Predictive Customer Analytics

UCSD MGT 100 Week 09

Kenneth C. Wilbur and Dan Yavorsky

Predictive Customer Analytics

- Importance of Customer Acquisition
- Market size: How many customers experience the core need?
- Diffusion: How does the served market change over time?
  - Model based on Bass (1969)
- CLV: How profitable are customer relationships?

Importance of Customer Acquisition

- Einav et al. (2021) analyzed all US Visa CC transaction data

```
- >$1T spent in 32B transactions by 428MM cards at 1MM stores from 2016-19
- Assume card~=customer
```

$$\text{BrandRev.} = \sum spend \equiv \sum \frac{stores}{1} \frac{cards}{stores} \frac{transactions}{cards} \cdot \frac{spend}{transactio}$$

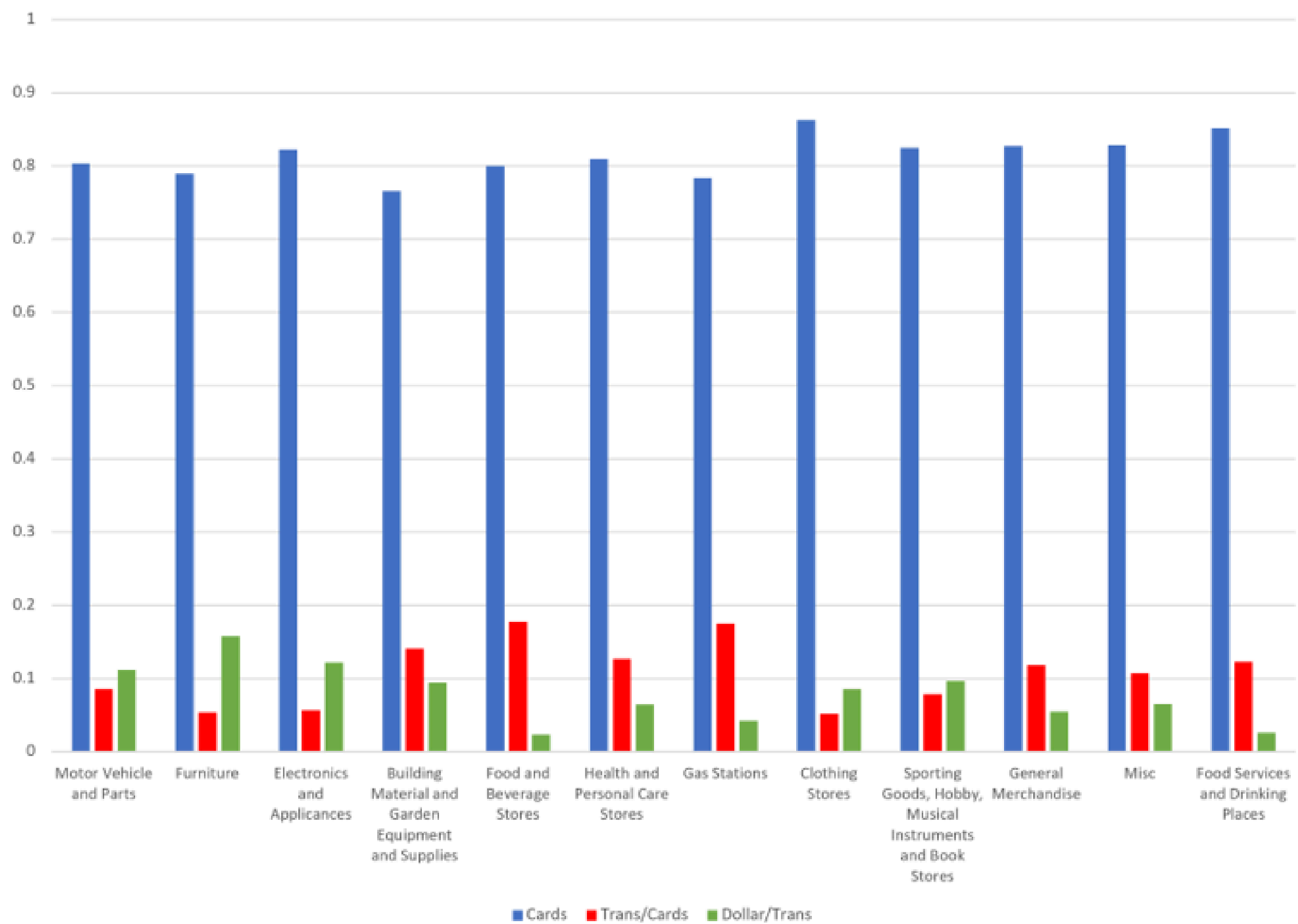
- Research question: How well does each factor explain brand revenue?
- Regressed log revenue on log RHS with merchant and year fixed effects

Table 2: Decomposing Sales in Offline Retail

	Stores	Cards/Store	Trans/Card	Dollar/Trans
A. Across Merchants	0.098	0.689	0.036	0.177
($N = 939,304$)	[0.123]	[0.761]	[0.231]	[0.547]
B. Within Merchants over Time	0.164	0.681	0.098	0.056
($N = 3,846,501$)	[0.804]	[0.975]	[0.942]	[0.971]
C. Across Stores within Merchants		0.841	0.077	0.082
($N = 1,926,714$)		[0.972]	[0.809]	[0.933]
D. Within Stores over Time		0.817	0.134	0.049
($N = 7,746,218$)		[0.995]	[0.969]	[0.987]

Note: All standard errors are less than 0.001. R-Squared values are reported in square brackets. Across Merchant and Across Store within Merchant decompositions are based on 2019 data. Across Merchant regressions include NAICS fixed effects. Within Merchants over Time and Within Stores over Time are based on 2016–2019 data. Within Merchants over Time regressions include merchant and year fixed effects. Across Stores within Merchants regressions include merchant fixed effects. Within store over Time regressions include store and year fixed effects. See Online Appendix B for robustness with respect to a longer panel of merchant/store data.

Figure 1: Decomposing Merchant Sales Growth by Industry



Note: This figure displays the coefficients of the "Within Merchant over time" decomposition by industry. The regressions are run with Visa data from 2016 through 2019, and include merchant and year fixed effects.

Predictive Analytics

- Weather forecasts
  - What to do, what to wear
- Stock prices
  - Buy or sell, how much, when
- Safety
  - Where to live, how to transport
- Lifespan
  - Schooling, savings/spending, work/retire
- Product quality
  - Buy now/later, return, warranty, insurance

Predictive Analytics

- Correlations alone can enable powerful predictions
- Causal drivers will improve predictions

```
  - But we do not need to know all the causal drivers to make predictions

  - E.g., future behavior often predicted using past behavior
```

- Predictive analytics are not prescriptive or diagnostic

```
  - Predictive analytics are oft misused & overinterpreted
```

- Predictive Analytics typically have wide error bars

```
  - Often ignored by those who don't understand ... and those who do

  - Accounting for data variability, estimation error and model uncertainty
.. CIs get wide
```

Meet Dray

Market Size

- Market size (): # of people who might pay to address the core need in a given time period

```
- Alternatively measured in $, units or volume
- Noisy but helps inform potential returns to investments
- Typical investor's first question: How big is the market?
  $100B market is viewed differently than a $100MM market
- How will you know if you got the right answer?
- What happens if you overestimate market size?
```

- “Marketing myopia:” Neglecting nontraditional competitors, e.g. Zoom v. Uber or Carnival v. Whistler

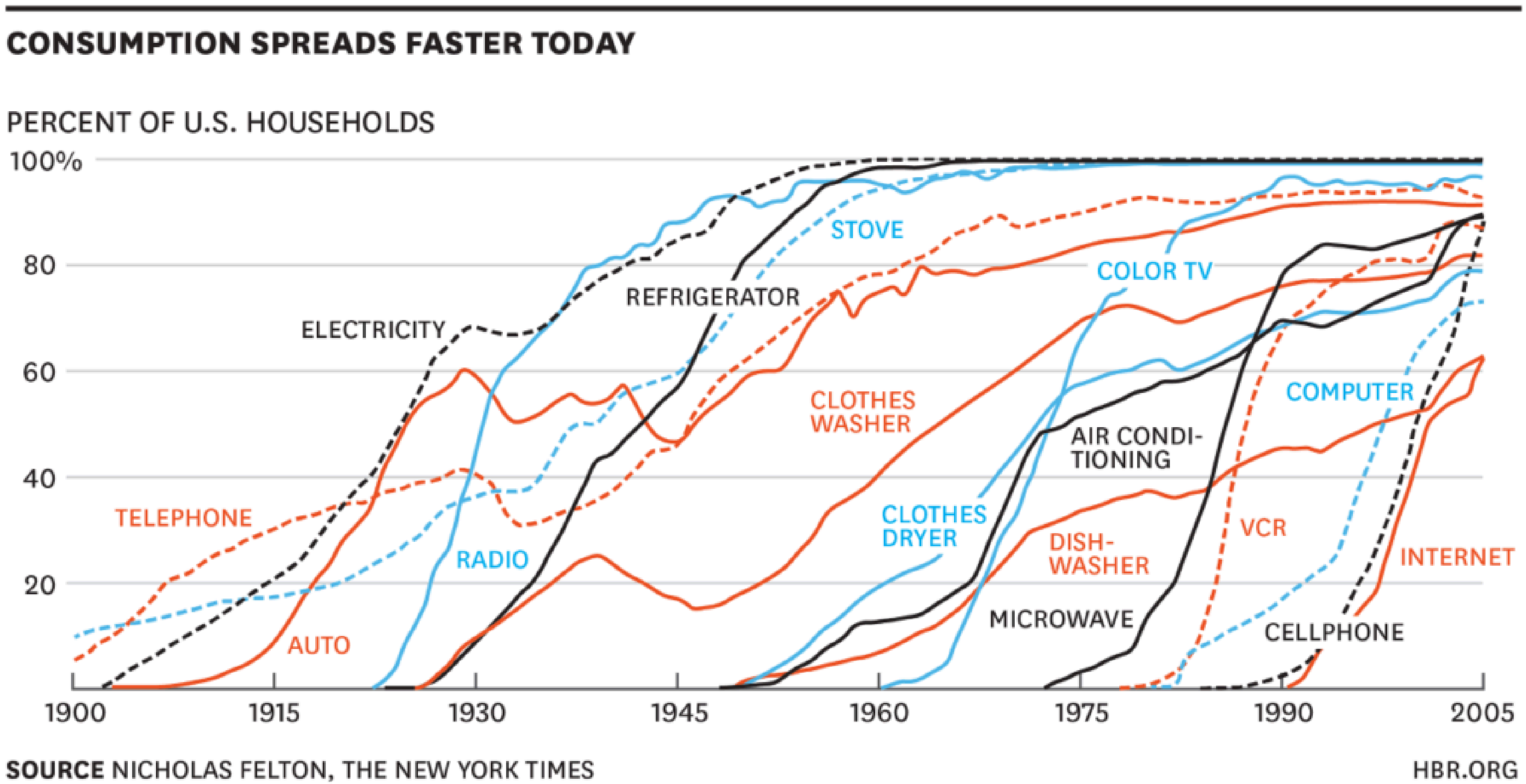
Market Size

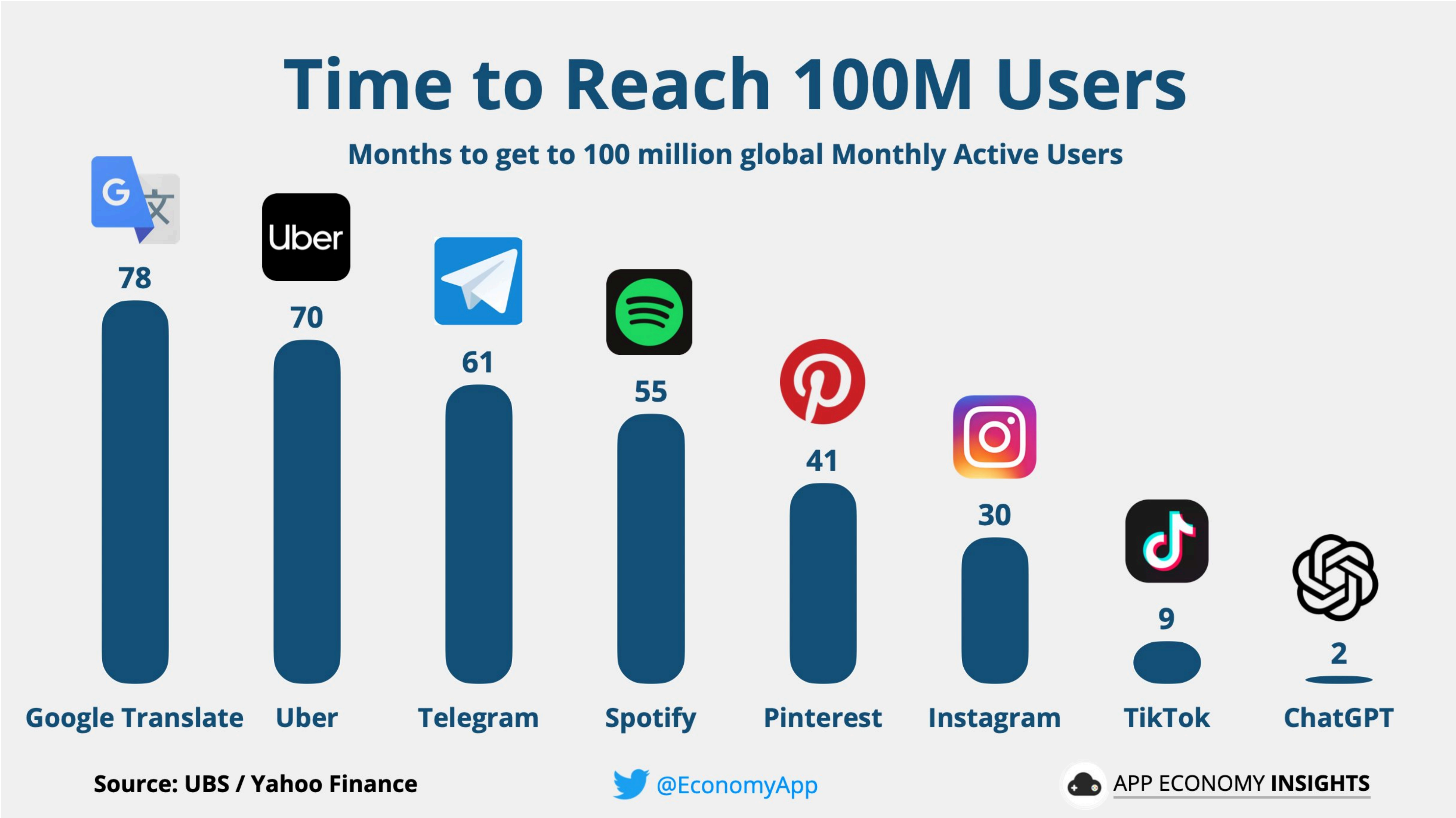
- 3 ways to estimate:
  - “Top Down” Total Addressable Market (TAM) :
    How many people have the core need?
  - “Bottom up” Served Available Market (TAM):
    How many people currently pay to solve the core need?
    - TAM=SAM+Unserved
  - Analyst estimates
- Best practice: Use all three, triangulate, gauge sensitivity

Case study: US Mattress Market

- USA population : ~340 million
- Assumption : (why? pros, cons?)
- Assumption : Avg mattress lasts 7 years (pros, cons?)
- Market size 47.1 million people annually
- Average mattress price : $283, across all bed sizes
- Market size $13.3B/year
- Let’s check Grand View Research & ISPA

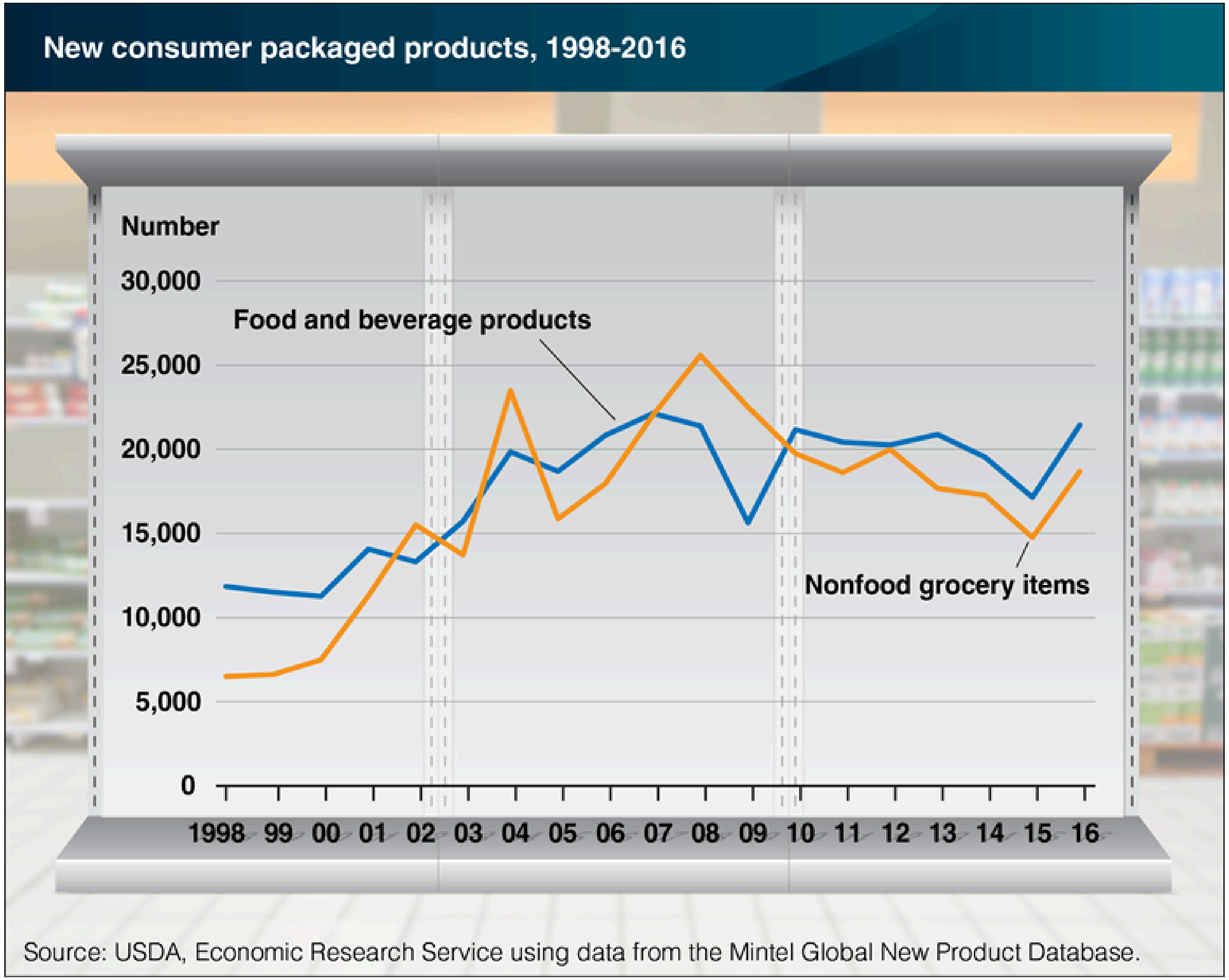
Diffusion curves





Time to Reach 100M Users
Months to get to 100 million global Monthly Active Users
78
Google Translate
Uber
70
Uber
61
Telegram
55
Spotify
41
Pinterest
30
Instagram
9
TikTok
2
ChatGPT
Source: UBS / Yahoo Finance
@EconomyApp
APP ECONOMY INSIGHTS

New Products by Year



Predicting Diffusion: Bass (1969)

- : Market size (we'll estimate this)
- : Time periods
- : Accumulated sales before time

```
- AKA "installed base"
- A(0)=0 by assumption
```

- : number of new adopters in time
- : Remaining customers who have yet to adopt,

- Bass (1969) proposed:

- : “coefficient of innovation”
- : “coefficient of imitation”
  - p and q assumed constant

Estimating Bass model via NLLS

- This is a first-order diffEQ with analytic solution
- If you have sales data by time, you can use Nonlinear Least Squares to estimate , and , i.e. choosing parameters to minimize square errors

Estimating Bass model via OLS

- Or, notice that
- We can regress on a quadratic in installed base

```
- If you want, recover p, q & M from the parameter estimates
- (3 equations with 3 unknowns)
```

- Extensions: Multiple markets, hazard models, types of “influence”

Models:estimators aren’t 1:1?

- Consider 3 OLS estimators:

```
"In theory, there's no difference between theory & practice. In practice,
there is."
Some theorical models offer multiple estimators. Some have no estimators
Estimates often differ due to assumptions and numerical properties
Subfield that invents estimators and studies their properties: "econometrics"
```

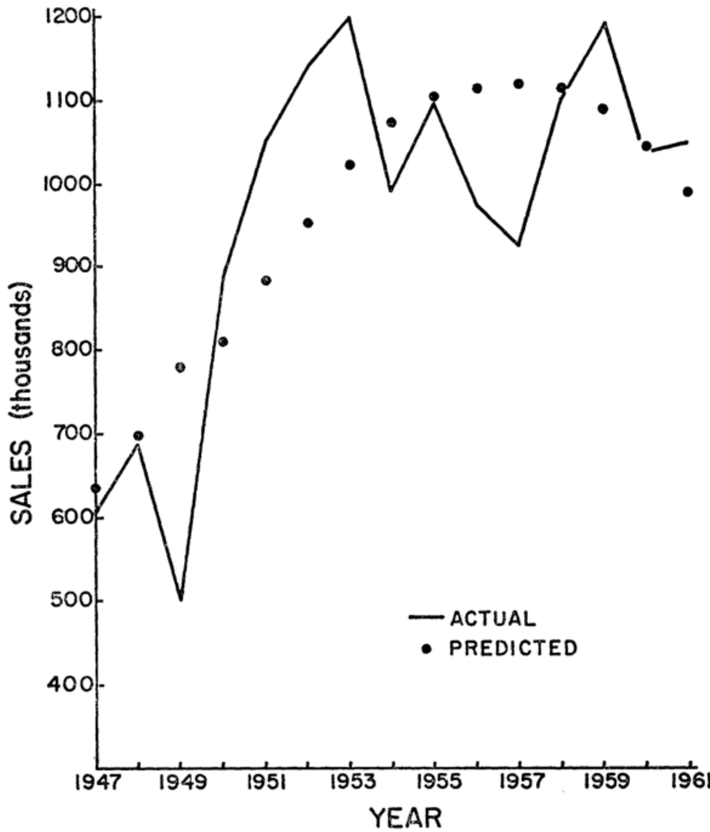


FIG. 5. Actual sales and sales predicted by regression equation (home freezers)

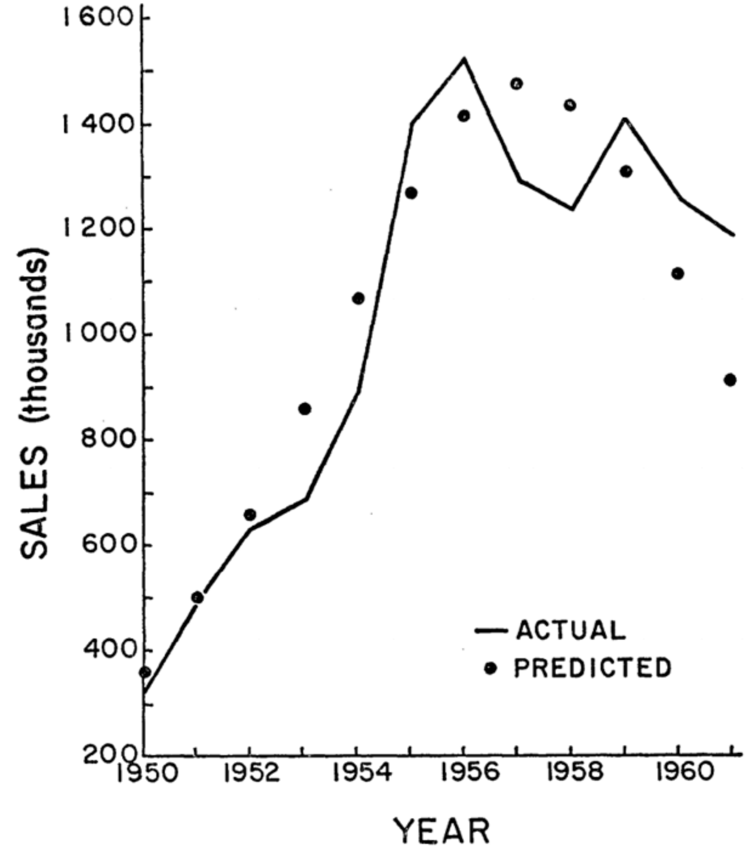
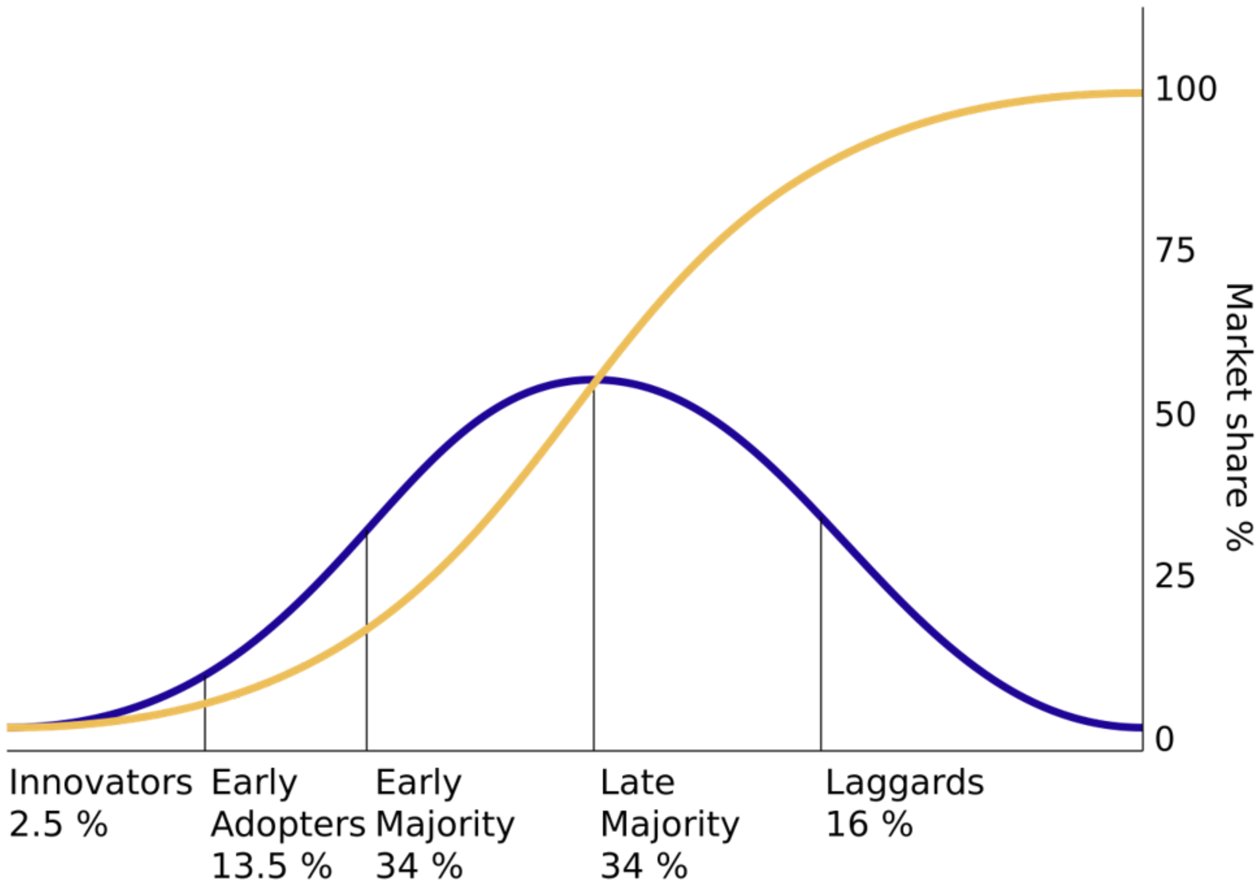
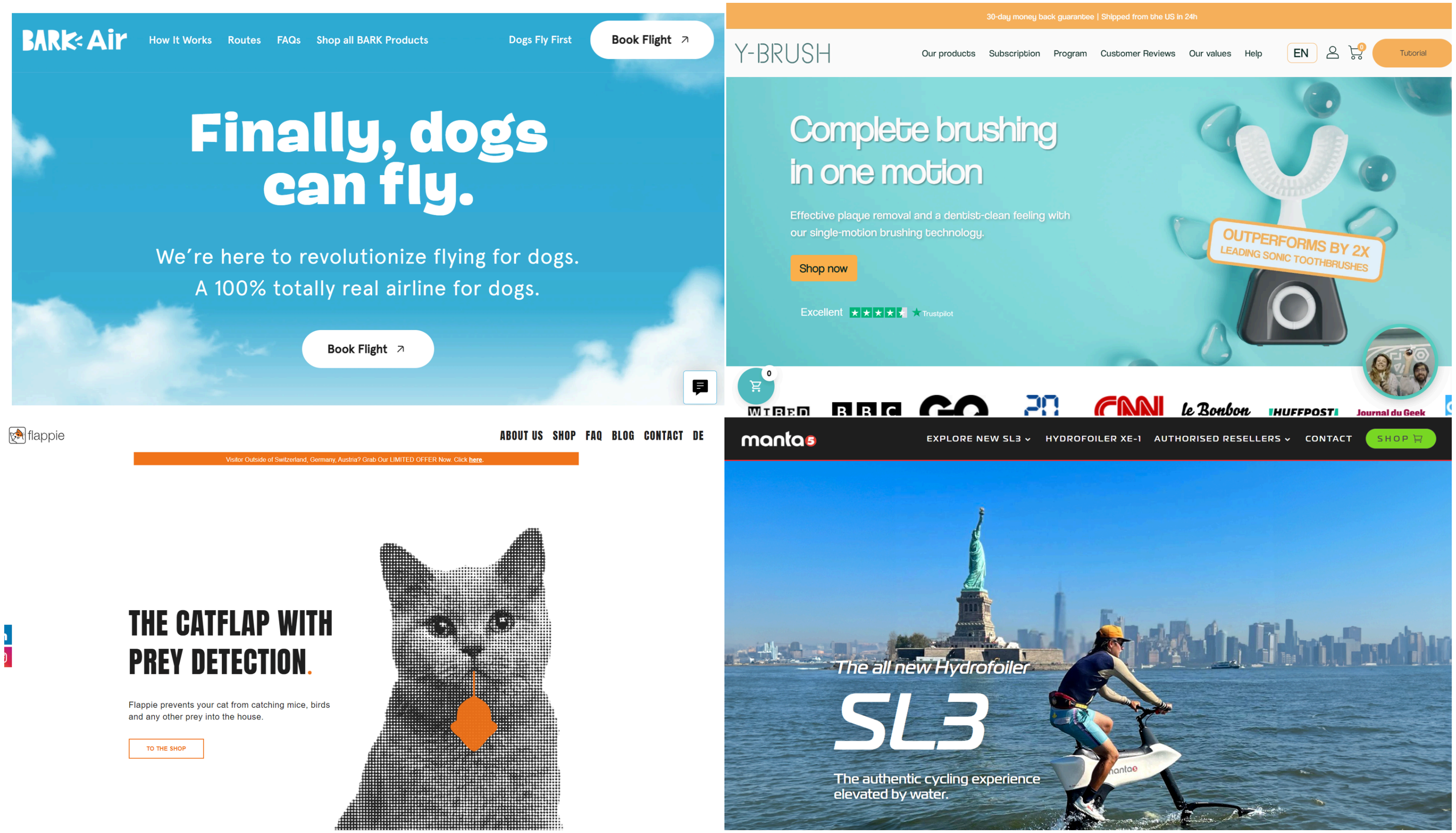


FIG. 8. Actual sales and sales predicted by model (clothes dryers)

Inspired by epidemiology



Which new products will catch on?



Rogers' ACCORD Framework (2003)

- Diffusion depends on Relative Advantage, Perceived Complexity, Compatibility, Observability, Risk, Divisibility (aka "Trialability")

```
- Summarized 40 years of research, incredibly influential on practice
- Provided diagnostics to interpret Bass (1969)'s predictive analytics
- E.g., a prototype could be evaluated on these 6 dimensions then modified
- Early example of HARKing but likely useful
```

More explanations for diffusion curves

- Heterogeneity might drive adoption timing

```
- Adoption driven by consumer preferences, needs, income, risk attitudes?
```

- Markets typically evolve after introduction

```
- Production becomes more efficient & reliable; costs fall; price may fall
- New features, technology generations, safety improves
- Competitors introduce variants targeting unserved customers
- Network effects, e.g. smartphone compatibility with chargers or
accessories
- Complementors, e.g. Verizon stores, iFixIt, Genius Bar
- Consumer preferences, e.g. expected reliability rises over time
```

anything, they're going to hold me to it.

Menu +

I think my idea on my team is to have the worst ideas so that everyone else can have slightly better ideas. What's one of your worst ideas that you would just never do that's obviously a bad idea?

I don't know. I'm sure I have a lot of bad ideas. I just don't know which ones are bad before I try them.

Fair enough.

The creative process is an interesting thing. We find this a lot at the company when we build features. If you're trying to innovate, a lot of the time, things that end up working seem like bad ideas upfront, and oftentimes, they seem like bad ideas to people who are even really good at innovating. I tell this story that, when we were originally building Discord back in 2015, we were a team of maybe 12 people at the time. Half of the people working on it thought Discord was a bad idea in a 12-person startup. It turned out to be a good idea.

It's really hard to tell beforehand whether innovation is going to work. It's really important to have space to try things, react to it, and innovate and iterate to see where it takes you. Sometimes magic comes out the other end, and a lot of times, you just get duds. But we don't ship those. We try not to ship them.

Market Size & Career Choice

- Your first job includes a bet on a market

```
- Mature market: Big, reputable employers, established career tracks
- But markets typically decline at some point

- Growing market: Exciting, high risk, high reward, more opportunity
- But market may not take off as you expect

- Will your skillset enable a transition if needed?
```

- Today’s safe option might not be safe!

```
- Mature markets will decline
- New markets will grow
- Consider diffusion trends, not just current market size
```

Core Needs Data: Google Trends

- Best available indicator of unserved market
- Google Trends reports search volume indices by keyword, place, time, service

```
- Also identifies keyword topics, trending terms & related queries
- Samples the query database, reports estimates not actual counts
- Requires a minimum query volume for reporting
- Free, so it could get sunsetted
```

Customer Lifetime Value

- CLV is the most powerful customer analytics metric

```
- Expresses the firm's value of an individual customer relationship as the
net present value of expected future customer profits
- Pioneered by catalogue retailers in the 1980s
- Has spread widely, but not yet everywhere
- CLV metrics enable quantification & serious discussion of new policies
```

CLV Example: Housing First

- “In 2005, Utah set out to fix a problem that’s often thought of as unfixable: chronic homelessness. The state had almost 2,000 chronically homeless people. Most of them had mental-health or substance-abuse issues, or both. Utah started by just giving the homeless homes…
- The cost of shelters, emergency-room visits, ambulances, police, and so on quickly piles up. Lloyd Pendleton, the director of Utah’s Homeless Task Force … said that the average chronically homeless person used to cost Salt Lake City more than $20,000/year. Putting someone into permanent housing costs the state just $8,000 [including case managers]…
- Utah’s first pilot program placed 17 people in homes scattered around Salt Lake City, and after twenty-two months not one of them was back on the streets. In the years since, the number of Utah’s chronically homeless has fallen by 74%.”

- Source

Housing First: Looking deeper

- Housing First has certainly not solved homelessness

```
- "Chronic" means disabled and unhoused for 1+ yrs, or 4x in 3 yrs
- ~28% of CA homelessness is chronic (2019)
- UT originally claimed 90% reduction, then revised their metric
definitions
- Reliable efficacy metrics are rare
- Housing First has been implemented haphazardly
- UT built new apartments. CA cities mostly use shelters, SROs, vouchers
- Key Q: Require wraparound services? E.g. Addiction treatment, etc
- Key Q: Does Housing First somehow encourage homelessness?
```

- I claim: Quantification enables bold policy shifts

```
- U.S. HUD adopted Housing First as preferred approach to homelessness in
2014
- CLV quantifies policy costs and benefits & enables ex-post evaluations
- We then can use data to refine CLV estimates and policies
```

Calculating CLV

- : planning horizon
- : contribution margin of serving customer in time
- : retention probability that customer buys in
- is the cost of capital
-
- and observable in past data; future values are predictions

CLV Example

- A tennis club charges an annual fee of $300
- The average club member spends $100 a year at the club (concessions, etc.)
- The average contribution margin on these additional expenditures is 60%
- Historically, 80% of the members rejoin the club in any given year. The club's cost of capital is 15%
- What is the club's CLV over a 1-year horizon?
- What is the club's avg. CLV over a 2-year horizon?
- What is the club's avg. CLV over a 3-year horizon?

Using CLV for Customer Acquisition

- Marketing campaigns should be profitable if Avg. Customer Acquisition Cost (CAC) CLV

```
- Caveat: So long as acquired customers have CLV>=avg CLV of existing
customers
- Often, managers impose a "fudge factor" as a speedbump
```

- Suppose the tennis club has a chance to pay $20k for a billboard. It will be seen by 100K people with an expected conversion rate of .1%. Should we do it?

- Similar “break-even” calculations possible for

```
- partnerships, opening new stores, price promotions, etc. Anything that
requires an upfront outlay to potentially acquire new customers, increase
current customer retention, or develop current customer spending
```

CLV Metrics in Practice

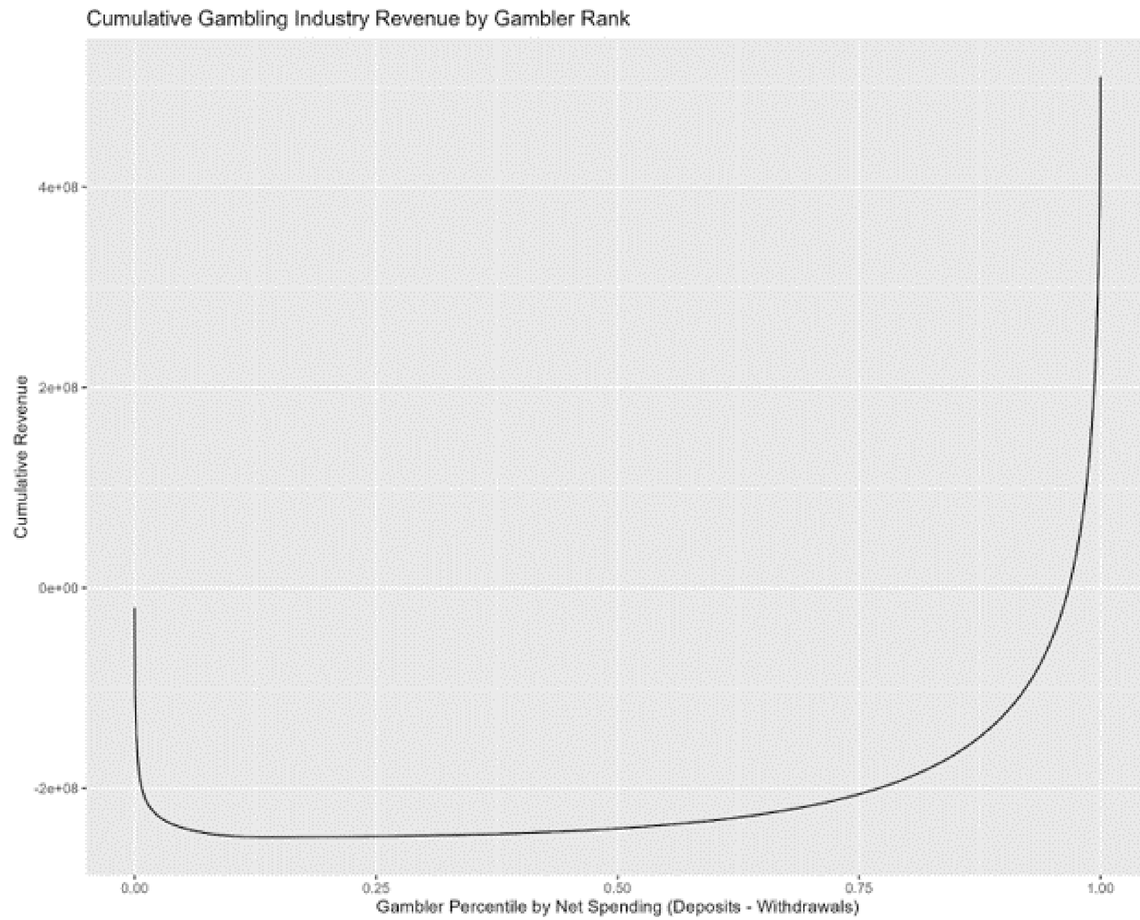
- CLV popularity rose alongside CRM data systems

```
- e.g. Oracle, SAP, Salesforce
- Services and retailers used CLV to set customer experiences: high-CLV
flyers got upgraded, high-CLV lodgers got better rooms, high-CLV shoppers
got sterling return service and attention, high-CLV callers got shorter
wait times and more consideration
```

- In the 00s, consulting firms published claims that 20-30% of customers were unprofitable. Many firms tried to “fire” unprofitable customers

```
- American Express offered some cardholders $300 to cancel their cards.
Best Buy stopped notifying some shoppers about upcoming promotions. Banks
used minimum balances and teller fees to drive away some accountholders
```

- Customers talk to each other; firing customers is a brand risk


Cumulative Gambling Industry Revenue by Gambler Rank
Cumulative Revenue
4e+08
2e+08
0e+00
-2e+08
0.00
0.25
0.50
0.75
1.00
Gambler Percentile by Net Spending (Deposits - Withdrawals)

CLV Cautions & Risks

- CRM data may be incomplete, disconnected, error prone

```
- E.g., can you connect customer identity across different credit card #s?
- Think about measurement error in retention rate
```

- CLV models are predictive analytics, not prescriptive or diagnostic

```
- CLV contains no customer purchase model
- Heterogeneous CLV-based policies may be self-fulfilling:
Treat someone as unprofitable, they may act that way
```

- Guiding Principles

```
- Firm & customer both can benefit from higher customer value. CLV metrics
for value creation benefits everyone
- Firm & customer have opposing interests in price. CLV for margin
extraction may generate perverse customer incentives
- Customer dissatisfaction may reveal CLV flaws; requires careful attention
- Hence many firms now measure point-of-sale satisfaction
```

Example: CLV for Pricing

- Assume price is $100, margin $50, retention 50%, discount rate 10%, horizon T=1

```
CLV = 50 + .5 * 50 / 1.1 = $72.72
```

- Suppose you consider increasing price to $120, holding all else constant

```
Then, CLV = 70 + .5 * 70 / 1.1 = $101.82
```

- Should you raise the price?

```
What is this analysis missing?
How could we resolve that problem?
Hint: CLV is predictive analytics, not diagnostic or prescriptive
```

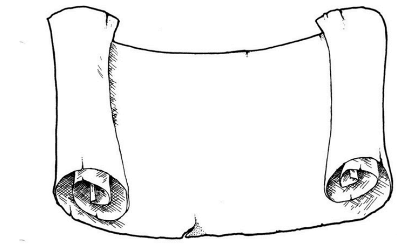
- Suppose you tutor a high school student for $250/month. There is a 10% chance the student will move and find a new tutor in any given month. There are six months remaining in the school year, at which point the tutoring arrangement will definitely end.

- What missing information do you need to calculate the current value of the tutoring arrangement?

Wrapping up

Class script

- Let's estimate the Bass model



Recap

- Customer acquisitions best predict revenues
- Market size estimates how many people share a core need
- Diffusion models predict how served market changes
- CLV metrics can quantify and enable novel policies

RECAP

Going further

- Innovation Diffusion and New Product Growth Models (Peres et al. 2009)
- Customer-Base Valuation in a Contractual Setting with Heterogeneity (Fader & Hardie 2010)
- Exploring the Distribution of Customer Lifetime Value (Fader & Hardie 2017)
- Firing your best customers (Avery & Fournier 2012)