Information Frictions and the Value of Online Reviews

Cayruã Chaves Fonseca

CEMFI

July 2021

Outline

Introduction

Motivation: Big Picture

- Study the impacts of **online reviews** on **consumers** and **firms**
- Focus on the **value** of online reviews as a source of **information**
  - Consumers have incomplete information about firms
  - Acquiring more information possible but costly
  - Reviews are an easy to reach source of information consumers can use
- Empirical application: **restaurant** industry
  - Data: online reviews and activity status (entry / exit)
  - Idea: check whether exposure to online reviews helps (harms) high (low) quality restaurants

Research Questions

1. Do online reviews allow consumers to make better choices?
2. Once we allow firms to optimally choose prices and whether to serve the market, what are the welfare consequences of online reviews?

Related Literature

Related Literature

- **Online Reviews and Consumer Information:**
  - **Online reputation:** Saeedi (2019), Vellodi (2018)
  - **Welfare:** Reimers and Waldfogel (2020)
  - **Consumer learning:** Fang (2019), Luca (2016)
- **Rational Inattention:**
  - **Theory:** Matějka and McKay (2015), Caplin, Dean, and Leahy (2019)
  - **Empirical:** Brown and Jeon (2020), Joo (2020), Bertoli, Moraga, and Guichard (2020), Porcher (2019),
- **Dynamic Oligopoly:**
  - **Theory:** Weintraub, Benkard, and Van Roy (2008)
  - **Empirical:** Iacovone et al. (2015), Qi (2013)

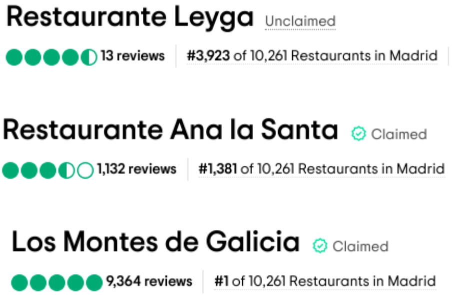
Static Model

Static Model

- **Supply:** heterogeneous restaurants (firms)
  - Online reputation: **n**ot rated, rated **b**elow or **a**bove median
  - Quality: **h**igh or **l**ow
  - Pricing decisions
- **Demand:** rationally inattentive consumers
  - Freely observe online reputation
  - Beliefs about the probability of each restaurant having high quality (depend on online reputation)
  - Can spend resources on gathering information to reduce uncertainty
- **Consumers' decision**:
  - How much and which information to acquire
  - Whether to eat at a restaurant and at which one
  - **Goal:** maximize expected value minus the information cost

Online Reputation

- **Online reputation:** not rated, below median, above median
  - Number of reviews and average rating (easily accessible info)
  - Not rated: zero or few reviews
  - Rated below or above: once reviewed, bad or good rating
  - Note: already have another approach with finer classes



Environment

- There are M restaurants indexed by i, each endowed with online reputation r_i and quality q_i
- Consuming from i delivers $v_i = q_i - p_i$ (outside option offers v_0)
- Online rating is observable and its three values (not rated, below, above) are denoted by $r_i = \{n, b, a\}$
- Quality is not observed by consumers and it is high or low $q_i = \{l, h\}$
- The probability that a firm of online reputation r has high quality is ϕ_r
- Consumers don't observe actual prices set firms but know the pricing rule they use

Environment

- Restaurants have full information: observe quality of all restaurants
- They simultaneously choose prices to maximize profits

$$\max_{p_i} \left(p_i - c\right) DP_i(q_i - p_i\, ;\, q_{-i} - p_{-i}) \tag{1}$$

- D is the number of consumers in the market
- $P_i(q_i - p_i, q_{-i} - p_{-i})$: probability that a consumer chooses firm i
  - Conditional on $i's$ quality, price, as well as qualities and prices of all other restaurants
- We need to look at the problem of the rationally inattentive consumers to solve for P_i

Consumer Priors

- Let $\boldsymbol{v}$ be the vector of payoffs of all alternatives (market state)
- Market state is just one but consumers don't know which one
  - Example: one restaurant with reputation r and the outside option
  - Consumers don't know if $\boldsymbol{v} = (v_0,\ l - p_l)$ or $\boldsymbol{v} = (v_0,\ h - p_h)$
  - They assign probabilities $1 - \phi_r$ and ϕ_r to each case respectively
- The same reasoning applies for a larger number of restaurants
- If $\boldsymbol{k}$ is a vector of qualities for all M restaurants, then

$$G(\boldsymbol{k}) = \prod_r (1 - \phi_r)^{M_r - H_{rk}} \phi_r^{H_{rk}} \tag{2}$$

- $G(\boldsymbol{k})$: probability of distribution of qualities actually being $\boldsymbol{k}$
- M_r: number of restaurants with online reputation r
- H_{rk}: restaurants of reputation r with high quality in $\boldsymbol{k}$

Information Processing

- **Rational inattention:** abstract way to model information processing
- Two stage decision
  1. Choose information strategy to refine prior $G(\boldsymbol{k})$
  2. Choose best restaurant (or eat at home) given posterior
- **Information strategy:** any joint distribution of signals and states
- The key aspect is the assumption on the cost of information:
  - A unit cost λ times the amount of information processed
  - Information processed = expected reduction in entropy from prior to posterior (convex function)
- **Trade-off** of rational inattentive consumer:
  - More information $\Rightarrow$ better expected choices
  - More information $\Rightarrow$ higher search costs

The Transformed Problem

- Problem can be framed as if consumers were selecting conditional choice probabilities (Matějka and McKay 2015)

$$\max_{P_i(\boldsymbol{k})} \sum_{k}\sum_{i=0}^{M} v_i(\boldsymbol{k})P_i(\boldsymbol{k})G(\boldsymbol{k}) \\ -\lambda\left[-\sum_{i=0}^{M} P_i \; log \; Pi + \sum_{\boldsymbol{k}}\left(\sum_{i=0}^{M} P_i(\boldsymbol{k}) \; log \; P_i(\boldsymbol{k})\right)G(\boldsymbol{k})\right] \tag{3}$$

- Subject to:

$$P_i(\boldsymbol{k}) \geq 0 \qquad \forall \; i \;, \; v \\ \sum_{i} P_i(\boldsymbol{k}) = 1 \quad \forall \; v \tag{4}$$

- $P_i = \sum_{\boldsymbol{k}} P_i(\boldsymbol{k})G(\boldsymbol{k})$ is the unconditional probability of choosing i

Demand: Conditional Choices

- Optimal information strategy induces choices that follow:

$$P_i(\boldsymbol{k}) = \frac{P_i e^{\left(q_i(\boldsymbol{k}) - p_i(\boldsymbol{k})\right)/\lambda}}{P_0 e^{v_o/\lambda} + \sum_i P_i e^{\left(q_i(\boldsymbol{k}) - p_i(\boldsymbol{k})\right)/\lambda}} \;, \quad \text{if} \quad P_i > 0 \tag{5}$$

- Logit "adjusted" by $P_i = \sum_{\boldsymbol{k}} P_i(\boldsymbol{k}) G(\boldsymbol{k})$
  - P_i is endogenous, not a parameter
  - As $\lambda \to 0$: back to a standard logit
- No general closed-form for P_i

$$\sum_{\boldsymbol{k}} \frac{e^{\left(q_i(\boldsymbol{k}) - p_i(\boldsymbol{k})\right)/\lambda}}{P_0 e^{v_o/\lambda} + \sum_i P_i e^{\left(q_i(\boldsymbol{k}) - p_i(\boldsymbol{k})\right)/\lambda}} G(\boldsymbol{k}) \leq 1 \;, \quad \forall \quad i > 0 \tag{6}$$

$$\text{with equality if} \quad P_i > 0$$

Static Equilibrium

- Look for a BNE:
  - Pricing rule $\Rightarrow$ consumer beliefs $\Rightarrow$ firms have no incentives to deviate
- Find unconditional choice probabilities $\boldsymbol{P}$ and prices such that:
  1. Given $\boldsymbol{P}$ (and induced conditional choice probabilities), prices simultaneously maximize profits of every firm
  2. Given prices, $\boldsymbol{P}$ solves the consumer problem
- Given the type of heterogeneity across restaurants, in equilibrium, there will be just:
  - 3 unconditional choice probabilities: (P_r, P_b, P_a)
  - 6 conditional choice probabilities: P_{rq}, $r = n, b, a$ and $q = l, h$

Consumer Welfare

- Consumer welfare here differs from standard discrete choice models

$$W = \sum_{\boldsymbol{k}} \sum_{i} P_i(\boldsymbol{k}) v_i(\boldsymbol{k}) G(\boldsymbol{k}) - C(I) \tag{7}$$

- The cost of information is:

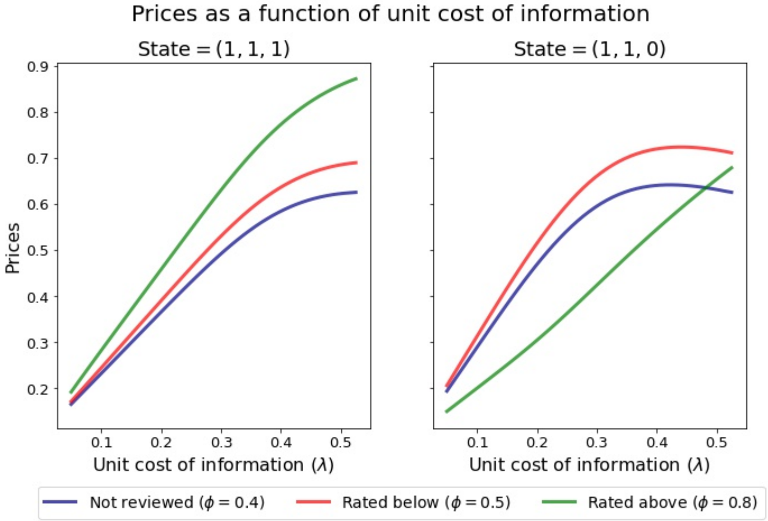
$$C(I) = \lambda \left[- \sum_{i} P_i \; log \; Pi + \sum_{\boldsymbol{k}} \left(\sum_{i} P_i(\boldsymbol{k}) \; log \; P_i(\boldsymbol{k})) \right) G(\boldsymbol{k}) \right] \tag{8}$$

- The intuition:
  - The more state specific are the conditional choices, the more the consumer must have spent processing information

Static Model: Comparative Statics

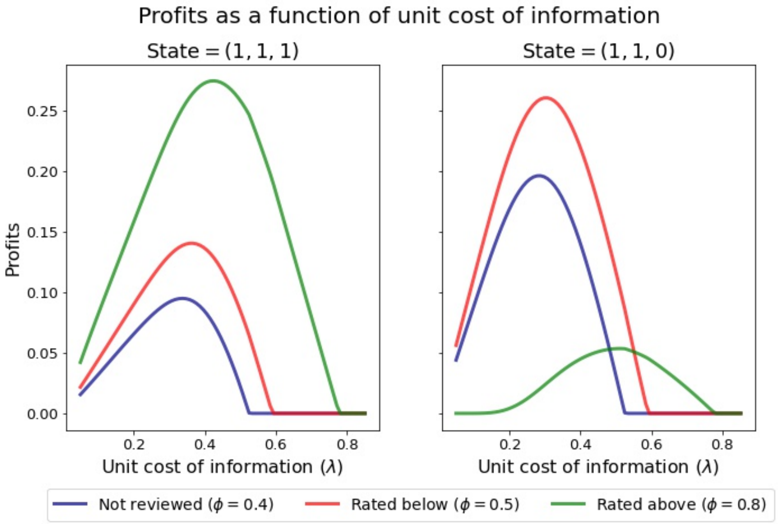
- Compute equilibrium numerically taking as given:
  - Value of outside option: $v_0 = 0$
  - Low quality: $l = 0$
  - High quality: $h = 1$
  - Restaurants' marginal cost: $c = 0.1$
  - Number of restaurants $M = 3$ (one of each online reputation type)
  - Consumer prior beliefs about quality $\phi = (0.4, 0.5, 0.8)$
- Focus is the effect of:
  - Unit cost of information: λ

Prices



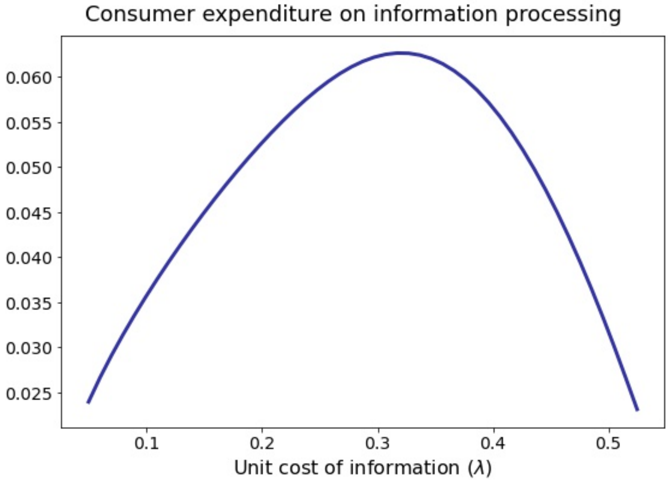
- Higher $\lambda \Rightarrow$ larger dispersion in payoffs

Profits

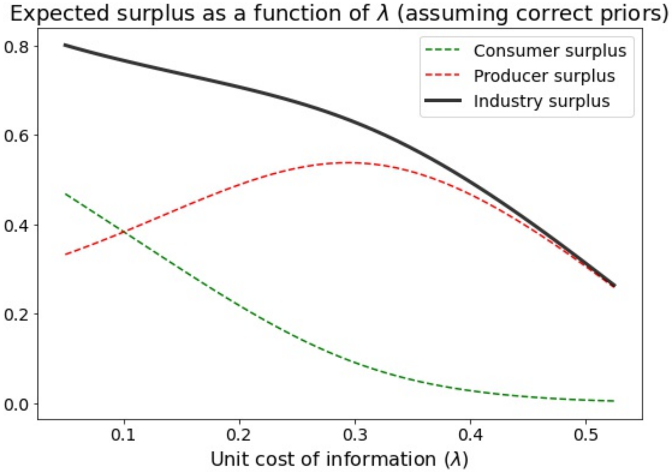


- Firms are better with intermediate values of λ

Consumers' Expenditure on Information



Expected Surplus



Dynamic Model

Model Setup

- **Basics:**
  - Embed previous static setting into an dynamic oligopoly model
  - Oblivious equilibrium: restricted degree of strategic interaction
- **Players and actions:**
  - Consumers: same as static model
  - Firms: incumbents may exit and potential entrants may enter
- At each t, the **timing** is:
  1. Incumbents observe sell-off value and make an exit decision
  2. Potential entrants decide whether to enter and pay the entry cost
  3. Incumbents make price decisions and receive profits (like before)
  4. Exiting firms exit and receive sell-off value
  5. New entrants enter and online reputation of incumbents may change
  6. State of the market updates and next period starts

Evolution of Online Reputation

- Evolution of online reputation $r = \{n, b, a\}$ conditional on quality q
- Exogenous transition rates to be estimated: $\gamma_q^{rr'} = Pr_q(r, r')$
- First approach:
  - Don't endogenize their relationship to demand
  - Dynamic pricing would make problem intractable
- Impose reasonable restrictions:
  - $\gamma_h^{nb} + \gamma_h^{na} > \gamma_l^{nb} + \gamma_l^{na}$: firms with high quality transitions faster from unknown to known (more reviews)
  - $\gamma_h^{na} > \gamma_h^{nb}$ and $\gamma_h^{ba} > \gamma_h^{ab}$: "correct" rating is more likely (the opposite signs if quality is low)
  - $\gamma_q^{nb} > \gamma_q^{ab}$ and $\gamma_q^{na} > \gamma_q^{ba}$: once there are many reviews, less likely to transition over ratings

Evolution of Online Reputation

- One alternative that satisfies the three requirements above is

$$\Gamma_h = \begin{bmatrix} 0.50 & 0.15 & 0.35 \\ 0.00 & 0.75 & 0.25 \\ 0.00 & 0.10 & 0.90 \end{bmatrix} \quad \Gamma_l = \begin{bmatrix} 0.60 & 0.28 & 0.12 \\ 0.00 & 0.90 & 0.10 \\ 0.00 & 0.25 & 0.75 \end{bmatrix} \tag{9}$$

- High quality more likely to transition out of being not rated
- Correct ratings are more likely
- **Not clear:** which quality type should have larger persistence once it transitions out of n
  - High quality receives more reviews: more likely to transition
  - High quality already accumulated more reviews: less likely to transition

Entry and Exit

- **Exit**
  - In each t, incumbents get private info. sell-off value $\psi_{it} \overset{iid}{\sim} Exp(K)$
  - Decide whether to exit (permanently)
- **Entry**
  - In every t, there is a large pool of potential entrants
  - Before entry, quality is uncertain
  - If enters, pay entry cost κ and with probability ω quality is high
  - Always start with online reputation being not reviewed ($r = n$)
  - Equilibrium entry rate will be determined by imposing zero expected profit condition
- Setup time: both decisions only take place in the end of the period

Equilibrium Concept: Oblivious Equilibrium

- Dynamic discrete choice game: standard is MPE
  - Symmetric strategies with all players best responding to each other
  - Strategies depends on current industry state
- **Approximation:** Oblivious Equilibrium (OE) (Weintraub, Benkard, and Van Roy 2008)
  - Intuition: many firms $\Rightarrow$ changes average out $\Rightarrow$ state $\approx$ constant
  - My sample has around 300 restaurants in a neighborhood
- "Close" to optimal decisions based on:
  - Own characteristics: online reputation and quality
  - Long-run average industry state: given an entry rate and competitors' exit behavior
- The industry state is a vector s_t with the number of incumbents of each online reputation and quality type

Long-Run Average State

- Let $\sigma_q(r)$ denote a cutoff exit strategy: exit if $\psi \geq \sigma_q(r)$
- Together with matrix Γ_q, $\sigma_q(r)$ determines "path" of firms
- One period transition: online reputation transition probability times continuation probability

$$Pr_{\sigma_q}(r, r') = \gamma_q^{rr'} \left[1 - e^{\left(-\frac{\sigma_q(r,\phi)}{K}\right)} \right] \tag{10}$$

- Let $Pr^w_{\sigma_q}(r, r')$ the w-period transition probabilities then the expected state in the long-run is:

$$\tilde{s}^q_{\sigma_q,\eta}(r) := \lim_{t \to \infty} \mathbb{E}[\boldsymbol{s}^q_t\,(r)] = \eta\, \omega_q \sum_{w=0}^{\infty} Pr^w_{\sigma_q}(n, r) \tag{11}$$

- $\tilde{s}^q_{\sigma_q,\eta}$ is long-run expected industry state given:
  - Exit strategy of incumbents
  - Entry rate and probability that entrants get low/high quality draw

Value Function

- Value of holding a restaurant of quality q, online reputation r, when competitors use exit strategy σ_q and entry rate is η

$$V_q(r \mid \sigma_l, \sigma_h, \eta) = \pi_q(r\,;\ \tilde{s}_{\sigma,\eta}) + \mathbb{E}_\psi \left[\max \left\{ \psi_{it}\,,\ VC_q(r \mid \sigma_l, \sigma_h, \eta) \right\} \right] \quad (12)$$

- Continuation value is:

$$VC_q(r \mid \sigma_l, \sigma_h, \eta) = \beta \mathbb{E}_{r'} \left[V_q(r' \mid \sigma_l, \sigma_h, \eta) \mid r \right] \quad (13)$$

- Note:
  - I use short-hand $\tilde{s}_{\sigma,\eta}$ to denote that the long-run average state depends on exit strategies of incumbents and the entry rate

Equilibrium Definition

1. Incumbents exit optimally

$$\begin{aligned} \sigma_h(r) &= VC_h(r \mid \sigma_h, \sigma_l, \eta) \\ \sigma_l(r) &= VC_l(r \mid \sigma_h, \sigma_l, \eta) \end{aligned} \tag{14}$$

2. Zero expected entry profits (or there is no entry)

$$\beta\Big[(1-\omega)V_l(n|\sigma_h,\sigma_l,\eta) + \omega V_h(n|\sigma_h,\sigma_l,\eta)\Big] \leq \kappa \\ \text{with equality if } \eta > 0 \tag{15}$$

3. Consumer beliefs are consistent with firm behavior

$$\phi_r = \frac{\tilde{s}_h(r)}{\tilde{s}_l(r) + \tilde{s}_h(r)} \quad , \quad r = n, b, a \tag{16}$$

Algorithm to solve for equilibrium

1. Pick guess for entry rate η
2. Pick guess for share with high quality in each online reputation ϕ_r
3. Given current guesses compute $\sigma(\eta, \phi)$ via value function iteration
    1. Pick a guess of $\sigma(\eta, \phi)$
    2. Compute continuation probabilities
    3. Compute expected industry state $\tilde{s}(\eta, \sigma(\eta, \phi))$
    4. A fixed-point until convergence of $\sigma^*(\eta, \phi)$
4. Compute expected industry state with converged $\tilde{s}(\eta, \sigma^*(\eta, \phi))$
5. Repeat from Step (2) until all three ϕ_r converge (consumer beliefs are consistent)
6. Repeat from Step (1) until zero expected entry profits is met

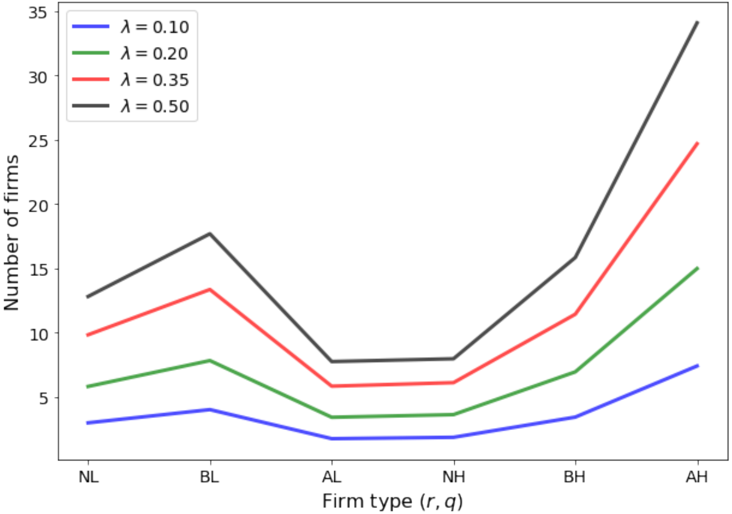
Simulate equilibrium: parameters

Table 1: Model parameters and values used for simulation

Parameter	Governs	Value
c	Marginal cost	1
D	Market size	50
v_0	Value of outside option	1
v_L	Value of low quality	4
v_H	Value of high quality	5
β	Discount factor	0.95
κ	Entry cost	23
ω	Prob. get high quality at entry	0.5
ψ	Mean scrape value	10
$(\gamma_l^{nb}, \gamma_l^{na})$	L-type transitions out of n	(0.20, 0.05)
$(\gamma_h^{nb}, \gamma_h^{na})$	H-type transitions out of n	(0.10, 0.40)
$(\gamma_l^{ba}, \gamma_l^{ab})$	L-type transitions over b, a	(0.15, 0.30)
$(\gamma_h^{ba}, \gamma_h^{ab})$	L-type transitions over b, a	(0.40, 0.20)

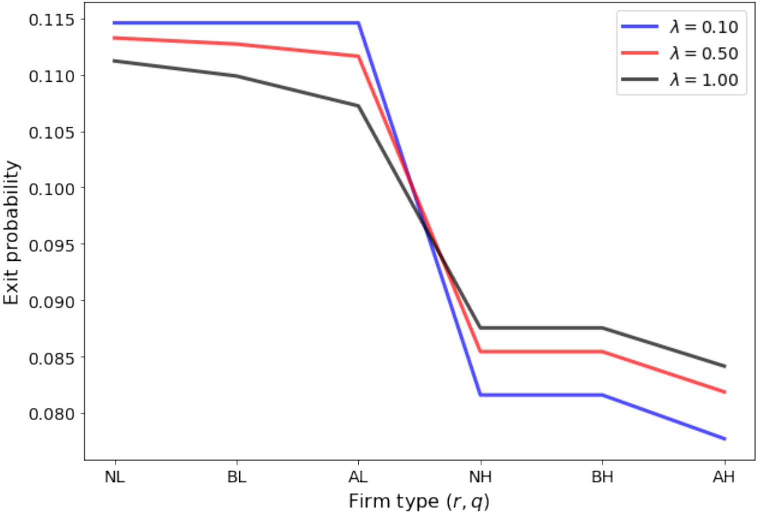
Equilibrium number of firms

Figure 1: Number of firms by type as a function of the cost of information



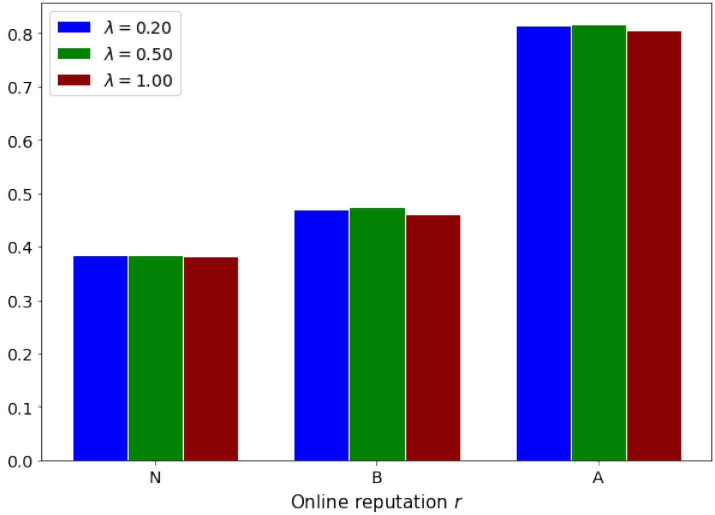
Exit probabilities

Figure 2: Exit probabilities by firm type as a function of the cost of information



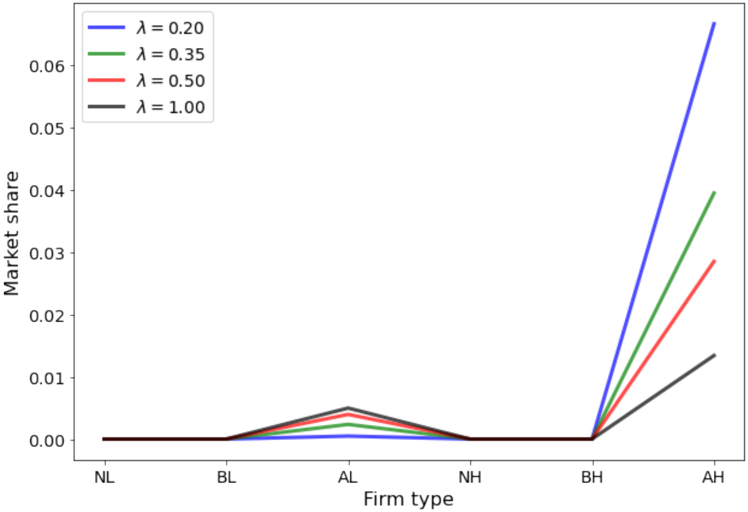
Share of high quality firms

Figure 3: Share of high quality firm by online reputation type



Firm size

Figure 4: Firm-level market shares by firm type and unit cost of information



Data and Empirics

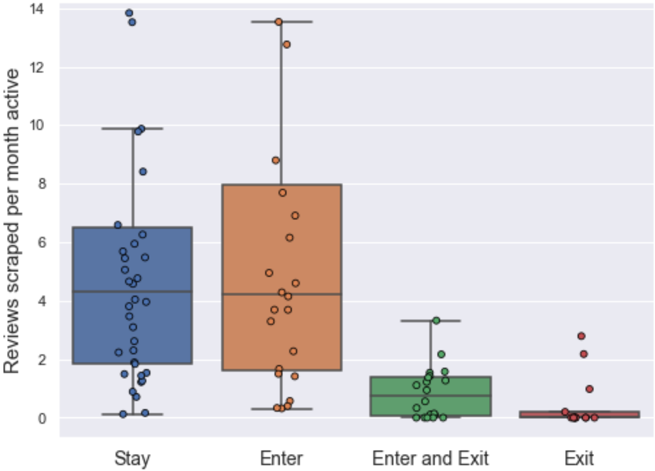
Panel of Restaurant Activity

- **Tripadvisor:**
  - History of reviews of all restaurants in Madrid listed in Jan/2020
  - Around $10,000$ restaurants and 1.2 million reviews
- **Municipal Census of Establishments:**
  - Addresses with a restaurant licence from 2014 to 2019
- **The goal:** build a panel of restaurant activity for the Centro
  - So far small sample with "real" entry/exit: 95 restaurants

Type of Move	Count	Share
Stay	38	0.40
Enter	24	0.25
Enter and Exit	20	0.21
Exit	13	0.14

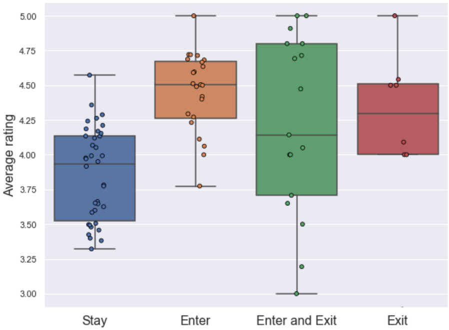
Online Activity by Move Type: Reviews

Figure 5: Restaurant reviews per month by type of move



Online Activity by Move Type: Ratings

Figure 6: Ratings by type of move: restaurant level (left) and rating level (right)

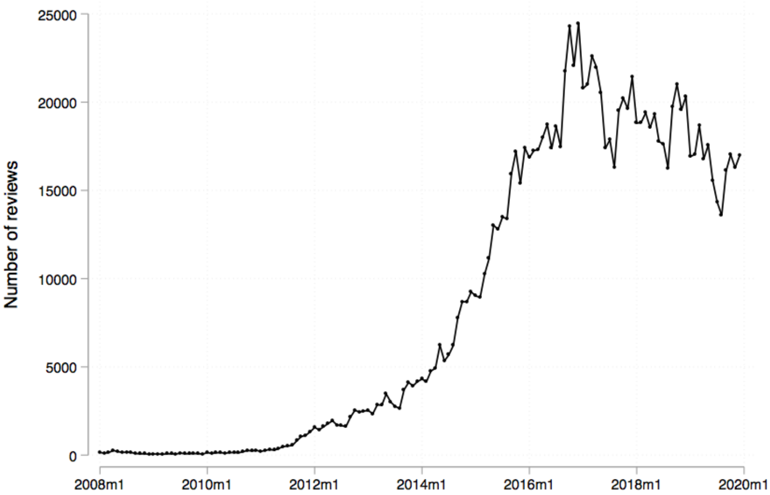


	Rating
Enter	0.501***
	0.017
Enter and Exit	-0.111*
	0.059
Exit	-0271
	0.063
Observations	23881

- Next, I show data from the full Tripadvisor sample

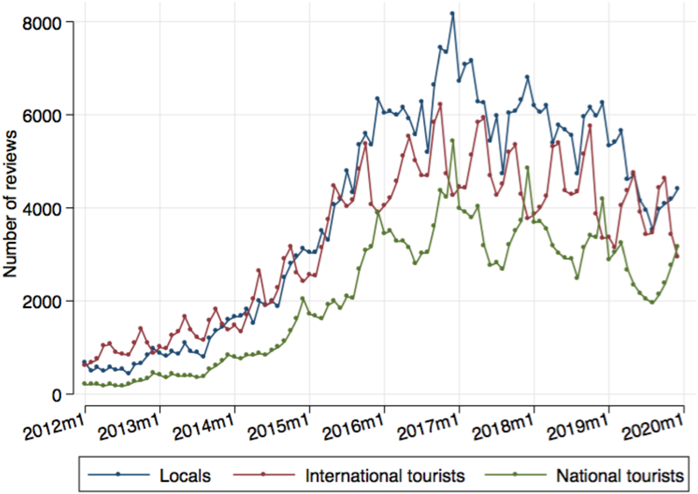
Number of reviews

Figure 7: Number of Tripadvisor reviews



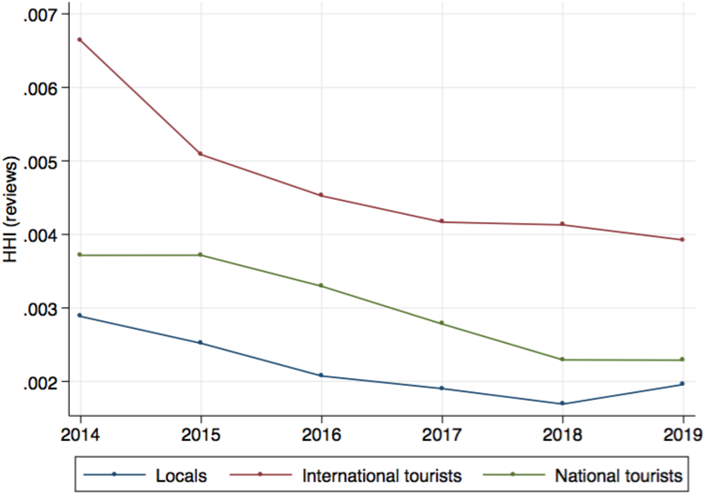
Number of reviews by user location

Figure 8: Total number of reviews by user location



Concentration of reviews

Figure 9: Review based HHI for restaurants in the "Centro" district



Survival and prices as function of online reputation

Table 2: Exit probability and prices as function of online reputation

	(Exit prob.)	(Price)
Reviewed: rated below median	-2.905***	5.984***
	(0.102)	(0.753)
Reviewed: rated above median	-3.240***	8.662***
	(0.144)	(1.045)
Few reviews (type *n* in the model)	-1.606***	17.039***
	(0.077)	(0.644)
Neighborhood FE	Yes	Yes
Year FE	Yes	–
Observations	25866	2902

Estimation / calibration

- The strategy I describe here is a standard nested-fixed point algorithm
- It "matches" model implied and observed transitions and entry rates
- Fix $v_L = 0$ and $v_H = 1$ and group all other parameters into vector θ
- **Entry rates:**
  - Model implied number of entrants per period is Poisson random variable with mean $\eta(\theta)$
  - Thus, prob. of observing k entrants is $p^e(k;\theta) = \frac{\eta(\theta)^k e^{-\eta(\theta)}}{k!}$
- **Incumbent transitions:**
  - Model implies a 6-dim vector of exit probabilities $p^x(\theta) = e^{-\frac{\sigma(\theta)}{K}}$
  - γ parameters are the model imposed transition probabilities over r
  - Probability of observing each transition to incumbents is $p^x(\theta)$ if it an exit and $\gamma_q(r, r')\,[1 - p^x(\theta)]$ if it is a continuation

Likelihood Function

- Contribution of having observed k_t entrants in period t is $p^e(k_t;\theta)$
- Contribution of whatever is observed in period t about incumbent i with quality q is $p^x_q(\theta)^{d_{it}}([1-p^x_q(\theta)]\gamma_q(r,r'))^{1-d_{it}}$
  - d_{it} equals one if firm i exited in period t
  - r and r' represents online reputation of firm i at t and $t+1$ respectively
- If I observed firm quality, then likelihood function would be:

$$L(\theta) = \prod_{i=1}^{N}\prod_{t=0}^{T-1}\Bigg(p^e(k_t;\theta)\ldots \\ \ldots\sum_{q=l,h}\mathbb{I}_i(q)\left[p^x_q(r_{it},\theta)^{d_{it}}\left[(1-p^x_q(r_{it},\theta))\gamma_q(r_{it},r_{it+1})\right]^{1-d_{it}}\right]\Bigg) \tag{17}$$

Likelihood Function

- Structure of the model: there is a ϕ_r probability of any given incumbent type r having high quality

$$L(\theta) = \prod_{i=1}^{N} \prod_{t=0}^{T-1} \Bigg(p^e(k_t;\theta) \\ \phi_{r_{it}}(\theta) \left[p_h^x(r_{it},\theta)^{d_{it}} \left[(1 - p_h^x(r_{it},\theta))\gamma_h(r_{it}, r_{it+1}) \right]^{1-d_{it}} \right] \\ \left(1 - \phi_{r_{it}}(\theta)\right) \left[p_\ell^x(r_{it},\theta)^{d_{it}} \left[(1 - p_\ell^x(r_{it},\theta))\gamma_\ell(r_{it}, r_{it+1}) \right]^{1-d_{it}} \right] \Bigg) \quad (18)$$

- Overall strategy:
  1. Pick guess of θ
  2. Solve for OE and obtain: $p^e(k;\theta), \phi_{r_{it}}(\theta), p_1^x(r,\theta)$
  3. Evaluate likelihood
  4. Repeat until convergence of likelihood function

Challenges and Next Steps

Curse of dimensionality

- Full-solution estimation approach
  - **Solve the entire model** for each guess of the parameter vector
  - Since I am using **OE**, my prior was that this was **possible**
  - In practice, with real data it turns out it is **not feasible**
- What's the bottleneck?
  - Time to solve static **RI** model: increases with number of firms
  - Execution time of **static model:**
    - $\boldsymbol{M} = (1, 1, 1)$: 1.5 seconds, $\boldsymbol{M} = (10, 10, 10)$: 32 seconds
    - $\boldsymbol{M} = (10, 22, 32)$: 3 minutes $\boldsymbol{M} = (50, 50, 50)$: 46 minutes
  - Execution time of **dynamic model:**
    - With parameter values that deliver $\boldsymbol{M} = (10, 22, 30)$: 40 minutes
- **Problem:** used OE to circumvent curse of dimensionality but RI static model has its own curse
- What to do about it?

Why do I need a dynamic model?

- Three types of models:
  1. Static: regress market shares on prices and other product characteristics
  2. Static Entry: zero profit (number of firms as function of market size)
  3. Dynamic: incumbents $\neq$ potential entrants, simultaneous entry/exit...
- Answering the question:
  - Static Model cannot be estimated: market share data not available
  - Static Entry: can be estimated but I think it is important to differentiate entrants from incumbents in my setting
- Why is it important to separate potential entrants from incumbents?
  - If consumers have limited information and are guided by reviews...
  - Then it's very different being an entrant (zero/few reviews) or an incumbent (already accumulated reviews)

Why do I need the RI model?

- **Warning:** I don't have a good answer
- In general, RI delivers empirically supported behavior that depart from standard models
  - **IO** (Brown and Jeon 2020): consumers' probability of choosing cheapest health insurance ("correct choice") varies with "stakes"
  - **Trade/Migration** (Bertoli, Moraga, and Guichard 2020): migrants unresponsiveness to shocks in wages in destination countries
- **Remark:** many of these features could also be rationalized by other models of incomplete information
  - Brown and Jeon (2020) start by showing that sometimes people choose the "wrong" health insurance plan
  - Incomplete information alone could generate that
  - However, they show that probability of choosing cheapest plan is U-shapped in stakes (variance of plan prices)

To which model should it be compared to?

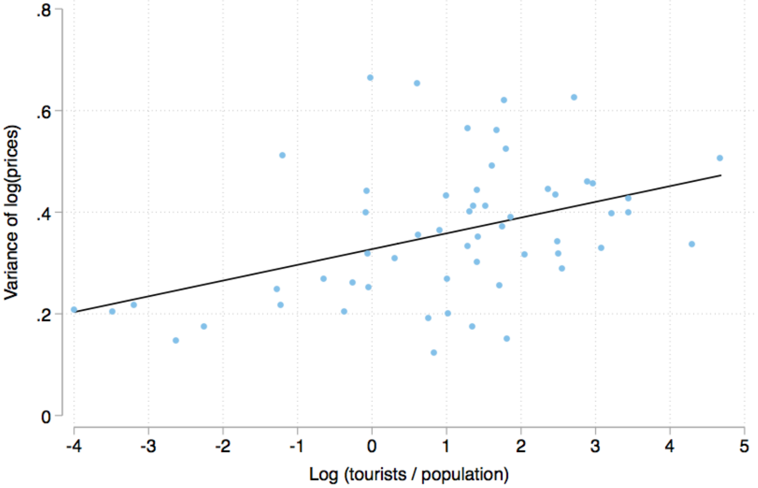
- Bertoli, Moraga, and Guichard (2020)
  - Compare it to a standard random utility model with logit taste shocks
  - Compare RI to a full-information model and argue that adding information frictions to migration decisions is important
  - Why not using another model with incomplete information?
  - They don't discuss it (tractability, ability to compare to other papers)
- Brown and Jeon (2020)
  - First compare to full information logit
  - Then to others models in the literature: "differential weight model"
  - Different coefficients on distinct aspects of price (premium and oop)
  - Contrary to RI information frictions are exogenous
  - Only in the RI model stakes affect information acquisition and deliver U-shaped relation of choice quality and stakes

Two questions to discuss

- **To discuss:**
  1. Looking for evidence of information frictions, what empirical patters are interesting to investigate?
  2. If not RI, what other models should I consider?
- With regards to the first point, I start with a quick look at prices
- Relationship between price dispersion and cost of information
- I use the density of tourists to proxy for cost of information
- With respect to price dispersion:
  - Consider neighborhood as a market
  - I first use the variance of raw prices
  - Then I regress price on review and rating to proxy for quality and look at the variance of quality adjusted prices

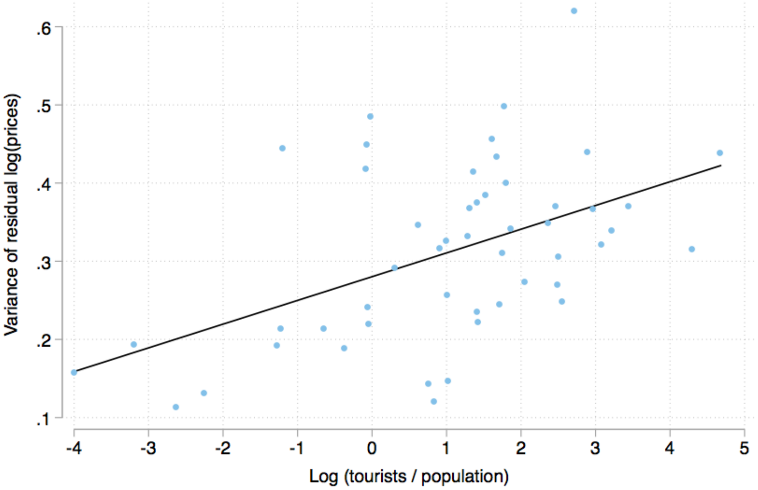
Variance of raw prices and tourist density

Figure 10: Variance of prices as a function of tourist density



Variance of quality adjusted prices and tourist density

Figure 11: Variance of residual prices as a function of tourist density



Variance of quality adjusted prices and tourist density

Table 3: Variance of residual prices and tourist density

	(1)	(2)	(3)	(4)
Log (tourists / population)	0.030***	0.040***	0.042***	0.045***
	(0.008)	(0.011)	(0.012)	(0.016)
Log (reviews / restaurant)		-0.026	-0.015	-0.014
		(0.020)	(0.033)	(0.038)
Log (restaurants in Trip)			-0.021	-0.020
			(0.054)	(0.062)
Log (restaurants)			0.007	0.008
			(0.056)	(0.074)
Socio-demographics	No	No	No	Yes
Observations	50	50	50	49

Note: socio-demographic controls are average house prices, share with university degree, average income, share aged between 20-39 and employment rate

A simpler model: no search

- With respect to point 2 (if not RI, which model) I start with a much simpler framework
  - No rational inattention or any other type of search
  - Standard logit demand as in Weintraub, Benkard, and Van Roy (2008)
- Utility of consumer i from eating at restaurant j at time t is

$$u_{ijt} = \beta \ln(q_{jt}) + \alpha \ln(Y - p_{jt}) + \epsilon_{ijt} \tag{19}$$

- In the standard model quality q_{jt} is assumed to be observed by consumers and firms
- Here I make a slight modification to this utility function

Online reputation and expected quality

- To be able to work with expected quality I assume consumer utility in linear in quality instead of log quality
- Moreover, I assume expected quality depends on online reputation: reviews and ratings

$$\mathbb{E}[q_{jt}|r_{jt}] = \phi_1 \ln(rev_{jt}) + \phi_2 \ln(rat_{jt}) + \phi_3 \ln(rev_{jt}) \ln(rat_{jt}) \tag{20}$$

- Consumers maximize expected utility, which is

$$\begin{aligned} \mathbb{E}_q[u_{ijt}|r_{jt}] &= \beta \mathbb{E}_q[q_{jt}|r_{jt}] + \alpha \ln(Y - p_{jt}) + \epsilon_{ijt} \\ &= \theta_1 \ln(rev_{jt}) + \theta_2 \ln(rat_{jt}) + \theta_3 \ln(rev_{jt}) \ln(rat_{jt}) + \alpha \ln(Y - p_{jt}) + \epsilon_{ijt} \end{aligned} \tag{21}$$

- Where $\theta_n = \beta \phi_n$

Market shares and prices

- Under these assumption market shares are

$$s_{jt} = \frac{e^{\theta_1 \ln(rev_{jt}) + \theta_2 \ln(rat_{jt}) + \theta_3 \ln(rev_{jt}) \ln(rat_{jt}) + \alpha \ln(Y - p_{jt})}}{1 + \sum_k e^{\theta_1 \ln(rev_{kt}) + \theta_2 \ln(rat_{kt}) + \theta_3 \ln(rev_{kt}) \ln(rat_{kt}) + \alpha \ln(Y - p_{kt})}} \tag{22}$$

- Firms simultaneously choose prices to maximize the following profits

$$\pi_{jt} = \max_{p_{jt}} \ (p_{jt} - c) D s_{jt} \tag{23}$$

- FOC imply the following

$$Y - p^*_{jt} + \alpha(p^*_{jt} - c)(s^*_{jt} - 1) = 0 \tag{24}$$

- **Two points** to discuss:
  - Alternatives to this model...
  - Suggestions on how to model evolution of rev_{jt} and rat_{jt}