Dynamic Pricing Optimization in Real-Estate under financial and time constraints

Industrial Engineering and Operations Research
COLUMBIA | ENGINEERING

Pricing Optimization in Real Estate: Profit maximization under financial and time constraints

Underlying problem

Current pricing methods:

- **Rely only on expert knowledge,** without rigorous quantitative assessments of market statistics

Resulting in:

- **Substantial mispricing** of real estate assets

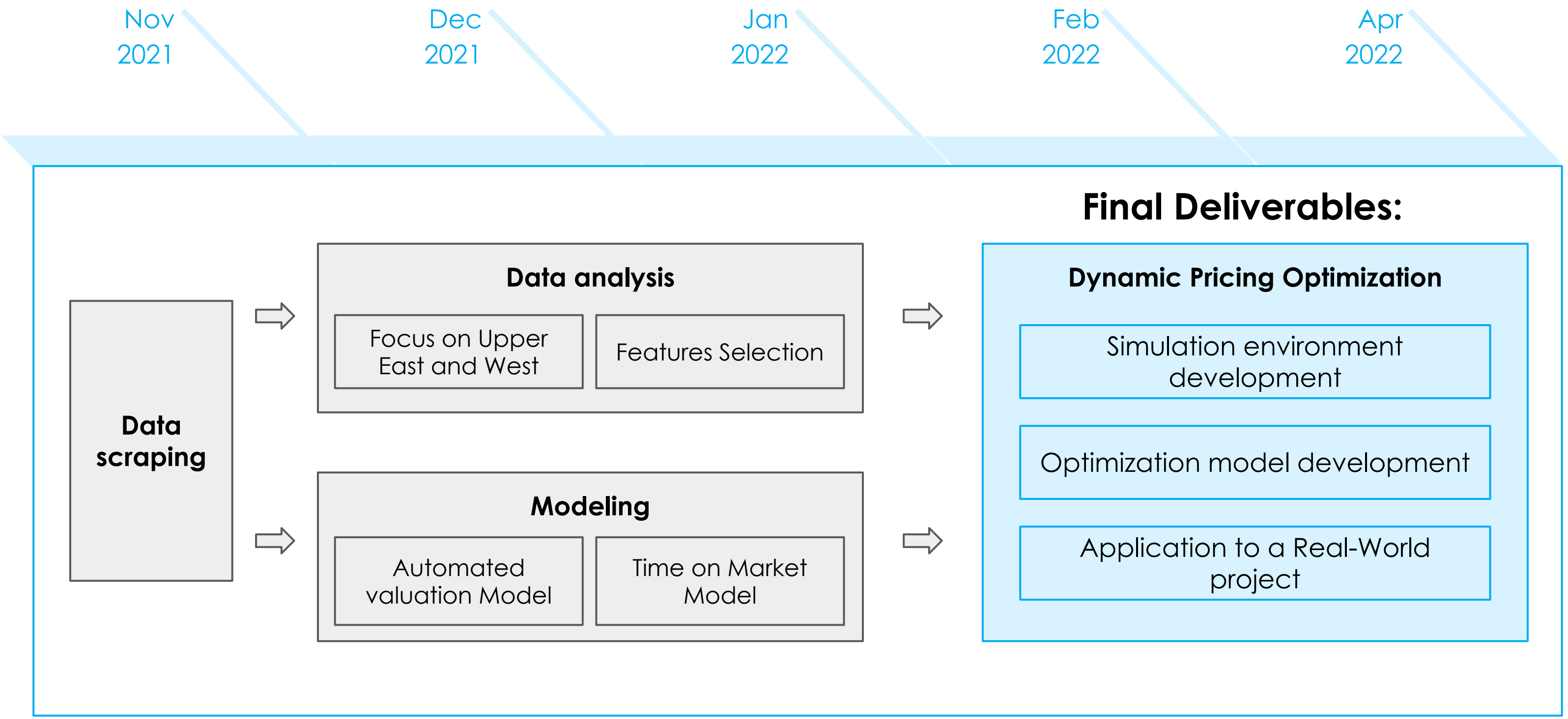
Key Deliverables

- **Dynamic Pricing Optimization algorithm** that maximizes returns on assets
- Application to a Real-World real estate project

Project Scope

- Location: **Manhattan**
- Real estate asset types: **Condominiums and Cooperatives**

A Three-Step Project - Focus on Pricing Optimization



Scraped data analysis suggests a focus on Upper East Side and Upper West Side Manhattan

Our data needs

- High importance of data quality
- Past sales data:
  - **Sale prices**
  - **Sale dates**
  - **Asset characteristics**
- Last year of data is required

Requires scraping online platforms

- Most efficient way of collecting recent data quickly
- Scraping twice during the year to update data
- Scraped information about around 6k sales on main marketplace platforms

Data-driven focus on UES and UWS

Neighborhood	Proportion of Manhattan[1] sales	Prices standard deviation
Upper East Side	23%	484
Upper West Side	23%	545
Midtown	14%	742
Downtown	34%	585
Financial District	6%	560

UES and UWS form a large enough homogeneous market

1. Excluding Harlem, not in the focus of the project

Full Workflow: 4 Models to Optimize Real Estate Prices

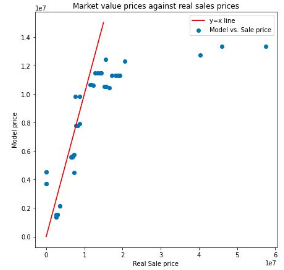
Key Assumptions

1. Assets have an 'intrinsic value' (AVM)
2. Sales ~= Demand (Simulation Model)
3. Selected market is homogeneous (DP Model)

a. Automated Valuation Model (AVM)

Determine intrinsic value of assets

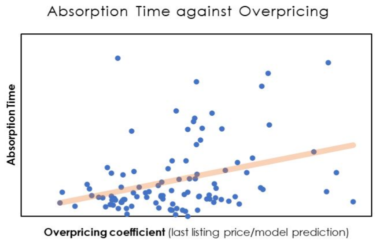
Highly accurate: OOS $R^2 = 0.90$



b. Time-on-Market Model (ToMM)

Determine link between price and time-to-sell

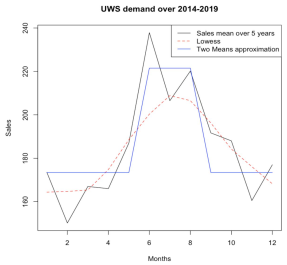
Low R^2 due to market inneficiencies



c. Demand Simulation Model

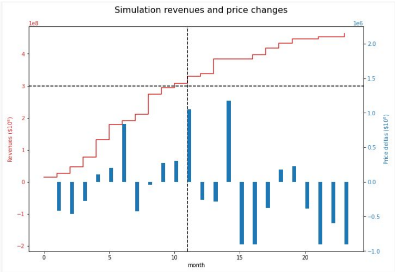
Sets up simulation environment for Dynamic Pricing backtesting

Poisson laws



d. Dynamic Pricing Model

Optimizes prices at each time interval to maximize profits



Determining market value of Real Estate assets using scraped data

Model Development

Feature engineering & selection

- From all available variables in the dataset:
  - Engineer potentially relevant variables
  - Select most important and statistically significant variables
  - Drop highly correlated variables

Model comparison & selection

- Build multiple models (Linear Regressions, Random Forests, Gradient Boosted Trees, various sets of features)
- Compare validation scores of each model
- Retain the best model and fine-tune it

Model Results & Possible Improvements

Most important features include

- Square footage
- Building age
- Floor

Observations

- Currently using: **Gradient Boosting algorithm**
- **NRMSE = 0.09**

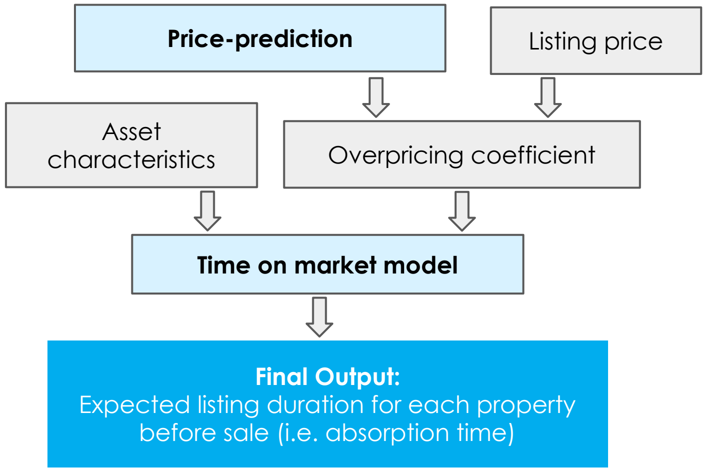
Possible improvements

- Include the market state (seasonality, mortgage rates) in the model

Time on Market model leverages price prediction to determine absorption time

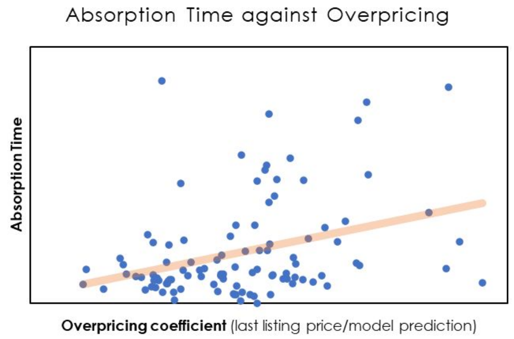
Model Overview

Feature Engineer an "Overpricing Coefficient" from the output of the previous model



Findings

Results: $R^2 = 0.511$ - positive but low correlation between overpricing and absorption time



- Low correlation explained by market inefficiencies
- The model's output can be used to infer optimization parameters

Creating a Simulation Environment for Optimization

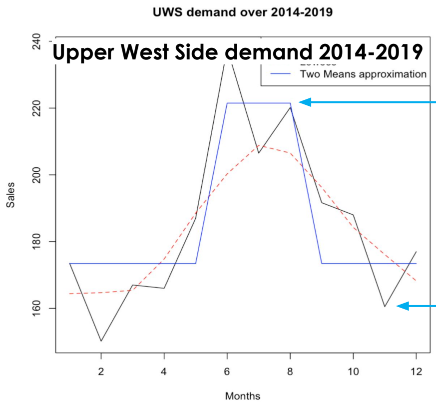
Simulating Demand in Real Estate

Demand in Real Estate Markets is complicated:

- 1) **Decentralized**, 2) **Heterogenous**, 3) **Dynamic**

Used aggregate historical sales as a proxy for demand; generated Poisson, 2-mean fixed distributions (#sales/week or month)

Generated different demand scenarios to account for dynamic market states



Two-means approx:

Summer months (June-Aug) have higher demand

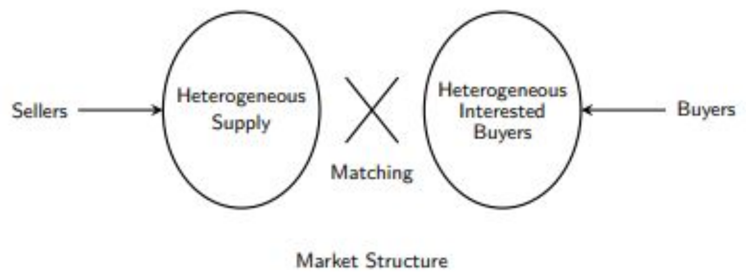
Other months have lower demand

Dynamic Pricing Optimization: Choice of Algorithm

Criteria: Robustness under real-estate context / constraints

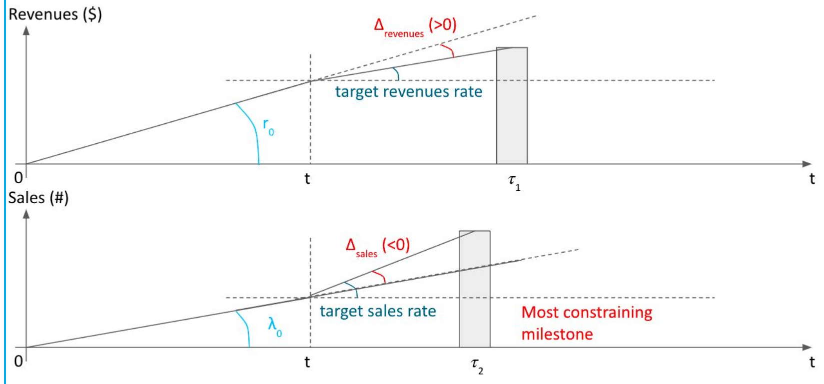
Took algo from Prof Maglaras' (CBS) paper: **"***Dynamic Pricing with Financial Milestones: Feedback-form Policies"*

- **Accounts for dynamism and decentralization of markets; does not require demand price response curve**



Pricing algorithm and application to a real-world project

Algorithm compares current and expected revenues and sales rates

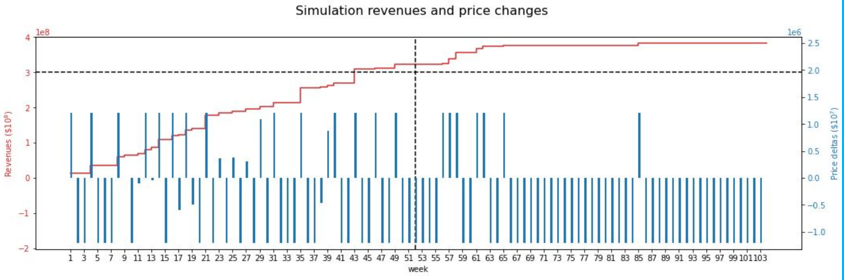


- Prices are decreased when revenues and sales overshoot targets
- Prices are increased otherwise, proportionally to the delta

O. Besbes and C. Maglaras, "Dynamic Pricing with Financial Milestones: Feedback-Form Policies," *Management Science,* vol. 58, pp. 1715-1731, 2012.

DP Model's Simulated Revenues and Price Matrix:

Targets: 1) Revenues should reach $300MM within 12 months , 2) All units should be sold in 24 months



Initial Pricing Matrix

$52,903,403					
$49,401,042					
$41,402,192					
$16,753,273			$18,791,208		
$13,149,426			$17,539,531		
$12,783,043			$19,357,942		
$13,369,707			$17,975,062		
$13,165,891			$15,866,984		
$11,514,335			$17,489,585		
$11,757,088			$14,744,795		
$10,684,427			$15,896,692		
$9,332,805		$7,494,402		$7,796,935	
$7,967,100		$6,585,948		$7,014,912	
$7,896,463		$6,593,283		$6,234,413	
$7,968,873		$7,646,335		$7,533,054	
$2,477,712	$3,161,664	$3,284,688	$2,798,627	$2,757,690	$3,282,577
$2,532,429	$3,341,479	$3,145,801	$2,752,490	$2,808,048	$2,723,861

Revenues = $0

Mid-sales Pricing Matrix

$39,185,890					
$39,202,366					
$26,921,075					
$11,269,502			$13,788,030		
$11,942,387			$13,294,191		
$15,209,799			$12,132,180		
$9,347,212			$13,147,145		
$9,559,750			$11,843,956		
$8,638,185			$14,132,298		
$10,946,656			$12,217,679		
$7,424,966			$10,166,432		
$6,182,306		$5,036,294		$4,836,481	
$6,795,868		$5,947,659		$5,134,332	
$6,499,565		$5,119,339		$5,319,328	
$6,926,029		$5,128,803		$5,424,721	
$1,639,823	$2,858,644	$1,977,385	$2,868,383	$2,023,294	$1,965,590
$1,672,632	$2,967,067	$2,083,282	$2,650,370	$2,532,083	$1,918,790

Revenues = $322,199,580

End of Sales Pricing Matrix

$40,958,895					
$40,688,615					
$27,424,409					
$11,967,504			$18,470,879		
$11,518,935			$12,985,214		
$14,202,473			$13,604,229		
$8,813,753			$12,138,869		
$9,821,344			$12,492,928		
$8,975,452			$13,741,518		
$10,686,685			$13,604,505		
$7,445,758			$10,398,586		
$5,628,231		$5,057,449		$5,249,377	
$6,510,407		$5,492,914		$6,534,866	
$6,666,614		$5,123,307		$5,230,855	
$6,899,973		$5,170,705		$5,617,101	
$1,698,495	$2,687,358	$2,000,402	$2,870,727	$1,960,182	$2,213,821
$1,680,057	$2,713,379	$2,306,357	$2,997,986	$2,724,213	$2,516,113

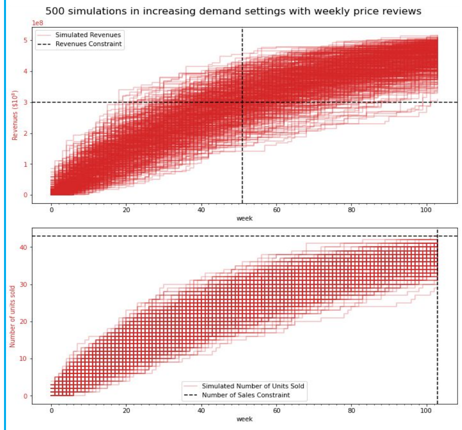
Revenues = $381,925,080

Numbers in these matrices are mock numbers

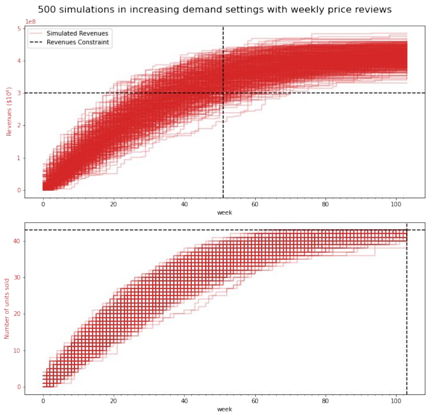
Results and Application to a Real-World Project

DP Model’s Simulated Revenues and Sales vs. Static Pricing:

Static Pricing:
500 simulations revenues and sales over time



Dynamic Pricing:
500 simulations revenues and sales over time



Fulfilment rate of Revenue targets under different demand scenarios

Demand	Weekly repricing	Monthly repricing	Static prices
Decreasing	59.2%	52.6%	38.4%
Stable	78.6%	56.0%	49.4%
Increasing	87.6%	63.8%	56.0%

Ave Fulfilment: 47.9% → 75.1%

Ave Revenues: $291.9m → $335.3m

Compared to a static price strategy, dynamic pricing gives **significantly better results** across all demand scenarios.