

**What's a
customer
worth?**

Etsy

Roberto Medri
Etsy

@paradosso



Winning with data

Five things about me

1. **On the Internet for the first time at age 19**
2. **Never wears black**
3. **Majored in Ancient Greek, Latin, and Film**
4. **Played billiards competitively with a two-time world champion at age 16**
5. **Former management consultant**

Five things about me

1. **On the Internet for the first time at age 19**
2. **Never wears green**
3. **Majored in Ancient Greek, Latin, and Film**
4. **Played billiards competitively with a two-time world champion at age 16**
5. **Former management consultant**

One of these is a lie.

Some Etsy numbers

\$525M GMS in 2011

800,000 active sellers

16,000,000 listings

90M visits per month

2.8M iPhone app downloads

>25% open rate for marketing emails

What do data scientists do at Etsy

We help PMs make good decisions.

We help marketing get the best bang for their buck.

We help make Etsy products better.

This talk

What Customer Lifetime Value (CLV) **is**

A stochastic approach to **estimating** CLV

How we **act** against CLV at Etsy

What CLV is

What CLV is Customer Lifetime Value in the open

“I know the customer lifetime value of my business is \$200, so that’s how much I can spend to acquire a customer”

What CLV is Customer Lifetime Value in the open

“I know the customer lifetime value of my business is \$200, so that’s how much I can spend to acquire a customer”

How many things are wrong with this?

What CLV is

Let me count the ways

1. Not all customers have the same CLV

What CLV is

Let me count the ways

1. **Not all customers have the same CLV**
2. **CLV is a forward-looking concept, you can't *know* how much it is**

What CLV is

Let me count the ways

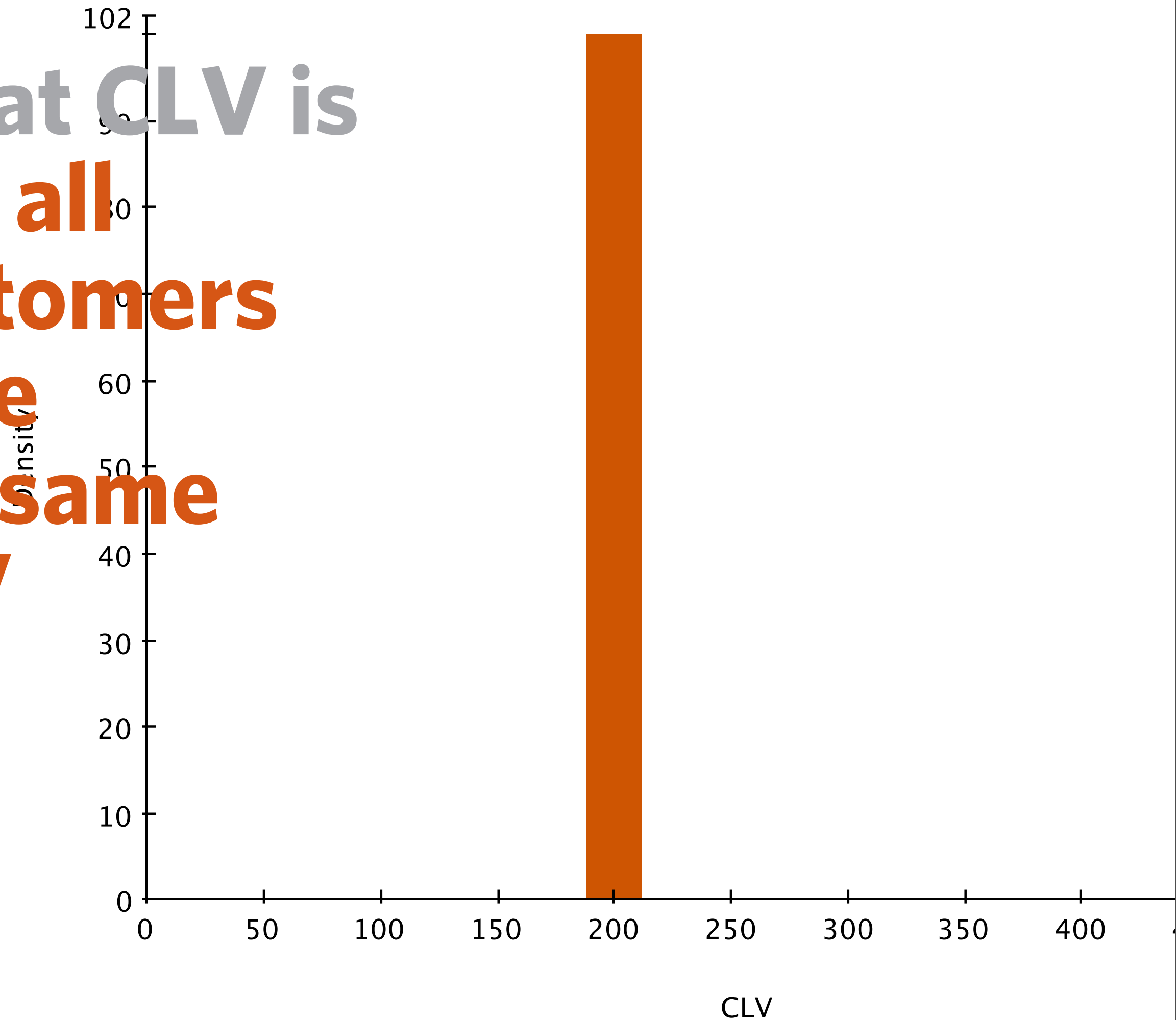
1. **Not all customers have the same CLV**
2. **CLV is a forward-looking concept, you can't *know* how much it is**
3. **What we are really interested in is Residual Lifetime Value (RLV), not past spend**

What CLV is

Let me count the ways

1. **Not all customers have the same CLV**
2. **CLV is a forward-looking concept, you can't *know* how much it is**
3. **What we are really interested in is Residual Lifetime Value (RLV), not past spend**
4. **Comparing a future, uncertain quantity (CLV) to a current, certain one**

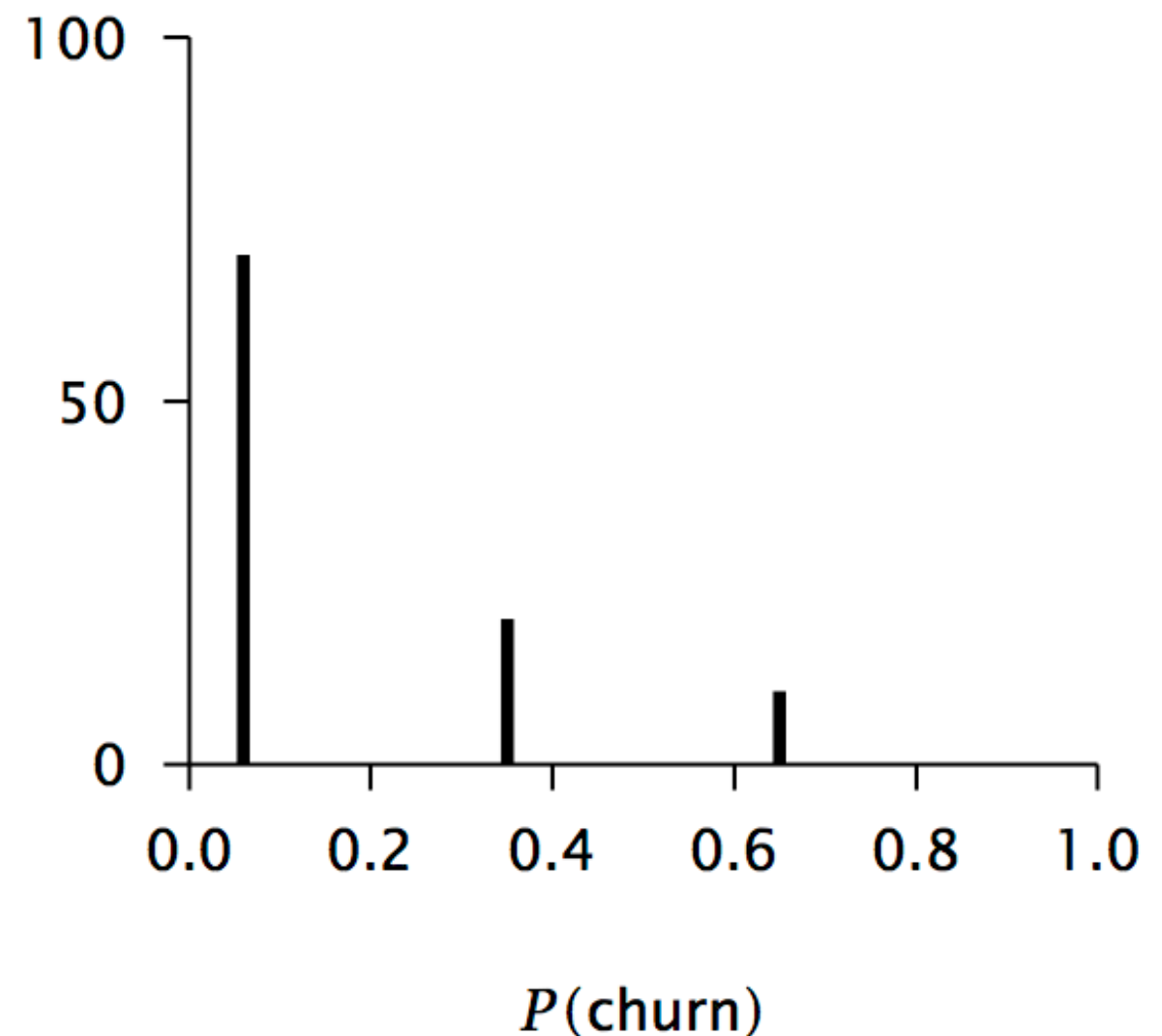
What CLV is
Not all
customers
have
the same
CLV



What CLV is

Not all customers have the same CLV

Cluster	$P(\text{churn})$	%CB
Low risk	0.06	70
Medium risk	0.35	20
High risk	0.65	10



What CLV is

Not all customers have the same CLV

Continuous Distribution: every customer has her own CLV



What CLV is Segmentation

Demographics?

What CLV is Segmentation

Demographics?



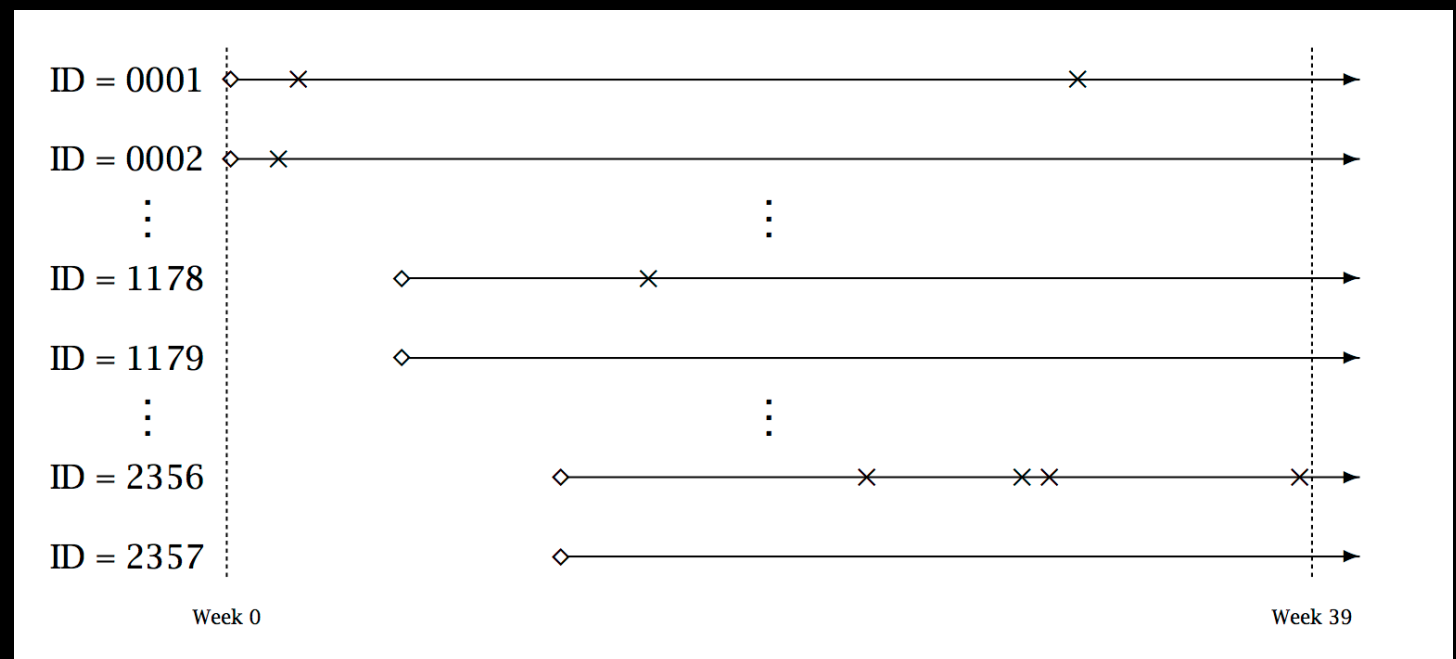
What CLV is

Segmentation: Behavior

Transaction log

Customer ID	Transaction Date	\$ Amount
1001	2012-05-01	\$30
1002	2012-05-01	\$20
1003	2012-05-02	\$50
1002	2012-05-03	\$70
1001	2012-05-04	\$10
1001	2012-05-05	\$10
1004	2012-05-06	\$100
...

Frequency and Recency



What CLV is

A definition

The present value of the sum of expected cash flows of an individual customer.

Individual-level estimates:

E(transactions) over next k years

E(\$) over next k years

P(alive)

P(returning)

Estimating CLV

Estimating Lifetime Value

A checklist

1. **What's the objective or the decision?**
2. **What do we want to model?**
3. **How do we model the behavior?**
4. **Deriving the mixture model**
5. **Fit the model to existing data**
6. **Predict**
7. **Check**
8. **Act upon your findings**

Estimating Lifetime Value

1. Objective

Model predictive statistics around *future* spend (RLV), in order to drive, e.g.:

Acquisition/Retention resource allocation

Individually targeted actions

Estimating Lifetime Value

2. What we want to Model

Future spend at the individual level

How many transaction/money will David Barton spend over the next 2 years?

Estimating Lifetime Value

3. How we want to model it

At every moment, customer flips two coins:

The first coin determines if the customer lives or dies.

The second coin determines if she buys or not.

Estimating Lifetime Value

4. How we want to model it

Customers have *their own, individual* live/die probabilities

Customers have *their own, individual* buy/not buy probabilities

Everyone has two unique coins.

Estimating Lifetime Value

4. Nasty Math

Survival process: exponential (one parameter!), latent parameter distributed gamma (two parameters)

Transaction process: Poisson process, latent parameters distributed gamma (two parameters)

Four parameters: fits in Excel.

Estimating Lifetime Value

5. Fit the model

Solve the double integral
MLE (maximize log likelihood)

$$L(r, \alpha, s, \beta \mid x, t_x, T)$$

$$= \frac{\Gamma(r+x)\alpha^r\beta^s}{\Gamma(r)} \left\{ \left(\frac{s}{r+s+x} \right) \frac{{}_2F_1(r+s+x, s+1; r+s+x+1; \frac{\alpha-\beta}{\alpha+t_x})}{(\alpha+t_x)^{r+s+x}} \right. \\ \left. + \left(\frac{r+x}{r+s+x} \right) \frac{{}_2F_1(r+s+x, s; r+s+x+1; \frac{\alpha-\beta}{\alpha+T})}{(\alpha+T)^{r+s+x}} \right\}, \text{ if } \alpha \geq \beta$$

$$L(r, \alpha, s, \beta \mid x, t_x, T)$$

$$= \frac{\Gamma(r+x)\alpha^r\beta^s}{\Gamma(r)} \left\{ \left(\frac{s}{r+s+x} \right) \frac{{}_2F_1(r+s+x, r+x; r+s+x+1; \frac{\beta-\alpha}{\beta+t_x})}{(\beta+t_x)^{r+s+x}} \right.$$

$\beta - \alpha$

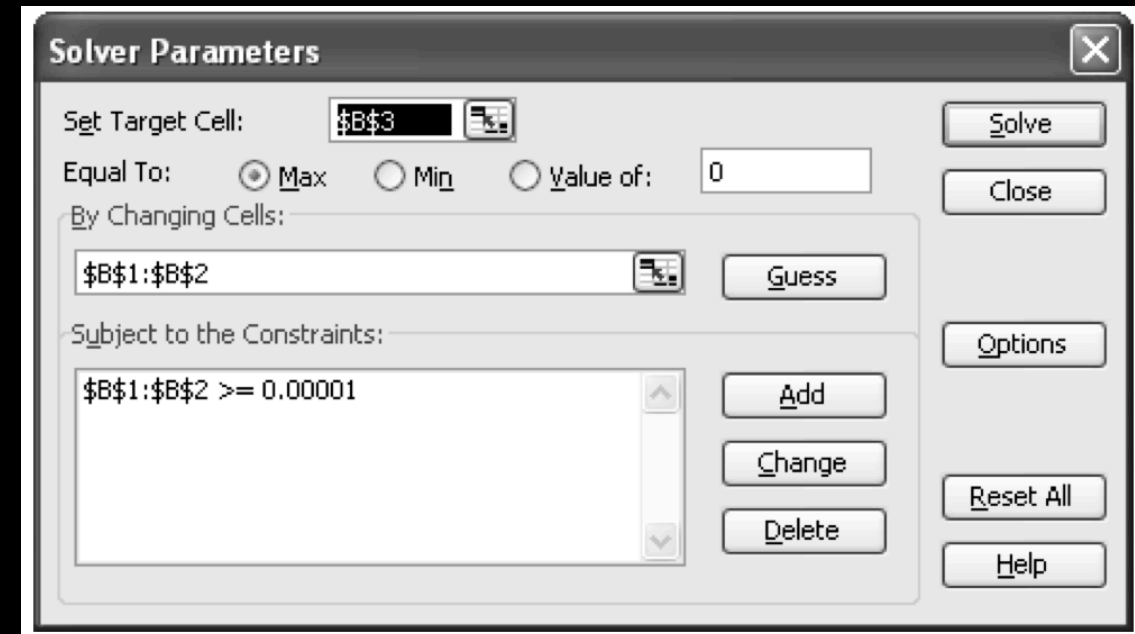
Estimating Lifetime Value

Fit the model

Excel

R

C / FORTRAN / Cython



Estimating Lifetime Value

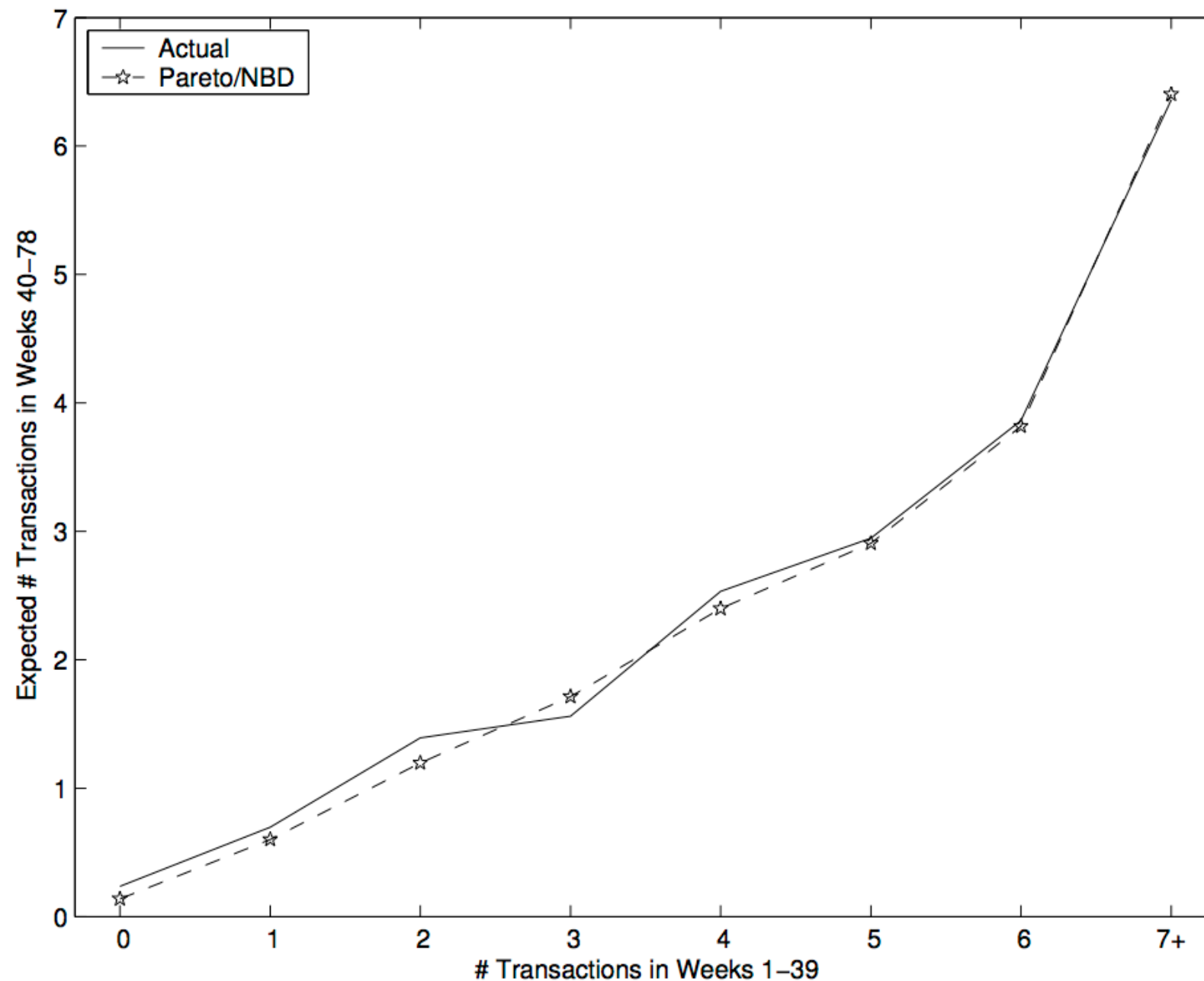
Predict

Pareto/NBD predicts the discounted expected number of future transactions

Multiply by the average monetary value of transaction and margin to get RLV

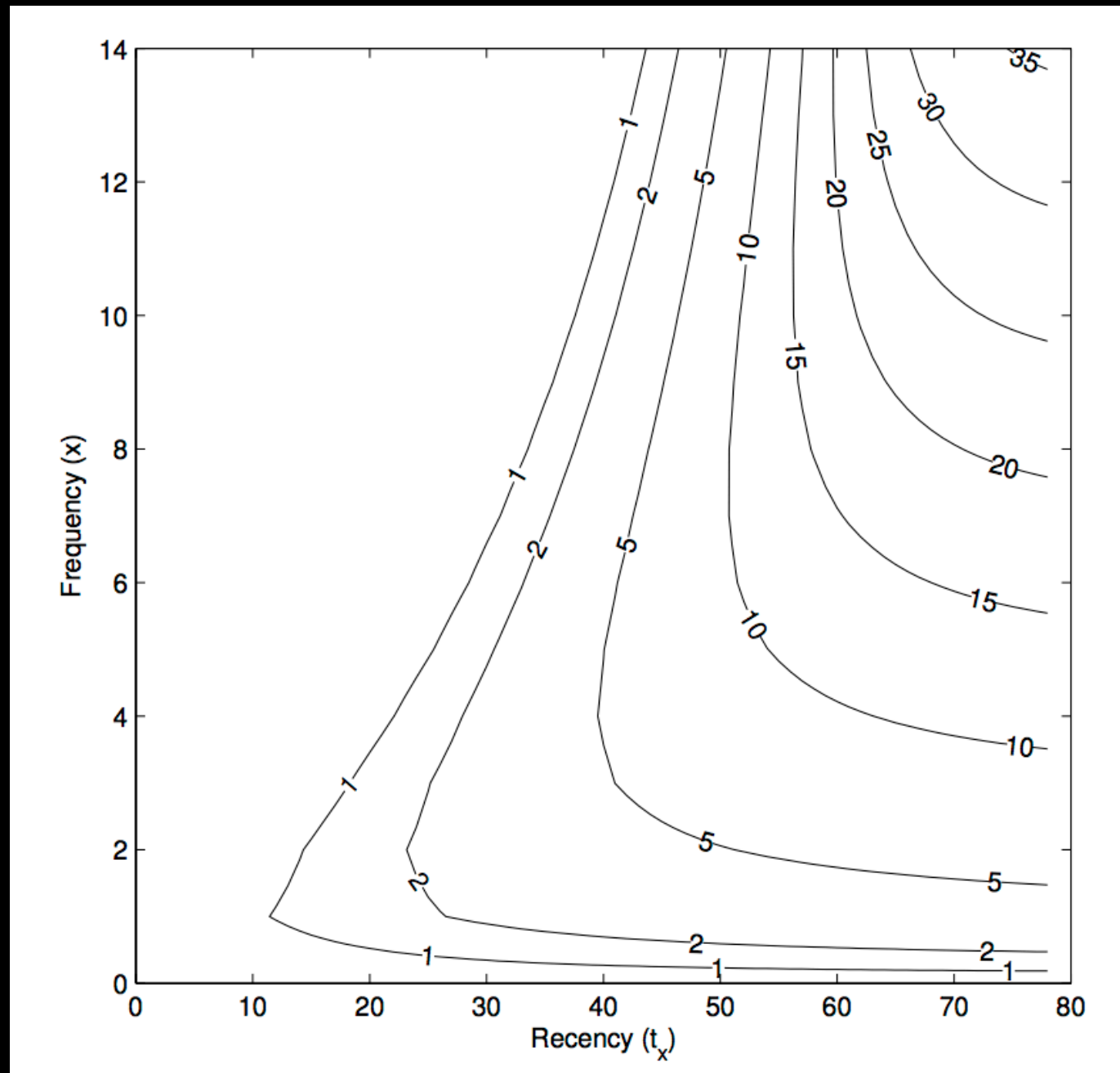
Estimating Lifetime Value

Conditional Expectations



Estimating Lifetime Value

Isovalue RF Curves



Estimating Lifetime Value

Segmentation and Allocation

Bin recency and frequency for behavioral-based segmentation

		Recency			
	Frequency	0	1	2	3
M=0	0	\$4.40			
M=1	1		\$6.39	\$20.52	\$25.26
	2		\$7.30	\$31.27	\$41.55
	3		\$4.54	\$48.74	\$109.32
M=2	1		\$9.02	\$28.90	\$34.43
	2		\$9.92	\$48.67	\$62.21
	3		\$5.23	\$77.85	\$208.85
M=3	1		\$16.65	\$53.20	\$65.58
	2		\$22.15	\$91.09	\$120.97
	3		\$10.28	\$140.26	\$434.95

		Recency			
	Frequency	0	1	2	3
M=0	0	\$53,000			
M=1	1		\$7,700	\$9,900	\$1,800
	2		\$2,800	\$15,300	\$17,400
	3		\$300	\$12,500	\$52,900
M=2	1		\$5,900	\$7,600	\$2,300
	2		\$3,600	\$26,500	\$25,800
	3		\$500	\$37,200	\$203,000
M=3	1		\$11,300	\$19,700	\$3,700
	2		\$7,300	\$45,900	\$47,900
	3		\$1,000	\$62,700	\$414,900

Estimating Lifetime Value

Check your model

Don't use (only) in-sample fit

Out of sample fit

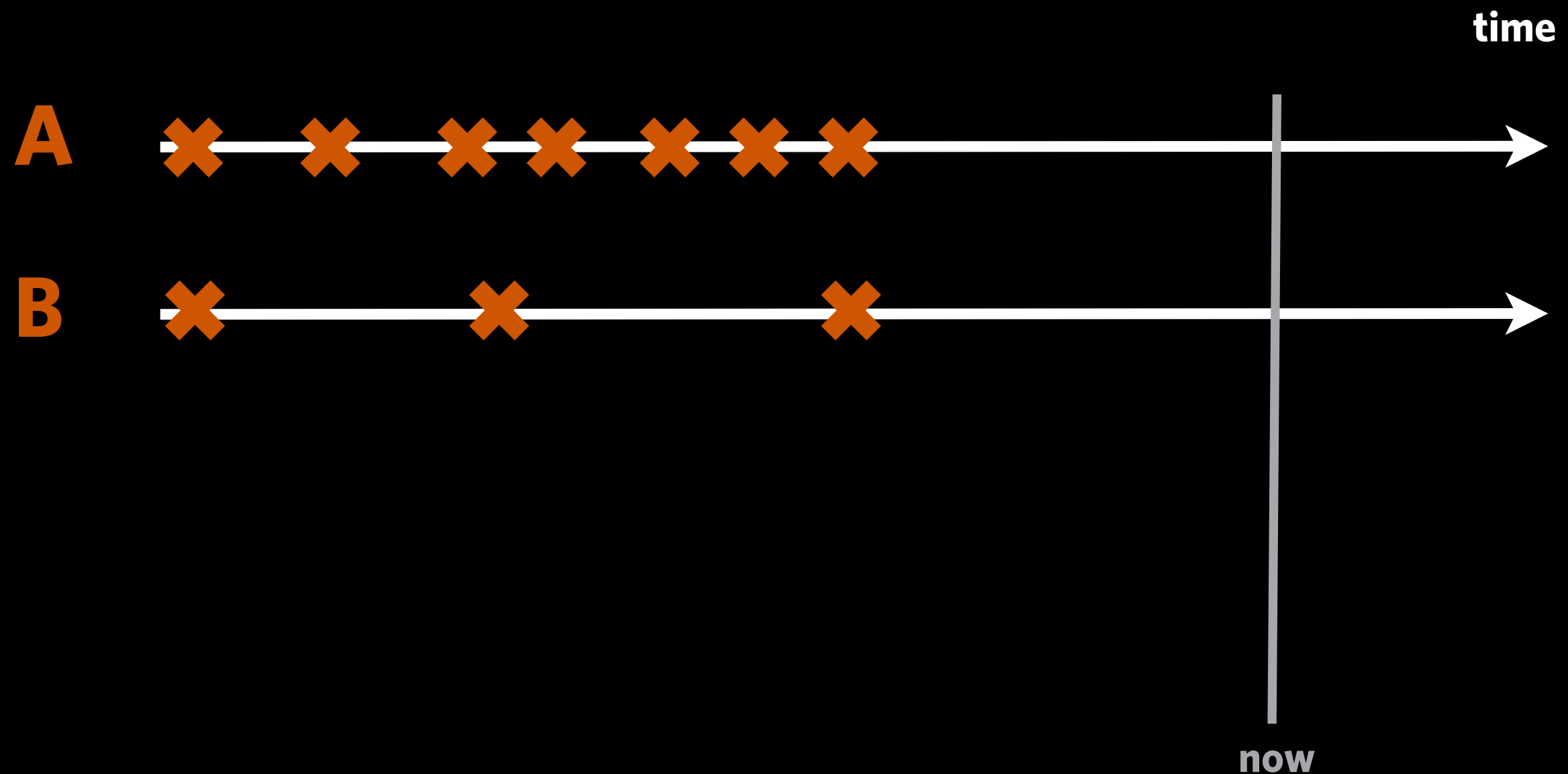
Divvy up your training set

Remove random observations

Simplicity and Consistency of the story

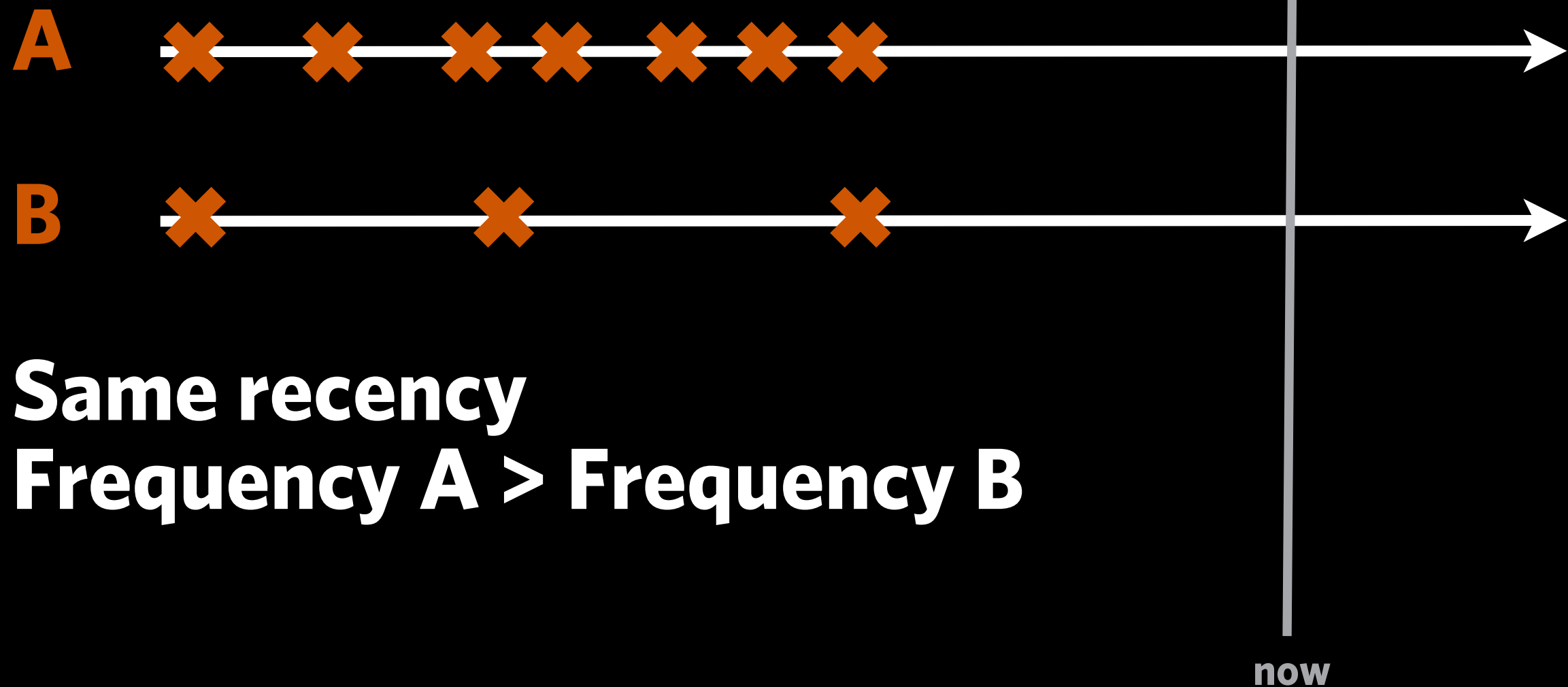
Estimating Lifetime Value

Who has the highest RLV?



Estimating Lifetime Value

Who has the highest RLV?



Estimating Lifetime Value

Who has the highest RLV?



Estimating Lifetime Value

Covariates & Context

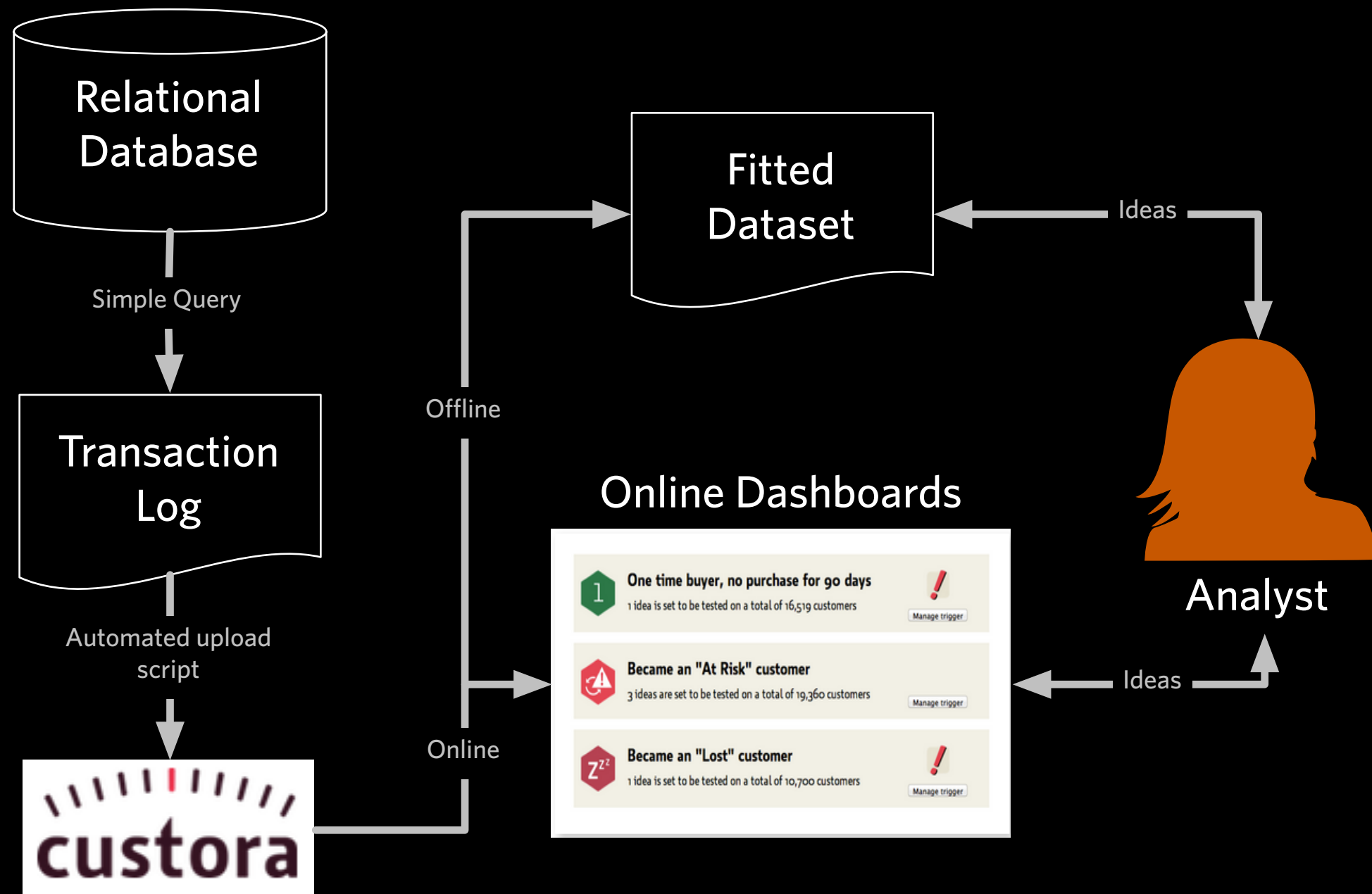
Other information we have about customers can be brought in as a covariate

Is the customer mobile?
What category they buy in?
Seasonality
Media/PR event
Etc.

Acting on CLV

Acting on CLV

Etsy Lifetime Value Stack



Acting on CLV

Customers of interest

Who are your best customers?

Recent high-profile customers

**Old-time best customers about to churn
or just churned**

Acting on CLV

Retention campaigns

Sent an email to 7.5M customers who hadn't bought in 60 days or more

Set aside a 5% control group

Emailed customer bought 11c more over the next 15 days vs. control ($p < .05$)

Made \$800K GMS directly, plus raised their RLV, total expected benefit \$4-6M GMS

Acting on CLV

Future: individualized, vectorized retention

E.g. send email when $p(\text{returning})$ falls
below threshold

Takes into account each customer's
intrinsic order frequency/spam fatigue

Acting on CLV

Future: Longitudinal customer treatments

**Keep track of all treatments/controls at
the customer level, together with their
purchasing patterns**

**Trying out bandit strategies in
marketing (explore/exploit)**

Acting on CLV

Fun stuff with fitted CLV datasets

Sum all RLV expectations across customers to value the company.

Look up and talk to your best customers.

Acting on CLV

Behavioral customer segmentation

One-time customers: **reinforcement, nudges**

Repeat customers: **premium services**

Top customers: **recognition**

Three things to remember from this talk

Customers have their own individual CLV. There's no one figure.

Don't use in-sample fit to judge a marketing model.

Who are your best customers? Really, what are their names?

Resources

Most of this content was developed by marketing scientists in the 60s.

Pareto/NBD was developed by Schmittlein, Morrison and Colombo in 1987.

Peter Fader, Wharton School at Upenn

Bruce Hardie, London Business School

<http://www.brucehardie.com/>

<https://www.custora.com/>



Three things to remember

Customers have their own individual CLV. There's no one figure.

Don't use in-sample fit to judge a marketing model.

Who are your best customers? Really, what are their names?

QUESTIONS?

Roberto Medri
@paradosso

Etsy