INDEX: a status report

November 1999

Pravin Varaiya, Hal Varian

Richard Edell, Karyen Chu, Walter Beckert

EECS, SIMS, Economics at UC Berkeley

Motivation

Internet use currently has zero marginal cost for use

- Result: congestion (see next slide)

Internet currently uses one quality of service (QoS)

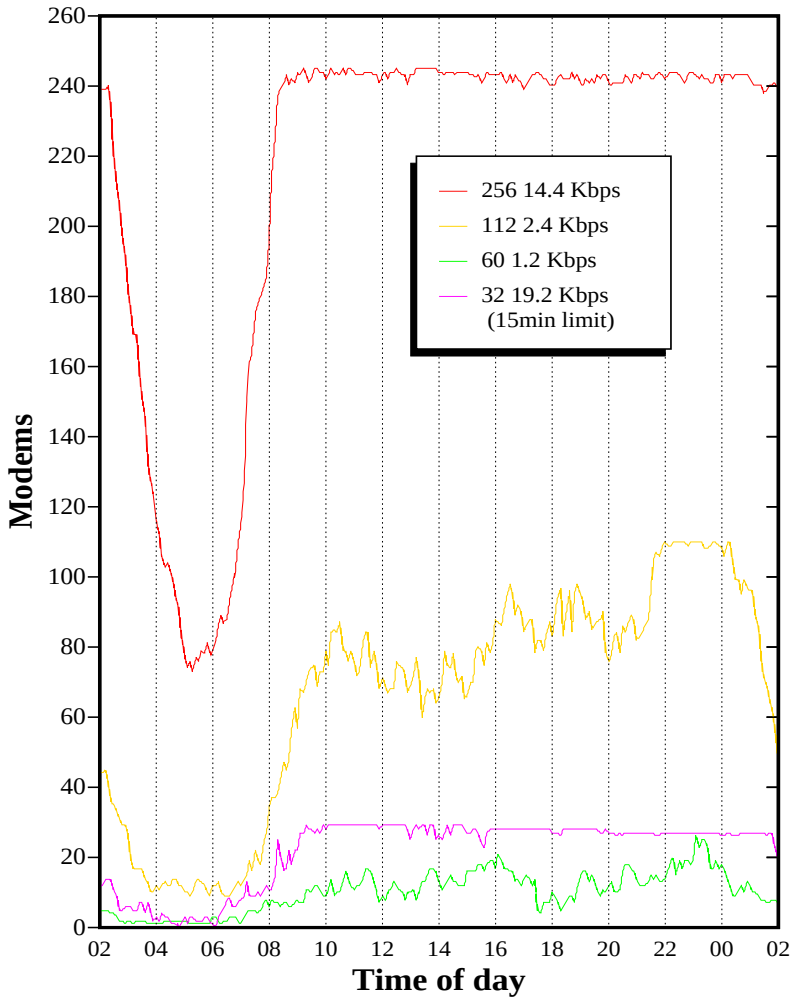
- Differential quality of service requires differential pricing

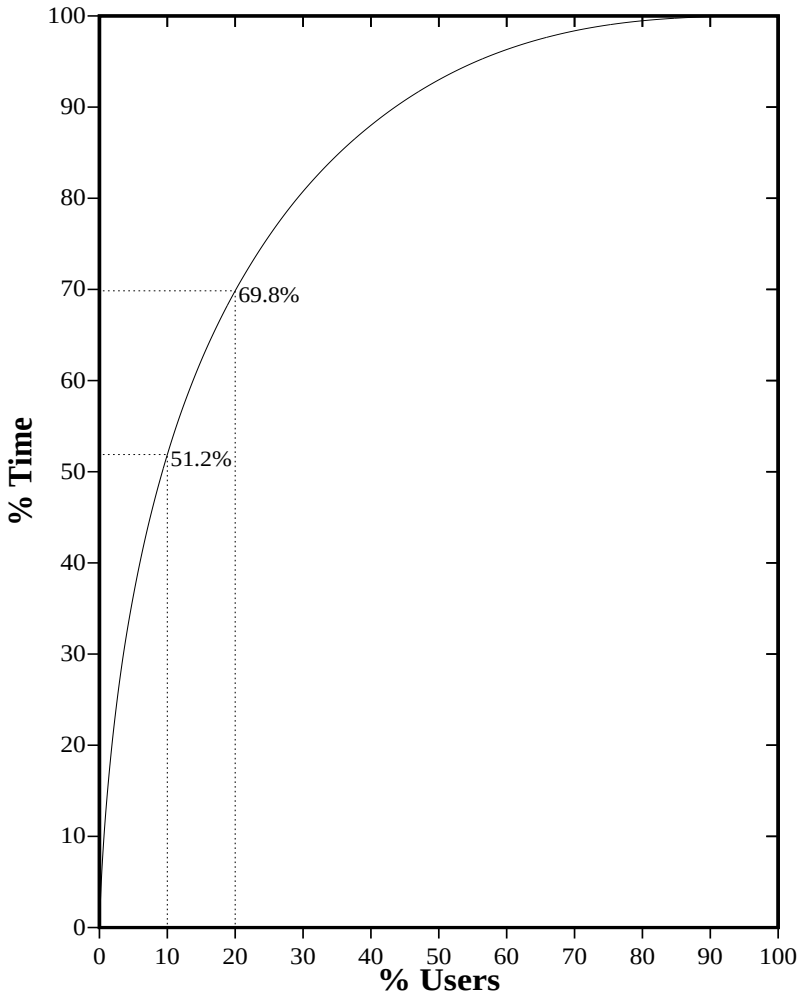
Research on QoS

- How to engineer QoS: a lot of work
- What QoS costs: a little work
- How users value QoS: nothing

Goal of INDEX project: measure how users value QoS

Modem use by day and user





Information about INDEX

INDEX = Internet DEmand eXperiment

Web page: **http://www.index.berkeley.edu**

Support: Cisco Systems, Pacific Bell, Hewlett Packard, and National Science Foundation.

What is our goal?

Goal: Measure how much users are willing to pay for Internet Quality of Service (QoS).

Dimensions of QoS

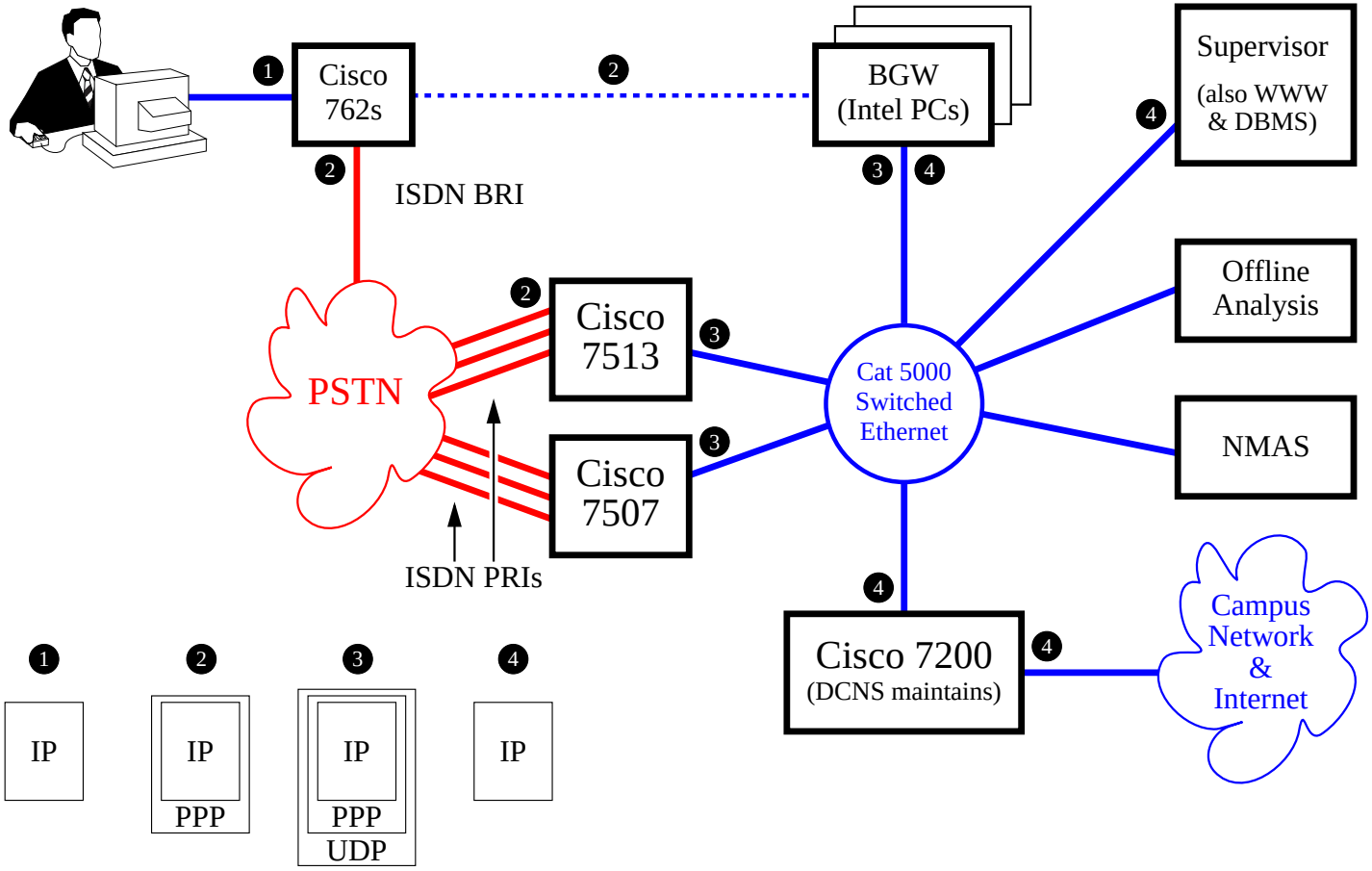
- bandwidth (symmetric and asymmetric)
- volume
- congestion
- delay to connect
- etc.

How do we do it?

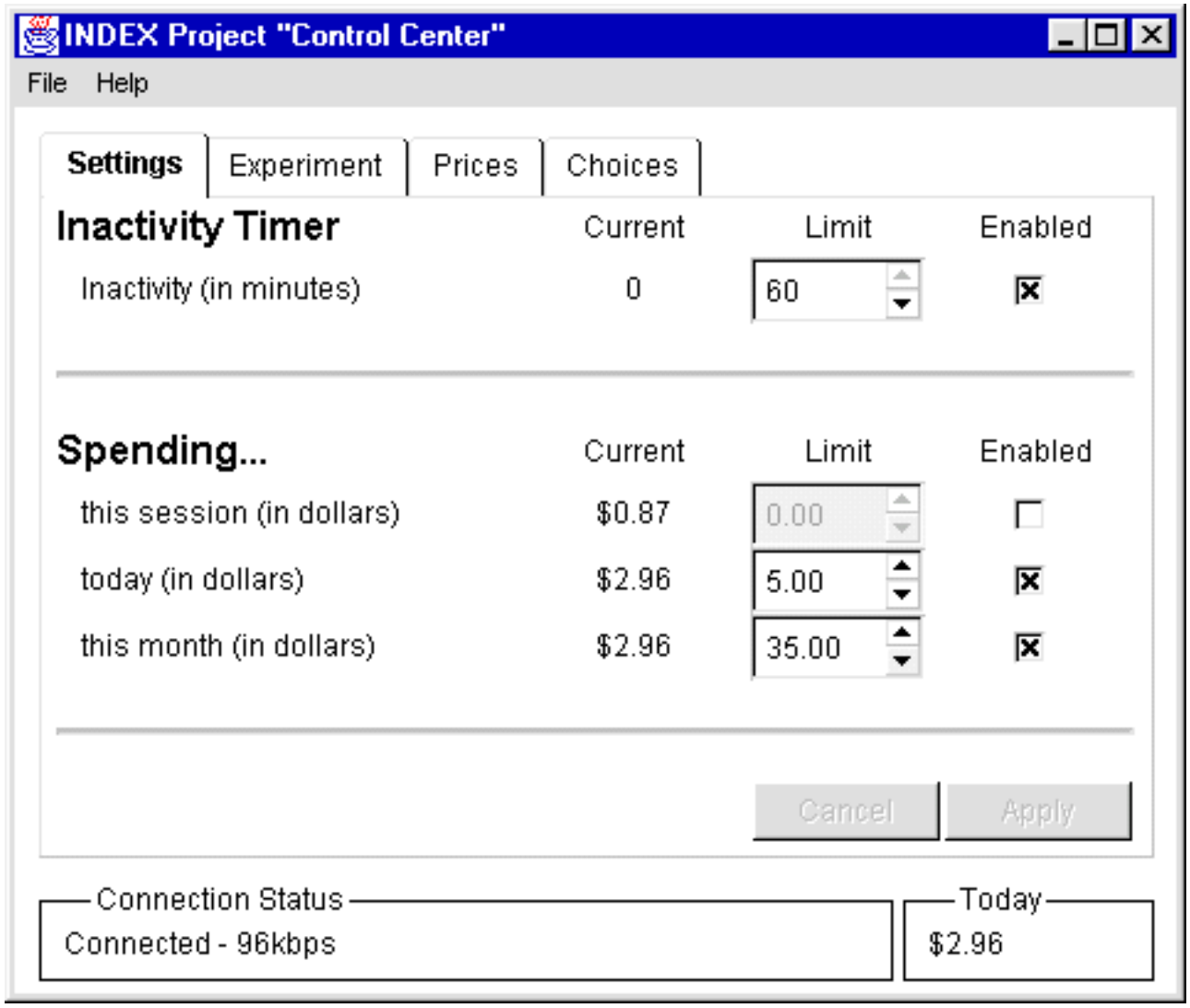
- Give 150 Berkeley people "free" ISDN service
  - hardware and setup costs are free
  - we cover PacBell's monthly charges
- Offer them different QoS for different prices
- Simulate the QoS
  - same QoS choices for several weeks
  - different prices each week

- Note: *real* money, *simulated* service

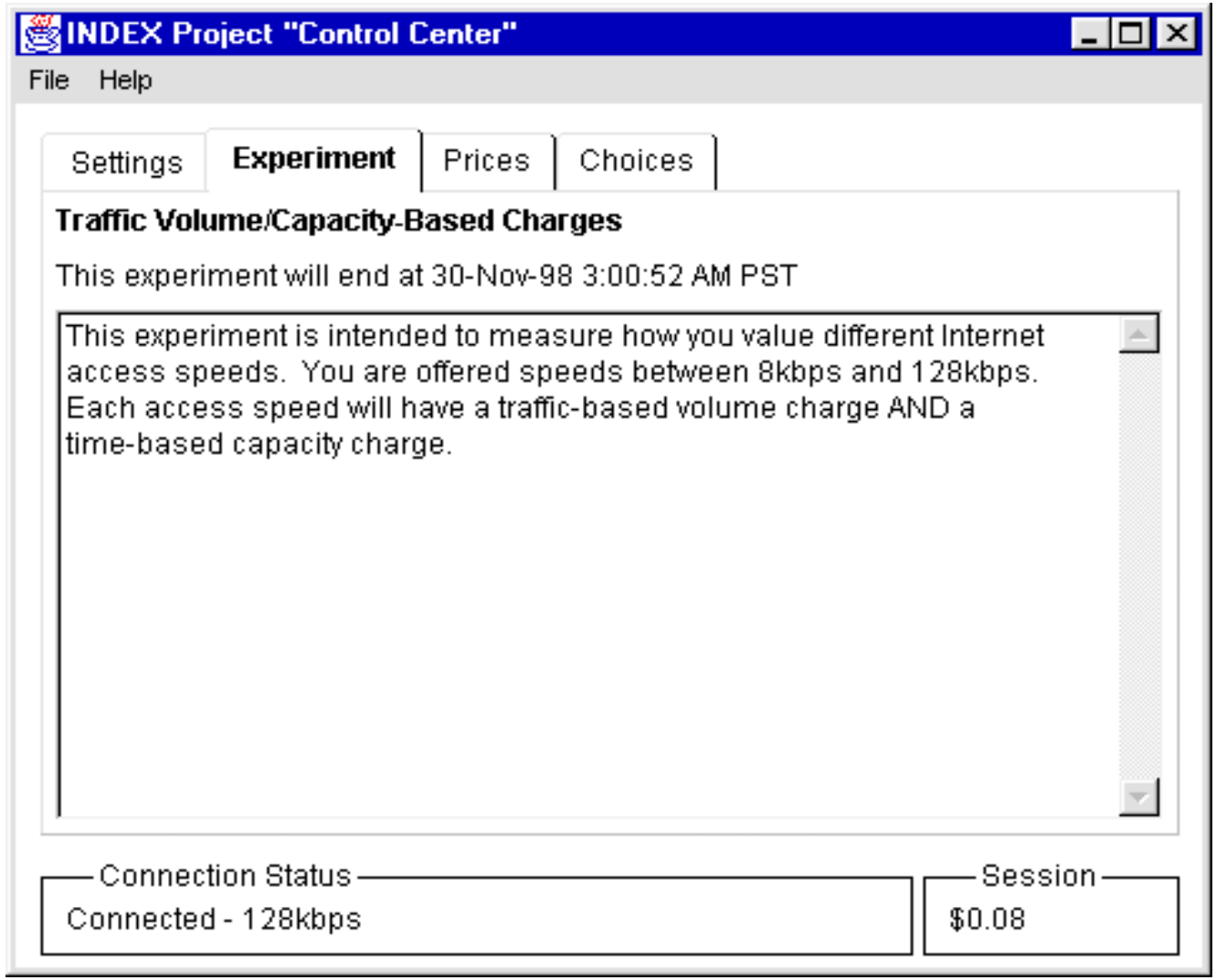
Architecture



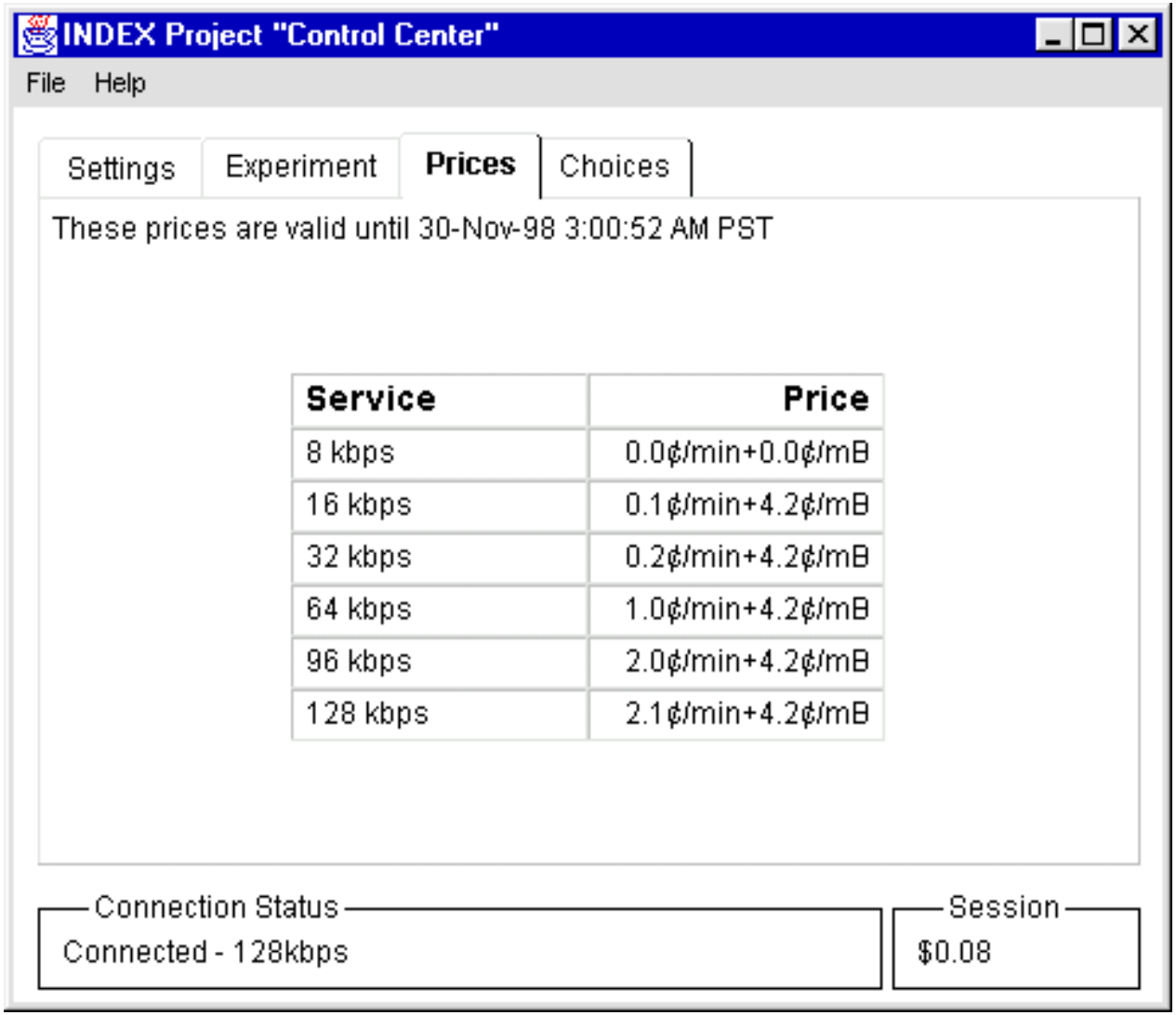
Interface for INDEX: settings



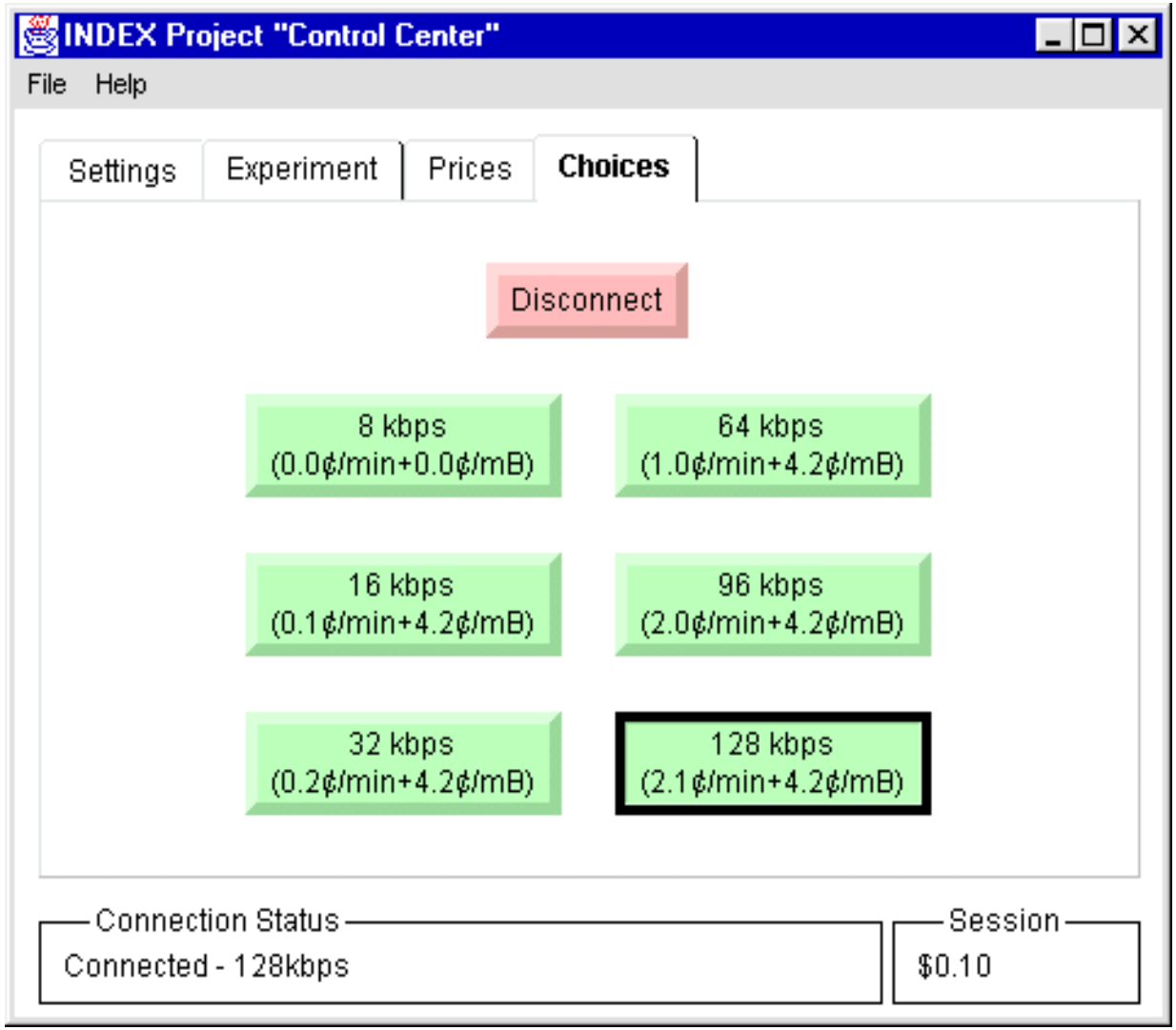
Interface: experiment



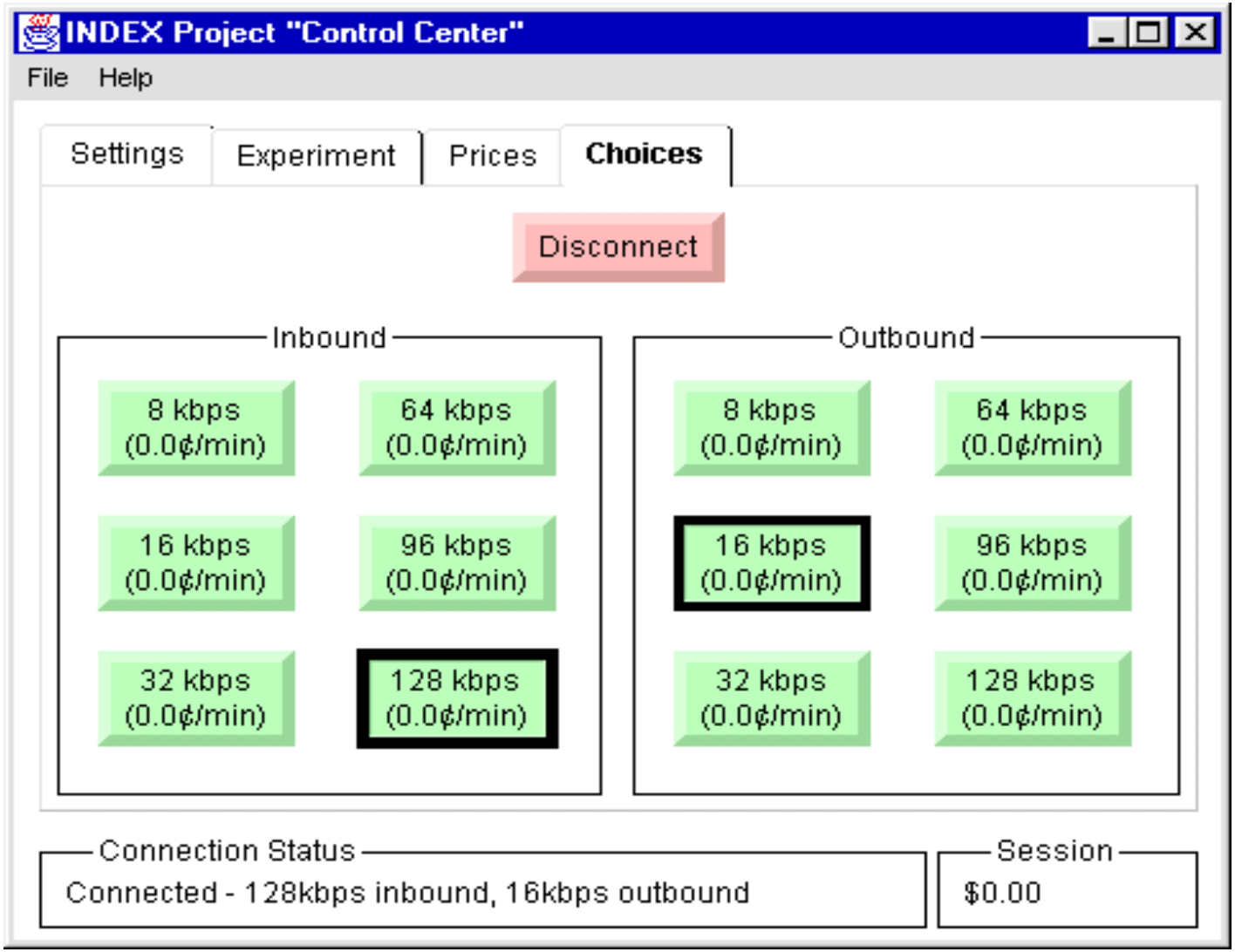
Interface: pricing



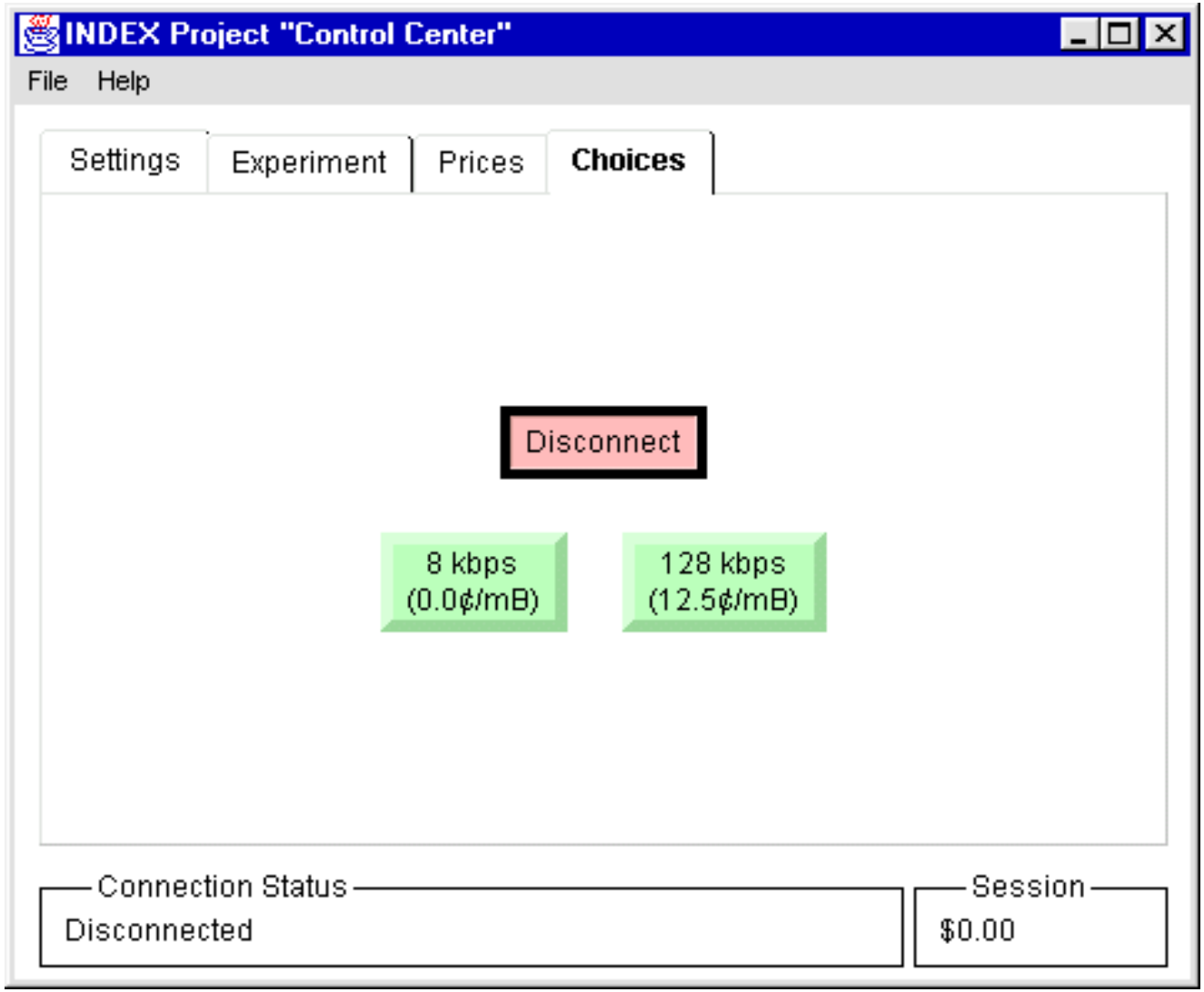
Interface: choices



Interface: asymmetric BW



Interface: volume pricing



Interface: billing

09/15/1998	128kbps service	23:13:08-23:13:22		
	0.000 megabyte	0.001	0.0000	0.0000
09/15/1998	128kbps service	23:14:28-23:14:38		
	0.000 megabyte	0.001	0.0000	0.0000
09/15/1998	128kbps service	23:20:20-09/16 00:08:15		
	0.152 megabyte	0.001	0.0002	0.0002
09/16/1998	128kbps service	12:58:38-13:07:00		
	8.35 minute	0.008	0.0668	
	0.607 megabyte	0.060	0.0364	0.1032
09/16/1998	128kbps service	13:09:51-13:09:57		
	0.10 minute	0.008	0.0008	
	0.000 megabyte	0.060	0.0000	0.0008

Where we are

Started providing service April 1, 1998

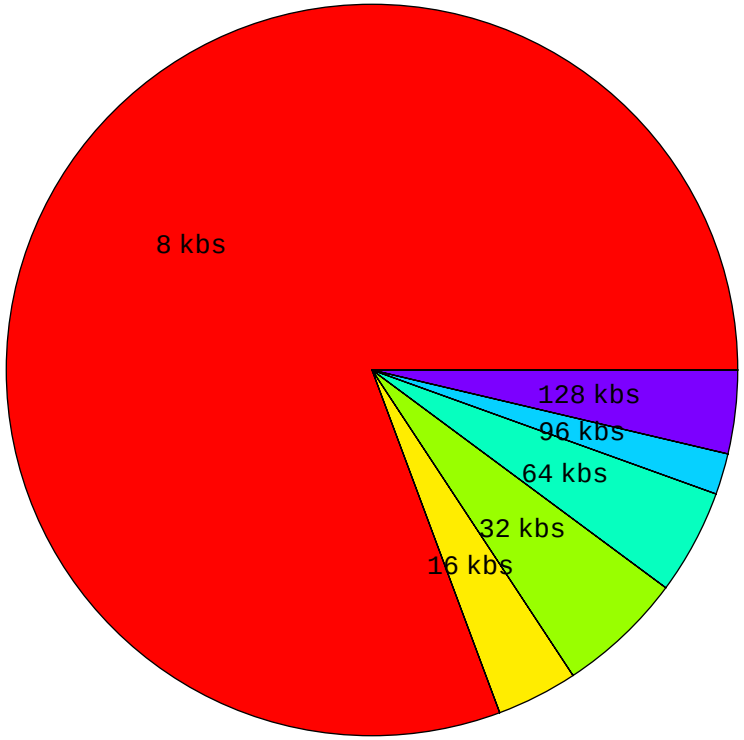
- have about 70 subjects
- subjects proceed through experiments asynchronously

Have run

- symmetric bandwidth
- asymmetric bandwidth
- volume pricing
- bandwidth + volume mixture (predetermined + self-selected)
- pay upfront for flat pricing

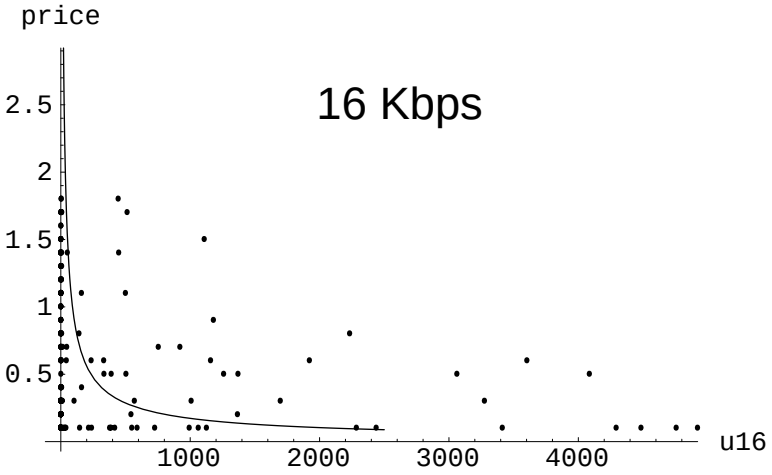
Symmetric bandwidth

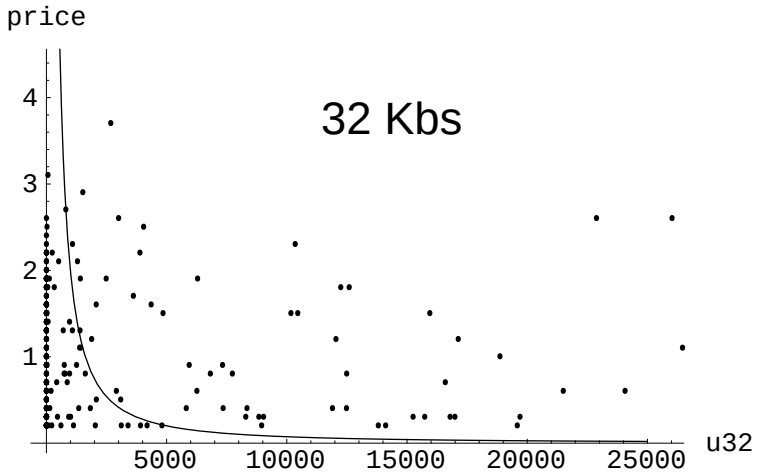
Usage overall:



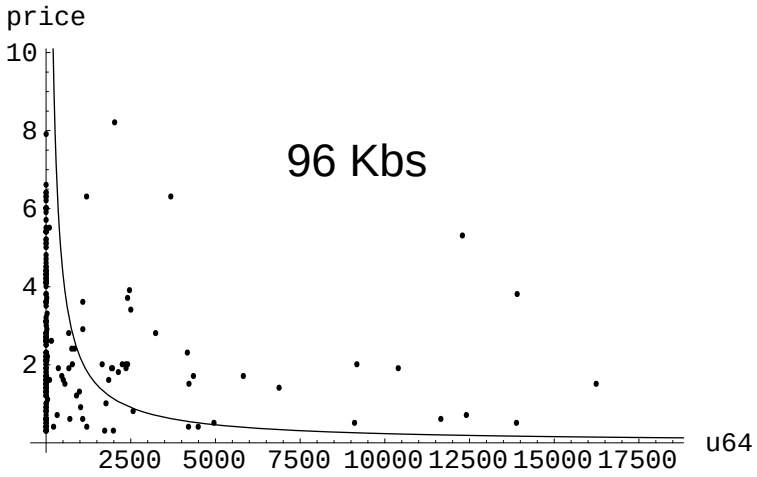
Note: 8 kbs is free

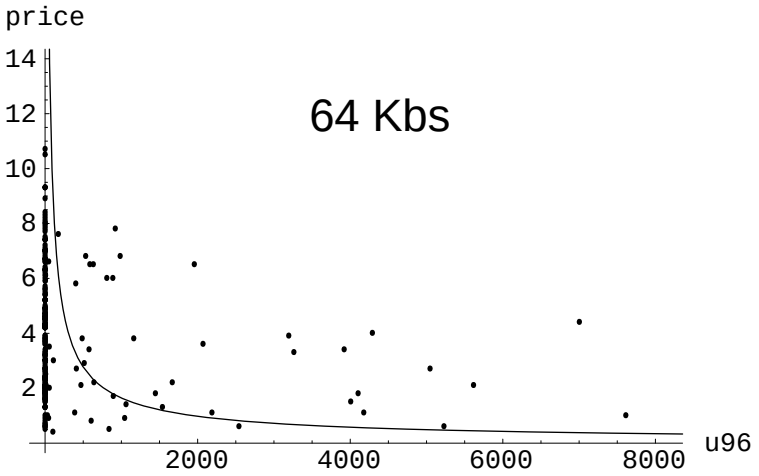
Scatter plots

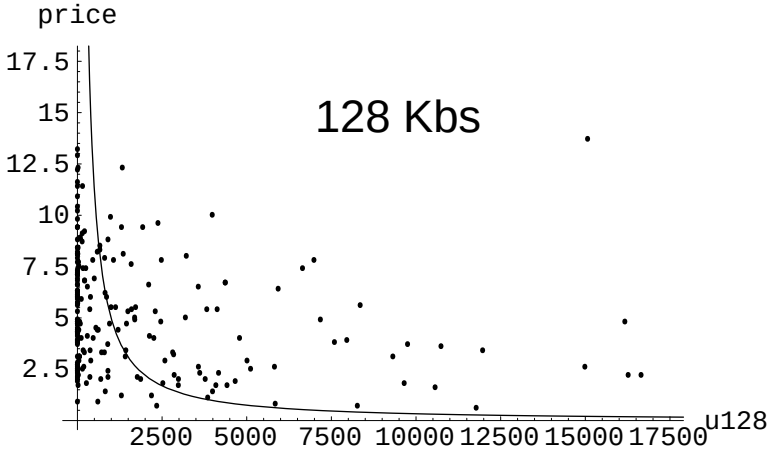




Scatter plots, continued







Demand estimates

Reduced form: estimate demand as a function of price

amount consumed = function of prices and individual characteristics

Structural: estimate parameters of utility function

choice of bandwidth depends on value of time, urgency, etc. as well as price and characteristics

First look at reduced form, then examine some options for structural estimates

Log Regressions

```
u128=-2.0p128 +.80p96 +.25p64 -.02p32 -.16p16
u96 =+1.7p128 -3.1p96 +.43p64 +.19p32 +.18p16
u64 =+.77p128 +1.8p96 -2.9p64 +.59p32 +.21p16
u32 =+.81p128 -1.0p96 +1.0p64 -1.4p32 +.15p16
u16 =+0.2p128 -.29p96 +.04p64 +1.2p32 -1.3p16
```

Red = negative own price effect

Light = not statistically significant

Lesson: *large* negative own price effect, positive cross price effect

How good is the fit?

	R^2	
	ISE	no ISE
u128	.95	.11
u96	.93	.25
u64	.92	.18
u32	.95	.14
u16	.90	.17

ISE = "individual specific effect"; interpretation

Conclusion: very good fits!

Structural estimates

Notation

b = bandwidth chosen

x = bits transferred

t = time at bandwidth $b = x/b$

$p(b)$ = price of bandwidth b

Utility

$u(x) - [c+p(b)]t$

c is value of time; varies with circumstances

$f(c)$ is probability distribution of c

Choice model

Let *(x,b*)* be choice. Optimization requires:

$$u(x) - [c + p(b^*)]\frac{x}{b^*} \geq u(x) - [c + p(b)]\frac{x}{b} \quad \text{for all } b.$$

Rearrange to find

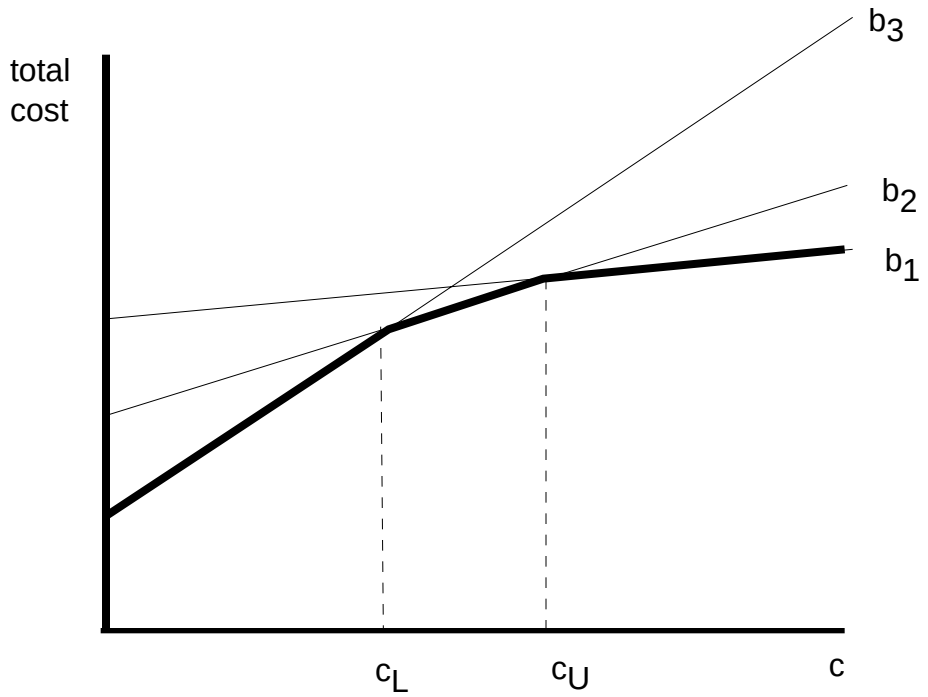
$$\min_{b^* < b} \frac{p(b^*)b - p(b)b^*}{b^* - b} \geq c \geq \max_{b^* > b} \frac{p(b^*)b - p(b)b^*}{b^* - b}$$

Note that this gives upper and lower bounds on cost of time.

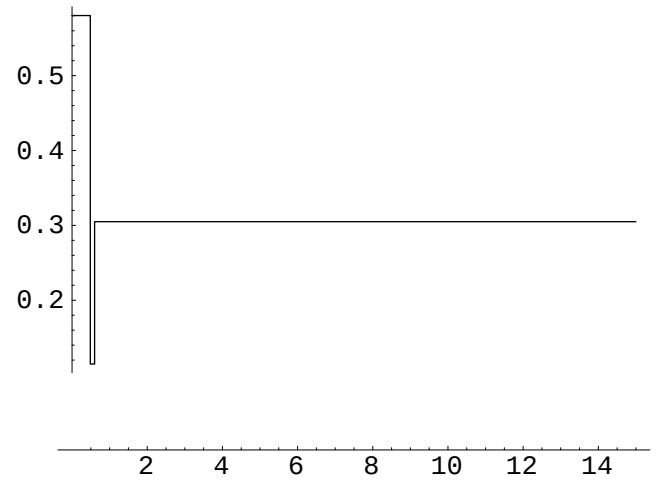
Picture of estimation method

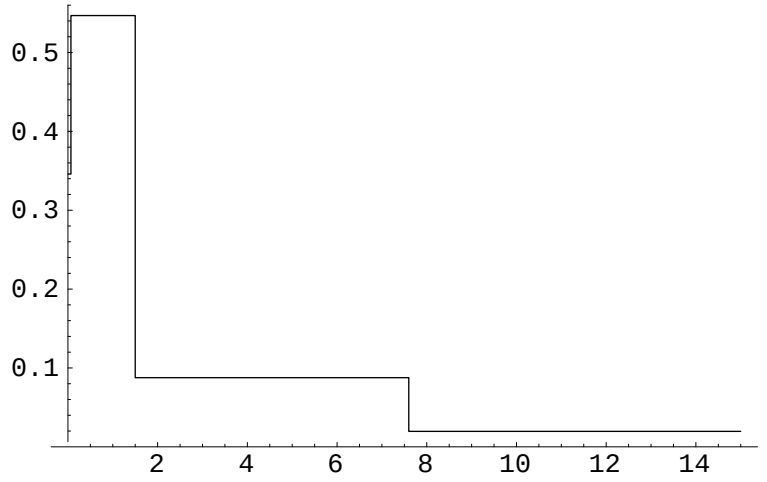
Plot $K(c) = [c + p(b_i)]\ (1/b_i)$ for $i = 1,\ldots,n$

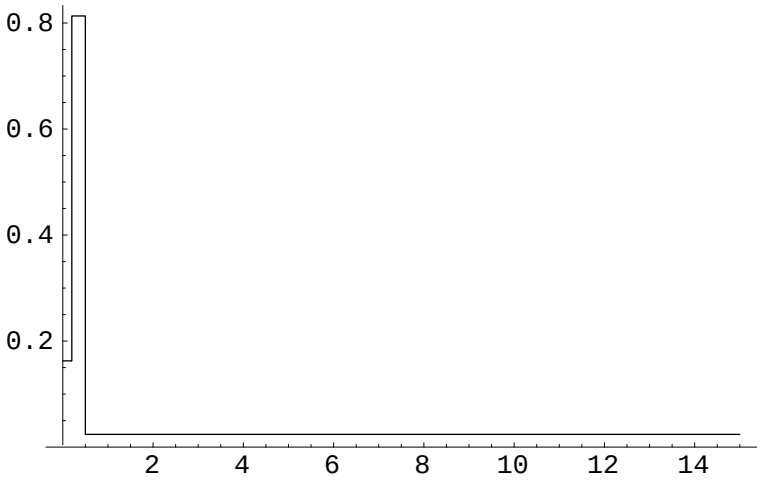
Observe (c_U^i, c_L^i, f^i) for $i = 1,\ldots,n$

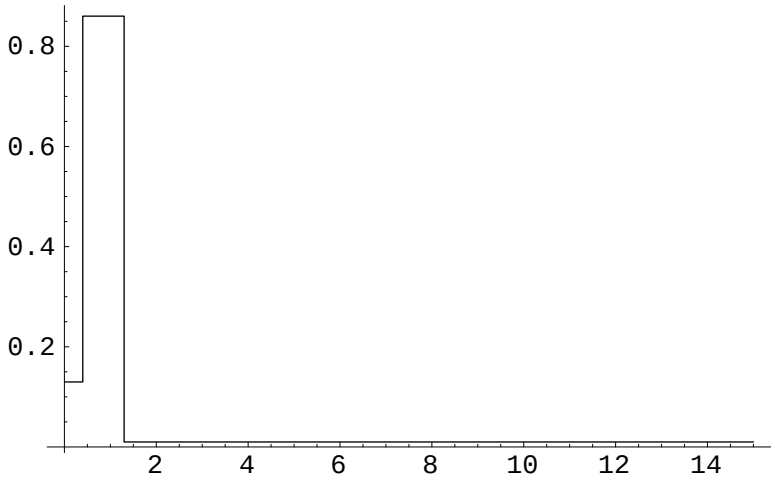


Histograms for one user









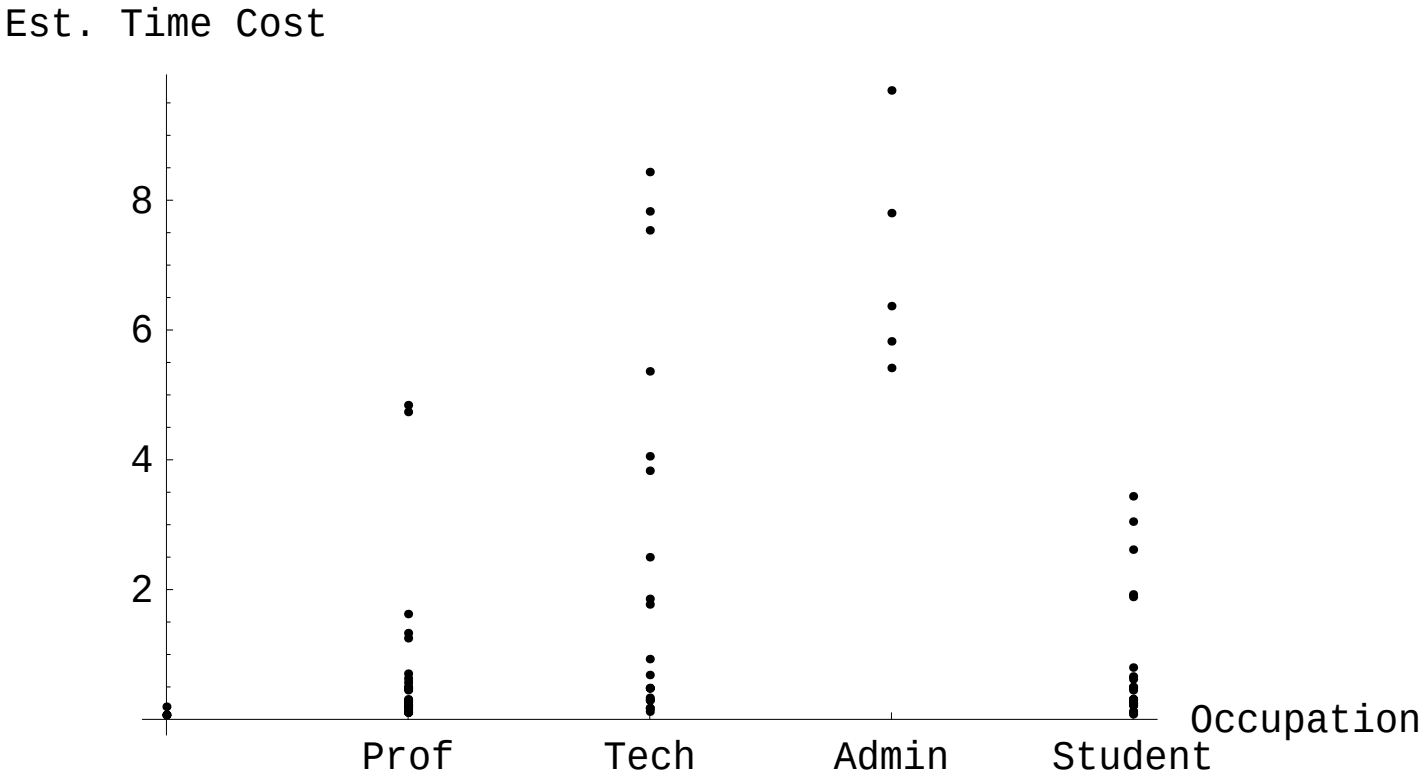
Distribution of time cost

Can calculate average time cost for each individual using these frequency distributions.

Table shows number of subject with average estimated time cost that falls in indicated range:

Range	1	2	3	4	5	6	7	8	9	10	11	12
Upper bound	39	8	3	4	1	2	2	1	2	0	3	0
Lower bound	63	3	0	0	1	0	0	0	0	0	0	0
Average	47	7	2	3	3	3	1	3	1	1	0	0

Who has high WTP?



Average cost of time

Very low WTP for time (= 1/2 cent per minute)

- very small charges induce large changes in behavior
- lot of low-value traffic out there
- admin/technical have highest WTP

WTP for population as a whole

cLowMean = 0.02
cUpMean = 0.40
cLowStd = 0.07
cUpStd = 1.11

Parametric fit

Suppose frequencies drawn from prob distribution *p(c,β)*

Find distribution that is close to all the frequencies

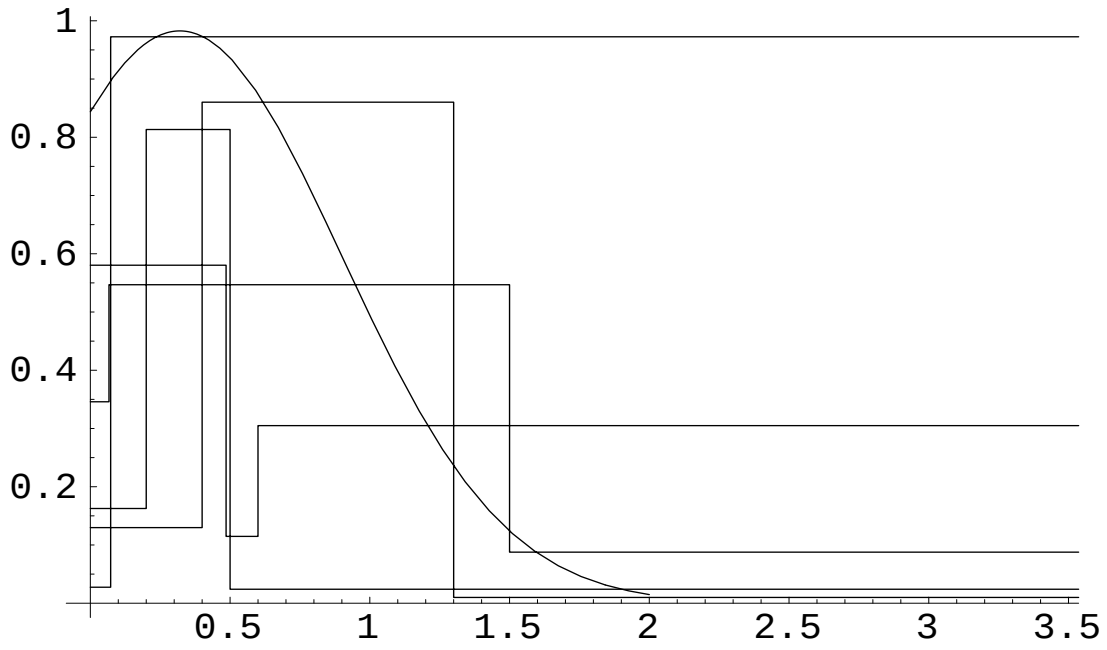
Closeness = Kullback-Leibler entropy measure, $\Sigma f \log p$

$$\max_{\beta} \sum_{i=1}^{n} f_i \log \int_{c_L^i}^{c_U^i} p(x,\beta)\, dx$$

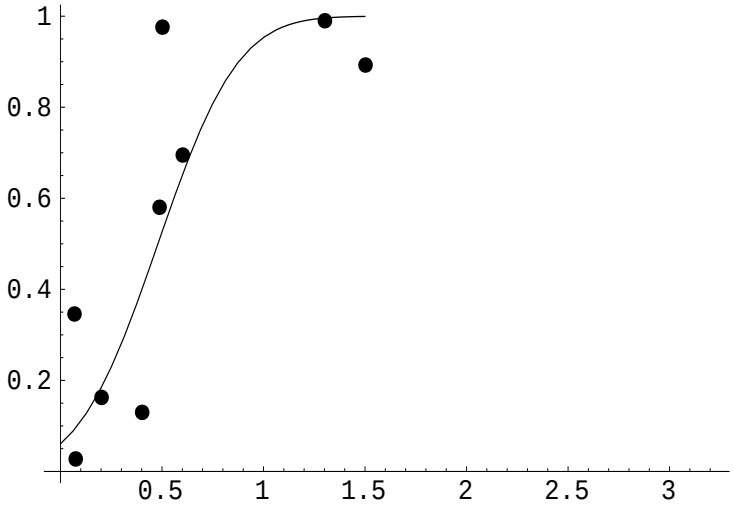
Truncated Normal

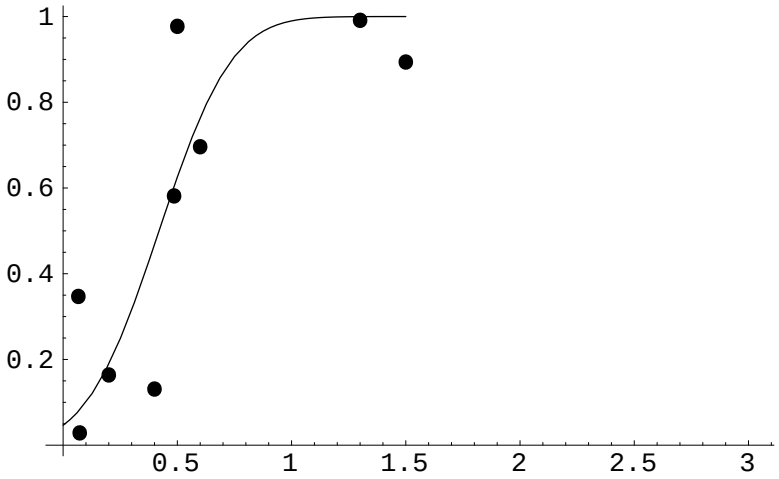
μ=.487, σ=.32

mean = .52



Picture of CDF





μ=.48, σ=.32

max entropy

μ=.42. σ=.25

min SSR

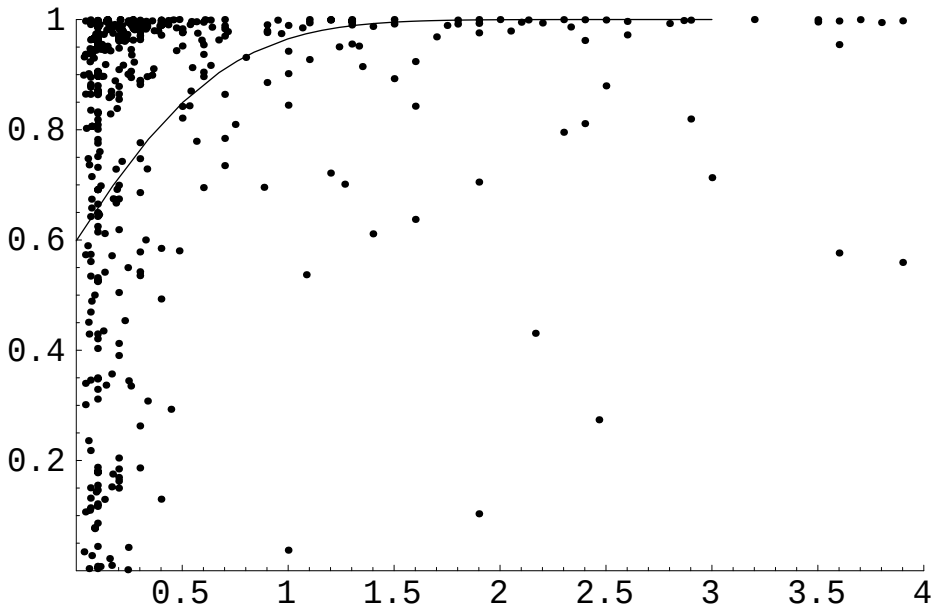
Entire Population

What about population as a whole?

cLowMean = 0.02, cUpMean = 0.40

cLowStd = 0.07, cUpStd = 1.11

Estimated parameters: μ=-.16, σ=.64, mean time cost = .18



Why such a low WTP?

Possible explanations

1. Our population is not representative

   can look at demographics, e.g., income effects

2. We only get lower bound on WTP for 128 kbs

   but this is only about 1/6 of use

3. We can't control QoS outside of Berkeley network

   high bandwidth isn't worth much if congestion is elsewhere

4. We can only measure value of *existing* applications

Volume experiments

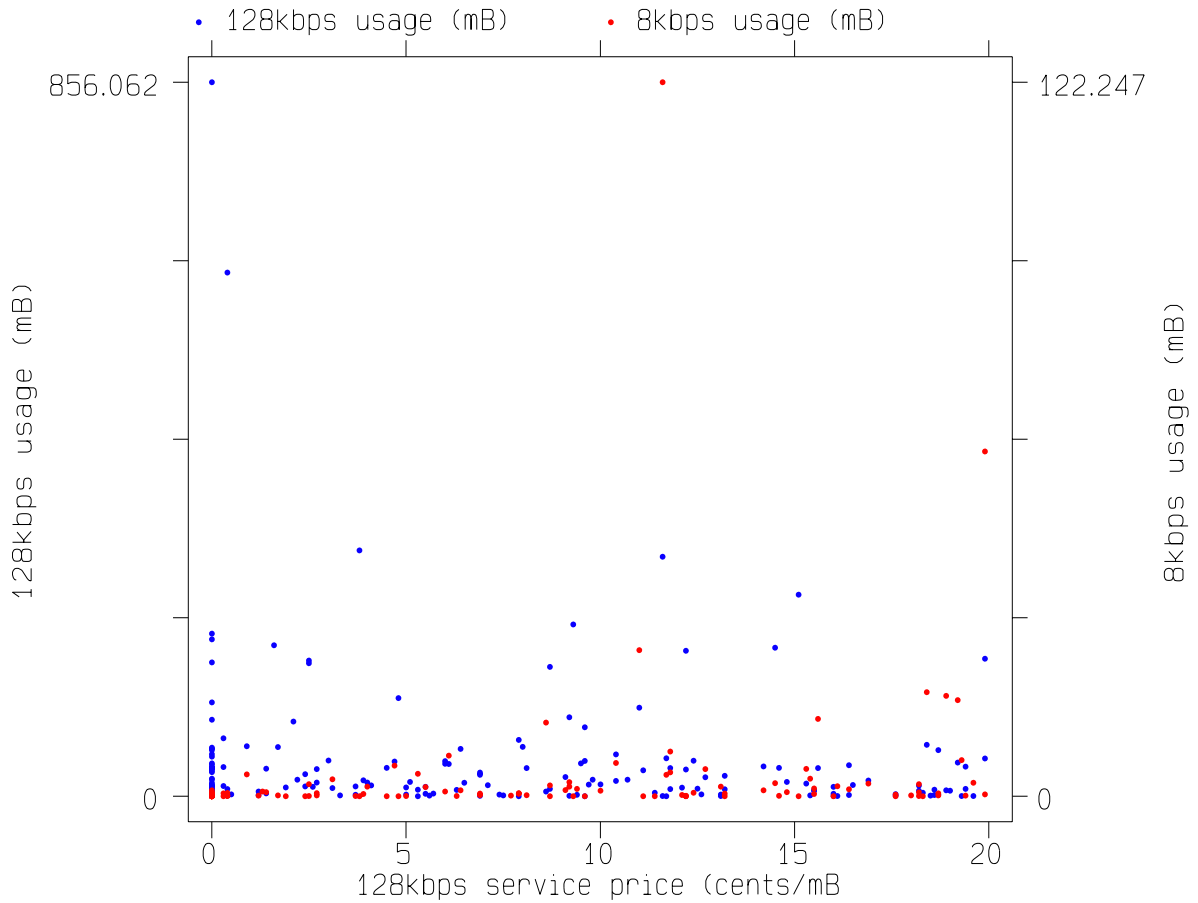
Two bandwidths

- 8 Kbs for free
- 128 Kbs for money

Price volume downloaded

- prices range from 1 to 20 cents per megabyte
- monthly expenditures similar in bandwidth and volume experiments

Scatterplot



Log Regressions

no ISE: u128 = -0.29 p128 $R^2 = 0.02$

ISE: u128 = -0.24 p128 $R^2 = 1.00$

Interesting finding:

amount transmitted is not sensitive to price

entire effect is individual-specific effect

Why? Probably choice offered is too extreme...

What's next?

Current experiment: pay flat fee on Sunday to opt out of metered pricing

Move to ADSL or cable modem?

What are *you* interested in?