Can Willingness to Pay be Identified without Price Variation?

What Big Data Can (and Cannot) Tell Us

Cheng Chou [1] Vineet Kumar [2]

[1]University of Leicester, UK

[2]Yale University, USA

Washington University
January 2020

Subscription business

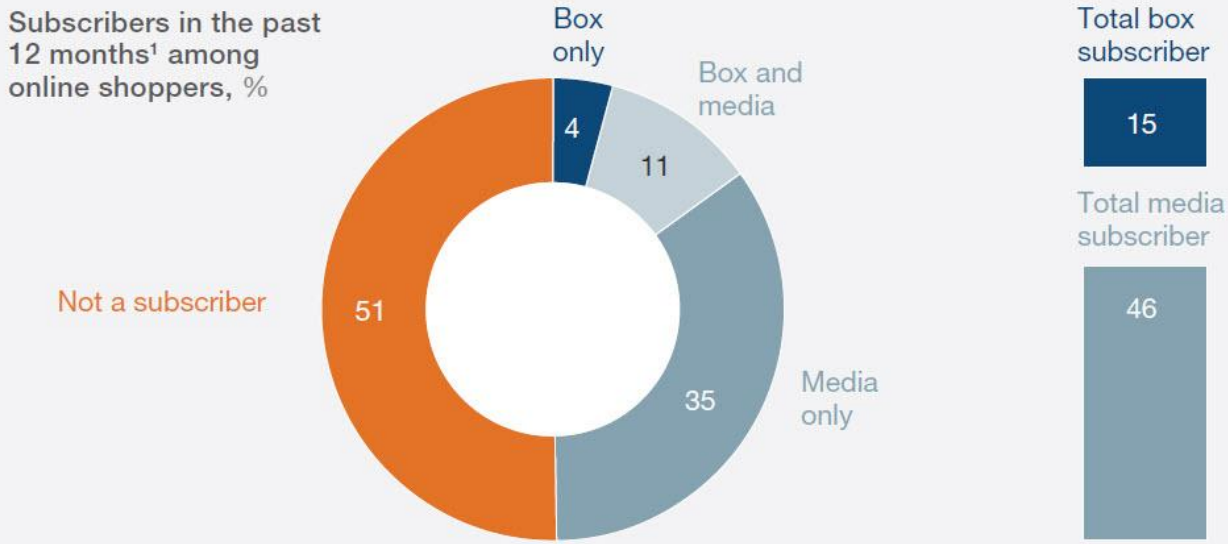
- Subscription market is fast growing and potentially huge
  - Growth rate $>$ 100% each year in the past 5 years
  - Multibillion revenue per year
  - Across a wide range of product categories (digital $+$ physical)
  - **Pay upfront and consume over time**

Subscription Services (Pay in advance with (un)limited usage)

Industry	Product or Service	Price ($)	Period	Total subscribers
Media & Entertainment	Netflix	9.99	Monthly	23 million (US)
	Spotify	9.99	Monthly	70 million (World)
	New York Times	3.75	Weekly	4 million (US)
	MoviePass	19.95	Monthly	2 million
	Kindle Unlimited	9.99	Monthly	–
	Apple News	9.99	Monthly	36 million
Software-as-a-Service	Microsoft Office 365	9.99	Monthly	120 million
	Adobe Creative Cloud (One App)	20.99	Monthly	15 million
	Dropbox Premium	9.99	Monthly	>11 million
Membership Clubs	Costco (Basic)*	60	Annual	94 million
	Amazon Prime	119	Annual	90 million
	24 hour fitness (Gym)	40	Monthly	4 million
eCommerce	Harry's	35	Monthly	–
	Birchbox	15	Monthly	2 million
	Rent the Runway	159	Monthly	6 million
Transportation	Public Transit Pass (MTA)	121	30-days	–
	Uber Ride Pass*	14.99	Monthly	–
	Jetblue "All You can Jet" Pass	699	Monthly	–

Subscription Services

Subscriptions are an increasingly common way to buy products and services online.



Note: Figures may not sum to 100%, because of rounding.

[1]Which of the following have you purchased or subscribed to in the past 12 months? % of those selecting online subscription-box service that delivers products regularly (eg, Blue Apron, Dollar Shave Club, Ipsy, Stitch Fix), subscription-based media (eg, ClassPass, Hulu, Netflix, Spotify), both, or neither.

McKinsey&Company | **Source:** McKinsey analysis

Subscription Services

E-commerce subscriptions generally fall into one of three categories.

E-commerce subscriptions, %		Key consumer value	Description	Example companies
Subscribe for replenishment	32	Save time and money	Replenish the same or similar items Primary categories are commodity items such as razors, vitamins	Amazon Subscribe & Save, Dollar Shave Club, and Ritual
Subscribe for curation	55	Be surprised by product variety	Receive a curated selection of different items, with varying levels of consumer decision making required Primary categories are apparel, food, beauty products	Birchbox, Blue Apron, and Stitch Fix
Subscribe for access	13	Gain exclusive access	Membership provides access and can convey additional “VIP” perks Primary categories are apparel, food	JustFab, NatureBox, and Thrive Market
	100%			

McKinsey&Company | **Source:** McKinsey analysis

Subscription business

- Design product + pricing in subscription markets
- Interested in the distribution of **willingness to pay (WTP)** for subscription service.
  - Demand curve
  - Elasticities of the WTP to product changes
  - Counterfactual analysis

Related Research

- WTP has been a topic of interest in marketing and economics
- Conjoint typically helps in figuring out valuation or part-worths for attributes (Green and Rao, 1971)
- Revealed preference stream uses transaction data for demand estimation, with individual data (Guadagni and Little 1983) or aggregate data (Berry 1994, BLP 1995)
- Comprehensive Survey: Breidert (2007)
- All these cases have price variation!

All these cases have price variation!

Contribution

- Main contribution: a novel method to identify & estimate semiparametrically the conditional distribution of WTP given customer characteristics and product features **when only usage variation is present**.
- No existing research that demonstrates how to obtain the WTP distribution in the **absence of price variation**.

Research questions

1. Without price variation, is it possible to identify distribution of WTP?
2. What demand responses and profits to counterfactual product and pricing choices by the firm can be determined?
3. How to measure the impact on valuations of product improvements?
4. Is there *additional* value in having price variation?

Big usage data of YBOX, a music streaming service

- YBOX is a music streaming service targeting Southeast Asia.
- 1 million users data (Jan 2015–Feb 2017):
  - subscription history
  - daily # of seconds listening music with the service
  - basic demographics (age and gender)
- No price variation for monthly music streaming service over time
- Average daily listening hours range from 45 mins to $>$ 6 hours
- Average monthly listening hours range from less than 1 hour to more than 150 hours.

When there is price variation

Cross section data with price variation.

Notation

- i indicates a consumer
- Subscription decision: $S_i = 1$ (sub) and $= 0$ (not).
- WTP: W_i
- Price: P_i

- Decision rule:

$$\underbrace{W_i - P_i}_{\text{money-metric utility of service}} \quad \text{vs} \quad \underbrace{\mu = 0}_{\text{money-metric utility of outside option}} \quad \Rightarrow$$

$$S_i = \begin{cases} 1, & W_i > P_i \\ 0, & W_i \leq P_i. \end{cases}$$

or $S_i = \mathbb{I}(W_i > P_i)$.

- When $W_i \perp\!\!\!\perp P_i$, for any w in the support of P_i

$$\underbrace{\Pr(W_i > w)}_{\text{Parameter: prob WTP } > w \text{ in the entire pop}} = \underbrace{\Pr(S_i = 1 \mid P_i = w)}_{\text{Data: Mkt shr in the pop facing price } w}.$$

Can Usage Help

Key insight from a simple example Consider a service like Spotify (or YBOX) as an example application

- How does a firm recover the WTP distribution without Price variation?
- Usage variation $\Rightarrow$ variation in **price of one unit of usage**

User	# Songs Per Month	Monthly Fee	"Price" Per Song
1	10	\$10	\$1
2	20	\$10	\$0.5

- (WTP for Spotify) = (WTP per Song) $\times$ (# Songs).
- Subscribe if (WTP per Song) $>$ ("Price" Per Song).
- Market share is informative for distribution of (WTP per Song).

- Suppose only two usages 10 and 20 songs in population with equal probabilities.

Users listening 10 songs	Users listening 20 songs
"Price" per song $= \$1$	"Price" per song $= \$0.5$
Retention rate $= 60\%$	Retention rate $= 80\%$
$P(\text{WTP per song} < \$1) = 0.4$	$P(\text{WTP per song} < \$0.5) = 0.2$

- $P(\#\ \text{Songs} = 10) = P(\#\ \text{Songs} = 20) = 0.5$
- $P(\text{WTP for Spotify} < \$10) = 0.5 \times P(\text{WTP per song} < \$10/10) + 0.5 \times P(\text{WTP per song} < \$10/20) = 0.3.$

Why is it challenging?

Toy example above shows this simple idea. However, it's not easy because we need to deal with:

1. Heterogeneity: (a) WTP per song, and (b) usage (# of songs)
2. Correlation between WTP per song and usage
3. Unobserved usage by nonsubscribers
4. Selection: nonsubscribers are different from subscribers in unobservable ways

Our approach

Deals with the above issues. We need to accommodate and separate out consumers who may have:

- (A) high WTP but listen to few songs
- (B) low WTP and listen to lots of songs

Steps in finding the distribution of WTP in general

Step 1:

Market share or retention rate ⇒ Distribution of WTP per song

Step 2:

Usage data ⇒ Distribution of usage (# Songs)

Step 3:

Combine the distribution of WTP per song and usage ⇒ Distribution of WTP for Service

Step 4:

Estimate elasticities, demand curve, and do counterfactual.

Steps in finding the distribution of WTP in general

Step 1:

Market share or retention rate $\Rightarrow$ Distribution of WTP per song

Step 2:

Usage data $\Rightarrow$ Distribution of usage (# Songs)

Step 3:

Combine the distribution of WTP per song and usage $\Rightarrow$ Distribution of WTP for Service

Step 4:

Estimate elasticities, demand curve, and do counterfactual.

Steps in finding the distribution of WTP in general

Step 1:

Market share or retention rate $\Rightarrow$ Distribution of WTP per song

Step 2:

Usage data $\Rightarrow$ Distribution of usage (# Songs)

Step 3:

Combine the distribution of WTP per song and usage $\Rightarrow$ Distribution of WTP for Service

Step 4:

Estimate elasticities, demand curve, and do counterfactual.

Steps in finding the distribution of WTP in general

Step 1:

Market share or retention rate $\Rightarrow$ Distribution of WTP per song

Step 2:

Usage data $\Rightarrow$ Distribution of usage (# Songs)

Step 3:

Combine the distribution of WTP per song and usage $\Rightarrow$ Distribution of WTP for Service

Step 4:

Estimate elasticities, demand curve, and do counterfactual.

Model setup: simplest case

Notation:

- i indicates a consumer
- WTP for **service**: W_i
- WTP **per unit of usage**: α_i
- Expected usage: Q_i^*
- Decision: $S_i = 1$ (sub) and $= 0$ (not)

Assumptions

1. $W_i = \alpha_i Q_i^* \Rightarrow \ln W_i = \ln \alpha_i + \ln Q_i^*$.
2. Q_i^* is observed for both subscribers and nonsubscribers.
3. $\alpha_i \perp\!\!\!\perp Q_i^*$.

▸ Microfoundation of usage?

Decision rule: $S_i = \mathbb{I}(P < W_i) = \mathbb{I}(P < \alpha_i Q_i^*)$.

Usage Model: Microfoundations

- Earlier with exogenous Q_i^* - the usage model was not specified.
- Each subscription setting might have a different micro-model
- Spotify Example: Consumers solve time allocation problem between listening to music and the outside option.
  - $u(q, \nu, T; \theta) = \theta_1 q \underbrace{\nu}_{\text{shock}} + \theta_2 q^2 + (\underbrace{T}_{\text{Time budget}} - q)$
  - Obtain the usage decision rule: $q(\nu, T; \theta)$
  - $Q_i^* = E_\nu [q(\nu, T; \theta)]$

Back to simplest model

Usage Model: Microfoundations

- Earlier with exogenous Q_i^* - the usage model was not specified.
- Each subscription setting might have a different micro-model
- Spotify Example: Consumers solve time allocation problem between listening to music and the outside option.
  - $u(q, \nu, T; \theta) = \theta_1 q \underbrace{\nu}_{\text{shock}} + \theta_2 q^2 + (\underbrace{T}_{\text{Time budget}} - q)$
  - Obtain the usage decision rule: $q(\nu, T; \theta)$
  - $Q_i^* = E_\nu \left[q(\nu, T; \theta) \right]$

Back to simplest model

Usage Model: Microfoundations

- Earlier with exogenous Q_i^* - the usage model was not specified.
- Each subscription setting might have a different micro-model
- Spotify Example: Consumers solve time allocation problem between listening to music and the outside option.
  - $u(q, \nu, T; \theta) = \theta_1 q \underbrace{\nu}_{\text{shock}} + \theta_2 q^2 + (\underbrace{T}_{\text{Time budget}} - q)$
  - Obtain the usage decision rule: $q(\nu, T; \theta)$
  - $Q_i^* = E_\nu [q(\nu, T; \theta)]$

◂ Back to simplest model

Usage Model: Microfoundations

- Why do we model microfoundations (starting from individual preferences)?
- Typically, we would want to do counterfactuals
- Issue to consider: Does θ (parameters of micromodel) depend on the counterfactual?
  - If it does not: we can use the same micromodel in counterfactuals. Don't really need to estimate θ.
  - if it does: we need to estimate the micromodel parameters θ
- Does the counterfactual obtain the same *distribution of usage*?

Back to simplest model

Usage Model: Microfoundations

- Why do we model microfoundations (starting from individual preferences)?
- Typically, we would want to do counterfactuals
- Issue to consider: Does θ (parameters of micromodel) depend on the counterfactual?
  - If it does not: we can use the same micromodel in counterfactuals. Don't really need to estimate θ.
  - if it does: we need to estimate the micromodel parameters θ
- Does the counterfactual obtain the same *distribution of usage*?

◂ Back to simplest model

Usage Model: Microfoundations

- Why do we model microfoundations (starting from individual preferences)?
- Typically, we would want to do counterfactuals
- Issue to consider: Does θ (parameters of micromodel) depend on the counterfactual?
  - If it does not: we can use the same micromodel in counterfactuals. Don't really need to estimate θ.
  - if it does: we need to estimate the micromodel parameters θ
- Does the counterfactual obtain the same *distribution of usage*?

◂ Back to simplest model

Usage Model: Microfoundations

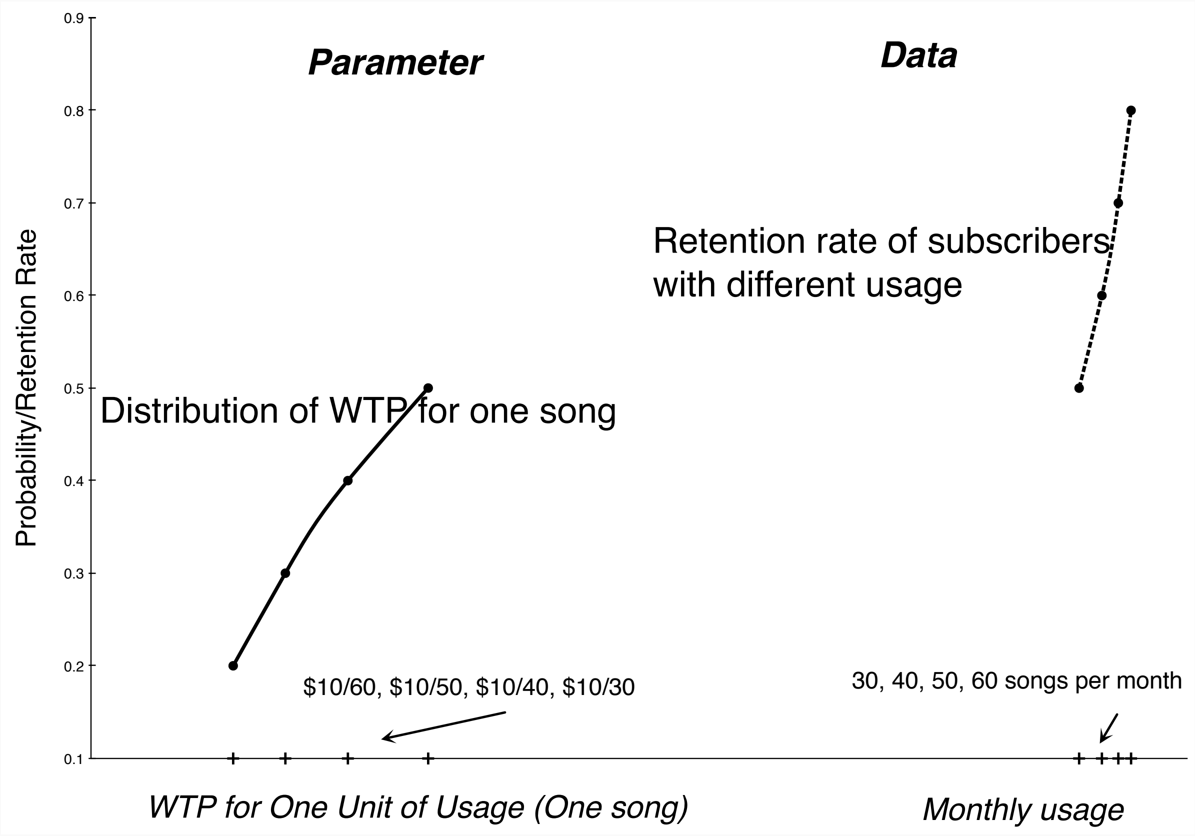
- Why do we model microfoundations (starting from individual preferences)?
- Typically, we would want to do counterfactuals
- Issue to consider: Does θ (parameters of micromodel) depend on the counterfactual?
  - If it does not: we can use the same micromodel in counterfactuals. Don't really need to estimate θ.
  - if it does: we need to estimate the micromodel parameters θ
- Does the counterfactual obtain the same *distribution of usage*?

◂ Back to simplest model

Nonparametric estimation of $F_\alpha(\alpha_i)$ (*Step 1*)

Decision rule $S_i = \mathbb{I}(\ln P < \ln \alpha_i + Q_i^*)$ implies that

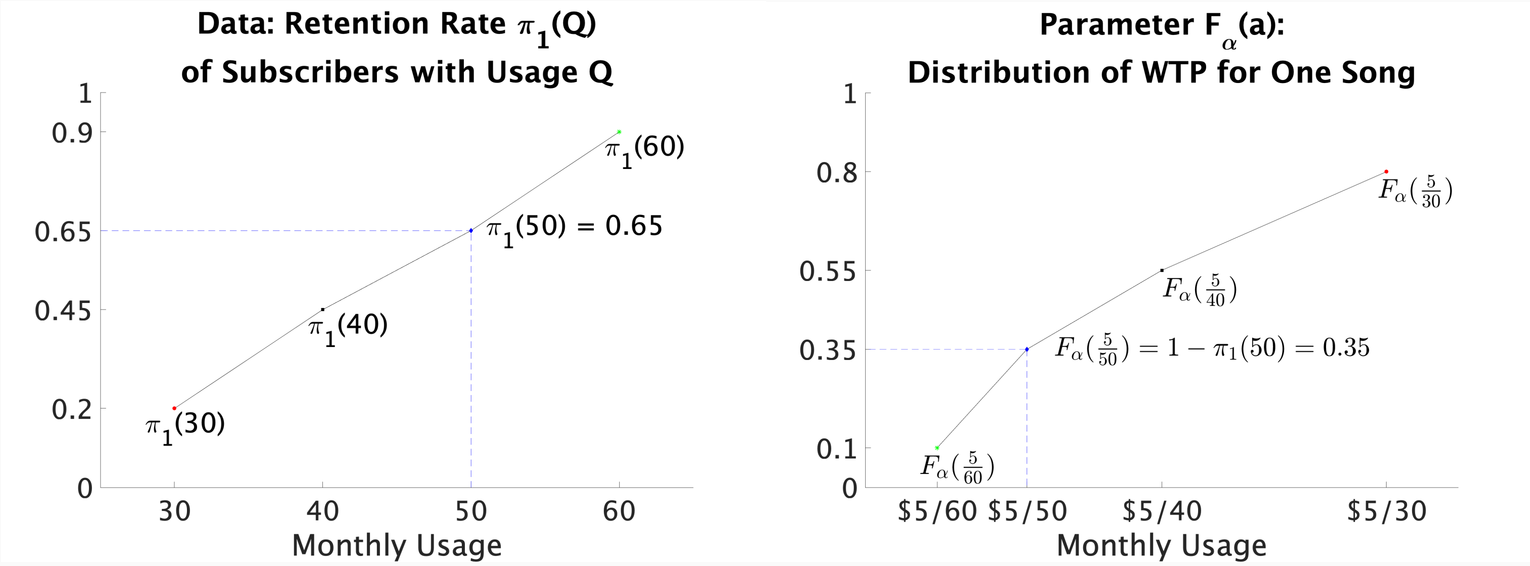
$$\Pr\big(\alpha_i \le P/\exp(Q)\big) = 1 - \Pr(S_i = 1 \mid Q_i^* = Q).$$



Nonparametric estimation of $F_\alpha(\alpha_i)$ (**Step 1**)

Decision rule $S_i = \mathbb{I}(P < \alpha_i Q_i^*)$ implies that

$$F_\alpha(P/Q) \equiv \Pr(\alpha_i \leq P/Q) = 1 - \underbrace{\Pr(S_i = 1 \mid Q_i^* = Q)}_{\pi_1(Q)} \quad \Rightarrow \quad \boxed{F_\alpha(a) = 1 - \pi_1\left(\frac{P}{a}\right)}$$



WTP: From usage $\Rightarrow$ subscription (*Step 2 and 3*)

- We obtain $F_\alpha(\alpha_i)$ by non-parametric estimation as earlier
- If:
  - Q_i^* is observed, so $F_Q(Q_i^*)$ is known, (*Step 2*)
  - $\alpha_i \perp\!\!\!\perp Q_i^*$,

  the distribution $F_w(W_i)$ can be easily calculated.
- Calculation uses $\ln W_i = \ln \alpha_i + \log Q_i^*$.

Example: Usage takes two values 10 and 20

$$\Pr(Q_i^* = \ln 10) = 1/2 \quad \text{and} \quad \Pr(Q_i^* = \ln 20) = 1/2.$$

$$F_w(W_i < w) = 0.5F_\alpha\left(\alpha_i < \frac{w}{10}\right) + 0.5F_\alpha\left(\alpha_i < \frac{w}{20}\right). \quad (\textit{Step 3})$$

WTP: From usage ⇒ subscription (*Step 2 and 3*)

- We obtain $F_\alpha(\alpha_i)$ by non-parametric estimation as earlier
- If:
  - Q_i^* is observed, so $F_Q(Q_i^*)$ is known, (*Step 2*)
  - $\alpha_i \perp\!\!\!\perp Q_i^*$,

  the distribution $F_w(W_i)$ can be easily calculated.
- Calculation uses $\ln W_i = \ln \alpha_i + \log Q_i^*$.

Example: Usage takes two values 10 and 20

$$\Pr(Q_i^* = \ln 10) = 1/2 \quad \text{and} \quad \Pr(Q_i^* = \ln 20) = 1/2.$$

$$F_w(W_i < w) = 0.5 F_\alpha \left(\alpha_i < \frac{w}{10}\right) + 0.5 F_\alpha \left(\alpha_i < \frac{w}{20}\right). \quad (\textit{Step 3})$$

WTP: From usage $\Rightarrow$ subscription (**Step 2 and 3**)

- If
  - $F_\alpha(\alpha_i)$ is known,
  - Q_i^* is observed, so $F_Q(Q_i^*)$ is known, (**Step 2**)
  - $\alpha_i \perp\!\!\!\perp Q_i^*$,

  the distribution $F_W(W_i)$ is calculated by using $W_i = \alpha_i Q_i^*$.

$$F_W(w) = \int_0^\infty F_\alpha(w/q)\,\mathrm{d}\,F_{Q^*}(q).$$

- It can shown that

$$\boxed{F_W(w) = 1 - \mathrm{E}\Big(\pi_1\Big(\frac{PQ_i^*}{w}\Big)\Big).}$$

Estimation for Simple case

Algorithm

Step 1: estimate $\pi_1(q) \equiv \Pr(S_i = 1 \mid Q_i^* = q)$;

Step 2a: estimate $\hat{F}_\alpha(a) = 1 - \hat{\pi}_1(P/a)$.

Step 2b: estimate $\hat{F}_W(w) = 1 - n^{-1} \sum_{i=1}^{n} \hat{\pi}_1(PQ_i^*/w)$.

Conditional distribution of WTP $F_w(W_i \mid X_i)$

- X_i: observed consumer and product characteristics
- $F_w(W_i \mid X_i) \Rightarrow$ demand curve and elasticities
- **Demand curve**: (*Step 4*)

$$\underbrace{\Pr(S_i = 1 \mid P = p^c)}_{\text{mkt shr for new price } p^c} = \Pr(W_i > p^c) = 1 - F_w(W_i = p^c).$$

- **Elasticities** (we derived explicit formulas): (*Step 4*)

$$\mathrm{E}\left(\frac{\partial \mathrm{E}(\ln W_i \mid X_i)}{\partial X_{j,i}}\right), \text{ e.g., } X_{j,i} \text{ is waiting time to reach cust service}$$

$$\mathrm{E}\Big[\mathrm{E}(\ln W_i \mid X_{other}, \text{Play music offline} = \text{Yes}) - \mathrm{E}(\ln W_i \mid X_{other}, \text{Play music offline} = No)\Big]$$

Two dimensional heterogeneity (*Step 1 and 2*)

- Heterogeneity in WTP per song.
- Project unobserved $\ln \alpha_i$ on X_{1i}:

$$\ln \alpha_i = \beta' X_{1i} + U_i.$$

- Estimate β and $F_u(U_i)$.
- No parametric assumption required for U_i.

Theorem

Under sufficient support conditions, β can be consistently estimated with OLS of

$$\frac{S_{it} - \mathbb{I}(Q^*_{it} - \ln P > 0)}{\underbrace{f_Q(Q^*_{it} \mid X_{1i})}_{\textit{Nonparametric Estimator}}} = \beta X_{1i} + u_{it}$$

Why is it challenging?

Toy example above shows this simple idea. However, it's not easy because we need to deal with:

1. Heterogeneity: (a) WTP per song, and (b) usage (# of songs)
2. Correlation between WTP per song and usage
3. Unobserved usage by nonsubscribers
4. Selection: nonsubscribers are different from subscribers in unobservable ways

Required Assumptions

Assumptions

1. $$\ln Q^*_{it} = \gamma' X_{2it} + V_i = \gamma'_a X_{2i,a} + \gamma'_b X_{2i,b} + \gamma'_c X_{2it,c} + V_i, \quad (1)$$

   where the unobserved fixed effect V_i can be correlated with U_{it} in the specification of WTP per unit usage (α_{it}). Among X_{2it}, $X_{2it,c}$ are time varying covariates, $X_{2i,a}$ and $X_{2i,b}$ are time invariant, and $X_{2i,b}$ does not belong to X_{1it}, but $X_{2i,a}$ belongs to X_{1it}.

2. The observed actual usage Q_{it} when $S_{it} = 1$ is given by:

   $$\ln Q_{it} = \ln Q^*_{it} + \varepsilon_{it},$$

Required Assumptions

Assumptions

1. (Big support of usage). The support of $\bar{Q}_{it} \equiv \gamma_c' X_{2it,c} \mid (X_{1it}, Z_{it})$ covers the support of $\ln(P/\alpha_{it}) - V_i \mid (X_{1it}, Z_{it})$;
2. (Independence of usage). $X_{2it,c} \perp\!\!\!\perp (U_{it}, V_i) \mid (, Z_{it})$;
3. (Valid and relevant IV). $\mathrm{E}(Z_{it}\eta_{it}) = 0$, $\mathrm{E}(Z_{it}Z_{it}')$ is nonsingular, and $\operatorname{rank} \mathrm{E}(X_{1it}Z_{it}') = \dim(X_{1it})$.
4. (Finite mean) $\mathrm{E}(\eta_{it}) < \infty$.

Two dimensional heterogeneity (*Step 1 and 2*)

- Heterogeneity in usage (# of songs), unobserved usage by nonsubscribers, and selection

$$Q^*_{it} = \gamma' X_{2it} + \underbrace{V_i}_{\text{Selection on unobservable}} \quad \text{(nonsubscribers + subscribers)}$$

$$Q_{it} = Q^*_{it} + \varepsilon_{it} \quad \text{for subscribers}$$

V_i and U_{it} (hence α_{it}) are correlated (selection)

- **Panel data with Selection:** Observed log usage Q_{it} for subscribers, and X_{2it} for both subscribers and nonsubscribers.
- Fixed effect reg $\Rightarrow \gamma$ and **"predicted usage"** $\bar{Q}_{it} = \gamma' X_{2it}$ for both subscribers and nonsubscribers.
- Identify the distribution of $U_{it} + V_i$. This gets us to WTP.
- **Variable in X_{2it} impacts usage (need some that doesn't impact WTP)**

Two dimensional heterogeneity (*Step 1 and 2*)

- Heterogeneity in usage (# of songs), unobserved usage by nonsubscribers, and selection

$$Q_{it}^* = \gamma' X_{2it} + \underbrace{V_i}_{\text{Selection on unobservable}} \qquad \text{(nonsubscribers + subscribers)}$$

$$Q_{it} = Q_{it}^* + \varepsilon_{it} \quad \text{for subscribers}$$

V_i and U_{it} (hence α_{it}) are correlated (selection)

- **Panel data with Selection:** Observed log usage Q_{it} for subscribers, and X_{2it} for both subscribers and nonsubscribers.
- Fixed effect reg $\Rightarrow \gamma$ and **"predicted usage"** $\bar{Q}_{it} = \gamma' X_{2it}$ for both subscribers and nonsubscribers.
- Identify the distribution of $U_{it} + V_i$. This gets us to WTP.
- **Variable in X_{2it} impacts usage (need some that doesn't impact WTP)**

Two dimensional heterogeneity (*Step 1 and 2*)

- Heterogeneity in usage (# of songs), unobserved usage by nonsubscribers, and selection

$$Q^*_{it} = \gamma' X_{2it} + \underbrace{V_i}_{\text{Selection on unobservable}} \qquad \text{(nonsubscribers + subscribers)}$$

$$Q_{it} = Q^*_{it} + \varepsilon_{it} \quad \text{for subscribers}$$

V_i and U_{it} (hence α_{it}) are correlated (selection)

- **Panel data with Selection:** Observed log usage Q_{it} for subscribers, and X_{2it} for both subscribers and nonsubscribers.
- Fixed effect reg $\Rightarrow \gamma$ and **"predicted usage"** $\bar{Q}_{it} = \gamma' X_{2it}$ for both subscribers and nonsubscribers.
- Identify the distribution of $U_{it} + V_i$. This gets us to WTP.
- **Variable in X_{2it} impacts usage (need some that doesn't impact WTP)**

Two dimensional heterogeneity (*Step 1 and 2*)

- Heterogeneity in usage (# of songs), unobserved usage by nonsubscribers, and selection
- Reduced form of Q_{it}^*:

$$Q_{it}^* = \gamma' X_{2it} + \underbrace{V_i}_{\text{Selection on unobservable}} \quad \text{(nonsubscribers + subscribers)}$$

$$Q_{it} = Q_{it}^* + \varepsilon_{it} \quad \text{for subscribers}$$

V_i and U_{it} (hence α_{it}) are correlated (selection)

- **Panel data with Selection:** Observed log usage Q_{it} for subscribers, and X_{2it} for both subscribers and nonsubscribers.
- Fixed effect reg $\Rightarrow \gamma$ and **"predicted usage"** $\bar{Q}_{it} = \gamma' X_{2it}$ for both subscribers and nonsubscribers.
- Identify the distribution of $U_{it} + V_i$. This gets us to WTP.

Two dimensional heterogeneity (*Step 1 and 2*)

- Heterogeneity in usage (# of songs), unobserved usage by nonsubscribers, and selection
- Reduced form of Q_{it}^*:

$$Q_{it}^* = \gamma' X_{2it} + \underbrace{V_i}_{\text{Selection on unobservable}} \qquad \text{(nonsubscribers + subscribers)}$$

$$Q_{it} = Q_{it}^* + \varepsilon_{it} \quad \text{for subscribers}$$

V_i and U_{it} (hence α_{it}) are correlated (selection)

- **Panel data with Selection:** Observed log usage Q_{it} for subscribers, and X_{2it} for both subscribers and nonsubscribers.
- Fixed effect reg $\Rightarrow \gamma$ and **"predicted usage"** $\bar{Q}_{it} = \gamma' X_{2it}$ for both subscribers and nonsubscribers.
- Identify the distribution of $U_{it} + V_i$. This gets us to WTP.

Two dimensional heterogeneity (*Step 1 and 2*)

- Heterogeneity in usage (# of songs), unobserved usage by nonsubscribers, and selection
- Reduced form of Q_{it}^*:

$$Q_{it}^* = \gamma' X_{2it} + \underbrace{V_i}_{\text{Selection on unobservable}} \qquad \text{(nonsubscribers + subscribers)}$$

$$Q_{it} = Q_{it}^* + \varepsilon_{it} \quad \text{for subscribers}$$

  V_i and U_{it} (hence α_{it}) are correlated (selection)

- **Panel data with Selection:** Observed log usage Q_{it} for subscribers, and X_{2it} for both subscribers and nonsubscribers.
- Fixed effect reg $\Rightarrow \gamma$ and **"predicted usage"** $\bar{Q}_{it} = \gamma' X_{2it}$ for both subscribers and nonsubscribers.
- Identify the distribution of $U_{it} + V_i$. This gets us to WTP.

Obtaining the conditional WTP

Theorem: Conditional Distribution of WTP

- Obtain β as 2SLS estimator of regressing Y_{2it} on X_{1it} with IV Z_{it}

$$Y_{2it} = \frac{S_{it} - \mathbb{I}(\bar{Q}_{it} - \ln P > 0)}{f_{\bar{Q}}(\bar{Q}_{it} \mid X_{1it}, Z_{it})}.$$

  $\bar{Q}_{it} = \gamma X_{2it}$ is special regressor.

- Obtain the conditional WTP distribution:

$$\mathrm{E}(\ln W_{it} \mid X_{1it}, X_{2it}) = \beta' X_{1it} + \gamma' X_{2it} + \mathrm{E}(\eta_{it} \mid X_{1it}, X_{2it})$$

$$\eta_{it} = \mathrm{E}\left(\frac{S - \mathbb{I}(\ln P - \beta' X_{1it} - \bar{Q}_{it} \leq 0)}{f_{\bar{Q}_{it}}(\bar{Q}_{it} \mid X_{1it}, Z_{it})} \,\middle|\, X_{1it}, Z_{it} \right)$$

Obtaining the conditional WTP

Theorem: Conditional Distribution of WTP

- Obtain β as 2SLS estimator of regressing Y_{2it} on X_{1it} with IV Z_{it}

$$Y_{2it} = \frac{S_{it} - \mathbb{I}(\bar{Q}_{it} - \ln P > 0)}{f_{\bar{Q}}(\bar{Q}_{it} \mid X_{1it}, Z_{it})}.$$

$\bar{Q}_{it} = \gamma X_{2it}$ is special regressor.

- Obtain the conditional WTP distribution:

$$\mathrm{E}(\ln W_{it} \mid X_{1it}, X_{2it}) = \beta' X_{1it} + \gamma' X_{2it} + \mathrm{E}(\eta_{it} \mid X_{1it}, X_{2it})$$

$$\eta_{it} = \mathrm{E}\left(\frac{S - \mathbb{I}(\ln P - \beta' X_{1it} - \bar{Q}_{it} \leq 0)}{f_{\bar{Q}_{it}}(\bar{Q}_{it} \mid X_{1it}, Z_{it})} \,\middle|\, X_{1it}, Z_{it} \right)$$

From Conditional WTP Distribution

The conditional WTP distribution:

- allows us to obtain demand curve by integrating and elasticities
- is primitive for doing counterfactual analysis

Subscription Services

Step	Description	Details
0	Data Requirements	Obtain panel data of n consumers over T billing periods. Data include price P, subscription choices $S_{it} \in \{0,1\}$, usage Q_{it} when they subscribe, factors that impact WTP per unit usage X_{1it} and usage X_{2it}. When X_{1it} is endogenous, we also need instruments Z_{it} for X_{1it} that are uncorrelated with errors U_{it} and V_i. When X_{1it} is not endogenous, $Z_{it} = X_{1it}$.
1	Fixed Effect Estimation on Usage	Using the data of the subscribers, estimate the following panel data model $\ln Q_{it} = \gamma' X_{2it} + V_i + \varepsilon_{it}$ with fixed effect estimator. Let $\hat{\gamma}$ be the estimator of γ, and let $\hat{Q}_{it} = \hat{\gamma}' X_{2it}$ for all i (subscribers and non-subscribers) and t.
2	Nonparametric estimation of $f_{\tilde{Q}}(\tilde{Q}_{it} \mid X_{1it}, Z_{it})$	Obtain a nonparametric estimator $\hat{f}_{\tilde{Q}}(\cdot \mid X_{1it}, Z_{it})$ of the conditional PDF $f_{\tilde{Q}}(\tilde{Q}_{it} \mid X_{1it}, Z_{it})$ from the sample $(\hat{Q}_{it}, X_{1it}, Z_{it})$.
3	IV regression to estimate β	Estimate β by the 2SLS IV regression of $\hat{Y}_{it}$ on X_{1it} using IV Z_{it}, where $\hat{Y}_{it} = \frac{S_{it} - \mathbb{I}(\hat{Q}_{it} - \ln P > 0)}{\hat{f}_{\tilde{Q}}(\hat{Q}_{it} \mid X_{1it}, Z_{it})}$.
4	CCP Estimation of π	Estimate the CCP function $\pi(X_{1it}, Z_{it}, \tilde{Q}_{it}) = \mathrm{E}(S_{it} \mid X_{1it}, Z_{it}, \tilde{Q}_{it})$ from the sample $(S_{it}, X_{1it}, Z_{it}, \hat{Q}_{it})$.
5	Estimate CDF of WTP	Estimate the conditional CDF of WTP F_W and of unexplained heterogeneity in WTP F_η.

(Illustrative) Empirical Application

Descriptive statistics of YBOX

Table: Descriptive Stat YBOX Users

	Churners	Non-churners
Average daily	1.17	1.33
listening hours	(0.99)	(1.09)
Age	29.53	29.92
	(8.25)	(8.14)
Is male	0.51	0.55
n of obs	1211	7487

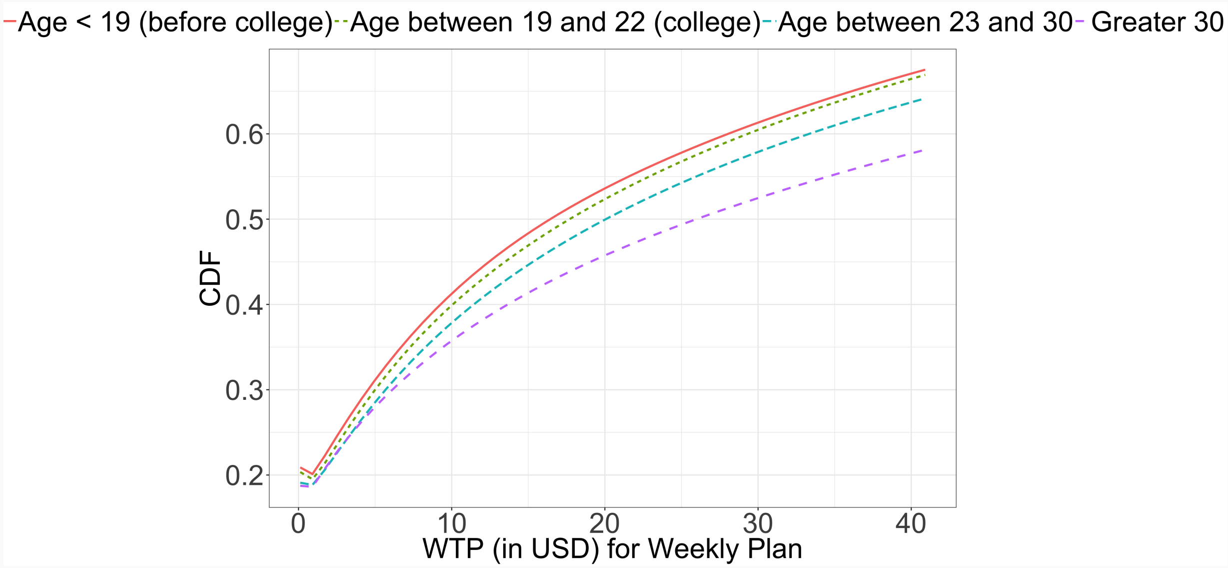
- The percentages of the three prices \$5, \$4, \$3 are 56.8, 17.7, and 23.9.

Counterfactual – Step 4)

- Design shorter subscription plan based on subscription and usage data
- Amazon for example is interested in introducing a short-term Prime plan (currently only annual)
  - If usage varies within the billing period (e.g., seasonality), WTP may be quite different
- What's the distribution of the WTP for the shorter plan?
- Evaluate the money-metric effect of product changes

Counterfactual (*Step 4*)

- Shorter subscription plan based on **higher frequency usage** data
  - Desire a plan of shorter length
  - What's the distribution of the WTP for the shorter plan?
- Evaluate the money-metric effect of product changes



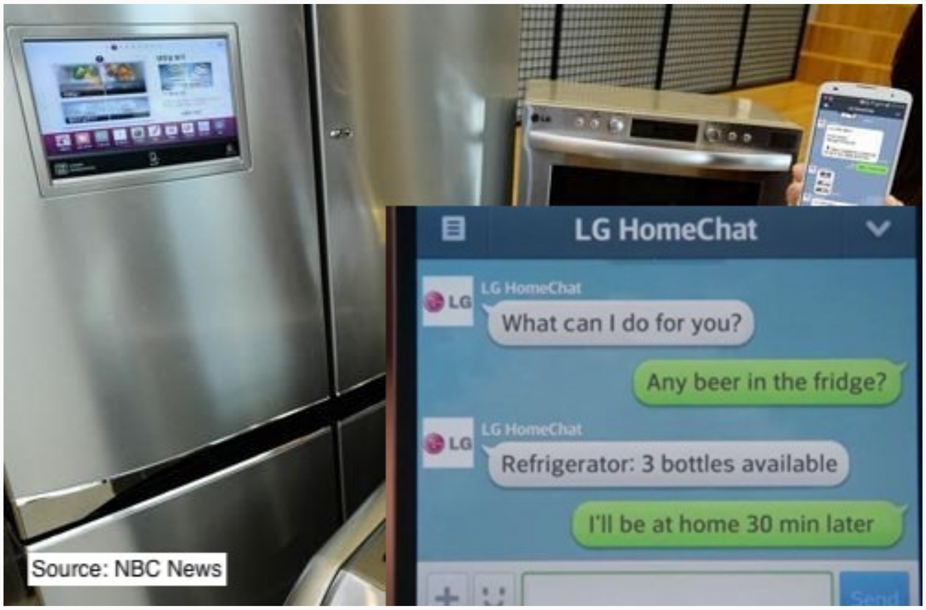
Conclusions

- Without Price variation, can we obtain WTP?
  - A: Qualified Yes.
- What big data on usage tracking **can tell** us?
  - The distribution of WTP under some restrictions (mainly, $\ln W_i = g(\ln \alpha_i, Q^*_{it}) = \ln \alpha_i + Q^*_{it}$)
- 3 Vs (Volume, Velocity and Variety) of big data might not be sufficient
  - Need "big variation" or "big support"
- Can design counterfactual products and pricing strategies

Conclusions

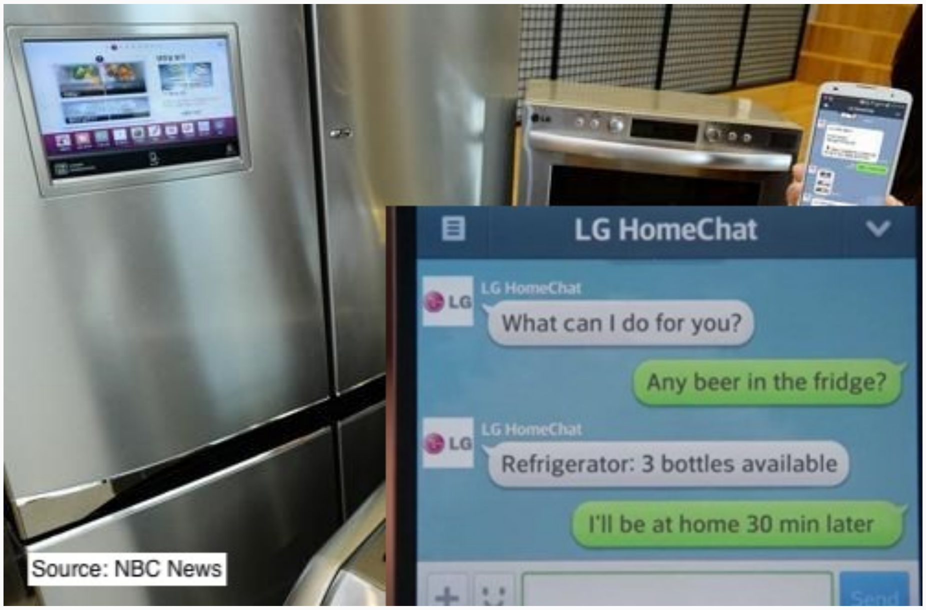
- What big data on usage tracking **can tell** us?
  - The distribution of WTP under some restrictions (mainly, $\ln W_i = g(\ln \alpha_i, Q_i^*) = \ln \alpha_i + Q_i^*$)
- **Cannot**
  - What is g?
  - Cannot tell us the value of outside option.
  - The role price variation, even limited, in identifying g.

A bigger picture (of a fridge)



- Essentially, we need the separation of purchase (subscription) and consumption (usage).
- Such separation also holds in packaged goods (beer)—but we did not track the usage.
- 5G and Internet of Things could enable such tracking.

A bigger picture (of a fridge)



- Essentially, we need the separation of purchase (subscription) and consumption (usage).
- Such separation also holds in packaged goods (beer)—but we did not track the usage.
- 5G and Internet of Things could enable such tracking.

Summary by R practice

- https://chengchou.shinyapps.io/WTPV1/

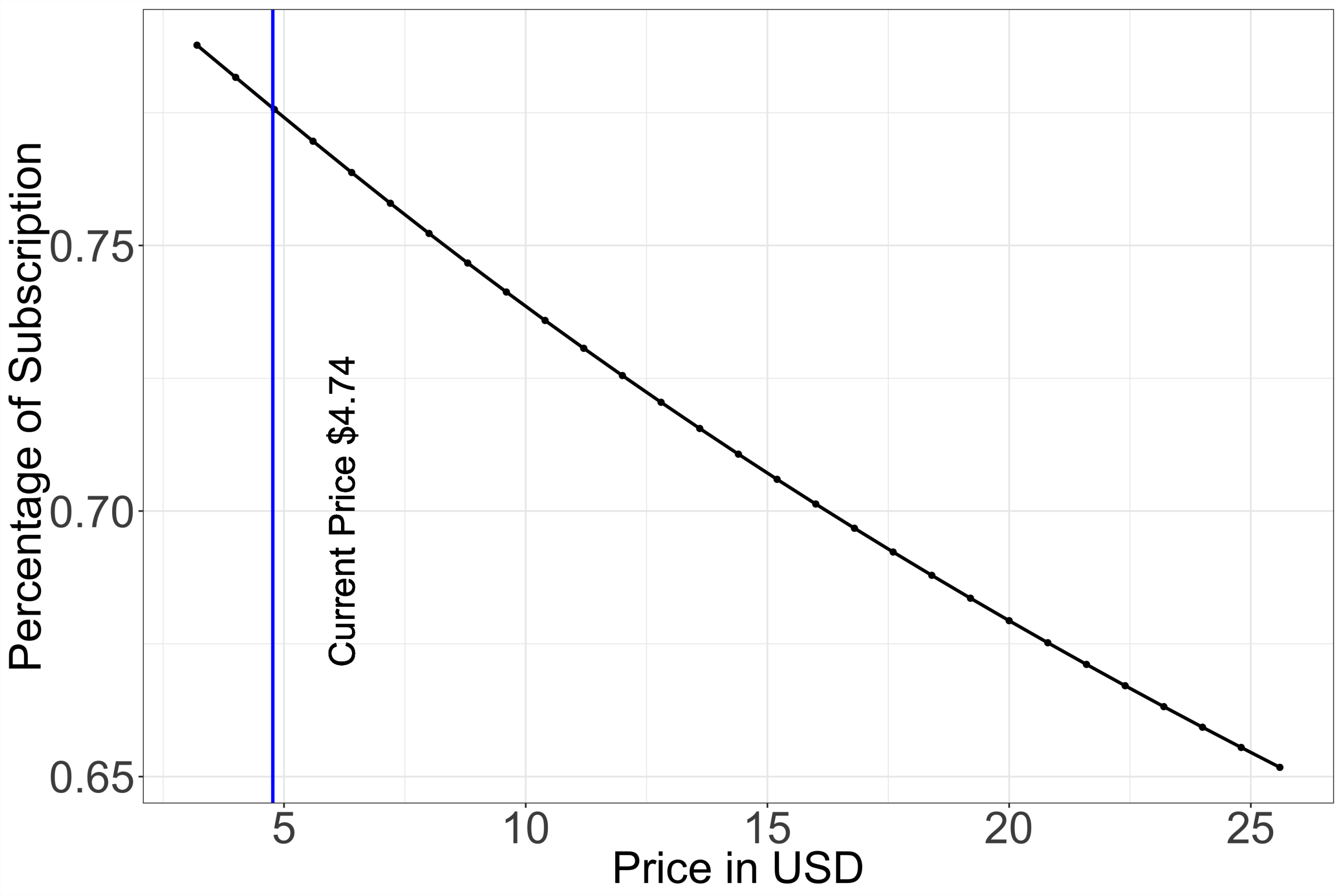
ADDITIONAL SLIDES

Estimated elasticities of WTP for YBOX monthly plan

- Age elasticity: 0.04 (0.01).
- Men are willing to pay 26.23% (18.09) more than women.

	Estimated Mean of Log WTP	
Age	Men	Women
before college	log($16)	log($16)
in college	log($24)	log($21)
23–30	log($33)	log($30)
> 30	log($49)	log($56)

Estimated demand curve



Is increasing the current price profitable?

When there is price variation

Cross section data with price variation.

Notation

- i indicates a consumer
- Subscription decision: $S_i = 1$ (sub) and $= 0$ (not).
- WTP: W_i
- Price: P_i

- Decision rule:

$$\underbrace{W_i - P_i}_{\substack{\text{money-metric} \\ \text{utility of service}}} \quad \text{vs} \quad \underbrace{\mu = 0}_{\substack{\text{money-metric utility} \\ \text{of outside option}}} \quad \Rightarrow$$

$$S_i = \begin{cases} 1, & W_i > P_i \\ 0, & W_i \leq P_i. \end{cases}$$

or $S_i = \mathbb{I}(W_i > P_i)$.

- When $W_i \perp\!\!\!\perp P_i$, for any w in the support of P_i

$$\underbrace{\Pr(W_i > w)}_{\substack{\text{Parameter: prob WTP} \\ > w \text{ in the entire pop}}} = \underbrace{\Pr(S_i = 1 \mid P_i = w)}_{\substack{\text{Data: Mkt shr in the pop} \\ \text{facing price } w}}.$$

What do you need $W_i \perp\!\!\!\perp P_i$?

- Without independence assumption, we have

$$\Pr(S_i = 1 \mid P_i = w) = \Pr(W_i > w \mid P_i = w) \neq \Pr(W_i > w),$$

  because W_i and P_i are correlated.
- From the above we only know $F_w(W_i = w \mid P_i = w)$ for any w in the support of P_i.
- We do not identify conditional distribution $F_w(W_i \mid P_i)$. For example, from the above one cannot identify

$$\Pr(W_i > w + \varepsilon \mid P_i = w),$$

  $\varepsilon \neq 0$.

Mapping market share to WTP distribution: price variation case

If P_i has enough variation, we know the entire distribution of WTP $F(W_i)$.