Information Acquisition and Consumer Choice

early draft … comments welcome

Tim Bresnahan
Tim Landvoigt
Pai-Ling Yin

January, 2011

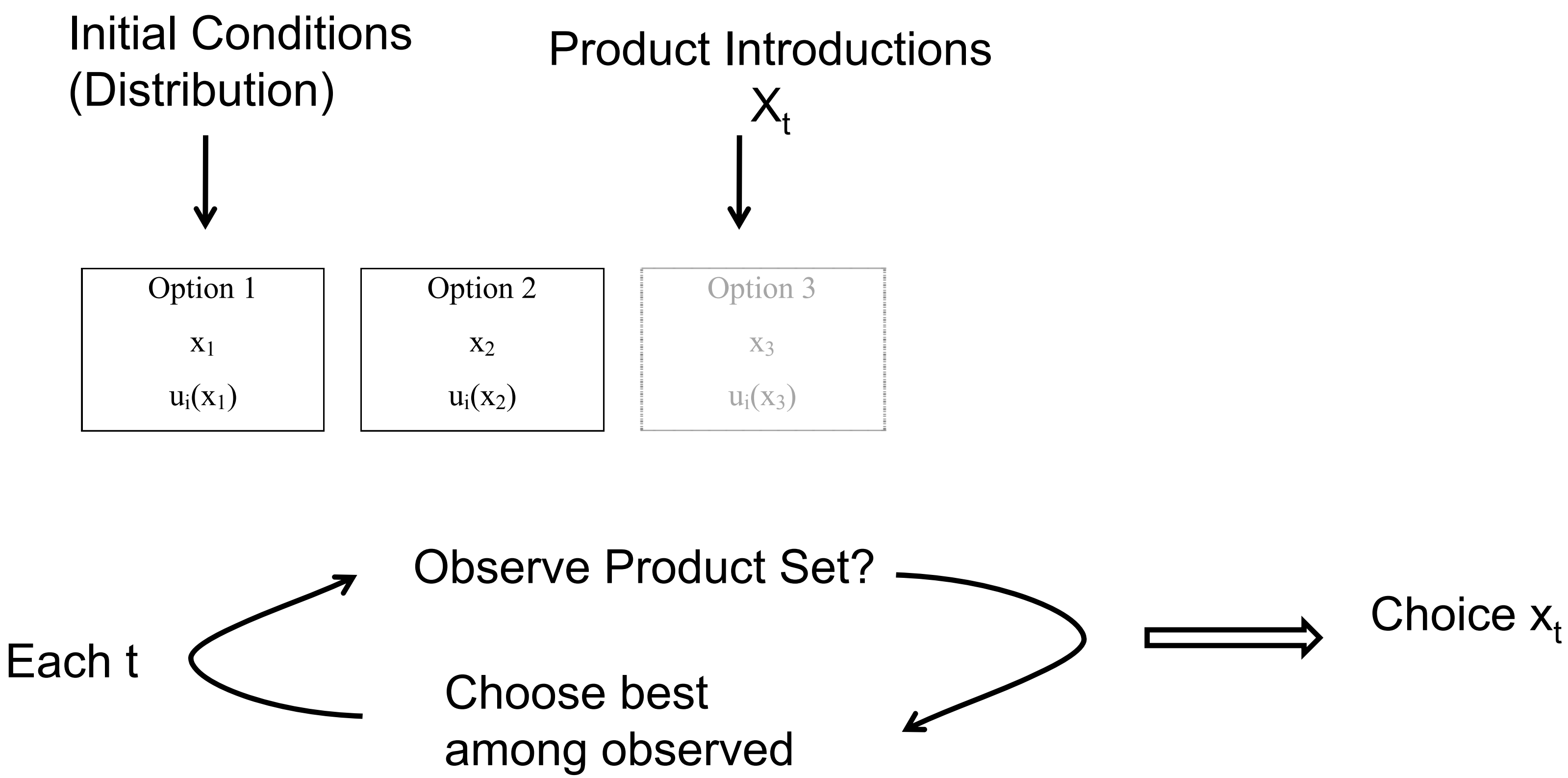
Overview and Motivation

- We estimate a model of consumers who
  - Might not know their choice set
  - Because new products are coming in (they know the process…)
  - In each period
    - Observe full choice set?
    - Potentially choose a new product | observation
- On a sample of late 1990s web browser users
- Because
  - “Distribution Convenience” very influential (B&Y 2005, others)
    - On brand choice NS vs IE
    - On updating to newest version
  - Why? Could be limited information. Could be download costs.
- We observe both consumer choice and (a trace of) information

  Estimate model for IE users on Windows OS

  Ignorance of choice set, not high access costs, can explain value of distribution

Framework


Initial Conditions
(Distribution)
Product Introductions
X_t
Option 1
x_1
$u_i(x_1)$
Option 2
x_2
$u_i(x_2)$
Option 3
x_3
$u_i(x_3)$
Each t
Observe Product Set?
Choose best
among observed
Choice x_t

Consumer's Optimization Problem (Model)

- Primitives
  - $u(x_t, z_i)$ flow utility
  - $\Phi(z_i)$ transactions cost of switching (e.g. download)
  - $\Psi(z_i)$ hazard for becoming informed
  - Knowledge of product introduction and improvement probabilities

- State variables

x_t	Technical level of product the consumer is using at the beginning of time t
τ_t	Number of periods since the last time the consumer observed the true state
$\tilde{x}_t$	$\tilde{x}_t = X_{t-\tau}$ -- Technical level at last observation
$\tilde{s}_t$	$\tilde{s}_t = s_{t-\tau}$ -- release status at last observation

Optimal Program
Value Function and Optimization

Value Function before observing	$W(x_t, \tilde{x}_t, \tilde{s}_t, \tau_t)$
Value function if observe	$V^O(x_t, s_t, X_t)$
	$V^O(x_t, s_t, X_t)=\max\{u(x_t)+\beta\, W(x_t, X_t, s_t, 0),\, u(X_t)- \varphi +\beta\, W(X_t, X_t, s_t, 0)\}$
.. if don't	$V^N(x_t, \tilde{x}_t, \tilde{s}_t, \tau_t)$
	$V^N(x_t, \tilde{x}_t, \tilde{s}_t, \tau_t)=u(x_t)+\beta\, W(x_t, \tilde{x}_t, \tilde{s}_t, \tau_t+1)$
Updating rule	$W(x_t, \tilde{x}_t, \tilde{s}_t, \tau_t)=(1- \psi)\, V^N(x_t, \tilde{x}_t, \tilde{s}_t, \tau_t)+ \psi E[V^O(x_t, s_t, X_t)\mid \tilde{x}_t, \tilde{s}_t, \tau_t]$

Sample and Data

- Survey microdata from GVU
  - We merge in web server logs
    - User fills out survey at GVU about web;
    - Web server captures browser / OS they are using
- Twice a year, 7 surveys
  - Fall 1995 to Fall 98
  - Begin right after IE 1 until last available
- 48,412 total observations
- We use IE on Windows users
  - 5556 observations

Dependent Variables

- *NEWEST*
  - Is user running (in fact) newest available to them
  - Client => Server (“user-agent”)
  - .402
- *RIGHT*
  - Does user correctly report *NEWEST*
  - Human user ⇔ “user-agent”
  - .561

GVU Sampling

- Oversamples ‘net-heads
  - (NS, Mac, etc. frequent)
- We keep only IE, Windows … undersamples ‘net-heads
- Blocks of missing responses (didn’t go to demographics page)
- Have many candidate z
- “Are you using the latest version of your browser” observed in only 3 waves

Empirical Realization: Observation Hazard Ψ_i

- We have two models
  - Ψ_i is a consumer attribute
  - Ψ_{it} depends on consumer (optimally chosen) search effort
    - Search cost is a consumer attribute
- Estimates today show first model (will be obvious why at end)

$$\psi_i = [1+\exp(-z_i\, \theta_p)]^{-1}$$

Empirical Realization: Adoption hazard | Observation

- If observation occurs at time t, adoption hazard is

$$\Pr([u(x_t)+\beta\ W(x_t,\ X_t,\ s_t,\ 0;\ \varphi_i,\ \psi_i) < u(X_t)-\ \varphi_i+\varepsilon_t+\beta\ W(X_t,\ X_t,\ s_t,\ 0;\ \varphi_i,\ \psi_i)]$$

- Call this

$$H(\varphi_i,\ \psi_i;\ \underline{x}_t,\ \underline{X}_t, s_t)$$

- Note both φ_i, ψ_i are functions of z_i

$$\psi_i = [1+\exp(-z_i\ \theta_p)]^{-1}$$

$$\varphi_i = \varphi_{max}\ [1+\exp(-z_i\ \theta_d)]^{-1}$$

- Model for *RIGHT*

$$\Pr(\mathit{RIGHT}_i) = \mathrm{logit}(\alpha_0\ W_i + \alpha_1\ z_i\ \theta_p + \alpha_2 \mathit{NEWEST}_i)$$

Empirical Realization: Initial Conditions

- Each user is assigned a range of dates for IC observation time
  - For each t in the range, knew X_t, etc.
  - Likelihood will be weighted over range of dates

- Some users' survey date and OS => positive probability newest browser came with their new computer: Pr(.)=*ADV_NEWEST*
  - Calculated from aggregate PC sales
- Survey date, OS, (self reported) time on web => range of initial conditions
  - Weighted by aggregated PC sales.

Almost done! Three (very important) Chores remain

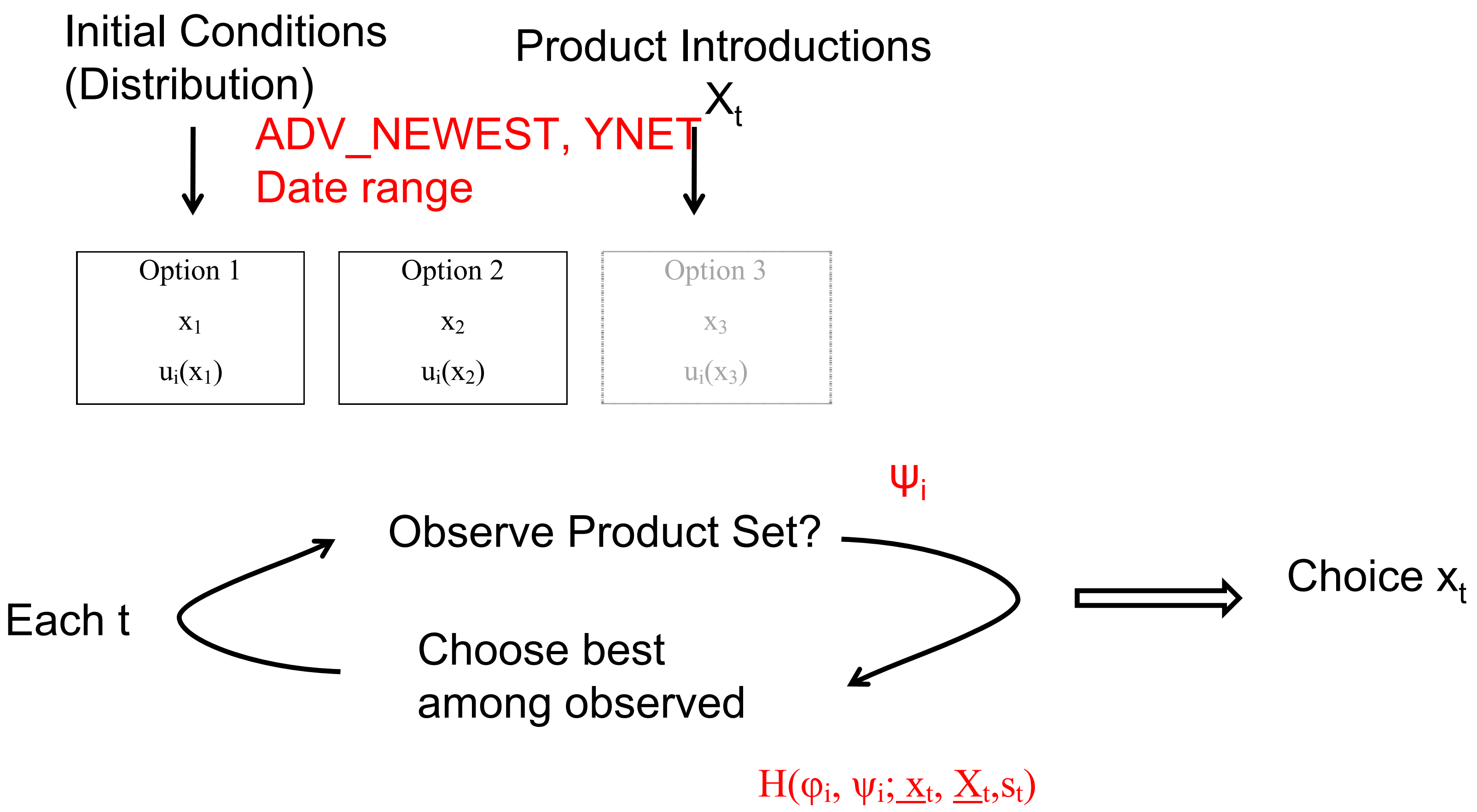
1. Solve the model for many combinations of φ_i, ψ_i

2. For each consumer, for each date in the initial conditions range, calculate the probability of NEWEST at time of observation
   accumulating observation hazard ψ_i, adoption hazard $H(\varphi_i, \psi_i)$

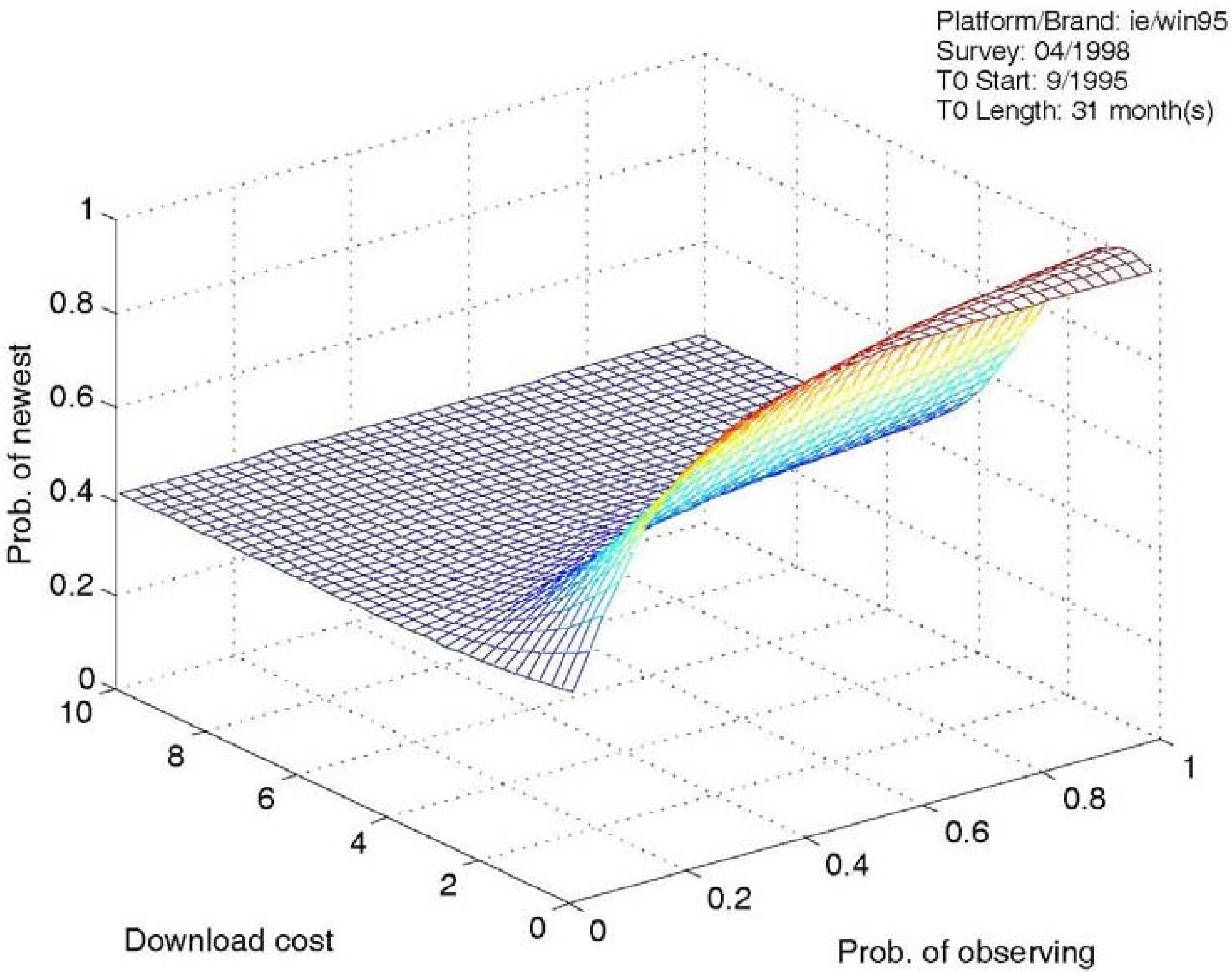
3. Weight across dates in the IC range

4. Approximate $\Pr(\text{NEWEST}; \varphi_i, \psi_i)$ as (long) polynomial

Revisit Framework



Pr(Newest; φ,ψ) for a case



Descriptive Stats of Data used in Estimation
(z chosen mostly to span space of tastes)

	Mean	Std Dev	Minimum	Maximum
NEWEST[1]	0.40266	0.49052	0.00000	1.00000
RIGHT[2]	0.56154	0.49628	0.00000	1.00000
IE	1	0	0	1
AWIN98	0.052556	0.22317	0.00000	1.00000
SURVEY	7.70644	1.70065	4.00000	10.00000
ADV_NEWEST	0.11009	0.16603	0.00000	0.62867
DHOURS[3]	0.081353	0.27340	0.00000	1.00000
I_HOURS	0.14147	0.12185	0.00000	0.50000
I_USE	1.14483	0.89117	0.00000	2.52000
I_PAYWORK_DK	0.25198	0.43419	0.00000	1.00000
DSPEED	0.95554	0.20613	0.00000	1.00000
LSPEED	3.76964	1.70072	0.00000	11.96582
OCCCOMP	0.27178	0.44492	0.00000	1.00000
MALE	0.71220	0.45278	0.00000	1.00000
DAGE	0.013499	0.11541	0.00000	1.00000
I_AGE	0.36144	0.13610	0.00000	0.83000
INCNS	0.13301	0.33962	0.00000	1.00000

Model estimates

	φ_i		ψ_i			*RIGHT*	
	Est	SE	Est	SE		Est	SE
CONST	0.906	0.629	-1.855	0.402	CONST	7.074	0.723
DHOURS	-0.521	0.307	-0.142	0.236	AWIN98	-1.354	0.233
HOURS	-0.743	0.728	0.491	0.642	SURVEY	-0.811	0.078
USE	-0.120	0.106	0.225	0.109	α	0.669	0.192
PAYWORK	-0.097	0.199	-0.713	0.212	δ	2.719	0.124
DSPEED	-1.965	0.709	-1.242	0.500			
LSPEED	-0.459	0.190	-0.102	0.043			
OCCCOMP	-0.247	0.171	0.265	0.174			
MALE	-0.023	0.185	0.827	0.266			
DAGE	-1.613	1.217	-1.329	0.654			
AGE	0.360	0.541	-1.290	0.600			
DINC	0.031	0.217	0.122	0.230			
INC	-0.184	0.228	-0.206	0.201			

Obs=5556
Ln(likelihood)=- 5222.969

For the columns headed φ_i, and ψ_i, what is presented are the estimates of θ and their estimated standard errors. For RIGHT, which has a new set of row labels, the columns are the probit estimates and their standard errors.

Predictions:

avg. Pr(NEWEST)	0.543	for users who have NEWEST
avg. Pr(NEWEST)	0.329	for users who don't have NEWEST
avg. Pr(RIGHT)	0.753	for users who are right
avg. Pr(RIGHT)	0.681	for users who are wrong

Predicting Adoption Hazard, Pt. 1

- Adoption Hazard is ψ_i*$H(\varphi_i, \psi_i; \underline{x}_t, \underline{X}_t, s_t)$
  informed*download if informed.

- Mean of H() .963
- Mean of H()*ψ_i .130
  - Most of non-adoption is non-informed-ness
  - Some variation in H() with z, but not much

Predicting Adoption Hazard, Pt. 2

Person	P= ψ_i H(φ_i, ψ_i)
	Unconditional Hazard for NEWEST
Mean	
Change in	both ϕ_i ψ_i
DHOURS	-0.011
HOURS	0.057
USE	0.023
PAYWORK	-0.054
DSPEED	-0.079
LSPEED	-0.008
OCCCOMP	0.028
MALE	0.084
DAGE	-0.082
AGE	-0.081
DINC	0.011
INC	-0.017

Findings

- Empirically, the determinants of having the newest IE are (in declining order)

1. Initial conditions (i.e., distribution)
2. Information about choice set
3. All other, including download speed and tastes

- Rational ignorance model fits well (slightly worse than descriptive probit in which some coeffs are impossibly large)
- Find information more important despite cross-equation restrictions

Conclusions

- Can model “rational ignorance” empirically
  - Particularly if observe a consumer error, e.g. *RIGHT*
- Ignorance of choice set, not high access costs, can explain value of distribution

- “Product placement” is valuable in many contexts
- “Opt in” is different from “Opt out” in many contexts
- When the difference is information (need to search for alternatives) our model is ready to go