Implications Dynamic Pricing: Experimental evidence

Przemysław Jeziorski  
UC Berkeley

Nadia Michelidaki  
Keybee Inc.

SICS 2019
Motivation

▶ E-commerce has given companies more opportunity to price discriminate, leading to “Dynamic pricing.”

*Ability to frequently change prices taking into account market conditions, such as timing, demand-level and competitors’ prices.*

▶ Particularly important in the industries with capacity constraints and stochastic demand.

**What is the impact of dynamic pricing on market efficiency, i.e. consumer and producer surplus?**

**Which forms of dynamic pricing are more effective?**
Contribution

- Clean identification of causal effects via field experiment
- We execute the project in 3 steps:
  - We designed field experiment that enables estimation of dynamic demand using simple logistic regression without extra instruments.
  - We use a structural model to compute counterfactual pricing regimes (flat pricing, dynamic pricing).
  - We test implications of dynamic pricing directly using a second field experiment (currently in the field).
- We take price endogeneity and selection seriously
Context

- Application: Short-term rental market (STR) – Airbnb, Homeaway, Booking.com
- Modal capacity is one
- Listing-days are priced depending on the date and days to reservation
- Other industries with similar characteristics: airlines, events with limited capacity, shared-economy, any industry with limited capacity selling out over time.
Findings

- Optimal intertemporal pricing increases firms profits by 6%, increases customer welfare by 26%.
- Optimal intra date price discrimination increases firms profits by 5%, but decreases customer welfare by 15%.
- There is an interaction between intertemporal and intradate price discrimination: price dispersion across dates is high for reservations far away, but decreases as the reservation gets closer.
  - Consumers arriving early and shopping for busy dates are worse off.
  - Consumers arriving late are better off for all dates.
Dynamic Pricing

Three reasons to do dynamic pricing:

▶ Variation in demand across reservation dates: intra-date price discrimination.
  ▶ Example: UC Berkeley graduation dates have higher prices.

▶ For a fixed reservation date, preferences of consumers vary depending on the number of days to the reservation:
  ▶ Example: Guests that try to book 3 months in advance are different than people trying to book for tomorrow.

▶ Hosts have an option value of selling the same reservation-date tomorrow.

▶ Other reasons outside of scope: price skimming, cost variation, bundling...
Profit and welfare impact

- Capacity of 1.
- Dynamic pricing may improve utilization and lower dead weight loss.
- Positive impact on profits in monopoly, but size unknown.
- Effect of welfare is ambiguous.
- Decomposition of the effect into different forms of dynamic pricing is unknown.
Why do we need an experiment. Part 1.

- Demand estimation:
  - Classical price endogeneity: zero marginal cost - hard to find credible cost shifters.
  - Remaining capacity is endogenous: capacity of 1 - being still on the market is endogenous leading to Heckman-style selection.
    - Example: If you are still on the market, demand yesterday was likely low. Thus, demand today is likely to be low, as well. Listings appear in the sample in lower than average demand periods.
  - Hard to find variables affecting attrition but not affecting demand (since attrition is literally driven by demand).

- How it was done in the past? Capacity was modeled using a structural model. We tackle this using randomization.
Why do we need an experiment. Part 2.

- Policy evaluation (currently in the field):
  - Evaluate causal impact of theoretically optimal dynamic pricing instead of relying purely on structural model. Structural model is still indispensable to:
    - Define optimal dynamic pricing
    - Compute theoretically optimal dynamic pricing
    - Provide micro-foundations to compute welfare
  - Test theoretical model
  - Provide managerial methodology
Why do we need a structural model?

- For industry application RNN approach could be immediately profitable.
- Need a model to learn about the mechanism.
  - Example: Model will reveal that we probably will need to include serial correlation of demand (possibly caused by correlated arrival of consumer types, or long-lived consumers).
- We would be able to decompose contribution different microeconomic mechanisms.
- We are able to measure welfare.
- When it stops working one day, it may be easier to debug.
The industry: Short-term rentals

- Short-term rental industry: Airbnb, Homeaway, Tripadvisor (Flipkey), Booking.com
- 100 million reservations
- 4 million listings
- 25b market volume
- 11% of Americans stayed in an STR
The firm: Keybee

- Uber for STR services: guest messaging, apartment turnover (cleaning, key delivery), pricing
- 500+ listings
- Entire US and Greece
- Currently doing ad-hoc dynamic pricing: collection of s-curves that adjust based on occupancy rates within a window centered on the rental date. If rental rates are low, apply more aggressive s-curve. Relies on the correlation in demand for neighboring dates. Manually calibrated.
Pricing experiment

- Random multiplicative noise to prices.
- Randomization on the listings-date-days to reservation level.
- Current noise is a classical instrument for current price.
- Lagged noise is an instrument for selection: it affects attrition probability and is uncorrelated with current demand.
- I’ll demonstrate that formally using a model in a few slides.
Data sample

<table>
<thead>
<tr>
<th>variable</th>
<th>mean</th>
<th>se(mean)</th>
<th>sd</th>
<th>min</th>
<th>max</th>
</tr>
</thead>
<tbody>
<tr>
<td>price</td>
<td>113.3744</td>
<td>.1319701</td>
<td>69.39279</td>
<td>30</td>
<td>1425</td>
</tr>
<tr>
<td>days</td>
<td>55.83587</td>
<td>.0653692</td>
<td>34.3726</td>
<td>0</td>
<td>151</td>
</tr>
<tr>
<td>bedrooms</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>booked</td>
<td>.007078</td>
<td>.0001594</td>
<td>.083833</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

- Sample of 39, 1 bedroom apartments across US
- 276,489 observations between 2019-01-01 and 2019-05-20
- Plus/minus 5%-15% multiplicative noise to prices.
Average pricing path
Demand model

- Guest: \( i \)
- Listing: \( j \)
- Reservation date: \( t \)
- Pricing day: \( s \)

We define a product as a pair \((j, t)\). Each product has capacity 1 and is offered for sale at different dates \(s\) (times to the reservation).

There is a population of potential guests with random utility

\[
    u_{ijts} = \alpha + \beta p_{jts} + \nu_{ijts}
\]

\( \nu_{ijtd} \) can be potentially serially correlated and have different distribution across \( j, t, \) and \( d \).
Consider a linear specification

\[ u_{ijts} = \alpha + \gamma d_{jts} + \beta p_{jts} + \xi_{jts} + \epsilon_{ijts} \]

- \( p_{jts} \) is likely to be correlated with \( \xi_{jts} \) (price endogeneity)
  - Listings with more demand are priced higher
  - Dates with more demand are priced higher

- \( d_{jts} \) is likely to be correlated with \( \xi_{jts} \) (selection)
  - Listings with more demand are shorter on the market
  - Dates with more demand are shorter on the market
Fixed-effects approach

- Cannot include $\alpha_{jts}$ fixed effects – one observation per fixed effect
- $\alpha_j$ and $\alpha_t$ fixed effects do not control for all listing-date heterogeneity. Pricing using utilization in neighboring date window causes endogeneity.
- Standard short-panel fixed-effects estimate is biased if demand is correlated over time.
  - $d_{jts}$ is correlated with $\xi_{jts-1}$
  - $\xi_{jts-1}$ is correlated with $\xi_{jts}$
Our data is generated by a following model

\[ u_{ijtd} = \alpha + \gamma d + \beta \tilde{p}_{jts} + \xi_{jts} + \epsilon_{ijtd} \]

where \( \tilde{p}_{jts} = \eta_{jts} p_{jts} \), \( \eta \) is random noise.

\( \eta_{jts} \) is a good instrument for \( \tilde{p}_{jts} \)

\( \eta_{jts-1} \) is a good instrument for \( d_{jts} \) – correlated with past demand but orthogonal with contemporaneous demand by construction.
Two-step control function estimation

- First step:
  - Regress \( \tilde{\eta}_{jts} \) on \( \eta_{jts} \) with listing fixed effects
  - Regress \( d \) on \( \eta_{jts-1} \) with listing fixed effects
  - Consider residuals \( \mu^p \), and \( \mu^d \)
  - Estimate the following model

\[
U_{ijts} = \alpha_j + \gamma d_{jts} + \beta \tilde{\eta}_{jts} + f(\mu^d_{jts}, \mu^p_{jts}) + \xi_{jts} + \varepsilon_{ijts}
\]
Second step estimation

- We face incidental parameter problem with respect to listing fixed effects.
- We would like to allow the error term to be correlated across listing and dates.
- Assume that marginal distribution of $\xi_{jts} + \epsilon_{ijts}$ is EV(1).
- Estimate using conditional pseudo-likelihood and cluster standard errors.
- Conditional likelihood eliminates the listing fixed effects.
- Pseudo-likelihood enables consistent (though not efficient).
- Correct for correlated error terms with clustered standard errors.
Demand results

<table>
<thead>
<tr>
<th></th>
<th>(1) booked</th>
<th>(2) booked</th>
<th>(3) booked</th>
<th>(4) booked</th>
<th>(5) booked</th>
<th>(6) booked</th>
</tr>
</thead>
<tbody>
<tr>
<td>Price</td>
<td>-0.0177***</td>
<td>-0.0504***</td>
<td>-0.0242***</td>
<td>-0.0122**</td>
<td>-0.0122**</td>
<td>-0.0133***</td>
</tr>
<tr>
<td></td>
<td>(0.000834)</td>
<td>(0.00134)</td>
<td>(0.00137)</td>
<td>(0.00477)</td>
<td>(0.00476)</td>
<td>(0.00476)</td>
</tr>
<tr>
<td>Remaining Days</td>
<td>-0.0336***</td>
<td>-0.0332***</td>
<td>-0.0480***</td>
<td>-0.0303**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00122)</td>
<td>(0.00123)</td>
<td>(0.0148)</td>
<td>(0.0149)</td>
<td>(0.0149)</td>
<td>(0.0152)</td>
</tr>
<tr>
<td>Residuals (price)</td>
<td></td>
<td></td>
<td>-0.0127***</td>
<td>-0.0125***</td>
<td>-0.0240***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.00486)</td>
<td>(0.00486)</td>
<td>(0.00538)</td>
<td></td>
</tr>
<tr>
<td>Residuals (days)</td>
<td></td>
<td></td>
<td>0.0147</td>
<td>-0.0200</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.0147)</td>
<td>(0.0152)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Listing FE</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Non-parametric CF</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>N</td>
<td>278654</td>
<td>278654</td>
<td>278654</td>
<td>278654</td>
<td>278654</td>
<td>278654</td>
</tr>
</tbody>
</table>

- Controlling for listing FEs removes part of classical price endogeneity.
### Demand results

<table>
<thead>
<tr>
<th></th>
<th>(1) booked</th>
<th>(2) booked</th>
<th>(3) booked</th>
<th>(4) booked</th>
<th>(5) booked</th>
<th>(6) booked</th>
</tr>
</thead>
<tbody>
<tr>
<td>Price</td>
<td>-0.0177***</td>
<td>-0.0504***</td>
<td>-0.0242***</td>
<td>-0.0122**</td>
<td>-0.0122**</td>
<td>-0.0133***</td>
</tr>
<tr>
<td></td>
<td>(0.000834)</td>
<td>(0.00134)</td>
<td>(0.00137)</td>
<td>(0.00477)</td>
<td>(0.00476)</td>
<td>(0.00476)</td>
</tr>
<tr>
<td>Remaining Days</td>
<td>-0.0336***</td>
<td>-0.0332**</td>
<td>-0.0480***</td>
<td>-0.0303**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00122)</td>
<td>(0.00123)</td>
<td>(0.0148)</td>
<td>(0.0149)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Residuals (price)</td>
<td>-0.0127***</td>
<td>-0.0125***</td>
<td>-0.0240***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00486)</td>
<td>(0.00486)</td>
<td>(0.00538)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Residuals (days)</td>
<td></td>
<td></td>
<td>0.0147</td>
<td>-0.0200</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.0147)</td>
<td>(0.0152)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Listing FE</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Non-parametric CF</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>N</td>
<td>278654</td>
<td>278654</td>
<td>278654</td>
<td>278654</td>
<td>278654</td>
<td>278654</td>
</tr>
</tbody>
</table>

▶ Controlling for “days to reservation” removes part of intertemporal price endogeneity.
▶ Early days have lower (and more inelastic) demand, and they are priced higher – opposite of Berry (94)
▶ Firms is not incorporating intertemporal unobserved heterogeneity explicitly, we obtain price endogeneity anyways.
Bimodal distribution of listing fixed-effects
Listing Fixed Effects

- Larger heterogeneity in demand across dates than across listings
Profit function

- Consider pricing matrix

\[ p = [p_{ts}]_{t=1,\ldots,T; s=1,\ldots,S} \]

- Profit of the listing is given by

\[
\sum_{t=1}^{T} E_p \sum_{s=1}^{S} I(\tau_t = s) p_{ts}
\]

- Recall: \( t \) is reservation date, \( s \) is pricing date
- \( \tau_t \) is stopping time of rental,
- If \( \tau_t = s \) the apartment rented \( s \) days before the reservation date
- Expectation is taken with respect to stopping time, conditional on \( p \)
Fully dynamic pricing

Choose unrestricted pricing matrix

\[ p = \begin{bmatrix} p_{ts} \end{bmatrix}_{t=1, \ldots, T; s=1, \ldots, S} \]

to maximize

\[ \sum_{t=1}^{T} E_p \sum_{s=1}^{S} I(\tau_t = s) p_{ts} \]

One can solve it using dynamic programming or simple maximization using Newton method
Restricted dynamic pricing

- Condition 1: Do not allow to price discriminate across reservation dates, i.e. allow only pricing matrices $p$ such that

\[ p_{ts} = \bar{p}_s \]

- Condition 2: Allow one price.

\[ p_{ts} = \bar{p} \]
Pricing paths

- Red line – restricted dynamic pricing
- Black line – common optimal price
## Counterfactual metrics

<table>
<thead>
<tr>
<th>Metric</th>
<th>Fully dynamic pricing</th>
<th>Only intertemporal</th>
<th>Flat price</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expected profits</td>
<td>$111.8</td>
<td>$106.8</td>
<td>$101.1</td>
</tr>
<tr>
<td>Consumer surplus</td>
<td>$107.6</td>
<td>$124.2</td>
<td>$97.6</td>
</tr>
<tr>
<td>Occupancy rate</td>
<td>74%</td>
<td>69%</td>
<td>62%</td>
</tr>
<tr>
<td>Average price</td>
<td>$157.2</td>
<td>$173.3</td>
<td>$161.8</td>
</tr>
</tbody>
</table>
Intertemporal price discrimination vs option value

- Two reasons for the price to change over time for the same date:
  - Option value of renting later
  - Different consumer arriving at different times
- To separate these effects we reshuffle customers across time randomly and recompute optimal price path
Random distribution of customers

- Red – baseline
- Green – random distribution of customers over time
- The time variation of optimal prices is muted
- We recompute profits using true distribution of customers over time:
  - Full intertemporal: $106
  - Random distribution of customers: $104
  - Flat price: $101
Competition

- Every firm is small
- Unilateral deviation of one firm is unlikely to trigger other firms to respond
- Coordinated change of behavior of all firms may matter
Stage 2 of the RCT

- Prices early on are too high
- Partly in the field, but need to fix pricing patterns before full rollout
- Demand is considered IID over time in the counterfactual, may cause the price to grow too fast
  - Ex: if demand is perfectly correlated the flat price is optimal
What’s next?

- We need serial correlation in demand and possibly learning (price skimming).
- When we get reasonable predictions we make a randomized rollout.
- Equilibrium effects? (Nonstationary oblivious equilibrium).
- New products:
  - Project fixed effects on the unstructured data from the listing.
  - Predict fixed effects for new listings.
- Endogenous reviews: higher price, lower review. We don’t want to skin the customer.
- Bundling (ex. can only book two days on weekend), Dynamic Bundling (cannot book short reservations early on).
Conclusion

- We show that dynamic pricing raises both profits and welfare.
- Intertemporal price discrimination is equally important and intraday price discrimination.
- Important interaction between intertemporal and intraday pricing.
- Option value is more important than difference in types of consumer that arrive over time.