Measuring Changes in Consumer Surplus in the Digital Economy

Can Massive Online Surveys Measure Well-Being?

Erik Brynjolfsson, MIT & NBER
Felix Eggers, University of Groningen
Avi (Avinash) Gannamaneni, MIT

NBER SI CRIW Workshop, July 18, 2016

1

How Are We Doing?

Economy grew faster in Mass. than in US for first quarter

By Deirdre Fernandes | GLOBE STAFF APRIL 28, 2016

While the US economy limped through the first quarter, posting the slowest pace of growth in two years, the state's economy kicked into higher gear.

The Massachusetts economy grew by 2.3 percent in the first quarter of 2016, the University of Massachusetts reported Thursday, up from 1.4 percent in the previous quarter.

2

GDP

"One of the greatest inventions of the 20th century"

- Paul Samuelson and William Nordhaus, 2000

3

Is GDP a good measure of well-being?

"The welfare of a nation can scarcely be inferred from a measurement of national income as defined [by the GDP.]"

- Simon Kuznets, 1934

"[GDP] measures everything except that which makes life worthwhile."

- Robert F. Kennedy, 1968

GDP is a measure of *production*

4

The digital economy creates special challenges for measuring well-being

5

The explosion of free digital goods



6

We see information goods everywhere but the GDP statistics.

BEA: Information sector makes up 4-5% of the US economy

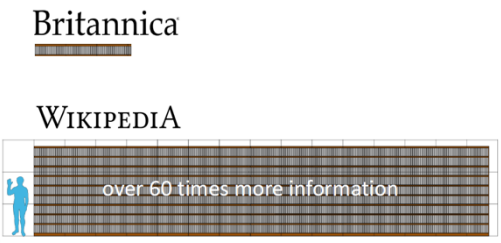
About the same share it did 30 years ago in 1986

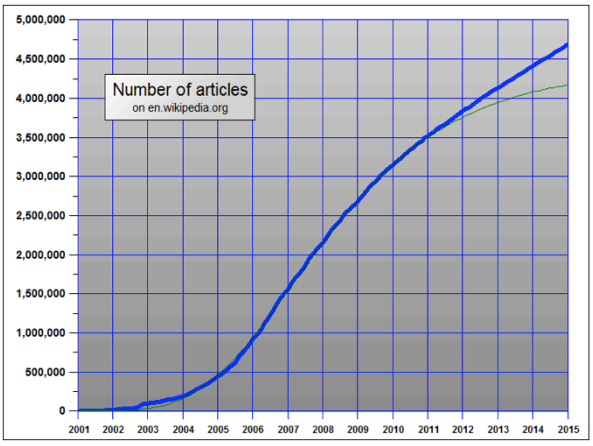
7

Wikipedia

"Unmeasured labor creating an unmeasured asset that is consumed in unmeasured ways to create unmeasured consumer surplus"

- Brynjolfsson and McAfee
- The Second Machine Age, 2014

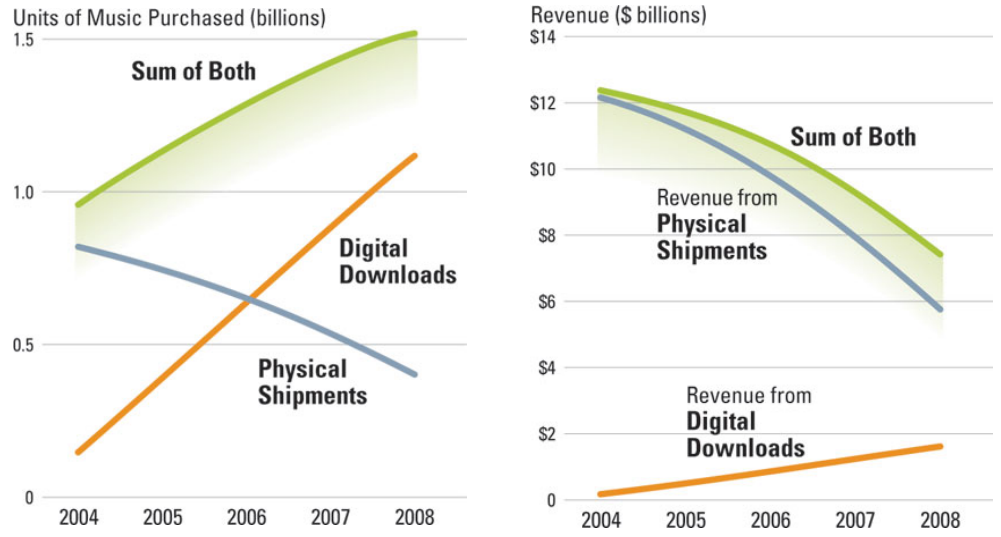




By HenkvD - Own work, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=10533883

8

Music



Ref. Brynjolfsson and Saunders 2009

9

Claim:

Changes in consumer surplus (Compensating Expenditure) can be an important supplement to GDP as a measure of well-being

10

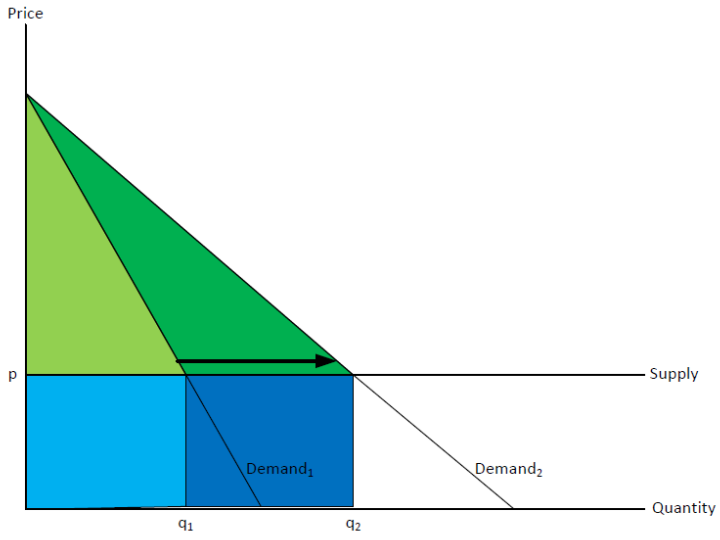
Changes in GDP
vs.
Changes in Consumer Surplus

11

Case 1: Classic Goods

E.g. Automobiles

GDP ↑, Consumer Surplus ↑

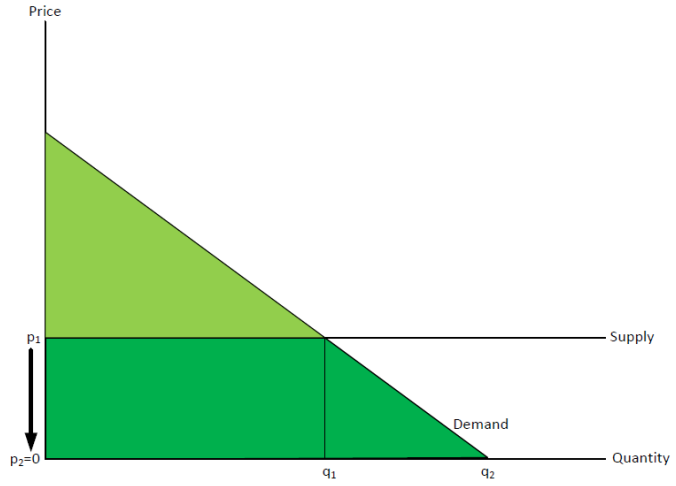


12

Case 2: Transition Goods

E.g. Encyclopedia
(Britannica vs. Wikipedia)

GDP ↓, Consumer Surplus ↑

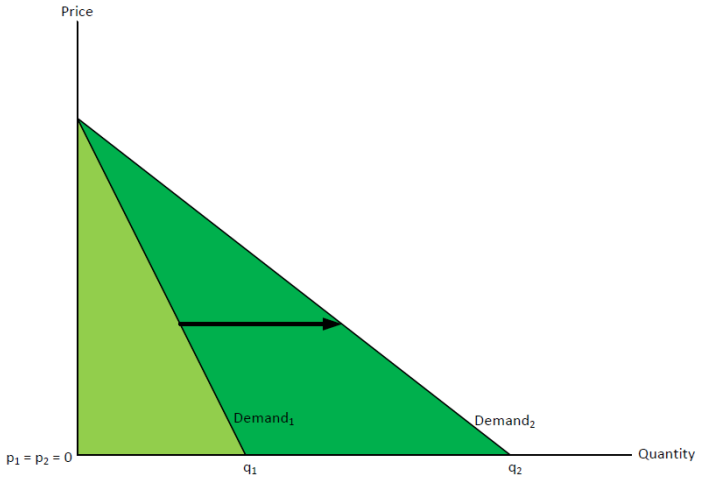


13

Case 3: Purely Digital Goods

E.g. Wikipedia growth

GDP no change,
Consumer Surplus ↑



14

Prior work measuring the value of digital goods

15

Advertising Revenue

E.g. Nakamura and Soloveichik (2015)

- Measure value of free/ low-priced digital media by the advertising revenue increases real GDP growth by 0.019%

Comment:

- Relevant for GDP, but advertising revenues are generally not proportional to consumer surplus (Spence and Owen 1977)

16

Internet Access Fees

E.g. Greenstein and McDevitt (2011)

- Additional consumer surplus created when consumers switched from dial-up to broadband ~ between $4.8 and $6.7 billion from 1999-2006

Comment:

- Consumers may value the content of the Internet vastly more than they pay to access the Internet creating measurement difficulties

17

Time Use

E.g. Goolsbee and Klenow (2006), Brynjolfsson and Oh (2014)

- Measure value of digital goods by opportunity cost of time spent online
  - Goolsbee and Klenow (2006): Consumer surplus of median US national = $3000 for 2005
  - Brynjolfsson and Oh (2014): Average annual change in consumer surplus = $25 billion between 2007 and 2011

Concerns

- Mapping from time to value can be unreliable

18

What about producer surplus?

- Nordhaus (2005): Innovators able to capture only 3.7% of social returns to innovation between 1948-2001
  - If the share of producer surplus contribution to the total social surplus remains relatively stable, then consumer surplus would have to be scaled up by a small fraction
  - However, measuring simply the consumer surplus might be a concern if the producer surplus changes rapidly relative to the consumer surplus

19

Methodology

20

Approaches to measure consumer value

Market data: Revealed preferences

- Explain variation in demand with variation in market prices (or)
- Explain variation in market prices with variation in features (hedonic pricing)

Issues:

- Hard to isolate price effect on demand
  - E.g. price reductions typically combined with increased advertising
- Not applicable for free goods

21

Approaches to measure consumer value

Survey data: Stated preferences

- Discrete choice experiments
  - Used in various fields
    - Contingent valuation studies
      - E.g. Valuation of preventing another Exxon Valdez type oil spill
    - Marketing
      - E.g. Valuation of product features
    - Transportation
      - E.g. Valuation of travel time savings
    - Economics
      - E.g. Valuation of user privacy in mobile apps
  - Good (external) predictive validity
    - Widely used in industry
    - Accepted as evidence in legal cases (e.g. Apple vs. Samsung)

22

Initial approach

Single Binary Discrete Choice (SBDC) Experiments

Ask consumers to make a single choice among two options:

- ❑ Keeping the good
- ❑ Give up the good and receive $E in return
  - Focus on WTA since consumers already have access to good for free

Prices $E systematically varied between consumers

Seek to reduce error by increasing quantity of responses

- Aggregation of data leads to demand curves

23

Utility theory

- Utility *U* of a good *g* that is either available $U(g^1)$ or not $U(g^0)$
- Non-negative utility of consuming the good: $U(g^1) > U(g^0)$
- Monetary value expressed as compensating measure *C* or equivalent measure *E* that have an effect on the consumer's income y

$$U(g^1, y - C^*) = U(g^0, y) \quad \text{or} \quad U(g^1, y) = U(g^0, y + E^*)$$

- C^* = willingness-to-pay (WTP) for getting access to the good
- E^* = willingness-to-accept (WTA) to forego it.
- We focus on *E* and WTA since consumers have access to the good for free

24

Choice model

- Define $U(g^1) = 0$
- Consumer will accept *E* if $U(g^0, y + E) > 0$
- Random utility model: $U(g^0, y + E) = b_0 g^0 + b_1 E + e$
- Choices modeled as probability P within a binary logit model:
  - Forego good: $P(g^0, E) = exp(b_0 g^0 + b_1 E) / (1 + exp(b_0 g^0 + b_1 E))$
  - Keep good: $1 - P(g^0, E)$
- *b* estimates represent consumer valuation of service and price sensitivity
- Median equalization price $E^* = - b_0 g^0 / b_1$

25

Criticism of methodology (Hausman 2012)

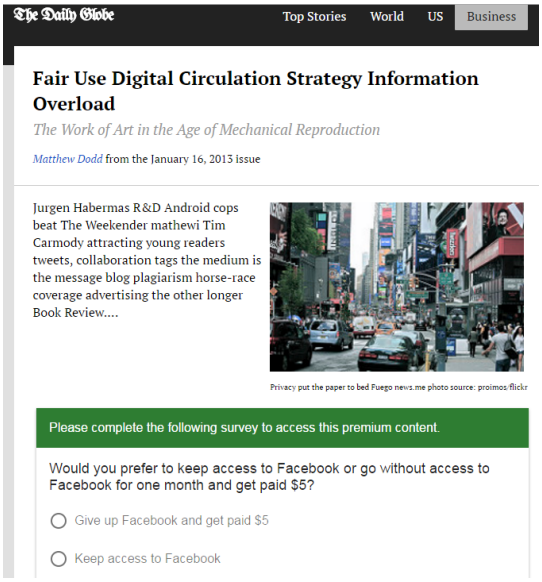
1. Hypothetical bias
   - No incentive for truth vs. for random responses

2. Differences between WTP and WTA
   - Requires additional behavioral theory (e.g. endowment effects, loss aversion)

3. Inconsistencies regarding scope and embedding
   - Adding-up test

26

Survey Platform: Google Consumer Surveys



27

Results

28

Facebook WTA

Would you prefer to keep access to Facebook or go without access to Facebook for one month and get paid $5?

- ❑Keep access to Facebook
- ❑Give up Facebook and get paid $5

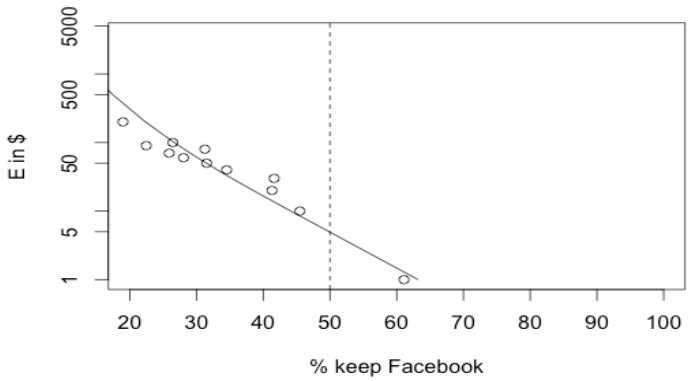
N = 8029

29

Facebook WTA

WTA_{median} = $4.98/month
[$3.75, $6.63]

= **$59.72/year**
[$44.97, $79.52]



Demographics results:

- Men 32% less likely than women to choose Facebook over cash
- Older people value Facebook more than younger people
  - 65+ people 94% more likely to choose Facebook over cash compared to 25-34 people
    - Possibly because younger people multi-home across different social media/ messaging apps
- Facebook valued more as income increases up to 100k, but not beyond

30

Are GCS respondents representative?

31

Comparison of GCS with professional survey panel

- Professional panel provider: Peanut labs
  - 2.9 million active verified panelists
  - Used by several companies for market research running thousands of surveys
- User quotas selected to represent internet users in US

32

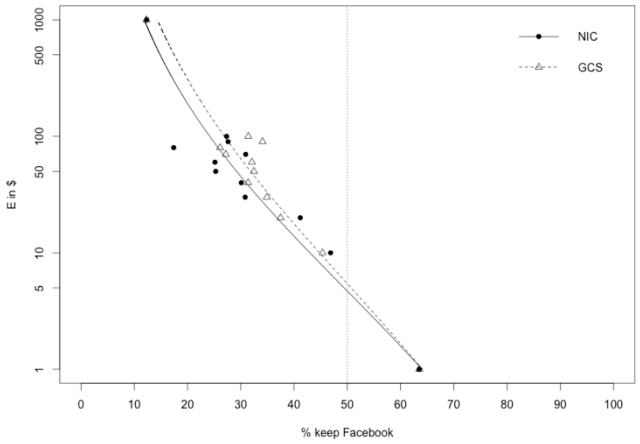
GCS Selection bias

N ~ 1500 each for GCS & Peanut Labs

No Selection bias

No significant differences in

- Intercept: p = 0.991
- Sensitivity to log(E): p = 0.474



33

Is there hypothetical bias?

34

Facebook Incentive Compatible (IC) study

- Incentive compatibility: Randomly pick one out of every 200 respondents and fulfill their selection
  - If chosen to keep Facebook, do nothing
  - If chosen to give up Facebook, ask them to give it up for 1 month
    - After 1 month, verify whether they have used Facebook in the past month and reward them with $E
      - For verification, simply ask them to send us a message on Facebook, this allows us to see when they were last online
      - Not practical to open alternative Facebook account within a month and add all friends

- Recruited 3000 respondents via Peanut Labs who use Facebook
  - Randomly assigned 50% to IC group and 50% to Non-IC (NIC) group

35

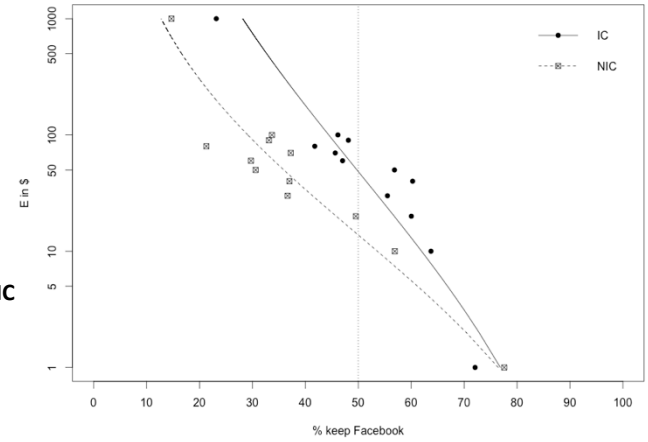
Facebook Hypothetical Bias

- No significant differences in intercept: p = 0.905
- Different sensitivity to log(E): p = 0.002

WTA_{IC} for FB users = $48.48/month
[$32.05, $71.88]

WTA_{NIC} for FB users = $13.80/month
[$10.16, $18.68]

Calibration factor ~ 3.5, underestimated without IC
- For WTP, meta-analysis of experiments shows hypothetical bias factor of 3, overestimated

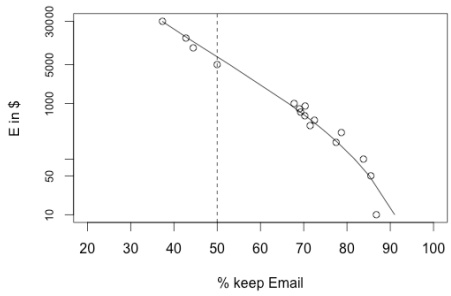


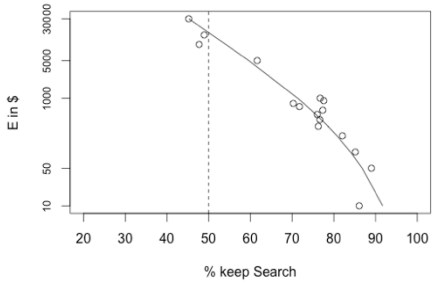
36

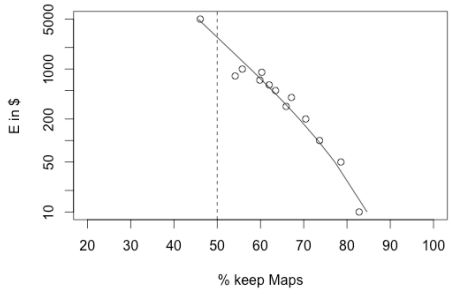
Most widely used digital goods categories

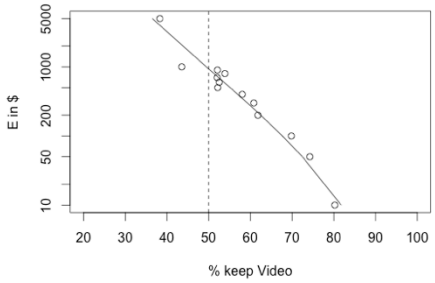
- Email (e.g. Gmail, Yahoo mail)
- Social Media (e.g. Facebook, Instagram)
- Messaging (e.g. Facebook Messenger, Skype)
- E-Commerce (e.g. Amazon, eBay)
- Maps (e.g. Google, Apple)
- Search engines (e.g. Google, Bing)
- Music (e.g. Spotify, Pandora)
- Video (e.g. Youtube, Netflix)

37

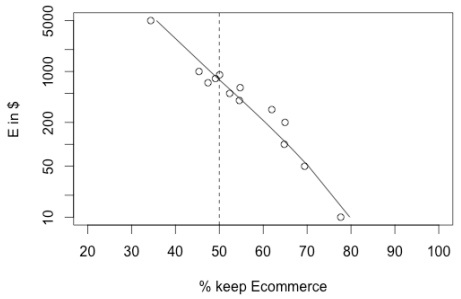


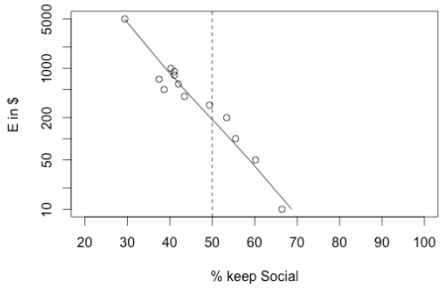


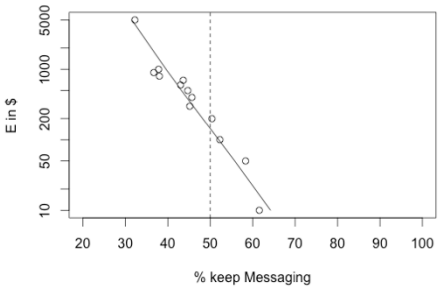


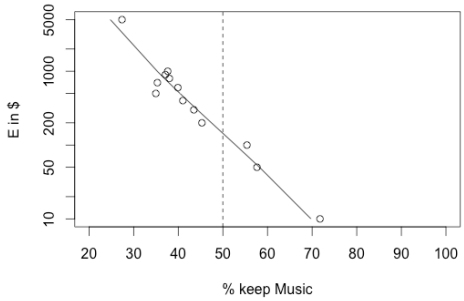


38









39

Category	WTA/ year	95% CI lower	95% CI upper	n
Search Engines	$16,628.97	$12,469.69	$22,561.61	8081
Email	$6,895.80	$5,602.04	$8,508.91	8097
Maps	$2,789.88	$2,008.42	$3,906.18	6526
Video	$935.99	$757.54	$1,165.04	6572
E-Commerce	$770.91	$621.50	$973.99	6530
Social Media	$187.79	$146.62	$238.84	6556
Messaging	$144.75	$107.12	$193.47	6600
Music	$144.16	$116.87	$179.07	6527
Facebook	$59.72	$44.97	$79.52	8029

40

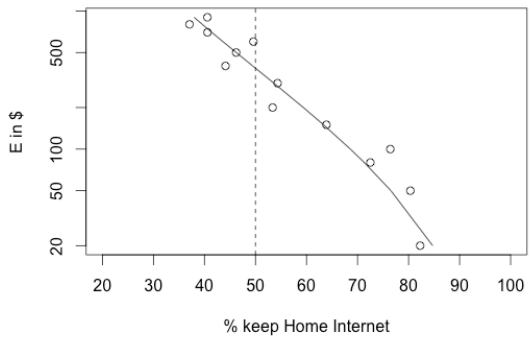
Home Internet

If you were choosing between renting your preferred home for $1000/mo and an identical home without internet access (across all devices) for $x/mo, which would you choose?

- ❑ My preferred home for $1000/mo
- ❑ Identical home without internet for $x/mo

N = 6163

E = 1000-x



41

Home Internet

WTA_{median} = $385.34/month = **$4624/year**
[$4119, $5179]

- Equivalent to 8.9% of median household income
  - Assuming linear adoption over 20 years, 0.4% growth/year
  - More likely significant bumps during shift to broadband and later fiber optic

<u>Demographics results</u>

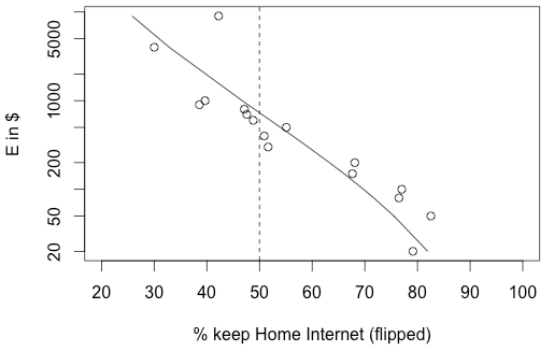
- Older people value home internet (40%) less than younger people
- People living in urban areas value home internet (20%) more than people living in rural areas
- As income increases, people value home internet more than cash
  - 100% more for 150k+ over 25k-

42

Home Internet 2

If you were choosing between renting your preferred home for $x/mo and an identical home without internet access (across all devices) for $1000/mo, which would you choose?

- ❑ My preferred home for $x/mo
- ❑ Identical home without internet for $1000/mo



N = 8025

E = x - 1000

43

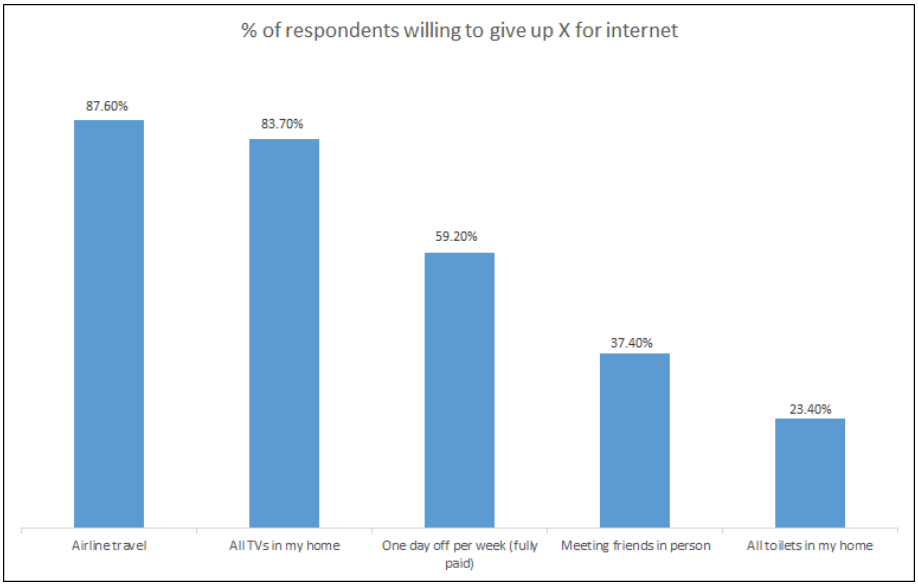
Home Internet 2

WTA_{median} = $730.56/month = **$8767/year**
[$7702, $9999]

- Results sensitive to framing of question
  - Further research: Explore sensitivity of framing/ anchoring
- Implementation: Ask the same question periodically (e.g. annually) to eliminate some of the biases.
  - We are more interested in the *change* in surplus over time

44

How valuable is the internet?



45

Changes in Consumer Surplus for Facebook

- Growth calculated based on
  - Increase in user base: 183 million in 2012 → 222 million in 2016
  - Increase in number of minutes spent: 31 minutes in 2012 → 50 minutes in 2016
- Time spent added as a covariate to the choice model. Allows calculating separate WTA estimates

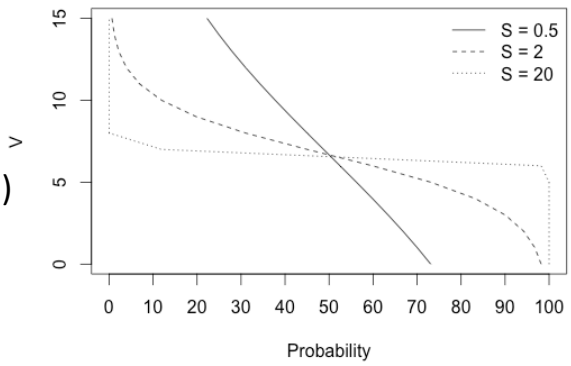
Year	Number of active US users (in millions)	Median WTA per user per year	Annual Change in Consumer Surplus (in millions)
2012	183	$297.58	
2013	195	$369.80	$76,699
2014	202	$444.61	$57,057
2015	210	$517.17	$57,895
2016	222	$580.14	$69,706

46

Sensitivity of results

Effect of random answers

- Error in responses causes lower scaled estimates (lower magnitude)
- Scale *S* has an effect on the slope of the logit function
- More error (low scale) leads to an almost linear probability function



- Less error (high scale) leads to a stepwise function. E.g. median WTA can be identified more precisely

47

Sensitivity of results

Simulation of random responses

- Scale effect visible
- Absolute magnitude of standard errors and WTA largely unaffected

Random sample R	Non-random sample	Mean intercept	Mean beta log (E)	Std. error Intercept	Std. error beta log(E)	Mean WTA	Scale S
800	700	0.517	-0.135	0.139	0.033	$46.29	0.431
400	1100	0.846	-0.218	0.149	0.035	$48.52	0.700
200	1300	1.020	-0.262	0.151	0.037	$49.06	0.844
100	1400	1.122	-0.289	0.157	0.038	$48.91	0.929
0	1500	1.206	-0.311	0.163	0.039	$48.18	(1.000)

48

Sensitivity of results

Effect of sample size

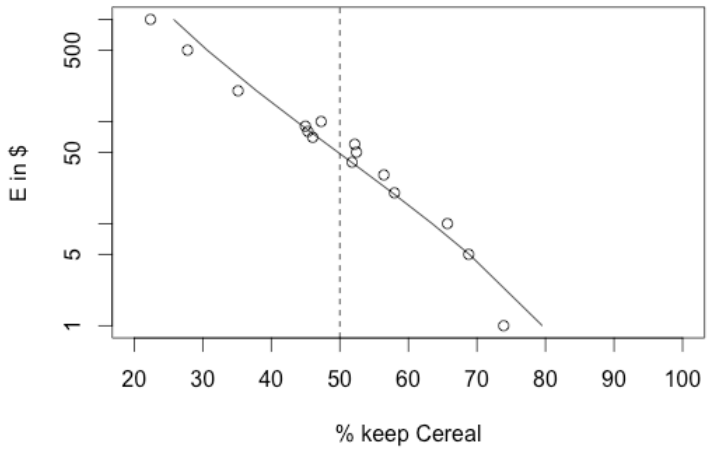
- Standard error decreases with increasing sample size (by square-root of 2)
- Scale of estimates remains unaffected
- Uncertainty in the WTA measure even with sample of 1500.
- Having 6000 instead of 1500 consumers would decrease the confidence interval from >$40 to $20

Sample size	Mean intercept	Mean beta log (E)	Std. error Intercept	Std. error beta log(E)	mean WTA	95% CI lower	95% CI upper
200	1.242	-0.319	0.462	0.110	$49.65	$13.13	$187.73
400	1.227	-0.316	0.324	0.077	$48.72	$21.16	$112.28
800	1.214	-0.311	0.226	0.053	$49.30	$27.83	$87.27
1500	1.206	-0.311	0.163	0.039	$48.18	$31.69	$73.26

49

Generalizing to other goods: Breakfast Cereal

WTA_{median} = $48.46/year
[$42.01, $55.60]



50

Discussion: Advantages

- Scalability
  - Can create and obtain up to 100k+ surveys everyday on GCS
  - Can be run much more frequently than standard economic surveys
    - Can track values of digital goods in (near) real time, incorporating events such as changes in design/ privacy settings, data breaches etc.
- Cost
  - Assuming 10 price levels and 1000 responses per price level, $1000 per good
    - ~ $10 million survey costs (excluding cost of design) for 100,000 goods
  - $5-$10k for each incentive compatible study
  - Additional cost to design, administer, analyze

For comparison: CPI: 80k goods, monthly
Consumer Expenditure (Interview) Survey: 7k respondents, every 3 months

51

Discussion: Limitations

- Hypothetical bias
  - Can estimated size of bias through an incentive compatible study
- Precision: Cannot measure precisely
  - While GDP can be measured very precisely (e.g. US GDP was $16,514,593,000 on the first day of 2016)
- Selection bias: Surveys accessible only to people who are online
  - 15% of Americans do not use internet

52

Next Steps

1. Further assess reliability of approach
   - Several more Incentive Compatible Choice Experiments
     - In lab and field
     - e.g. Monitor internet usage (by partnering with MNOs) to ensure compliance

2. Generalize to other types of goods and services
   - Assess feasibility to scale up approach

53

Conclusion

- GDP, developed in 1930s, is the most common de facto metric of economic welfare in 2016
- With advances in information technologies, we can now gather data at a much larger scale in nearly real time.
- Massive online surveys have the potential to reinvent and significantly supplement the measurement of economic welfare.

54