Essays in Reputation Games

Gökmen Öz

June 1, 2023

Outline

- Chapter 2: A Dynamic Setup
- Chapter 3: Empirical Research
- Chapter 1: A Static Setup

Chapter 2: A Dynamic Setup

Sustainable Reputations With A Biased Review Platform

Motivation

- A long-run player and a sequence of short-run players play the following game repeatedly:

	B	N
H	1,1	-1,0
L	2,-1	0,0

- Highest ever payoff to P1 is 1.5.

Sustainable Reputations With A Biased Review Platform

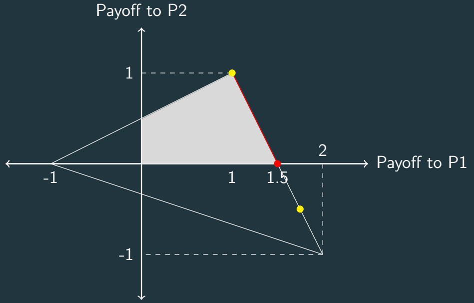
Motivation

Max/Sup payoff to P1	no censoring	censoring
no reputation effect	1	1
reputation effect	1	1.5

- Previous literature: Censoring helps the normal type P1 in the incomplete information game by yielding almost a payoff of 1.5.
- This paper: To show: $\lambda > 0$ is a prior belief that P1 is committed to $H \Rightarrow$ sup(payoff to normal P1)$=\frac{1.5-\lambda}{1-\lambda} > 1.5$

Sustainable Reputations With A Biased Review Platform

Motivation



- Why $\frac{1.5-\lambda}{1-\lambda}$?
- The best ever average payoff to $P1$ is 1.5 and in such equilibria $P2$ always buys:

$$\lambda + (1-\lambda)\bar{V} = 1.5.$$

which yields $\bar{V} = \frac{1.5-\lambda}{1-\lambda}$.

Sustainable Reputations With A Biased Review Platform

Model

- There are 2 actions available for each player, $a_1 \in A_1 = \{H, L\}$ and $a_2 \in A_2 = \{B, N\}$.
- Prior to time $t = 0$, nature chooses a type that is either commitment or normal for the long-run player: $w \in \Omega = \{c, n\}$.
- Let $\lambda = 1/3$ denote the probability that the long-run player is a commitment type ($\bar{V} = 1.75$).
- Commitment type always takes action H.

- Short-run players do not observe the past play and stage-game payoffs, but observe a review score $s \in S = \{1, ..., K\}$ announced by a review platform.
- A review platform R observes past actions and announces a public review score s from a finite set $S = \{1, ..., K\}$:
  - $R = (S, \Delta(S), P : A_1^\infty \times S^\infty \to S)$.
- Players move simultaneously.

Sustainable Reputations With A Biased Review Platform

Theorem

To show:

For any $\lambda > 0$ and $u < \bar{V}$;

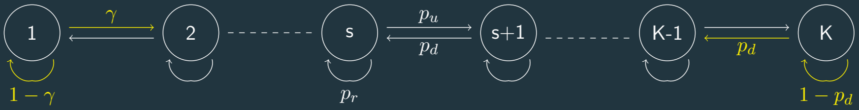
there exists a $\bar{K}$ such that for all $K > \bar{K}$,

there exists a $\bar{\delta}$ such that for all $\delta > \bar{\delta}$,

there is an R and a PBE where the payoff to the normal type of the long-run player is at least u after every history.

Sustainable Reputations With A Biased Review Platform

Step 1: Constructing a transition rule $P(a_1, s)$.



p_d, p_r, p_u	$s = 1$	$s \in \{2, ..., K-1\}$	$s = K$
$a_1 = H$	NA, 5/12, 7/12	1/6, 1/6, 4/6	1/6, 5/6, NA
$a_1 = L$	NA, 5/12, 7/12	1/6, 5/6, 0	1/6, 5/6, NA

Sustainable Reputations With A Biased Review Platform

Step 2: Equilibrium Strategies

- Let $\sigma_1(s)$ denote the probability that P1 plays H at score s.
- Let $\sigma_2(s)$ denote the probability that P2 plays B at score s.
- Punishment phase ($s = 1$): $(\sigma_1(1), \sigma_2(1)) = (0, 0)$.
- Reputation building phase ($s \in \{2, ..., K-1\}$):
  $(\sigma_1(s), \sigma_2(s)) = (\frac{0.5-\lambda(s)}{1-\lambda(s)}, 1)$ where

  $$\lambda(s) = \frac{\lambda\pi_c(s)}{\lambda\pi_c(s) + (1-\lambda)\pi_n(s)}$$

  where π_c and π_n are the steady-states induced by P, σ_1.
- Reputation exploitation phase ($s = K$):
  $(\sigma_1(K), \sigma_2(K)) = (0, 1)$.

Sustainable Reputations With A Biased Review Platform

Step 3: Upward drift

- Minimizing $\pi_n(1)$:

$$\begin{aligned}\pi_n(s+1)p_d &= \pi_n(s)p_u \\ p_d &= \frac{1}{6} \\ p_u &= \frac{2}{3}\sigma_1(s) \\ \Rightarrow \pi_n(s+1)\frac{1}{6} &= \pi_n(s)\frac{2}{3}\sigma_1(s)\end{aligned}$$

$$\sigma_1(s) > 0.25 \Rightarrow \pi_n(s+1) > \pi_n(s)$$

Sustainable Reputations With A Biased Review Platform

Step 4: Perturbing P for $\delta < 1$

$$
\begin{aligned}
V(s) &= 1 - \delta + \delta\Big[\frac{2}{3}V(s+1) + \frac{1}{6}V(s) + \frac{1}{6}V(s-1)\Big] \\
&= (1-\delta)2 + \delta\Big[\frac{5}{6}V(s) + \frac{1}{6}V(s-1)\Big]
\end{aligned}
$$

implies $V(s+1) - V(s) = \frac{1.5(1-\delta)}{\delta}$ and $V(s) = 1.75$, that is a contradiction.
Perturb P as

$$
\begin{aligned}
p_d(s) &= p_d(s) + \epsilon_s \\
p_r(s) &= p_r(s) - \epsilon_s \\
\gamma &= \gamma - \epsilon_1
\end{aligned}
$$

where $\epsilon_s = \frac{(K-s)(1-\delta)}{\delta}$.

Sustainable Reputations With A Biased Review Platform

Step 4: Perturbing P for $\delta < 1$

Plugging in the perturbed transition probabilities, we find:

$$V(s+1) - V(s) = 1.5\frac{1-\delta}{\delta}$$
$$\Rightarrow V(s) = 1.75 - \frac{1.5(K-s)(1-\delta)}{\delta}$$

Sustainable Reputations With A Biased Review Platform

Step 5: Checking if $(\sigma_1(s), \sigma_2(s))$ are optimal.

- Given $V(s)$, $\sigma_1(s)$ is optimal.
- Given $\sigma_1(s)$ and $\lambda(s)$, $\sigma_2(s)$ is optimal.
- To check: $\lambda(K) = 0.5$
  - The general representation of P is

$$p_u(H) = \frac{1}{\eta} \quad p_u(L) = 0 \quad p_d(s) = \frac{\eta - 1}{2\eta} + \epsilon_s$$

  - $\eta > 1$ for P to be well-defined.
  - $\eta < 2$ for the upward drift to exist.
  - Consider $\eta = 1$, $\eta = 2$ and use Intermediate Value Theorem.

Sustainable Reputations With A Biased Review Platform

Step 6: Equilibrium properties.

- An asymptotic closed-form solution for η can be found:

$$\begin{aligned} \lim V(s,\eta) &= \bar{V} \\ \lim \left\{ 2 - 0.5(\eta - 1) - \frac{1.5(K-s)(1-\delta)}{\delta} \right\} &= 1.75 \\ 2 - 0.5(\eta - 1) &= 1.75 \end{aligned}$$

Sustainable Reputations With A Biased Review Platform

Step 6: Equilibrium properties.

- $\frac{\pi_c(K)}{\pi_c(K-1)} \approx 4$, $\frac{\pi_n(K)}{\pi_n(K-1)} \approx 4\sigma_1(K-1)$ and $\frac{\pi_c(K)}{\pi_n(K)} = 2$.
- The posterior at any score s is

$$\lambda(s) \approx 0.5^{K-s+1}$$

- $\arg\min_{s \in \{2,\dots,K-1\}} \sigma_1(s) = K-1$:

$$\begin{aligned}\sigma_1(K-1) &= \frac{0.5 - \lambda(K-1)}{1 - \lambda(K-1)} \\ &\approx \frac{0.5 - 0.25}{1 - 0.25} \\ &\approx \frac{1}{3} > 0.25\end{aligned}$$

Sustainable Reputations With A Biased Review Platform

Step 7: Initial periods.

- We have constructed a review platform and found an equilibrium strategy profile for the long-run, that is if the game had started at $-\infty$.
- However, the beliefs will never hit the steady-states.
- Therefore, our equilibrium profile $(\sigma_1(s), \sigma_2(s))$ will not work.

Sustainable Reputations With A Biased Review Platform

Step 7: Initial periods.

- Choose $\pi^0(K-1) = 1$:

$$\lambda^1(K) = \frac{\lambda p_u(H)}{\lambda p_u(H) + (1-\lambda)\sigma_1(K-1)p_u(H)} = 3/5 > 0.5$$

- Choose a large N such that for $t < N$:

$$\sigma_1^t(s) = 0$$
$$\sigma_2^t(s) = \begin{cases} 1 & if \quad s = K \\ 0 & o.w. \end{cases}$$

Sustainable Reputations With A Biased Review Platform

Step 7: Initial periods

- P is as if P1 is playing $\sigma_1(s)$ for $t < N$:

$$p_u(L, s) = \frac{2}{3}\sigma_1(s)$$
$$p_r(L, s) = \frac{5}{6} - \frac{2}{3}\sigma_1(s) - \epsilon_s$$
$$p_d(L, s) = \frac{1}{6} + \epsilon_s$$

- Time-dependent strategies for $t \geq N$:

$$\sigma_1^t(s) = \frac{0.5 - \lambda^t(s)}{1 - \lambda^t(s)}$$

for $s \in \{2, ..., K - 1\}$.

Sustainable Reputations With A Biased Review Platform

Step 8: Transition with a jump

- In fact, P still does not work:

$$\begin{aligned} V(s+1) - V(s) &= 1.5\frac{1-\delta}{\delta} \\ V(1) &= \delta[\gamma\Big(V(1) + 1.5\frac{1-\delta}{\delta}\Big) + (1-\gamma)V(1))] \\ V(1) &= 1.5\gamma << 1.75 \end{aligned}$$

Sustainable Reputations With A Biased Review Platform

Step 8: Transition with a jump

- Given P, the strategies are not optimal.
- Our solution: jump from score 1 to 3:

$$V(1) = 3\gamma \approx 1.75$$

- The key: the effect of the jump on the steady-states is negligible when K is large relative to the size of the jump.

Chapter 3: Empirical Research Evidence From Yelp Reviews

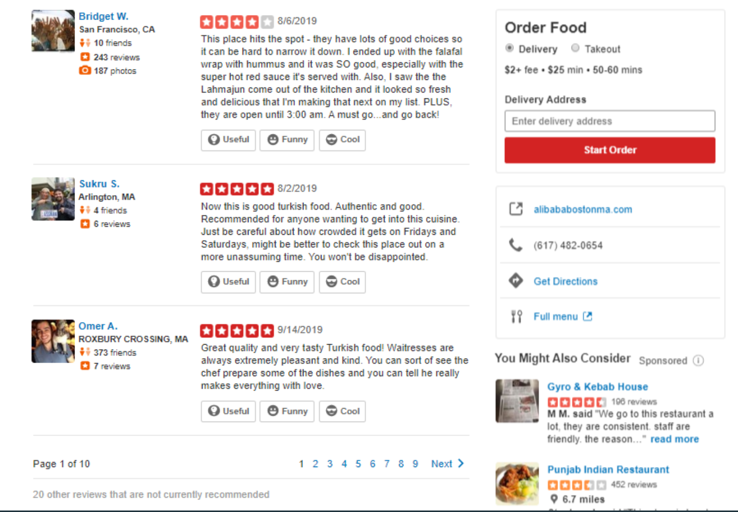
Evidence From Yelp Reviews

Motivation and Data

- Existing study on fake reviews relies on small data.
- The result is simply that better reputation leads to less incentive to produce fake reviews.
- This rationale ignores the actions of business owners (reputation exploitation).
- The goal of this study is to conduct an empirical analysis on the incentive to produce fake reviews with a scraped data set that is larger in both depth and breadth.

Evidence From Yelp Reviews

Motivation and Data



Evidence From Yelp Reviews

Motivation and Data

20 reviews for Ali Baba that are not currently recommended

Note: The reviews below are not factored into the business's overall star rating.

Margaux M.
Boston, MA
0 friends
1 review

5/7/2015

This place likes to say their motto is "We do, we do."

However, be warned.

They don't.

Khai N.
Boston, MA
161 friends
7 reviews
1 photo

3/22/2015

I ordered on a Sunday as was told it would not take more than 1 hour. Over an hour later I called and they said it would take another hour (2 hrs total). I live 10 minutes away... I told the guy that someone there said it wouldn't take more than an hour, and he basically wrote me off and said no one here would tell you that. OH, AND THE FOOD CAME COLD. So terrible lol. Very poor service. May be a place to order if you like being lied to or being called a liar.

Muhammed A.
Malden, MA
0 friends
2 reviews

1/22/2015

The owner in the rudest man I have ever seen so bad service no hospitality not clean place he act like gangster just want to fight told I should just take food and go don't talk at all beside the food was not cooked well rice is horrible

---

Helena L.
Quincy, MA
0 friends
6 reviews
4 photos

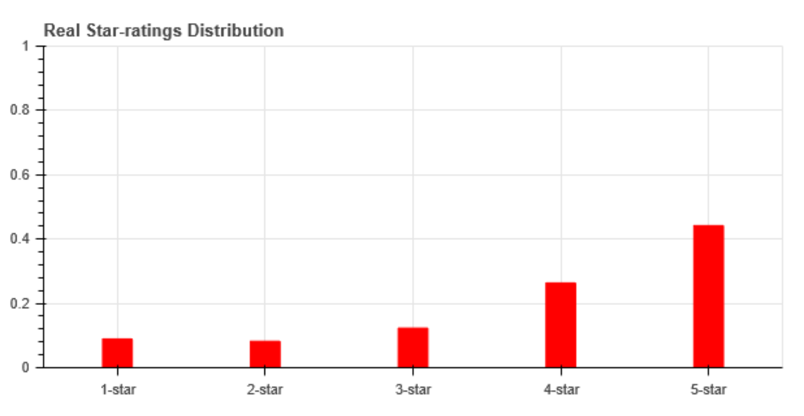
10/16/2017

I came here for dinner and was craving for some Turkish food. I got lahmacun and it tasted so bad. It was dried and the meat was crunchy and burnt. It was not the same as the one I had in Turkey. I am disappointed.

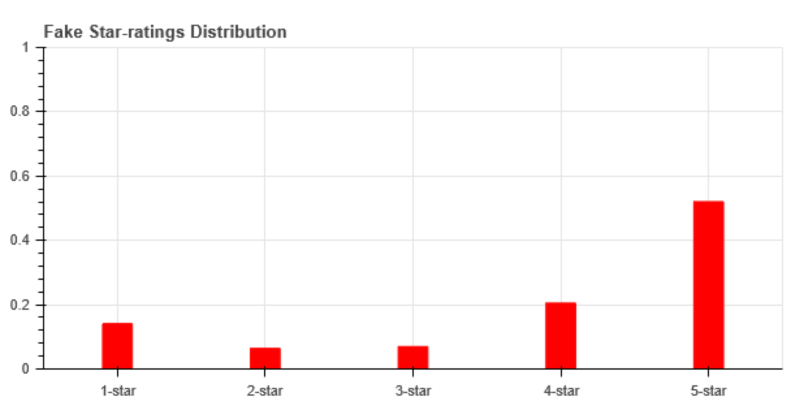
Evidence From Yelp Reviews

- 581407 uncensored, 81588 censored reviews.
- Some of the features are review length, whether reviewer has a profile picture, number of reviewer's friends, star-rating, average star-rating given by reviewer so far, the duration reviewer has been a member, etc.

Evidence From Yelp Reviews



Evidence From Yelp Reviews



Evidence From Yelp Reviews

Method

$$y^*_{it} = X_{it}\beta + b_i + \tau_t + \epsilon_{it} \tag{1}$$

y^*_{it} is the number of fake reviews, X_{it} is a covariate matrix, b_i is business effect, and τ_t is time effect.

y^*_{it} is not observable.

Evidence From Yelp Reviews
Method

Assumption:

- $Pr[Filtered|\neg Fake] = a_0$
- $Pr[Filtered|Fake] = a_0 + a_1$

Then,

$$\begin{aligned} y_{itk} &= \alpha_0(1 - y^*_{itk}) + (\alpha_0 + \alpha_1)y^*_{itk} \\ \sum_{k=1}^{n_{it}} y_{itk} &= \sum_{k=1}^{n_{it}} \left[\alpha_0(1 - y^*_{itk}) + (\alpha_0 + \alpha_1)y^*_{itk}\right] \\ y_{it} &= \alpha_0 n_{it} + \alpha_1 y^*_{it} \end{aligned} \tag{2}$$

Evidence From Yelp Reviews

Method

Substituting (1) to (2):

$$y_{it} = \alpha_0 n_{it} + \alpha_1 (X_{it}\beta + b_i + \tau_t + \epsilon_{it}) \tag{3}$$

Within estimator:

$$\begin{aligned} y_{it} + \bar{y} - \bar{y}_t - \bar{y}_i &= \alpha_0 (n_{it} + \bar{n} - \bar{n}_t - \bar{n}_i) \\ &\quad + \alpha_1 \beta (X_{it} + \bar{X} - \bar{X}_t - \bar{X}_i) \\ &\quad + \epsilon_{it} + \bar{\epsilon} - \bar{\epsilon}_t - \bar{\epsilon}_i \\ \ddot{y}_{it} &= \alpha_0 \ddot{n}_{it} + \alpha_1 \beta \ddot{X}_{it} + \ddot{\epsilon}_{it} \end{aligned} \tag{4}$$

Evidence From Yelp Reviews

Results

The effects we see:

- Higher rating $\Rightarrow$ reputation exploitation $\Rightarrow$ fake reviews next period
- Fake reviews (low effort) $\Rightarrow$ lower rating

Evidence From Yelp Reviews

Results

OLS Regression Results

Dep. Variable:	Fake 5 stars (t+1)	R-squared (uncentered):	0.725
Model:	OLS	Adj. R-squared (uncentered):	0.725
Method:	Least Squares	F-statistic:	1.477e+04
Date:	Wed, 01 Jan 2020	Prob (F-statistic):	0.00
Time:	10:38:52	Log-Likelihood:	-1.3908e+05
No. Observations:	33672	AIC:	2.782e+05
Df Residuals:	33666	BIC:	2.782e+05
Df Model:	6		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Total Reviews	0.1727	0.003	50.768	0.000	0.166	0.179
Rating	0.4706	0.179	2.628	0.009	0.120	0.822
Friends	2.444e-06	6.25e-07	3.912	0.000	1.22e-06	3.67e-06
Review Length	-0.0019	1.82e-05	-104.544	0.000	-0.002	-0.002
Review Count	-0.0004	1.51e-05	-24.790	0.000	-0.000	-0.000
Photo	-0.0667	0.012	-5.745	0.000	-0.089	-0.044

Omnibus:	30553.751	Durbin-Watson:	0.057
Prob(Omnibus):	0.000	Jarque-Bera (JB):	11373300.825
Skew:	3.548	Prob(JB):	0.00
Kurtosis:	92.756	Cond. No.	1.25e+06

Evidence From Yelp Reviews

Results

OLS Regression Results

Dep. Variable:	Real 5-stars	R-squared (uncentered):	0.487
Model:	OLS	Adj. R-squared (uncentered):	0.487
Method:	Least Squares	F-statistic:	5213.
Date:	Wed, 01 Jan 2020	Prob (F-statistic):	0.00
Time:	10:41:24	Log-Likelihood:	-60745.
No. Observations:	33004	AIC:	1.215e+05
Df Residuals:	32998	BIC:	1.216e+05
Df Model:	6		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Total Reviews	0.0206	0.002	11.932	0.000	0.017	0.024
Fake 5-stars	-0.0347	0.005	-7.303	0.000	-0.044	-0.025
Review Length	0.0001	2.65e-05	3.984	0.000	5.37e-05	0.000
Friends	6.698e-06	4.78e-07	14.010	0.000	5.76e-06	7.63e-06
Review Count	-0.0002	1.36e-05	-16.386	0.000	-0.000	-0.000
Photo	0.3065	0.003	100.217	0.000	0.301	0.313

Omnibus:	9230.628	Durbin-Watson:	1.537
Prob(Omnibus):	0.000	Jarque-Bera (JB):	651068.962
Skew:	-0.429	Prob(JB):	0.00
Kurtosis:	24.742	Cond. No.	1.34e+04

Chapter 1: A Static Setup

Media Bias Out Of The Blue

- Previous literature: Media bias originates either from supply or demand, i.e. exogenous asymmetry.
- This paper: Media bias can emerge out of the blue.
- How? I define bias over the difference between the numbers of journalists from two opposite political camps.

Media Bias Out Of The Blue

Environment

- A continuum of voters and a media outlet.
- A binary state of the world: $s \in S = \{L, R\}$.
- A voter has a belief θ on S where $\theta \sim U[0,1]$.
- A public signal $\eta \in S$:
  $Prob(\eta = L|S = L) = Prob(\eta = R|S = R) = \pi \in [0.5, 1]$.
- 2 types of information source (e.g. journalist):
  - One that is strong when $S = L$: $Prob(\mu = L|S = L) = 1$.
  - The other: $Prob(\mu = R|S = R) = 1$.

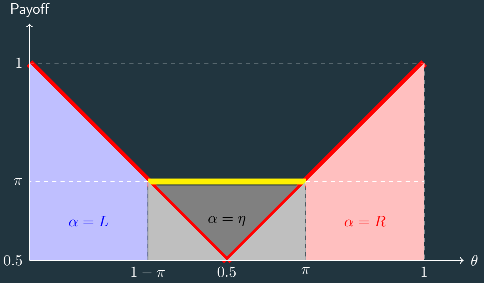
Media Bias Out Of The Blue

Actions and Payoffs

A voter takes action α and receives the following payoff:

$$u(\alpha, S) = \begin{cases} 1 \text{ if } \alpha = S \\ 0 \text{ otherwise} \end{cases}$$

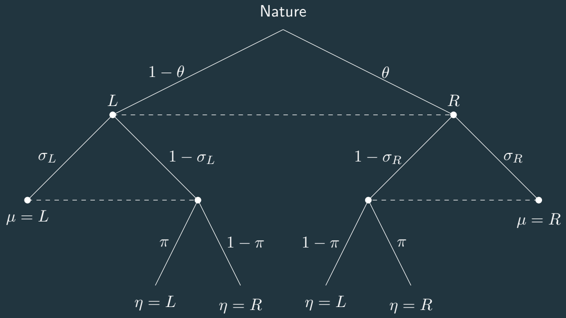
Expected payoff with respect to θ without media:



Media Bias Out Of The Blue

Actions and Payoffs

The media outlet chooses the probabilities of reaching the left-wing and right-wing information sources, σ_L and σ_R, given a budget constraint.



Media Bias Out Of The Blue

Payoffs and Actions

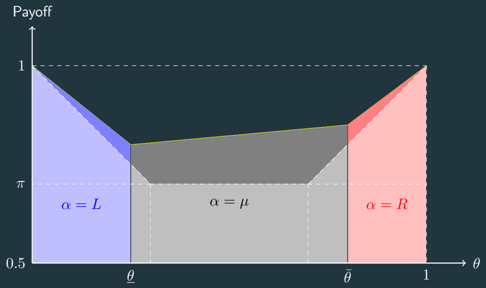
- If the media report conflicts with the public signal, then the true state has been identified by the media report.

Media	Public	Action
L	R	L
R	L	R
R	R	?
L	L	?

Media Bias Out Of The Blue

Results

Suppose $\sigma_R > \sigma_L$.



- The media profit Π is the sum of the dark areas.
- The optimal is $(0, \sigma)$ or $(\sigma, 0)$ where $\sigma > 0$ for any non-convex and a set of convex cost functions.

Media Bias Out Of The Blue

Extension

2 identical media outlets:

Media 1	Media 2	Public	Action
R	L	R	L
R	L	L	R
R	R	L	R
L	R	L	R
L	R	R	L
L	L	R	L
R	R	R	?
L	L	L	?

The optimal is $(\sigma, 0), (0, \sigma)$ where σ is bound by the common budget constraint.