

# Business Backtesting of ML Models:

A Case Study in Real Estate

QCon New York June 2017

Nelson Ray

**Who has run an A/B test before?**

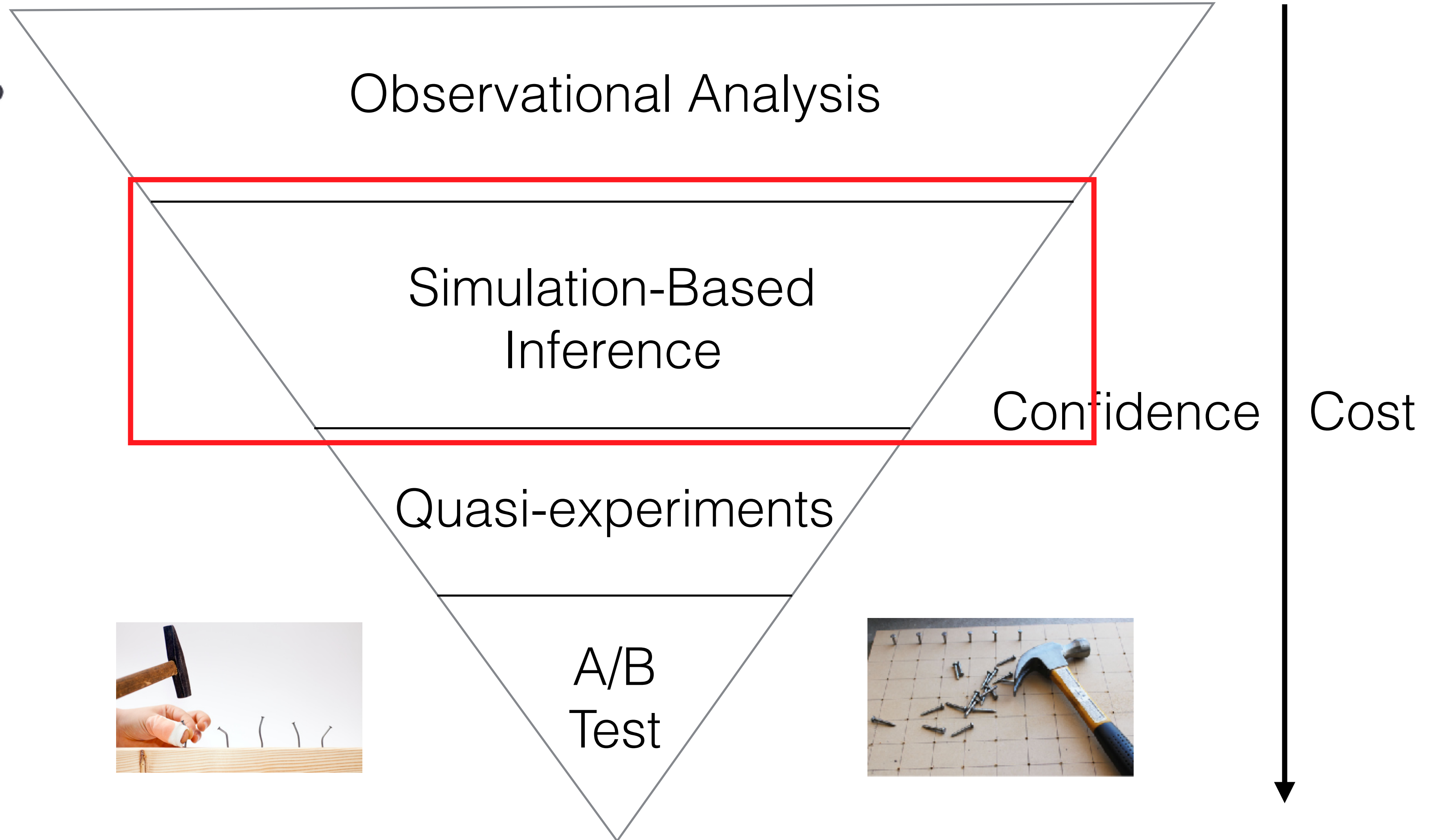
**Did it go off without a hitch?**

# Unguided A/B Testing





# Focus of this Talk



## Talk Structure

- Real Estate 101 for Home Buyers and Sellers
- The Opendoor Way
- Resale Risk
- Problems with A/B Testing
- How Simulation Helped in Real Estate
- A General Recipe for Simulation
- Team Info.

## Introduction

### Q A/B Testing Support Group



Hi, I'm Nelson!

I'm a recovering A/B testing user from such places as...

- Facebook
- Metamarkets
- Google
- Opendoor

I'll cover how to perform a business backtest of your ML models using simulations!

# Empower everyone with the freedom to move

---

\$25T of assets

63.5% of Americans are homeowners

#1 consumer expenditure (\$17,798/yr)

\$1.4T of annual transaction volume

\$100B in fees

CURRENT PROCESS

SELLERS

BUYERS

100+ day process with 14% failure rate

100+ day process with friction at each step

<b>Seller</b> Decide to move	<b>Realtor</b> Research online Receive bids Interview Choose	<b>Sale Ready</b> Improvements Yard work Cleaning Photographs	<b>List</b> MLS, Zillow, Trulia Open houses Showings Maintenance	<b>Contract</b> Offer Counter, Acceptance Inspection Financing period	<b>Closing</b> Final walkthrough Offline signatures Title records	<b>Contract</b> Offer Inspection Financing	<b>Search</b> Open houses Showings Viewings	<b>Realtor</b> Research online Receive bids Interview Choose	<b>Discovery</b> Finances Location Timing	<b>Buyer</b> Decides to move
<b>DAY 0</b>		\$1000's in upfront costs	90+ Days	4-5% in price drops and concessions	<b>DAY 121</b> 14% of deals fall-through 6-7% in fees	Fears of the home condition and financing	Months of viewing suboptimal listings		Months of research and gathering data	<b>DAY 0</b>

5.5M Americans per year buy and sell through this process



CURRENT PROCESS

SELLERS

100+ day process with 14% failure rate

BUYERS

100+ day process with friction at each step

<b>Seller</b> Decide to move	<b>Realtor</b> Research online Receive bids Interview Choose	<b>Sale Ready</b> Improvements Yard work Cleaning Photographs	<b>List</b> MLS, Zillow, Trulia Open houses Showings Maintenance	<b>Contract</b> Offer Counter, Acceptance Inspection Financing period	<b>Closing</b> Final walkthrough Offline signatures Title records	<b>Contract</b> Offer Inspection Financing	<b>Search</b> Open houses Showings Viewings	<b>Realtor</b> Research online Receive bids Interview Choose	<b>Discovery</b> Finances Location Timing	<b>Buyer</b> Decides to move
<b>DAY 0</b>		\$1000's in upfront costs	90+ Days	4-5% in price drops and concessions	<b>DAY 121</b> 14% of deals fall-through 6-7% in fees	Fears of the home condition and financing	Months of viewing suboptimal listings		Months of research and gathering data	<b>DAY 0</b>

5.5M Americans per year buy and sell through this process

SELLERS

**Fill out a short home profile** to ensure we can accurately price your home.

**Simply enter your address** to experience an automated, hassle-free sales process.

**And receive an offer in minutes** with a full report of your home's value.

**Selling your home is as easy as clicking next**



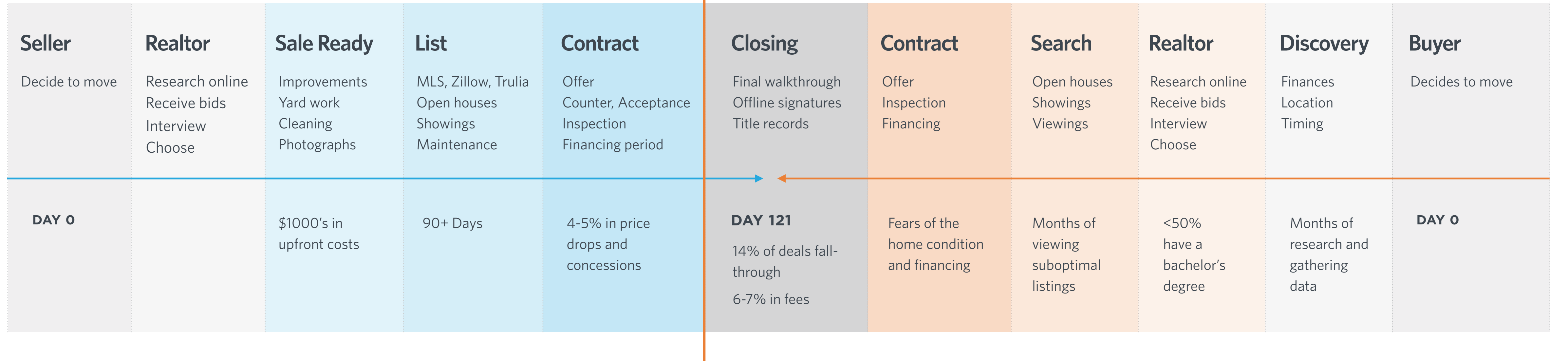
**CURRENT PROCESS**

**SELLERS**

**100+ day process with 14% failure rate**

**BUYERS**

**100+ day process with friction at each step**



**5.5M Americans per year buy and sell through this process**



# Buy a home you'll love — guaranteed.

Every Opendoor home includes our 30-day satisfaction guarantee, certified 180-point inspection, and a 2-year, extended coverage warranty from OneGuard.

[Learn more](#)

Thousands of buyers shop with us monthly

Searching and showings are self-service, on-demand

Our buyers have exclusive access to our inventory

All homes come with a money-back guarantee and a 2-year warranty

Buying a home is as easy as clicking next



Phoenix

[See listings](#)



Dallas-Fort Worth

[See listings](#)



**What is our risk in reselling a home?**



## Home 1

- Listed ~\$800k
- **6+ months on market**





## Home 2



- Listed ~\$300k
- **1 month on market**





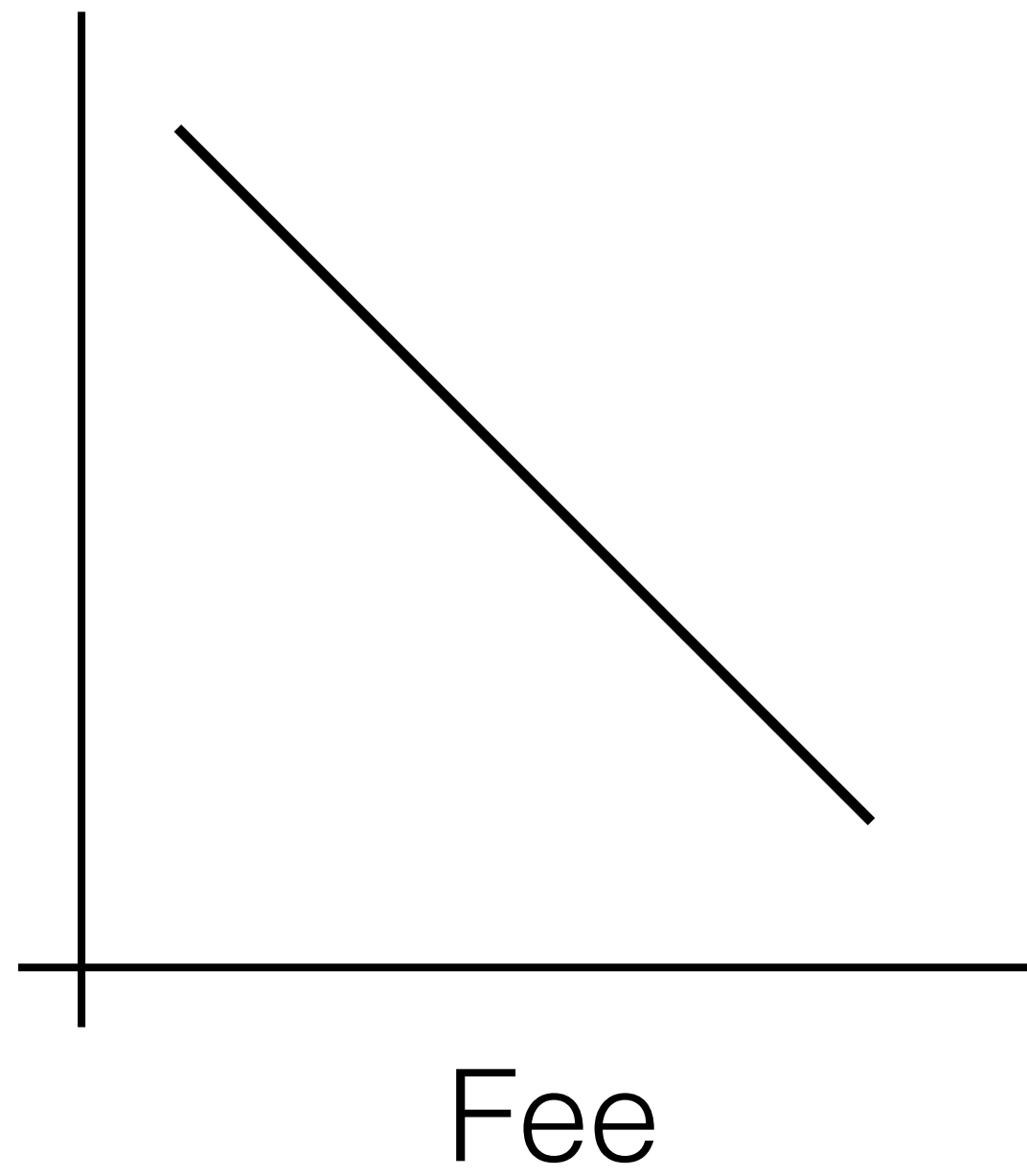
## Our Philosophy

- Opendoor bears risk in reselling the house
- Costs vary substantially by house
- Fair to each seller to charge based on their expected cost

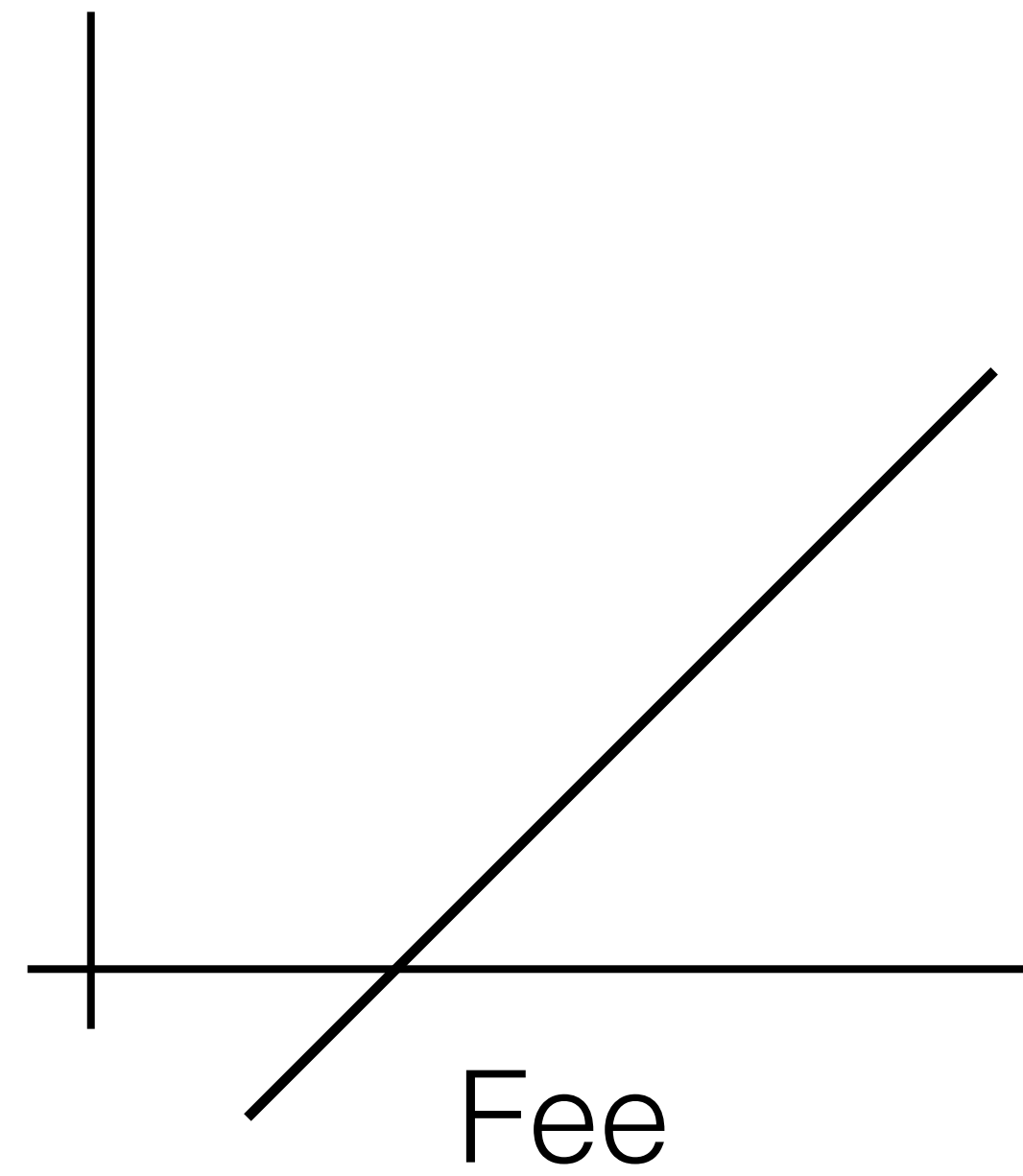
# Framing the problem

# House Economics

Conversion



Profit





## Formalization

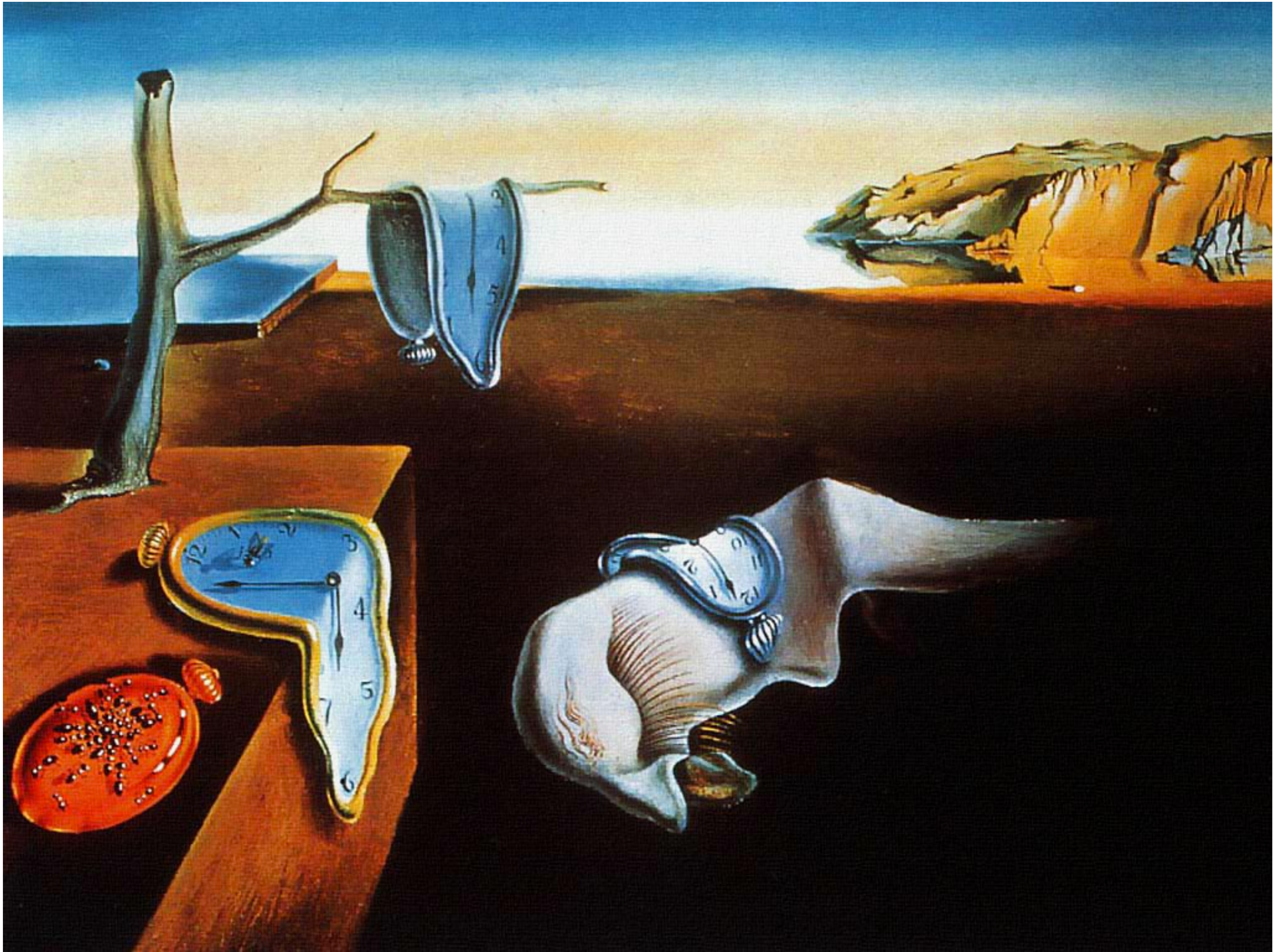
$\max_{f \in \mathcal{F}} \mathbb{E}[\text{number of houses acquired}]$  subject to expected profit  $> x$

- Infinite number of pricing models
- Assuming we even had a candidate  $f'$ , how do we test this?
  - A/B testing approach
    - randomize on offers:  $f$  vs  $f'$
    - evaluate  $\{\# \text{ of houses, profit}\}$

## Metric Measurement Lag

- Time to observe #
  - **days**
- Time to observe \$
  - **months**





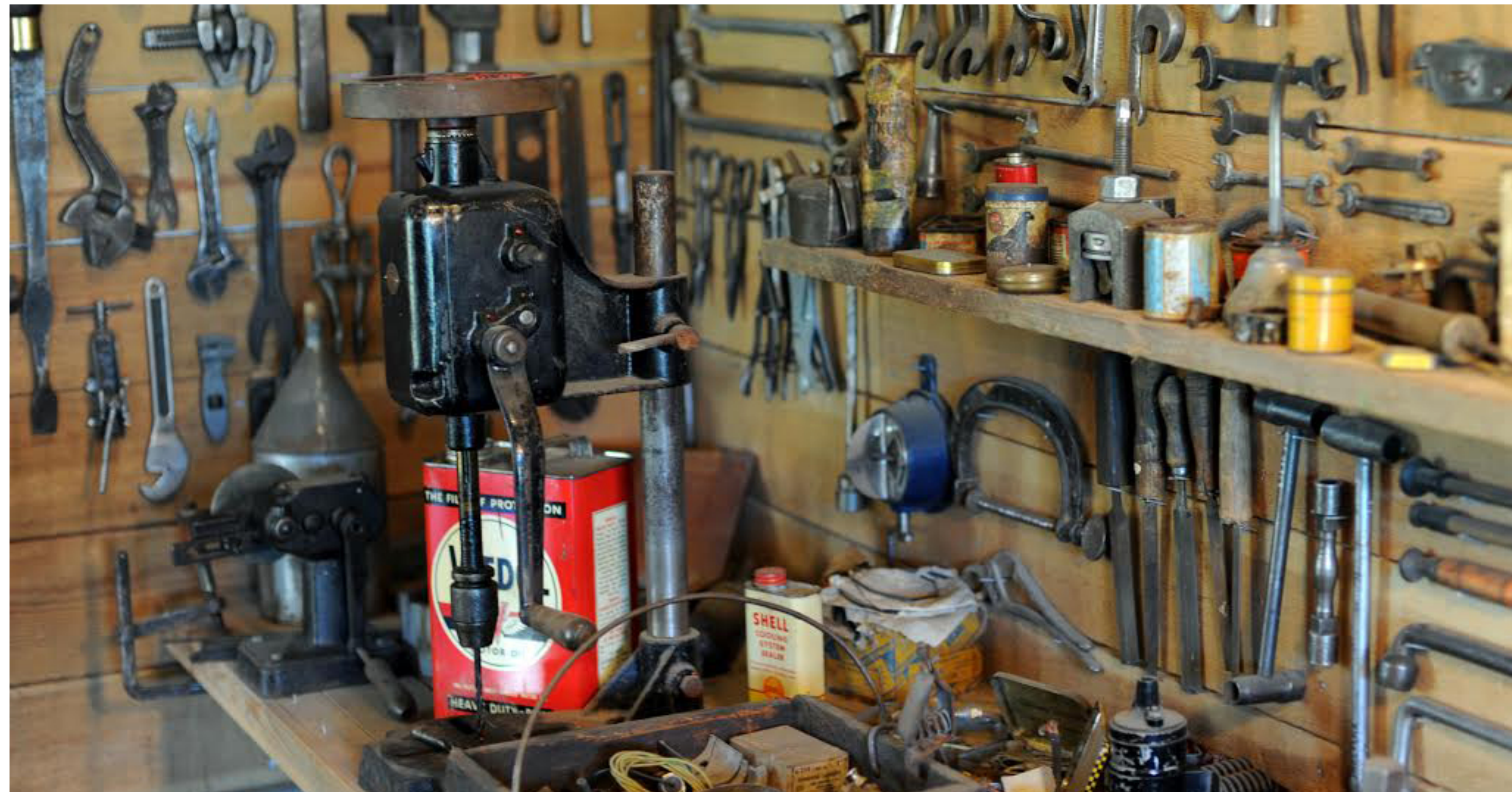


## Formalization

$\max_{f \in \mathcal{F}} \mathbb{E}[\text{number of houses acquired}]$  subject to expected profit  $> x$

- Infinite number of pricing models
- Assuming we even had a candidate  $f'$ , how do we test this?
  - A/B testing approach
    - randomize on offers:  $f$  vs  $f'$
    - evaluate {# of houses, profit}
- Many months of measurement lag







## Simulating Offers

- Historical transaction data
- House lists on the market
- Simulate our buying process
- Estimate our costs
- Observe actual outcome for house

## Simulating Offers



Actual resale cost: \$50k

Expected resale costs

- $f_{\text{under}}: \$10\text{k} \rightarrow \{P_{\text{accept}}: .9, \$: -40\text{k}\}$
- $f_{\text{base}}: \$55\text{k} \rightarrow \{P_{\text{accept}}: .1, \$: 5\text{k}\}$

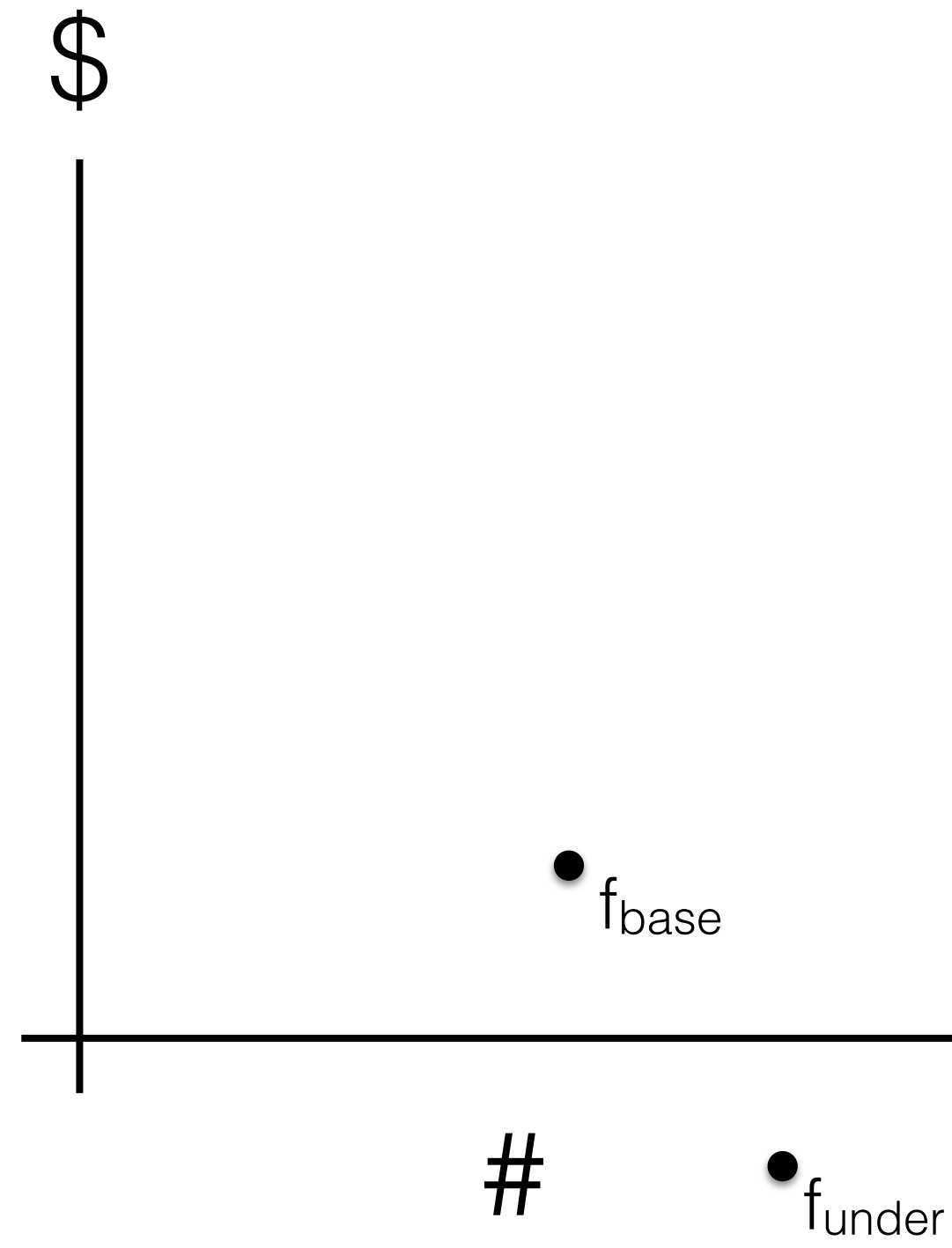


Actual resale cost: \$10k

Expected resale costs

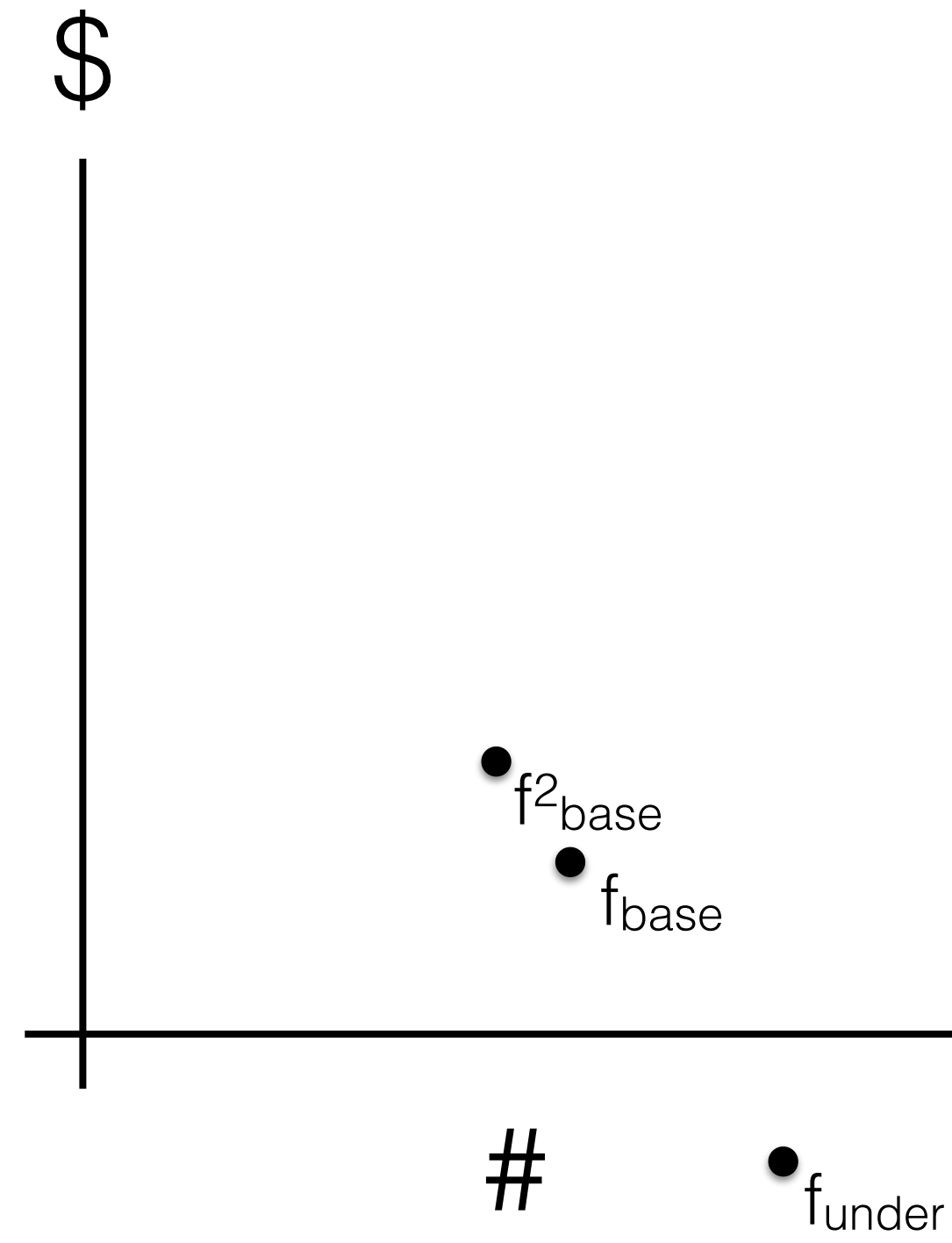
- $f_{\text{under}}: \$5\text{k} \rightarrow \{P_{\text{accept}}: .9, \$: -5\text{k}\}$
- $f_{\text{base}}: \$8\text{k} \rightarrow \{P_{\text{accept}}: .7, \$: -2\text{k}\}$

# Simulating Offers

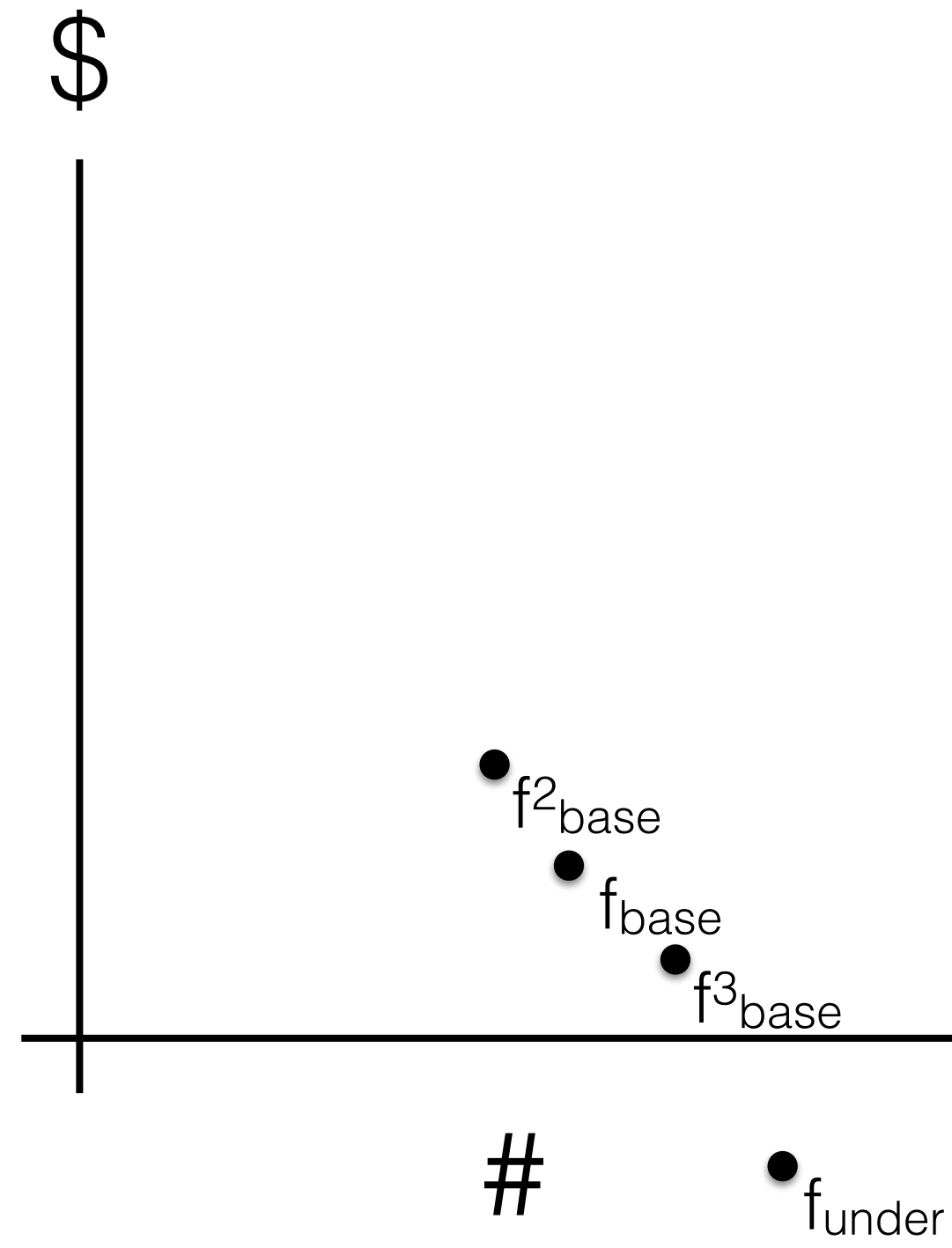




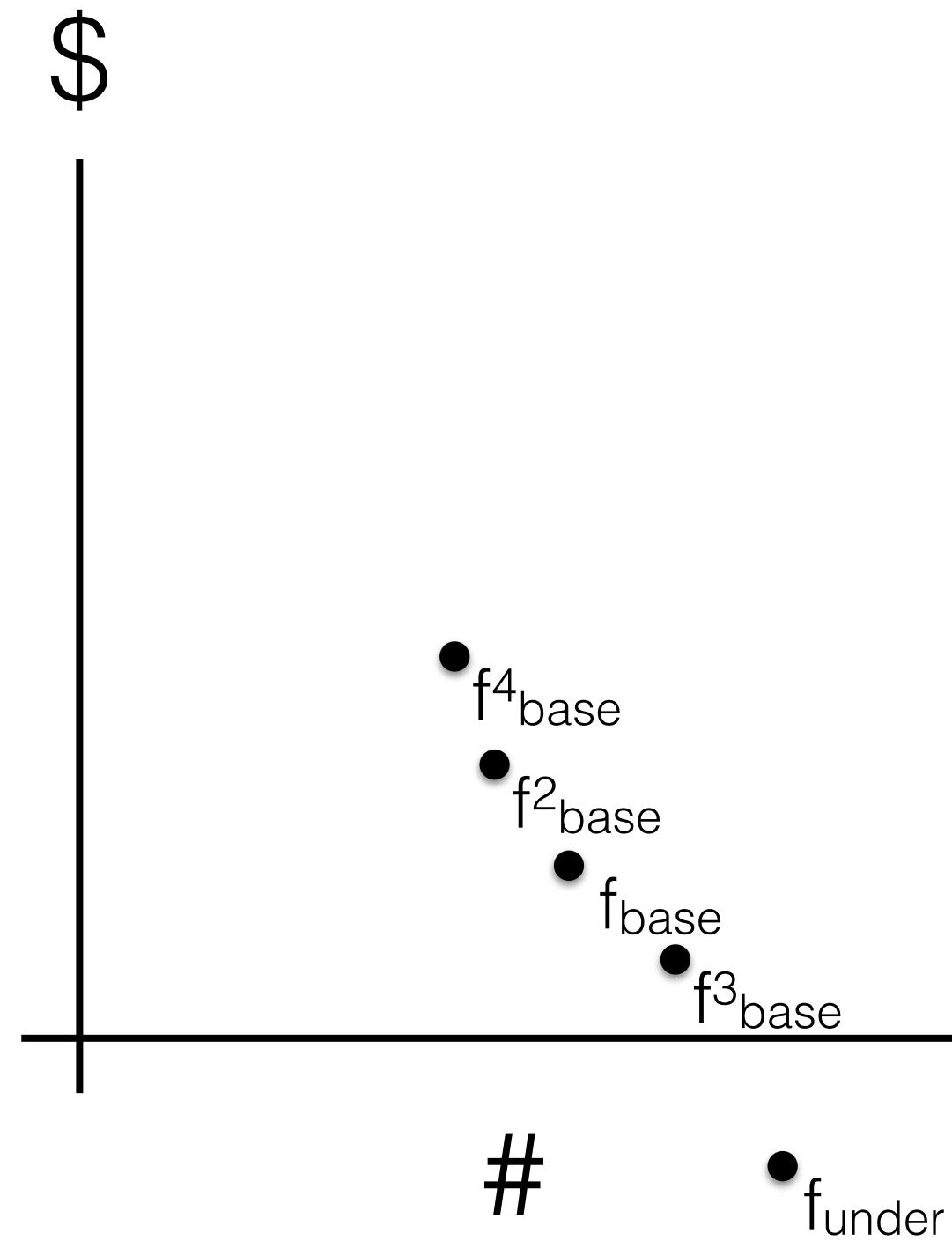
# Simulating Offers



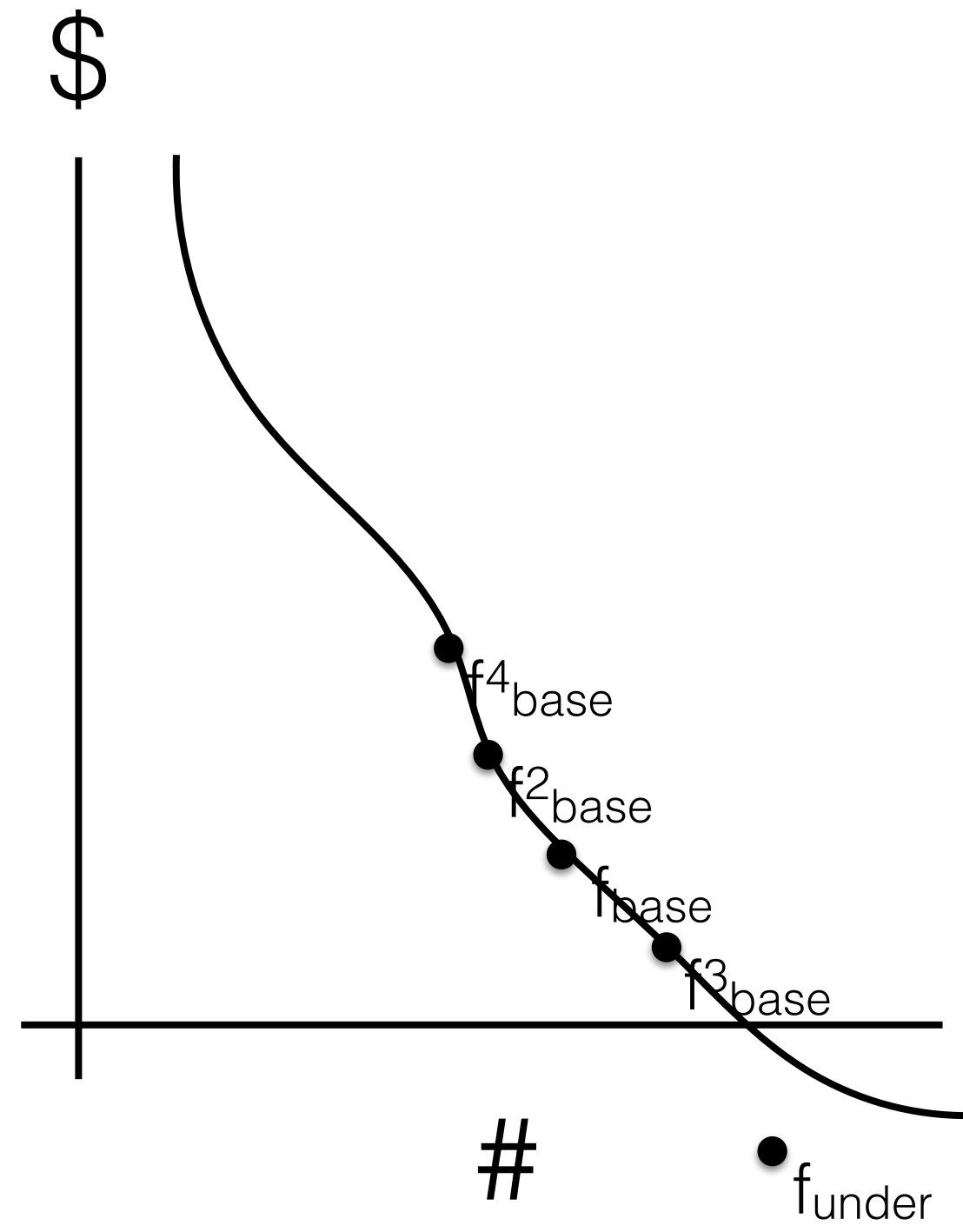
# Simulating Offers



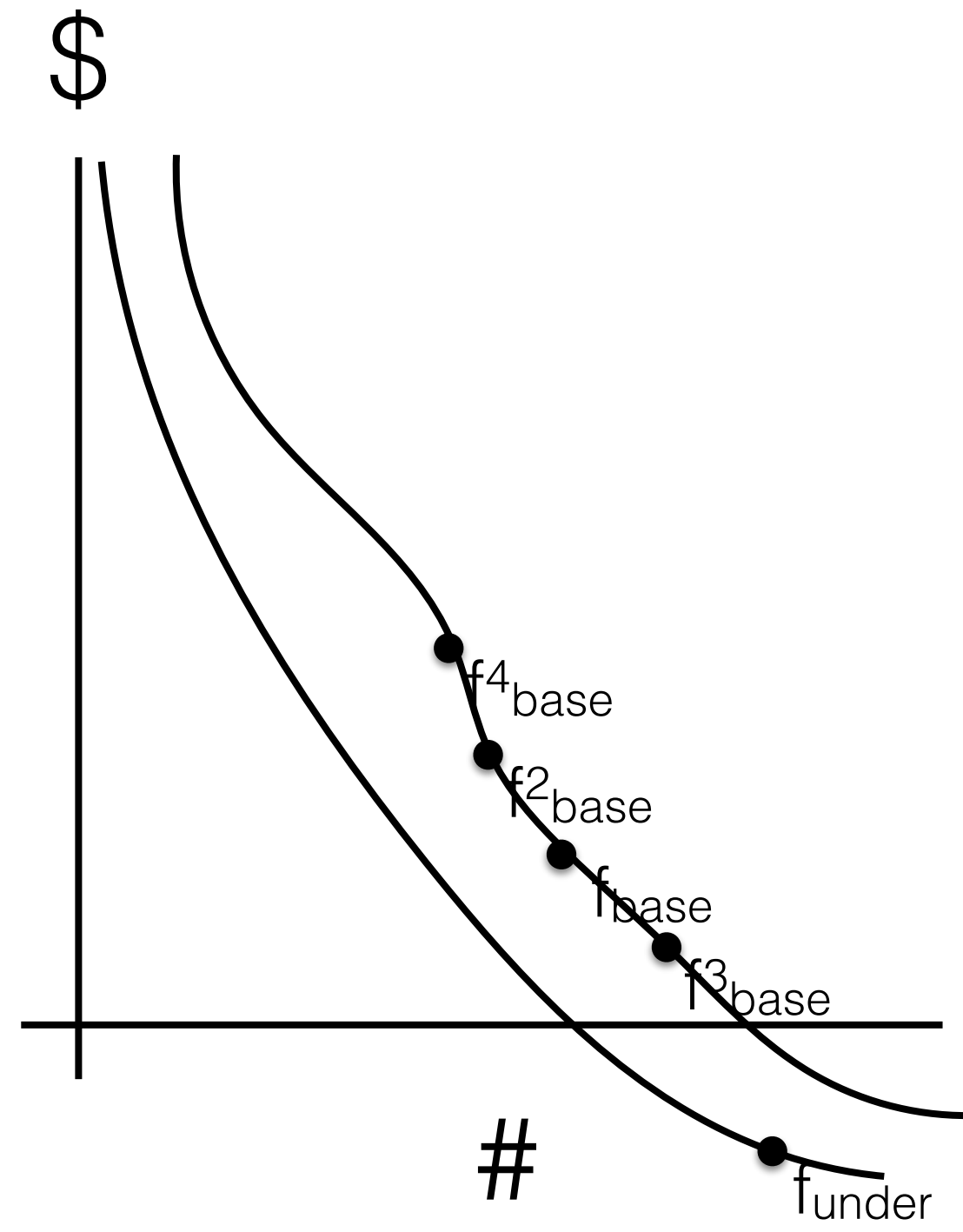
# Simulating Offers



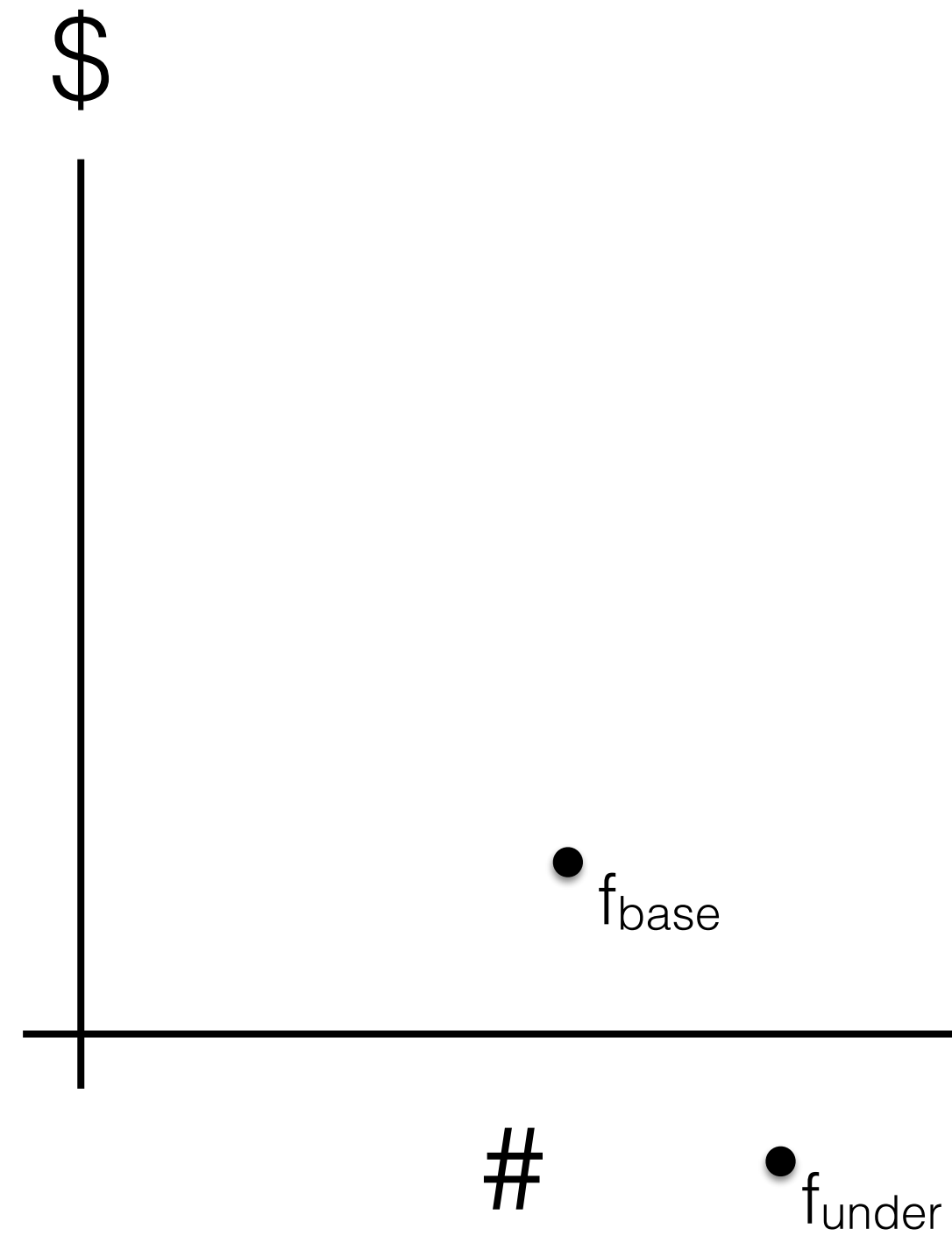
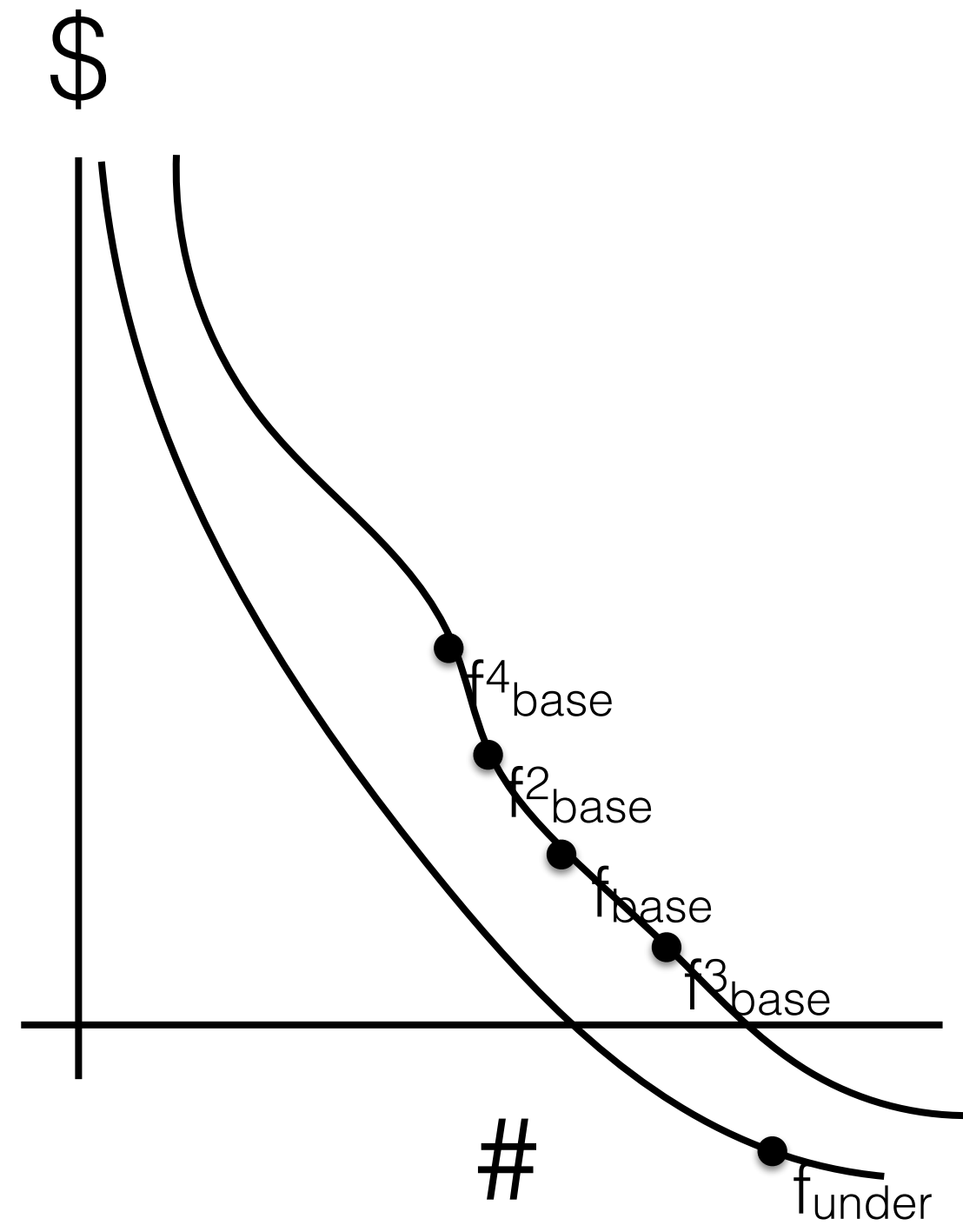
# Simulating Offers



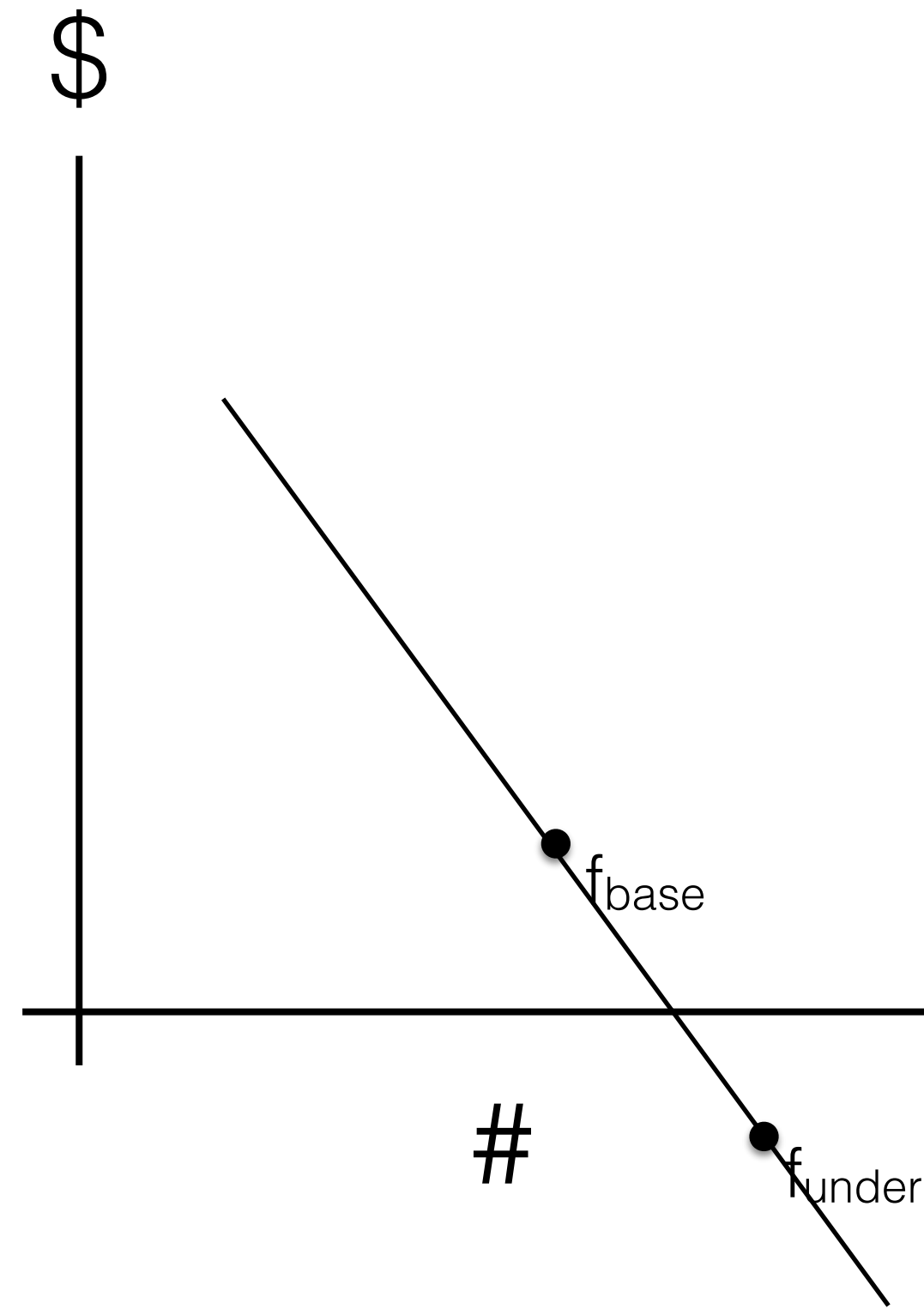
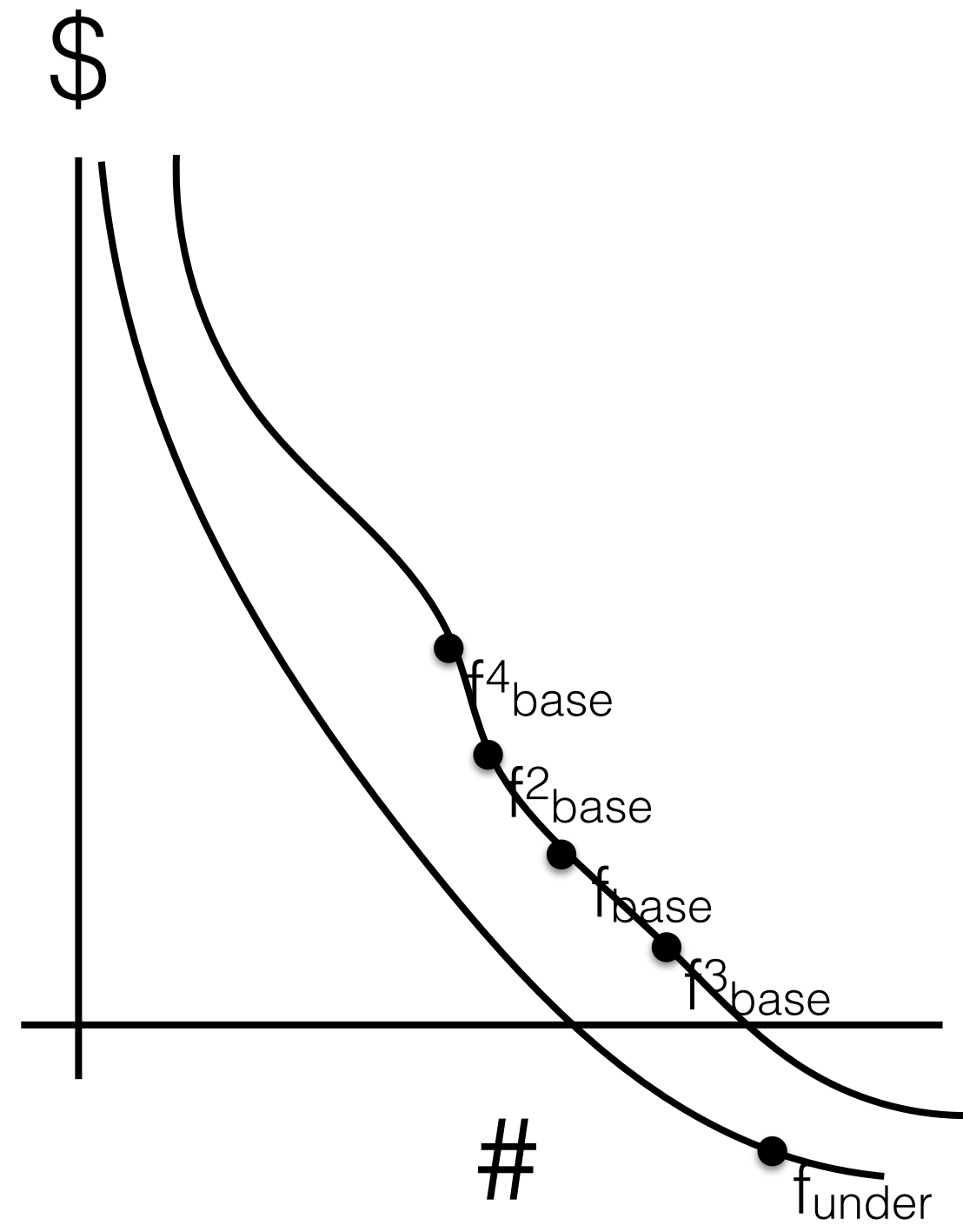
# Simulating Offers



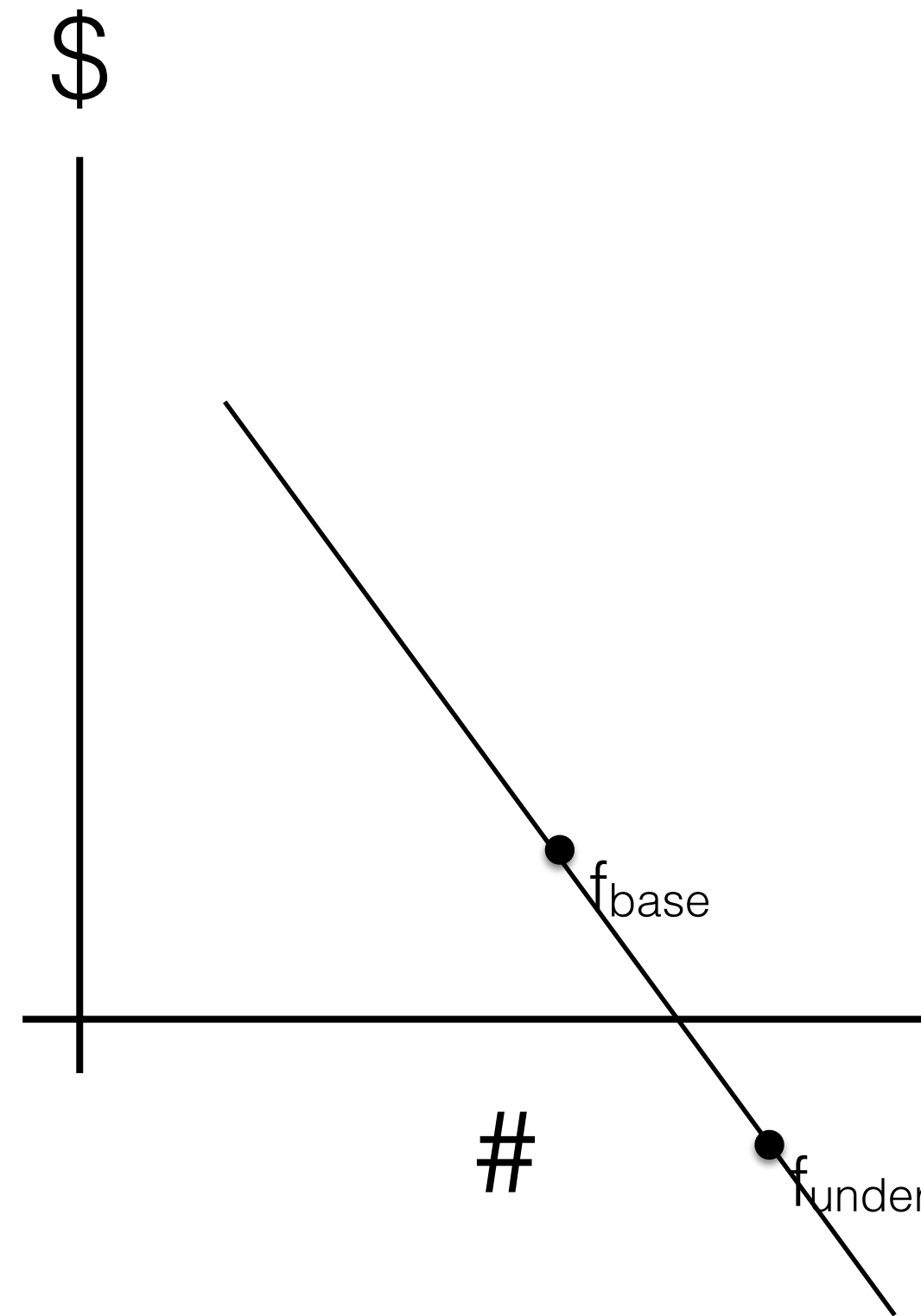
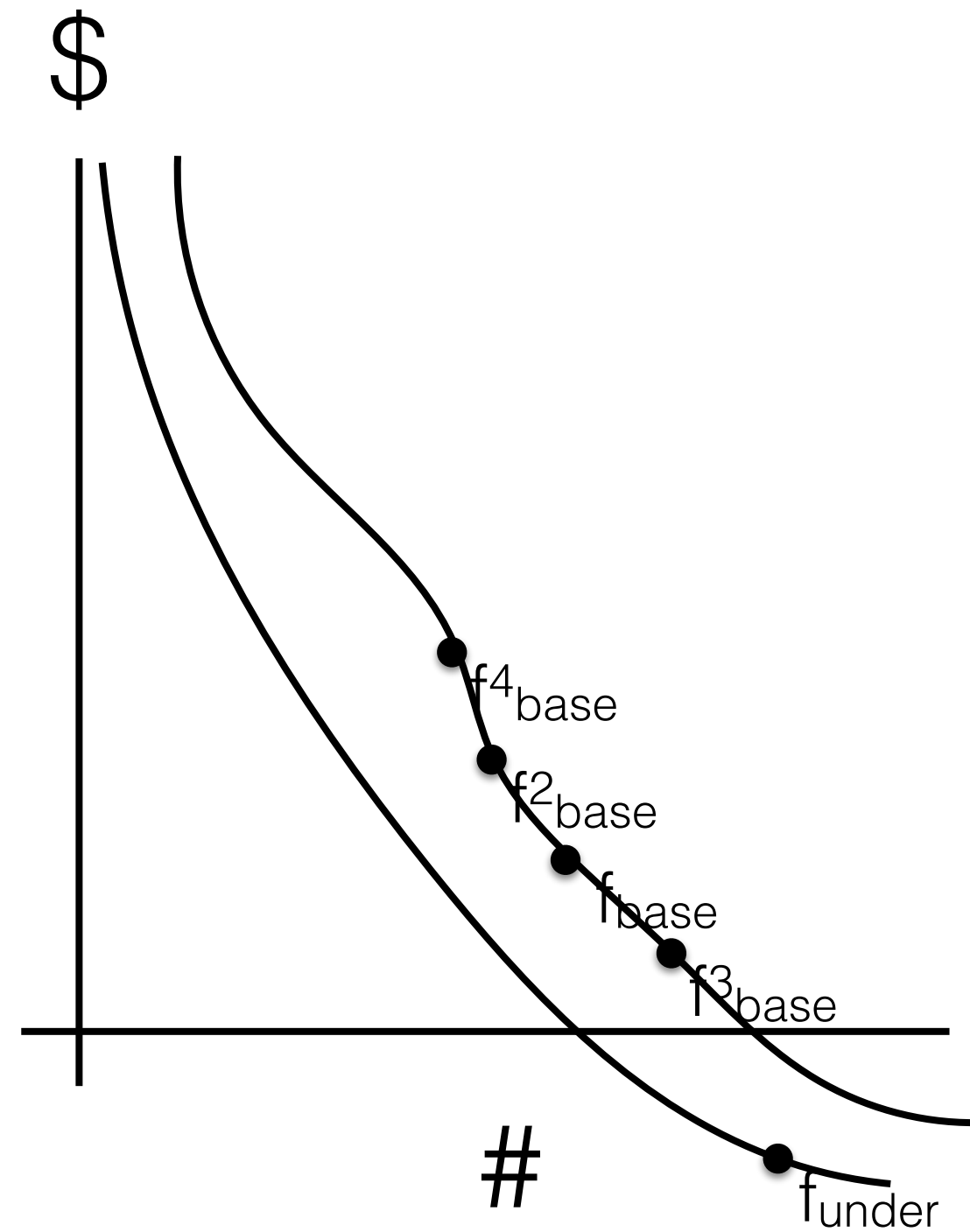
# Simulating Offers



# Simulating Offers



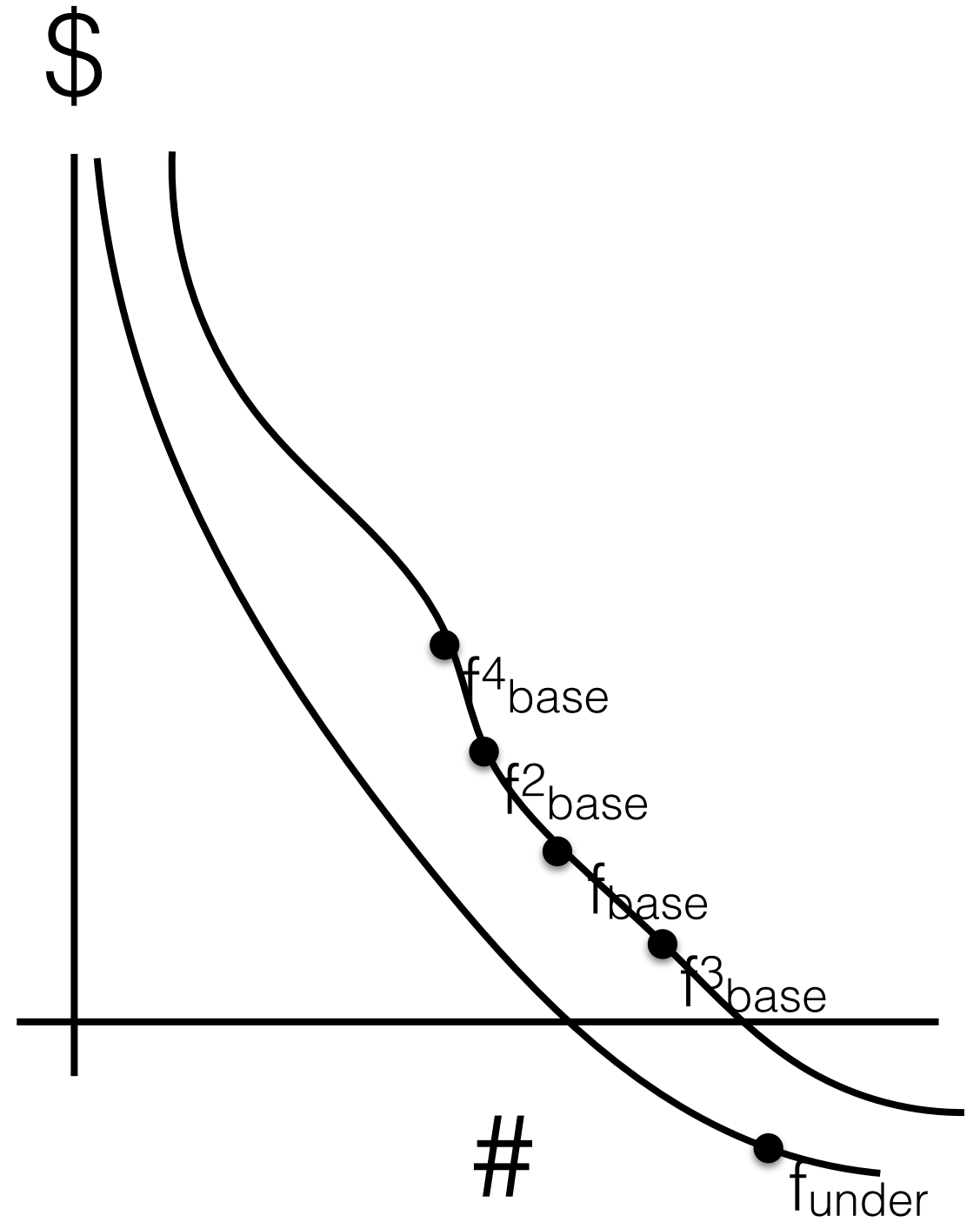
# Understanding Current Trade-Offs



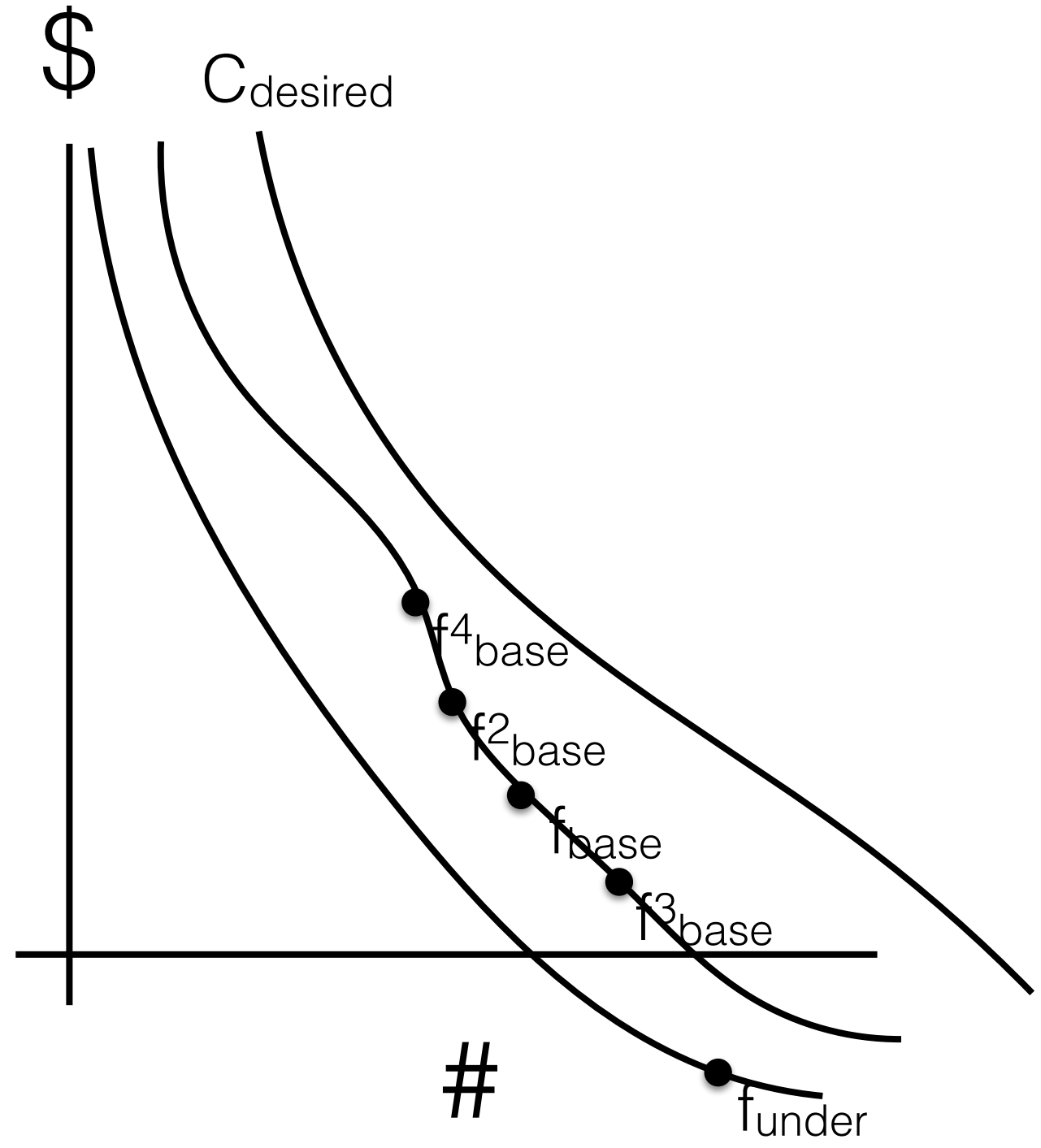
- Clarity into trade-offs
  - Identify suitable candidates
- Backtesting with business metrics
  - Seconds vs months
- Only cost is computational
  - Though quality dependent on simulation models



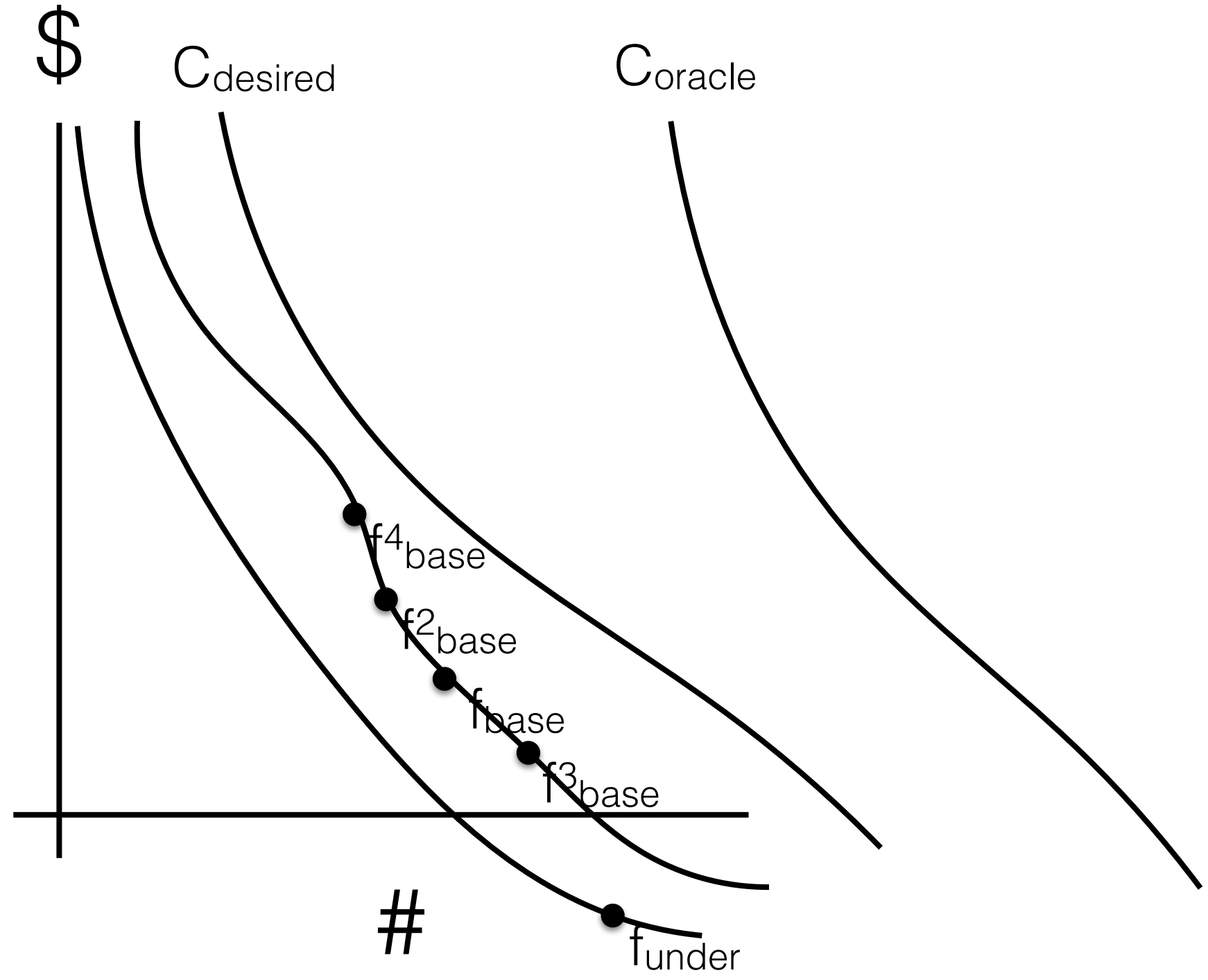
# Estimating Future Trade-Offs



# Estimating Future Trade-Offs



# Estimating Future Trade-Offs





## Oracle Performance



Actual resale cost: \$50k

Expected resale costs

- $f_{\text{under}}: \$10\text{k} \rightarrow \{P_{\text{accept}}: .9, \$: -40\text{k}\}$
- $f_{\text{base}}: \$55\text{k} \rightarrow \{P_{\text{accept}}: .1, \$: 5\text{k}\}$
- **$f_{\text{oracle}}: \$50\text{k} \rightarrow \{P_{\text{accept}}: .15, \$: 0\text{k}\}$**

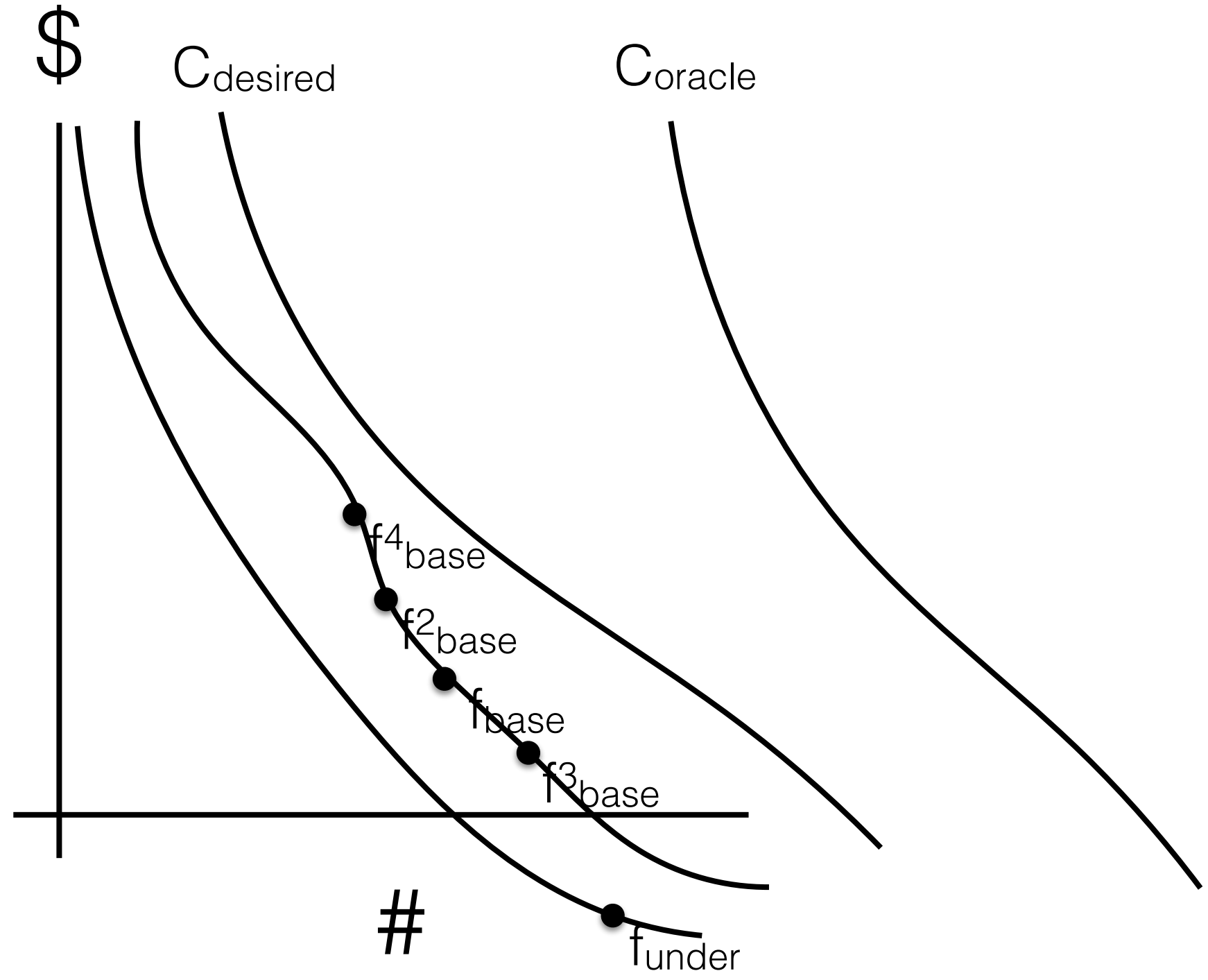


Actual resale cost: \$10k

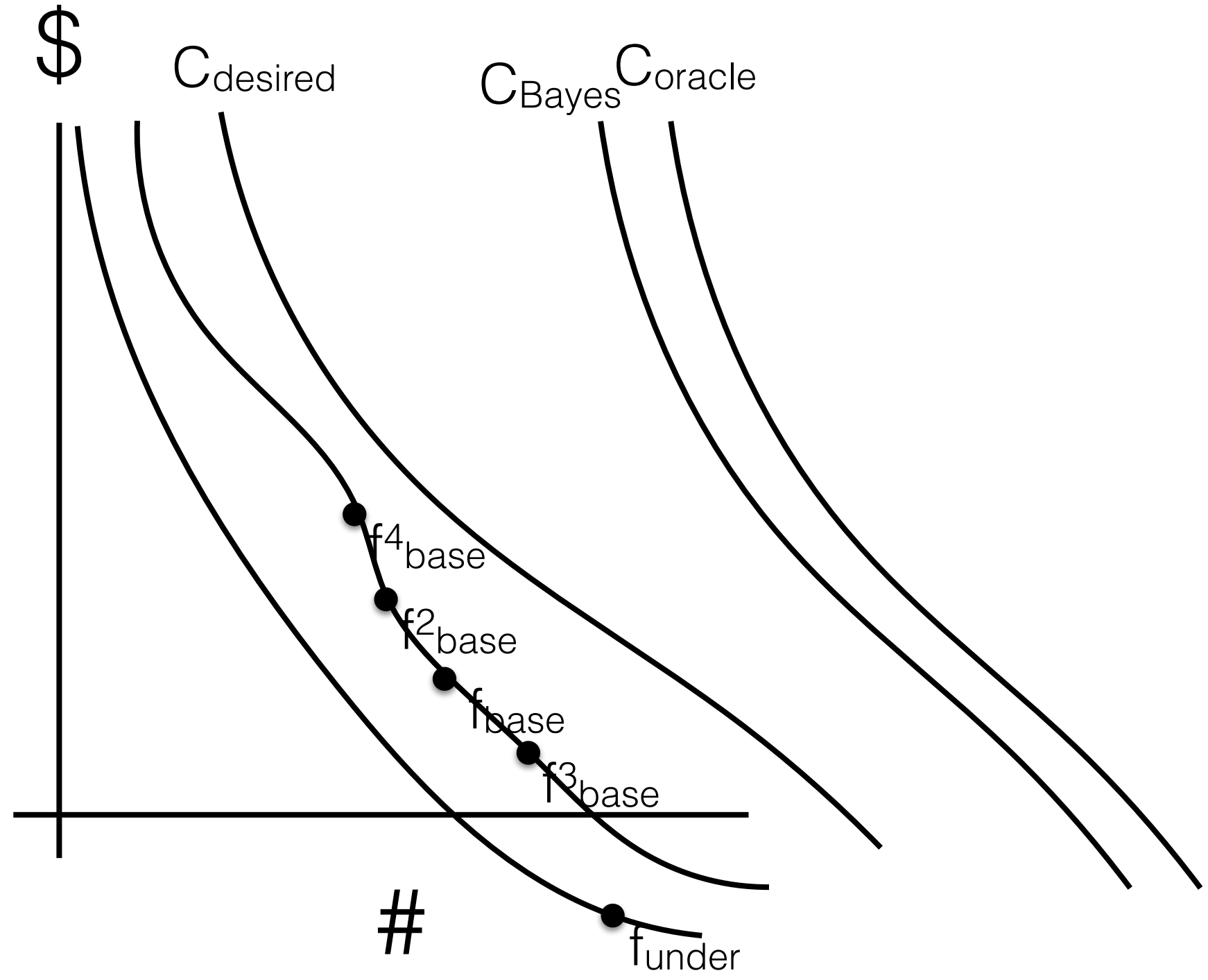
Expected resale costs

- $f_{\text{under}}: \$5\text{k} \rightarrow \{P_{\text{accept}}: .9, \$: -5\text{k}\}$
- $f_{\text{base}}: \$8\text{k} \rightarrow \{P_{\text{accept}}: .7, \$: -2\text{k}\}$
- **$f_{\text{oracle}}: \$10\text{k} \rightarrow \{P_{\text{accept}}: .65, \$: 0\text{k}\}$**

# Estimating Future Trade-Offs

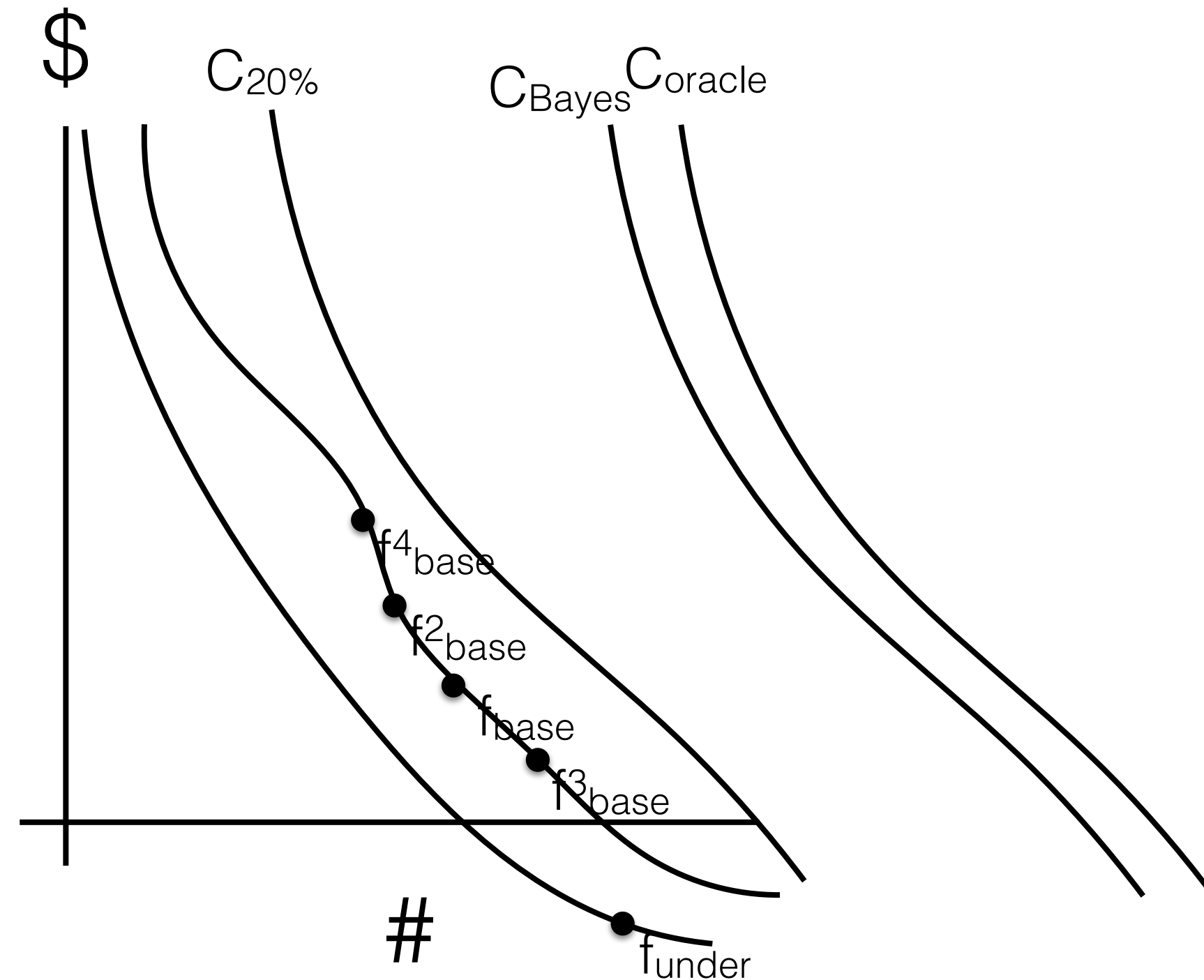


# Estimating Future Trade-Offs





# Estimating Future Trade-Offs



- Estimate what is theoretically achievable
- Set ML improvement goal
  - Translation into business trade-offs
  - “Easy” part is to hit ML target

# Simulation Accuracy

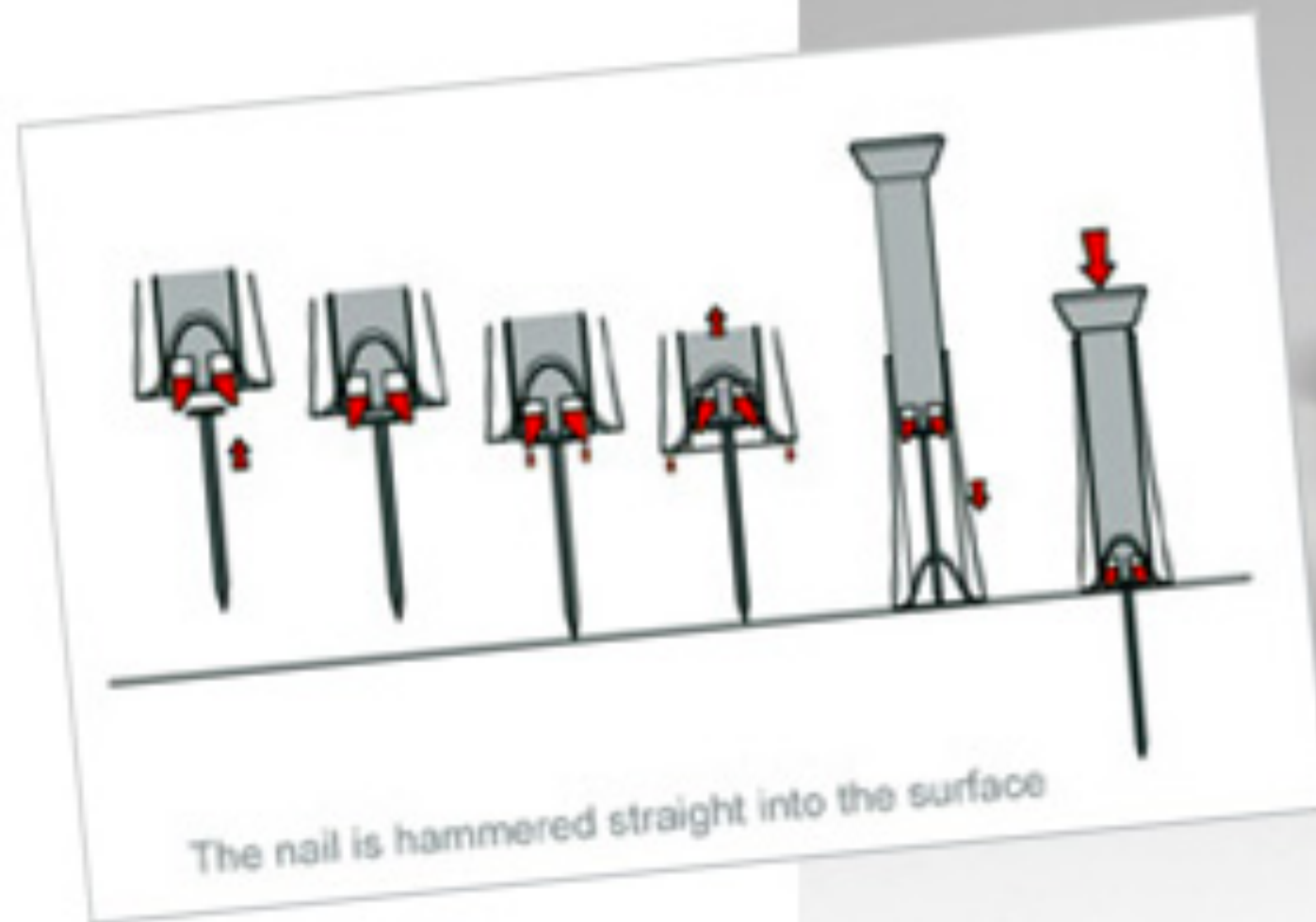


# Unguided A/B Testing





# The Guide



# Guided A/B Testing





# Pyramid of Causal Inference



Observational Analysis



Simulation-Based Inference

Confidence Cost

Quasi-experiments



A/B Test





# Recipe for guided testing

## Simulating Offers

- Historical transaction data
- House lists on the market
- Simulate our buying process
- Estimate our costs
- Observe actual outcome for house

Data generating process

User model

## Recipe: Data Generating Process

Simple version: replay historical data

- Home buying and selling
  - Past housing transactions
- Ridesharing services
  - Passenger app sessions
- Search engine ad auctions
  - Stock of potential ads



## Recipe: User Model

- Home buying and selling
  - $P(\text{sell} \mid \text{cost})$
- Ridesharing services
  - $P(\text{accept ride} \mid \text{price, ETA})$
- Search engine ad auctions
  - $P(\text{click} \mid \text{user features, ad features})$

**A/B test responsibly**

**Simulate before testing**



## Opendoor by the numbers

- Founded: March 2014
- Transactions / month: 500
- Number of employees: 300
  - 50 data scientists and engineers
- We're hiring!
- E-mail: [nelson@opendoor.com](mailto:nelson@opendoor.com)



# Acknowledgements





# Opendoor





**Q&A**