The Cost Of The Cold-Start Problem On Airbnb

Florian Dendorfer[1] and Regina Seibel[2]

January 09, 2025

[1]Department of Economics, University of Toronto, Email: florian.dendorfer@utoronto.ca
[2]Rotman School of Management, University of Toronto, Email: regina.seibel@rotman.utoronto.ca

The Cold-Start Problem

- Many platforms rely on peer-to-peer reviews: Amazon, Airbnb, Uber, Temu, Expedia...
- Faced with abundance of products, reviews help consumers to identify high quality.
- Purchasing and reviewing previously unreviewed products generates valuable information.
- However, consumers don't take this externality into account when making their purchase decision.

The Cold-Start Problem

- Many platforms rely on peer-to-peer reviews: Amazon, Airbnb, Uber, Temu, Expedia...
- Faced with abundance of products, reviews help consumers to identify high quality.
- Purchasing and reviewing previously unreviewed products generates valuable information.
- However, consumers don't take this externality into account when making their purchase decision.

$\Rightarrow$ Inefficiently low speed of social learning?

The Cold-Start Problem

Maybe, but also sellers respond to social learning:

- Prices: New sellers set low prices to accumulate first reviews
- Entry and Exit: Expected profits are influenced by social learning

The Cold-Start Problem

Maybe, but also sellers respond to social learning:

- Prices: New sellers set low prices to accumulate first reviews
- Entry and Exit: Expected profits are influenced by social learning

amazon Start Grow Services Resources Fees

How to price new products

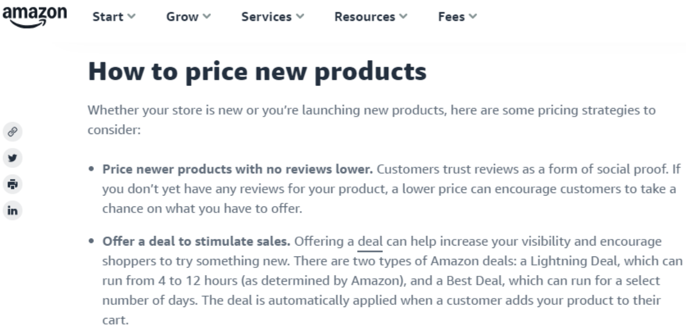
Whether your store is new or you're launching new products, here are some pricing strategies to consider:

- **Price newer products with no reviews lower.** Customers trust reviews as a form of social proof. If you don't yet have any reviews for your product, a lower price can encourage customers to take a chance on what you have to offer.
- **Offer a deal to stimulate sales.** Offering a deal can help increase your visibility and encourage shoppers to try something new. There are two types of Amazon deals: a Lightning Deal, which can run from 4 to 12 hours (as determined by Amazon), and a Best Deal, which can run for a select number of days. The deal is automatically applied when a customer adds your product to their cart.

The Cold-Start Problem

Maybe, but also sellers respond to social learning:

- Prices: New sellers set low prices to accumulate first reviews
- Entry and Exit: Expected profits are influenced by social learning



Should we incentivize consumers to explore more?

Contribution

We estimate a structural model of social learning on Airbnb

$\Rightarrow$ quantify the cost of the cold-start problem taking into account the endogenous supply side

Contribution

We estimate a structural model of social learning on Airbnb

$\Rightarrow$ quantify the cost of the cold-start problem taking into account the endogenous supply side

Theoretical literature on cold-start problem:

- Introducing exploration into recommender system increases welfare
  [Che and Hörner, 2018, Kremer et al., 2014]
- Incorporating endogenous seller pricing decision may alleviate underexploration and even lead to overexploration
  [Bergemann and Välimäki, 1996, Bergemann and Välimäki, 2000, Vellodi, 2022]

Contribution

We estimate a structural model of social learning on Airbnb
$\Rightarrow$ quantify the cost of the cold-start problem taking into account the endogenous supply side

Theoretical literature on cold-start problem:

- Introducing exploration into recommender system increases welfare
  [Che and Hörner, 2018, Kremer et al., 2014]
- Incorporating endogenous seller pricing decision may alleviate underexploration and even lead to overexploration
  [Bergemann and Välimäki, 1996, Bergemann and Välimäki, 2000, Vellodi, 2022]

Empirical literature on dynamic oligopoly:

- Price as investment decision:
  [Dubé et al., 2010, Besanko et al., 2019, Chen, 2016, Ching, 2010]

Model - Demand

- A consumer's indirect utility of renting listing $j \in \mathcal{J}_t$ in $t \in \{1, .., +\infty\}$ is

$$u_{jt} = \gamma \mathbb{E}[\omega_j | K_{jt}, N_{jt}] + \beta_{l(j)} + (1 + f)\alpha p_{jt} + \xi_{jt} + \epsilon_{jt}.$$

  - N_{jt} is j's number of reviews (reported successes & failures) in t,
  - K_{jt} is j's number of *good* reviews (reported successes only) in t,
  - ξ_{jt} captures the unobserved (to us) demand shock,
  - $l(j) \in \{1, 2, 3, 4\}$ is j's **type**.
  - p_{jt} is the price and f is a platform fee.

- As $\omega_j \sim Beta(a, b)$, $\mathbb{E}[\omega_j | K_{jt}, N_{jt}] = \frac{a+K_{jt}}{a+b+N_{jt}}$.

- Listings are **capacity constrained** and can be booked only once in t.

Demand function

Model - State Transitions

- $x_{jt} = (K_{jt}, N_{jt}, I(j))$ is j's state in t.

$$
\begin{tikzcd}[column sep=huge]
 & (K_{jt}, N_{jt}+1, I(j)) \\
(K_{jt}, N_{jt}, I(j)) \arrow[ur, dash, "\textit{bad review}"] \arrow[r, dash, "\textit{no review}"] \arrow[dr, dash, "\textit{good review}"'] & (K_{jt}, N_{jt}, I(j)) \\
 & (K_{jt}+1, N_{jt}+1, I(j))
\end{tikzcd}
$$

- $Pr(\text{good review}) = \text{demand} \times \text{review probability} \times \text{prior}$.
- $Pr(\text{bad review}) = \text{demand} \times \text{review probability} \times (1 - \text{prior})$.
- $Pr(\text{no review}) = 1 - Pr(\text{good review}) - Pr(\text{bad review})$.

Model - Supply

- Active hosts set rental prices and receive the per-period revenue.
- They observe the random, idiosyncratic cost of operating the listing in $t+1$ and decide whether to exit at the end of t.
- Inactive hosts observe random, idiosyncratic entry cost and decide whether to enter at the beginning of $t+1$.
- Hosts and consumer share the same public information
- Review outcomes are determined, and the **new state distribution** realizes.
- With endogenous market entry and exit, the model captures how the cold-start problem affects the number of active listings, as in [Vellodi, 2022].

Equilibrium

- We characterize the symmetric **oblivious equilibrium** to approximate the Markov perfect equilibrium [Weintraub et al., 2008].
- Hosts choose their strategies (rental price, exit, entry) based on their own state and knowledge of the long-run average industry state.
- They ignore strategic effects on competitors' entry and exit decision.

Data

- We use data from AirDNA on all Airbnb listings in Manhattan, NY, from 2016 to 2019.
- Entire apartments, ≤ 2 guests, no pets, 1 bathroom, 1 bedroom, ≥ 1 picture.
- After cleaning the data (and imputing missing values), we have rental prices, bookings, number of reviews, and ratings for 7,687 listings and 62,937 listing-months.
- K is the number of five-star reviews required to achieve the observed average rating, if $N - K$ reviews were one-star reviews.

Summary stats

Listing types

Calibrated Parameters

Demand estimation

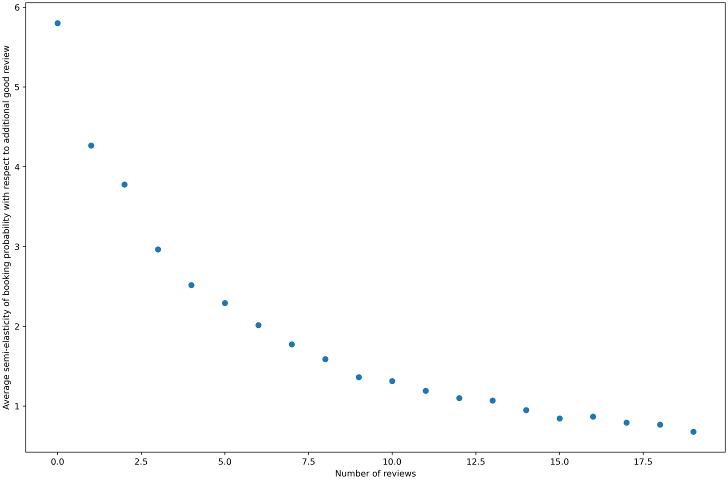
		w/o instruments		w/ instruments	
prior	ψ	-0.1810	(0.1463)	1.3171**	(0.7421)
	ι	1.9633***	(0.1233)	1.6212	(1.1995)
rental price	α	-0.0020***	(0.0001)	-0.0086***	(0.0015)
types	β_1	-10.6106***	(0.1109)	-10.5354***	(1.7158)
	β_2	-10.1166***	(0.1110)	-9.8218***	(1.7193)
	β_3	-9.6580***	(0.1116)	-9.4401***	(1.7221)
	β_4	-9.2443***	(0.1114)	-8.9099***	(1.7415)
expected quality	γ	1.6000**	(0.1358)	2.8607	(1.8711)
Observations		49,214		25,824	

Note: *p<0.1; **p<0.05; ***p<0.01. Robust standard errors in parenthesis.

- Prior estimates imply: $\mathbb{E}[\omega_j|0,0] \approx 4.15$ (a $=$ 3.99 & b$=$1.07)
- The average own-price elasticity is **-1.04**.

Instruments

Review Semi-elasticities of Demand



Cost Estimation

- We estimate the type-specific cost parameters by Maximum Likelihood targeting the state distribution.
- Based on the demand estimates, we repeatedly solve the model for different parameter values.
- Estimates imply the following average incurred costs:

Type	1	2	3	4
Average entry cost	\$1567	\$2756	\$3654	\$4899
Average operating cost	\$1339	\$2263	\$2984	\$3656

Model Fit

Counterfactual 1

- If we could impose *relative price changes* on hosts to change the speed of social learning and maximize welfare, would we
  - want increase social learning?
  - be able to achieve substantial welfare gains?

- To make this problem manageable: combination of review dependent per-booking subsidies/taxes with lump-sum transfer.

- Define 5 review intervals: [0-1], [2-5], [6-10], [11-15], [16-20]

- BUT: Avoid results being conflated with change in overall demand for Airbnb hosts!
  - This affects welfare for reasons beyond cold-start problem

  $\Rightarrow$ Constrain Taxes/subsidies to be zero on average

Counterfactual 1

	total	*listings within review interval*				
		[0,1]	[2,5]	[6,10]	[11,15]	[16,20]
Δ avg rental price	1.5%	-16.6%	-3.9%	2.4%	5.0%	9.7%
Δ avg occupancy rate	2.4%	13.9%	3.4%	-1.6%	-3.6 %	-6.4%
Δ # listings	4.6%	-18.1%	-9.4%	-1.4%	5.6%	54.0%

- Total welfare $\uparrow$ **$340,000**
- Consumer surplus $\uparrow$ \$350,000, Airbnb host surplus $\downarrow$ \$57,000, Airbnb revenue $\uparrow$ \$45,000.
- To put things in perspective: Total host revenue per month in Manhattan $\sim$ \$4 million.

Counterfactual 2

- We also run a counterfactual where *only entrants with no reviews* are forced to change their price.
- Other hosts are allowed to set individually optimal prices.
- The optimal price decrease ranges by entrant type but lies between 8% and 15%.
- The overall number of listings decreases which harms consumers.
- This harm is more than compensated by the gain from faster social learning.

$\Rightarrow$ Consumer gain around \$44,000 overall, but Airbnb hosts and Airbnb revenue are harmed.

$\Rightarrow$ Total welfare increase is only \$21,000.

Conclusion

- We have estimated an empirical model of social learning using Airbnb data.
- We find that the cold-start problem exists on Airbnb and is quantitatively significant.
- BUT: The supply-side response matters for the extent of the problem and the effectiveness of counterfactual interventions.
- Social learning matters for any platform that relies on peer-to-peer reviews.

Appendix

References I

Bergemann, D. and Välimäki, J. (1996).
Learning and strategic pricing.
Econometrica, 64(5):1125–1149.

Bergemann, D. and Välimäki, J. (2000).
Experimentation in markets.
The Review of Economic Studies, 67(2):213–234.

Berry, S., Levinsohn, J., and Pakes, A. (1995).
Automobile prices in market equilibrium.
Econometrica, pages 841–890.

Besanko, D., Doraszelski, U., and Kryukov, Y. (2019).
How efficient is dynamic competition? the case of price as investment.
American Economic Review, 109(9):3339–64.

Che, Y.-K. and Hörner, J. (2018).
Recommender systems as mechanisms for social learning.
The Quarterly Journal of Economics, 133(2):871–925.

Chen, J. (2016).
How do switching costs affect market concentration and prices in network industries?
The Journal of Industrial Economics, 64(2):226–254.

References II

Ching, A. T. (2010).
A dynamic oligopoly structural model for the prescription drug market after patent expiration.
International Economic Review, 51(4):1175–1207.

Dubé, J.-P. H., Hitsch, G. J., and Chintagunta, P. K. (2010).
Tipping and concentration in markets with indirect network effects.
Marketing Science, 29(2):216–249.

Kremer, I., Mansour, Y., and Perry, M. (2014).
Implementing the "wisdom of the crowd".
Journal of Political Economy, 122(5):988–1012.

Vellodi, N. (2022).
Ratings design and barriers to entry.
Working Paper.

Weintraub, G. Y., Benkard, C. L., and Van Roy, B. (2008).
Markov perfect industry dynamics with many firms.
Econometrica, 76(6):1375–1411.

Demand

- Number of arriving consumers is Poisson distributed with mean μ.
- Define $u_{jt} = \nu_{jt} + \epsilon_{jt}$.
- Demand $q(\mathbf{p}_t, x_{jt})$ for j in t is

$$q(p_{jt}, x_{jt}, P_t, s_t) = 1 - \exp\left(-\mu \frac{\exp(\nu(p_{jt}, x_{jt}))}{1 + \exp(\nu(p_{jt}, x_{jt})) + \sum_x^X (s_t(x) - \mathbb{1}_{x = x_{jt}}) \exp(\nu(P_t(x), x))}\right)$$

Back

Data Summary

	mean	std	min	25%	50%	75%	max
Rental price	$193.02	$60.13	$70.33	$150.75	$184.78	$270.95	$562.43
Occupancy rate	60.64%	33.57%	0.00%	33.33%	69.23%	100.00%	100.00%
Number of reviews	10.30	8.19	0.00	2.00	9.00	20.00	20.00
Rating	4.51	0.72	1.00	4.40	4.67	5.00	5.00
Monthly exit rate	3.21%	0.80%	1.57%	2.68%	3.21%	4.08%	5.55%
Monthly entry rate	4.40%	1.85%	0.39%	3.17%	4.36%	6.26%	9.83%
Lifespan (in months)	17.67	16.38	1.00	3.00	12.00	39.00	52.00

Back

Data

To construct types we regress demand on rental price, $K - N$, month-year, and listing fixed effects and divide the estimated listing fixed effect coefficients into 4 quartiles (types 1-4).

	Price	**Occupancy**	**Reviews**	**Rating**	**Lifespan**
type 1	$186.30	34.60%	9.89	4.39 stars	14.6 months
type 2	$192.83	53.47%	10.48	4.48 stars	20.2 months
type 3	$189.45	68.58%	11.04	4.56 stars	21.2 months
type 4	$204.40	80.92%	9.32	4.58 stars	14.7 months

Back

Calibrated Parameters

Parameter		**Value**
Discount factor	δ	0.995
Revenue fee	f	0.142
Arrival rate	μ	10,000
Review probability	υ_r	0.992
Maximum number of reviews	$\bar{N}$	20
Maximum number of listings	J	10,000

Back

Demand estimation

- We invert aggregated demand and estimate the demand parameters using GMM [Berry et al., 1995].
- We instrument the rental price with the average reservation length of the listing.
- We use the de-meaned lags of of the occupancy rates and the rating as instruments for the number of (good) reviews.

Back

Model Fit

	Data	Model
Average number of active listings	1,210	1,152
Share of unreviewed listings	15%	20%
Average rating	10.3	7.3
Average rental price	\$193	\$183
Average occupancy rate	61%	60%
Average exit rate	3.2%	13%

Back