **65**. Social cost $65 \times 4000 = 260,000$

Breass Paradox

4000 commuters



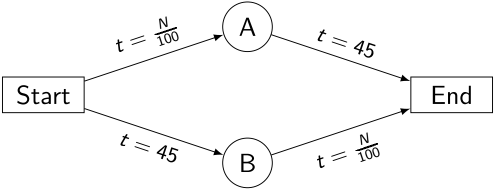
Equilibrium: 2000 go in each route
Social optimization: same behavior

Individual cost **65**. Social cost $65 \times 4000 = 260,000$

An individual who switches losses from the switch

Breass Paradox

4000 commuters



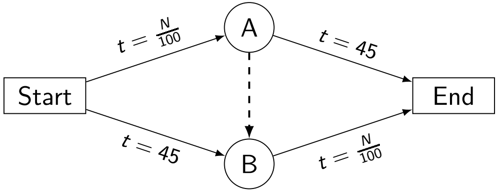
Equilibrium: 2000 go in each route
Social optimization: same behavior

Individual cost **65**. Social cost $65 \times 4000 = 260,000$

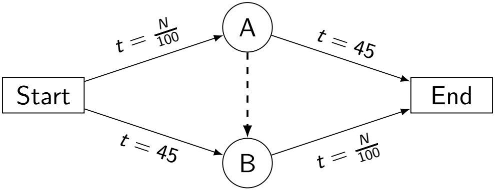
An individual who switches losses from the switch

For others, some win, some lose. The absolute changes coincide, **but** there are more losers than winners. This is more so when additional switches occur

New network

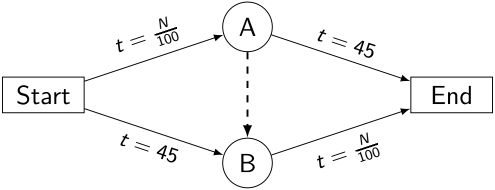


New network



If only one could use the bridge, he would gain $\approx 45 - 40 = 5$.

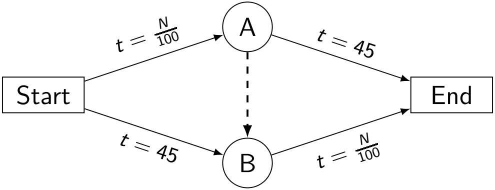
New network



If only one could use the bridge, he would gain $\approx 45 - 40 = 5$.

Equilibrium: All use the bridge
Social optimization: 500 use bridge, 1750 use original routes each

New network

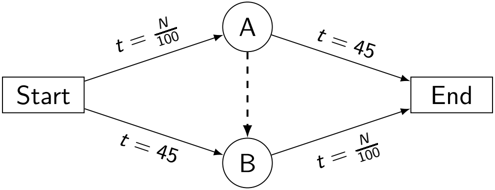


If only one could use the bridge, he would gain $\approx 45 - 40 = 5$.

Equilibrium: All use the bridge
Social optimization: 500 use bridge, 1750 use original routes each

Equilibrium: individual cost **80**. Social cost $80 \times 4000 = 320,000$
Social optimization: 500 suffer $45 < 65$, 3500 suffer $67.5 > 65$

New network



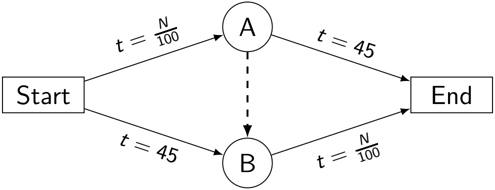
If only one could use the bridge, he would gain $\approx 45 - 40 = 5$.

Equilibrium: All use the bridge
Social optimization: 500 use bridge, 1750 use original routes each

Equilibrium: individual cost **80**. Social cost $80 \times 4000 = 320,000$
Social optimization: 500 suffer $45 < 65$, 3500 suffer $67.5 > 65$

Social cost in social optimality:
$500 \times 45 + 3500 \times 67.5 = 258,750 < 260,000 < 320,000$

New network



If only one could use the bridge, he would gain $\approx 45 - 40 = 5$.

Equilibrium: All use the bridge
Social optimization: 500 use bridge, 1750 use original routes each

Equilibrium: individual cost **80**. Social cost $80 \times 4000 = 320,000$
Social optimization: 500 suffer $45 < 65$, 3500 suffer $67.5 > 65$

Social cost in social optimality:
$500 \times 45 + 3500 \times 67.5 = 258,750 < 260,000 < 320,000$

Price of Anarchy (PoA):

$$\frac{320,000}{258,750} \approx \frac{5}{4} < \frac{4}{3} \text{ (theoretical bound)}$$

On average, we wait more than average

The model:

- $2n$ customers seek service in one out of two servers

On average, we wait more than average

The model:

- $2n$ customers seek service in one out of two servers
- all need to show up before 9:00am

On average, we wait more than average

The model:

- $2n$ customers seek service in one out of two servers
- all need to show up before 9:00am
- at each server, random order of service (no early birds)

On average, we wait more than average

The model:

- $2n$ customers seek service in one out of two servers
- all need to show up before 9:00am
- at each server, random order of service (no early birds)
- service lasts one minute

On average, we wait more than average

The model:

- $2n$ customers seek service in one out of two servers
- all need to show up before 9:00am
- at each server, random order of service (no early birds)
- service lasts one minute
- unobservable: each selects the server randomly

On average, we wait more than average

The model:

- $2n$ customers seek service in one out of two servers
- all need to show up before 9:00am
- at each server, random order of service (no early birds)
- service lasts one minute
- unobservable: each selects the server randomly

On average, we wait more than average

The model:

- $2n$ customers seek service in one out of two servers
- all need to show up before 9:00am
- at each server, random order of service (no early birds)
- service lasts one minute
- unobservable: each selects the server randomly

Q: What is the mean queueing time?

On average, we wait more than average

The model:

- $2n$ customers seek service in one out of two servers
- all need to show up before 9:00am
- at each server, random order of service (no early birds)
- service lasts one minute
- unobservable: each selects the server randomly

Q: What is the mean queueing time?

A: Uniform between 0 and n? $(n-1)/2$??

On average, we wait more than average

The model:

- $2n$ customers seek service in one out of two servers
- all need to show up before 9:00am
- at each server, random order of service (no early birds)
- service lasts one minute
- unobservable: each selects the server randomly

Q: What is the mean queueing time?

A: Uniform between 0 and n? $(n-1)/2$?? No!

On average, we wait more than average

The model:

- $2n$ customers seek service in one out of two servers
- all need to show up before 9:00am
- at each server, random order of service (no early birds)
- service lasts one minute
- unobservable: each selects the server randomly

Q: What is the mean queueing time?

A: Uniform between 0 and n? $(n-1)/2$?? No!

A: Correct above by $O(\sqrt{n})$? 'inspection paradox'? 'length bias'?

On average, we wait more than average

The model:

- $2n$ customers seek service in one out of two servers
- all need to show up before 9:00am
- at each server, random order of service (no early birds)
- service lasts one minute
- unobservable: each selects the server randomly

Q: What is the mean queueing time?

A: Uniform between 0 and n? $(n-1)/2$?? No!

A: Correct above by $O(\sqrt{n})$? 'inspection paradox'? 'length bias'? No!

On average, we wait more than average

The model:

- $2n$ customers seek service in one out of two servers
- all need to show up before 9:00am
- at each server, random order of service (no early birds)
- service lasts one minute
- unobservable: each selects the server randomly

Q: What is the mean queueing time?

A: Uniform between 0 and n? $(n-1)/2$?? No!

A: Correct above by $O(\sqrt{n})$? 'inspection paradox'? 'length bias'? No!

A: $(n-1)/2 + 1/4$??

On average, we wait more than average

The model:

- $2n$ customers seek service in one out of two servers
- all need to show up before 9:00am
- at each server, random order of service (no early birds)
- service lasts one minute
- unobservable: each selects the server randomly

Q: What is the mean queueing time?

A: Uniform between 0 and n? $(n-1)/2$?? No!

A: Correct above by $O(\sqrt{n})$? 'inspection paradox'? 'length bias'? No!

A: $(n-1)/2 + 1/4$?? Yes!

On average, we wait more than average

The model:

- $2n$ customers seek service in one out of two servers
- all need to show up before 9:00am
- at each server, random order of service (no early birds)
- service lasts one minute
- unobservable: each selects the server randomly

Q: What is the mean queueing time?

A: Uniform between 0 and n? $(n-1)/2$?? No!

A: Correct above by $O(\sqrt{n})$? 'inspection paradox'? 'length bias'? No!

A: $(n-1)/2 + 1/4$?? Yes!

Anti-paradoxically, the error is not a function of n and
equals (only) a quarter of a service time

Proof

Proof

"Proof" 1:

$$\frac{n+O(\sqrt{n})}{2n}\frac{n+O(\sqrt{n})}{2}+\frac{n-O(\sqrt{n})}{2n}\frac{n-O(\sqrt{n})}{2}=\frac{n}{2}+O1/\sqrt{n}$$

Proof 2: Tag a customer

- $(2n-1)/2$ others are expected in each line
- $(2n-1)/4=(n-1)/2+1/4$ services ahead

Proof

"Proof" 1:

$$\frac{n+O(\sqrt{n})}{2n}\frac{n+O(\sqrt{n})}{2}+\frac{n-O(\sqrt{n})}{2n}\frac{n-O(\sqrt{n})}{2}=\frac{n}{2}+O1/\sqrt{n}$$

Proof 2: Tag a customer

- $(2n-1)/2$ others are expected in each line
- $(2n-1)/4=(n-1)/2+1/4$ services ahead

Proof 3:

$$X \sim \text{Bin}(2n, 1/2)$$

Mean queueing time:

$$\text{E}(\frac{X}{2n}\frac{X-1}{2}+\frac{2n-X}{2n}\frac{2n-X-1}{2})=\frac{n-1}{2}+\frac{1}{4}$$

THANK YOU

Some facts

- The equilibrium arrival rate: $\lambda_e = \mu - \frac{C}{R}$
- The socially optimal arrival rate: $\lambda_s = \mu - \sqrt{\frac{C\mu}{R}}$
- Either rate is not a function of the (high) potential rate
- $$\lambda_s < \lambda_e \Rightarrow \text{long queues}$$
- The consumer surplus is zero in equilibrium.
  It is $(\sqrt{R\mu} - \sqrt{C})^2$ in social optimization
- No gain in equilibrium from extra service capacity.
  A gain under social optimization