Office of the
Chief Economist

Microsoft

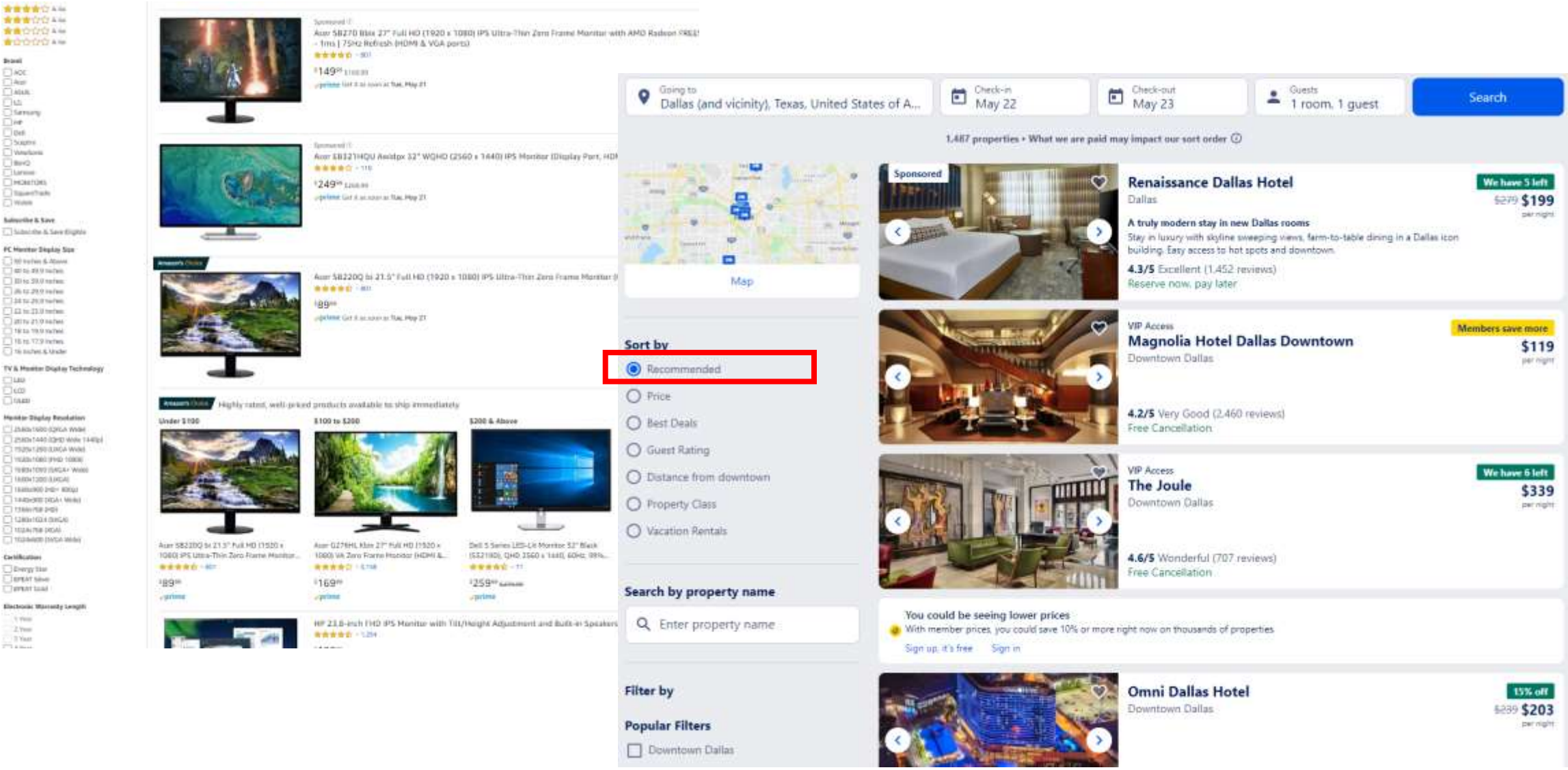
Online Search and Rank Optimization

May 2019

Will Wang

joint with Giovanni Compiani (Berkeley), Greg Lewis (Microsoft), Sida Peng (Microsoft)

Online search is ubiquitous



How to design online search platforms?

- Focus on search behavior and product rankings
- Long-run relationships with consumers: maximizing consumer surplus

- Understand search behavior:
  - Rich data in search patterns and choices
  - A general framework that subsumes several existing models (e.g. Weitzman 1979)

- Ranking optimization (under parametric assumptions):
  - Ranking affects salience and thus search patterns
  - Computationally tractable

Literature

- Theory: Stigler (1961), Diamond (1971), Butters (1977), Rothschild (1974), Salop and Stiglitz (1977), Weitzman (1979), Varian (1980), Stahl (1982), Burdett and Judd (1983), Wolinsky (1986), Milgrom and Roberts (1986), Mortensen and Pissarides (1994), Keller and Oldale (2003), Ellison (2005), Gabaix and Laibson (2006), Armstrong, Vickers and Zhou (2009), Athey and Ellison (2011), Branco, Sun and Villas-Boas (2012), Bordalo, Gennaioli and Schleifer (2013), Ke, Shen and Villas-Boas (2016)

- Empirics: Mehta, Rajiv and Srinivasan (2003), Hong and Shum (2006), Hortacsu and Syverson (2006), Sovinsky (2008), Kim, Albuquerque and Bronnenberg (2010), Abaluck and Gruber (2011), Chandra and Tappata (2011), Wildenbeest (2011), Jeziorski and Segal (2015), Blake, Nosko and Tadelis (2016), Honka and Chintagunta (2017), Abaluck and Adams (2017), Honka, Hortacsu and Vitorino (2017), Blake, Moshary, Sweeney and Tadelis (2018), Brancaccio (2018), Hodgson and Lewis (2018), Moraga-Gonzalez, Sandor and Wildenbeest (2018)

How do consumers search?

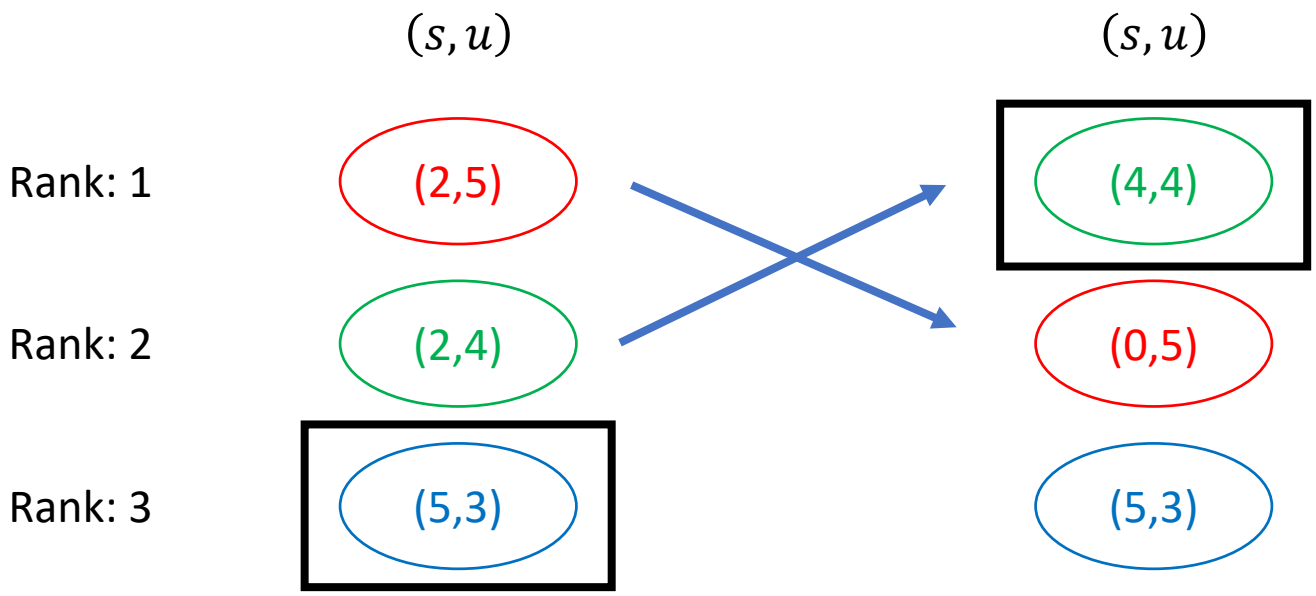
- A general double index framework (subsuming several popular search models)

- Each product has a search and utility index (s_j, u_j)
- The consumer knows the search indices and the utility of the outside option u_0
- The remaining utilities $\{u_j\}_{j=1}^{J}$ are unknown
- Consumer learns u_j by searching for product j

- Rank goods according to s: $s_1 \geq s_2 \geq \cdots \geq s_J$
- If $s_1 < u_0$, stop searching without purchase
- Else, search good 1, then
  - If $s_2 < \max\{u_0, u_1\}$, stop searching and maximize utility over goods 0 and 1
  - Else, search good 2, then
    - …

Why not rank by "product relevance"?

- How to best rank the alternatives to maximize consumer surplus?
- Consider an example: 3 products Red, Green, and Blue; $u_0 = 0$
  - Ranking affects search index on top of a "baseline" search index (based on other attributes)
- Further complicated with consumer heterogeneity



Consumer surplus

- The product that gets purchased is the one with the highest $\min\{u_{ij}, s_{ij}\}$ (modifying Choi, Dai, Kim, ECMA 2018) – i.e. both salience and utility matter for choice

- Introduce "double logit": Let $s_{ij} = \delta_j^S + \epsilon_{ij}^S$ and $u_{ij} = \delta_j^U + \epsilon_{ij}^U$, where ϵ's are distributed Type I EV

- Assume perfect correlation of errors: $\min\{u_{ij}, s_{ij}\} = v_{ij} = \delta_j^V + \epsilon_{ij}$, where $\delta_j^V = \min\{\delta^S, \delta^U\}$

- Ex-ante consumer surplus (i.e. prior to the realization of the common logit shock):

$$E(CS) = log\left(\sum_j \exp \delta_j^V\right) + \sum_{j:\delta_j^U > \delta_j^S} \frac{\exp \delta_j^V}{\sum_k \exp \delta_k^V}\left(\delta_j^U - \delta_j^S\right)$$

Rank optimization

- Allocate resource r from budget R to increase salience: $\delta_j^S = \delta_{j0}^S + r$
- Marginal benefit of rank promotion: let $q_j = \frac{\exp \delta_j^V}{\sum_k \exp \delta_k^V}$, when $\delta_j^U > \delta_j^S$,

$$\frac{\partial\big(E(CS)\big)}{\partial r} = q_j\big(\delta_j^U - \delta_j^S\big) - q_j \sum_k q_k\big(\delta_k^U - \delta_k^S\big)$$

- "Potential" (=$\delta^U - \delta^S$): Always only promote positive potential products
- Salience: Within the set of positive potential products, baseline market share matters
- Avoid two mistakes:
  - Don't promote products that will be ultimately disappointing
  - Don't promote products that have high potential but very low baseline salience

Rank optimization

- Allocate resource r from budget R to increase salience: $\delta_j^S = \delta_{j0}^S + r$
- Marginal benefit of rank promotion: let $q_j = \frac{\exp \delta_j^V}{\sum_k \exp \delta_k^V}$, when $\delta_j^U > \delta_j^S$,

$$\frac{\partial\big(E(CS)\big)}{\partial r} = q_j\big(\delta_j^U - \delta_j^S\big) - q_j \sum_k q_k\big(\delta_k^U - \delta_k^S\big)$$

- "Potential" (=$\delta^U - \delta^S$): Always only promote positive potential products
- Salience: Within the set of positive potential products, baseline market share matters
- Avoid two mistakes:
  - Don't promote products that will be ultimately disappointing
  - Don't promote products that have high potential but very low baseline salience

A greedy algorithm

- Discrete ranks require an algorithm

- For each rank $r = 1,2,\dots,J$, let $\delta_j^S = \delta_{j0}^S + f(r)$ if product j is ranked at r
  - If only one or no products with potential $(\delta_j^U - \delta_{j0}^S)$ greater than $f(r)$

    $\Rightarrow$ rank product j with highest potential at rank r ("Minimize waste")

  - If multiple products with potential greater than $f(r)$

    $\Rightarrow$ rank product j with highest $\frac{\partial(E(CS))}{\partial r}$ among products with potential greater than $f(r)$

Simulations

- Base case: $\delta^U, \delta^S \sim i.i.d. N(0,1)$
  - Ranking technology: $f(r) = A \cdot \exp(-r)$, where A=5

- Normalize CS from worst result (0) to best result (1) across all possible rankings

- Results hold for many products
  - Our algorithm is fast to compute when enumeration is infeasible

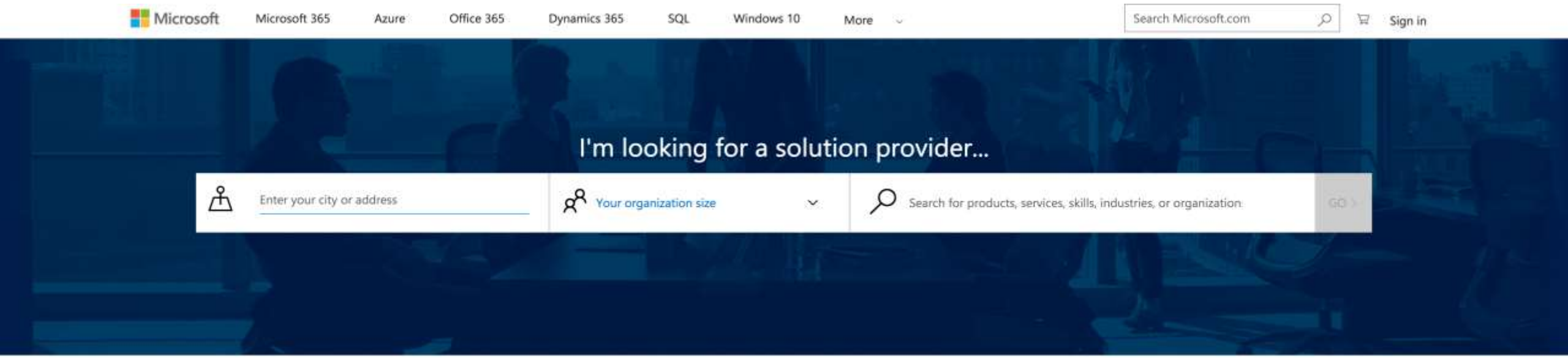
	Base	Strong ranking (A=15)	High utility / Weak ranking ($\bar{u}$=3; A=1)
Enumerating all rankings			
Best result	100%	100%	100%
Average result	30%	35%	58%
Worst result	0%	0%	0%
Ranking by			
Our algorithm	**95%**	**99%**	**99%**
Utility	89%	93%	93%
Potential: utility - base search index	96%	98%	88%

Empirical application: Microsoft Partner Center

- Launched in 2017 to help match Microsoft partners among themselves to exchange goods and services
- The Partner Center generates roughly 1000 matches per day with over 8000 global partners

- Our data contains 7638 unique customer visits and 14066 clicks on partner profiles (searches)
  - Half of the customers chose at least one partner (i.e. submitting a referral)
    - 42% of the customers search before choosing
    - 7% of the customers directly chose partners
  - The other half of the customers did not choose any partner:
    - 36% of the customers left after some searching
    - 16% of the customers left without any search

Microsoft Partner Center

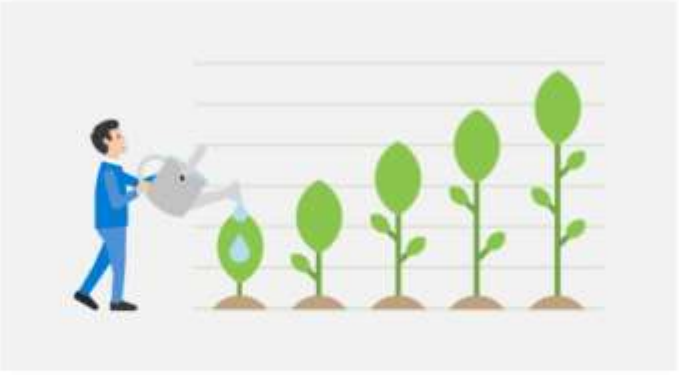


Connect with trusted Microsoft solution providers

Microsoft certified solution providers specialize in providing up-to-date Microsoft technology-based customer solutions all over the world.

Save time and money

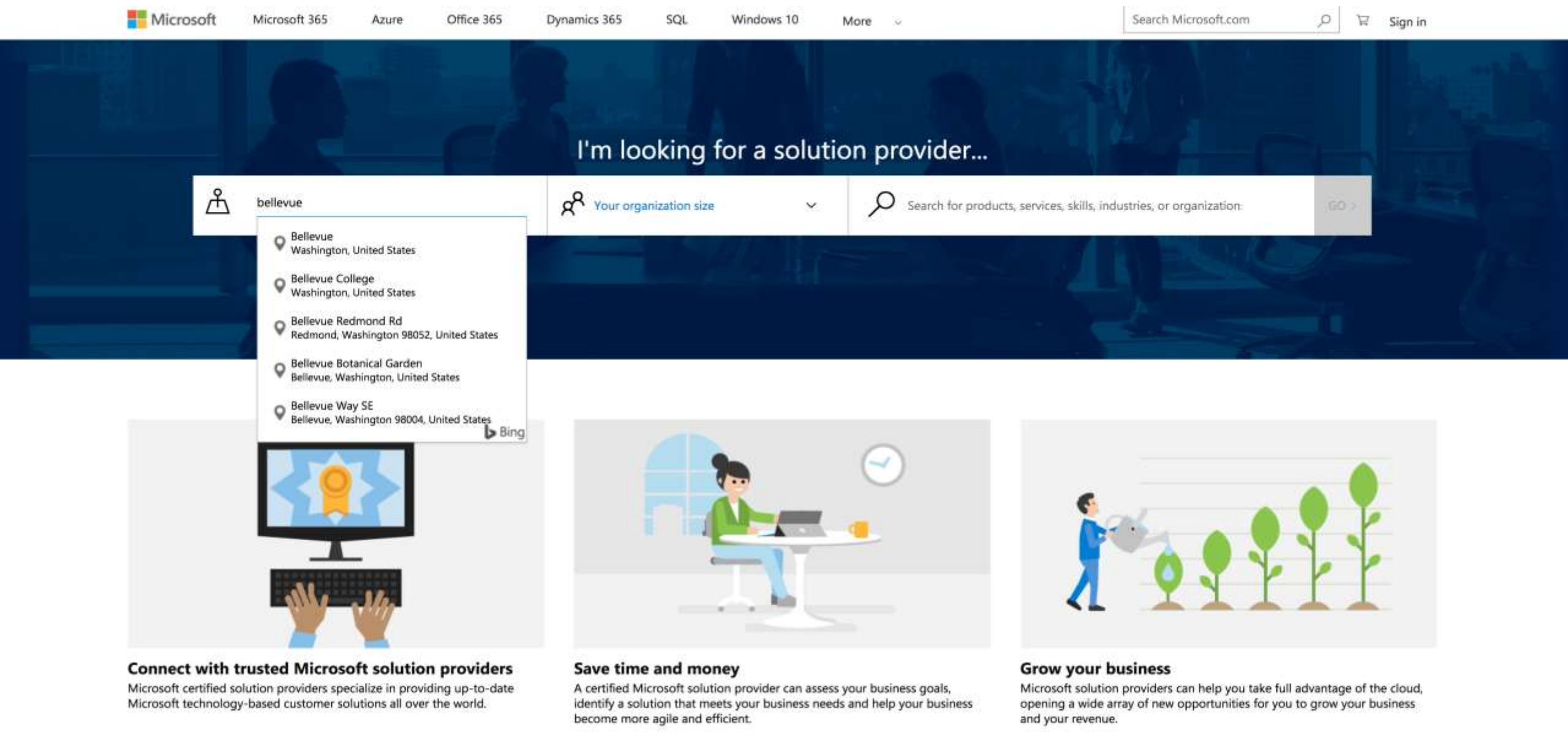
A certified Microsoft solution provider can assess your business goals, identify a solution that meets your business needs and help your business become more agile and efficient.



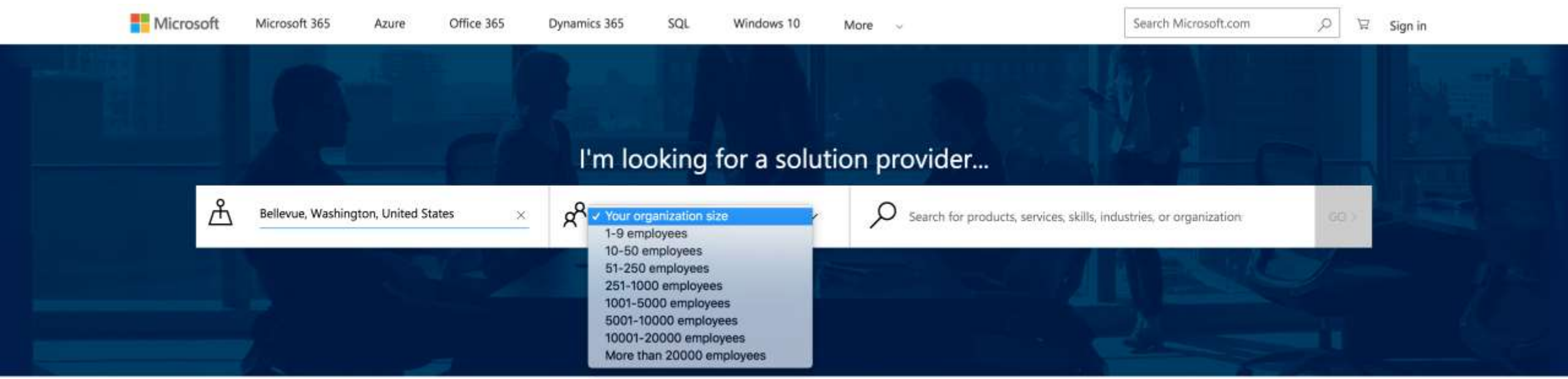
Grow your business

Microsoft solution providers can help you take full advantage of the cloud, opening a wide array of new opportunities for you to grow your business and your revenue.

Microsoft Partner Center



Microsoft Partner Center



Connect with trusted Microsoft solution providers

Microsoft certified solution providers specialize in providing up-to-date Microsoft technology-based customer solutions all over the world.

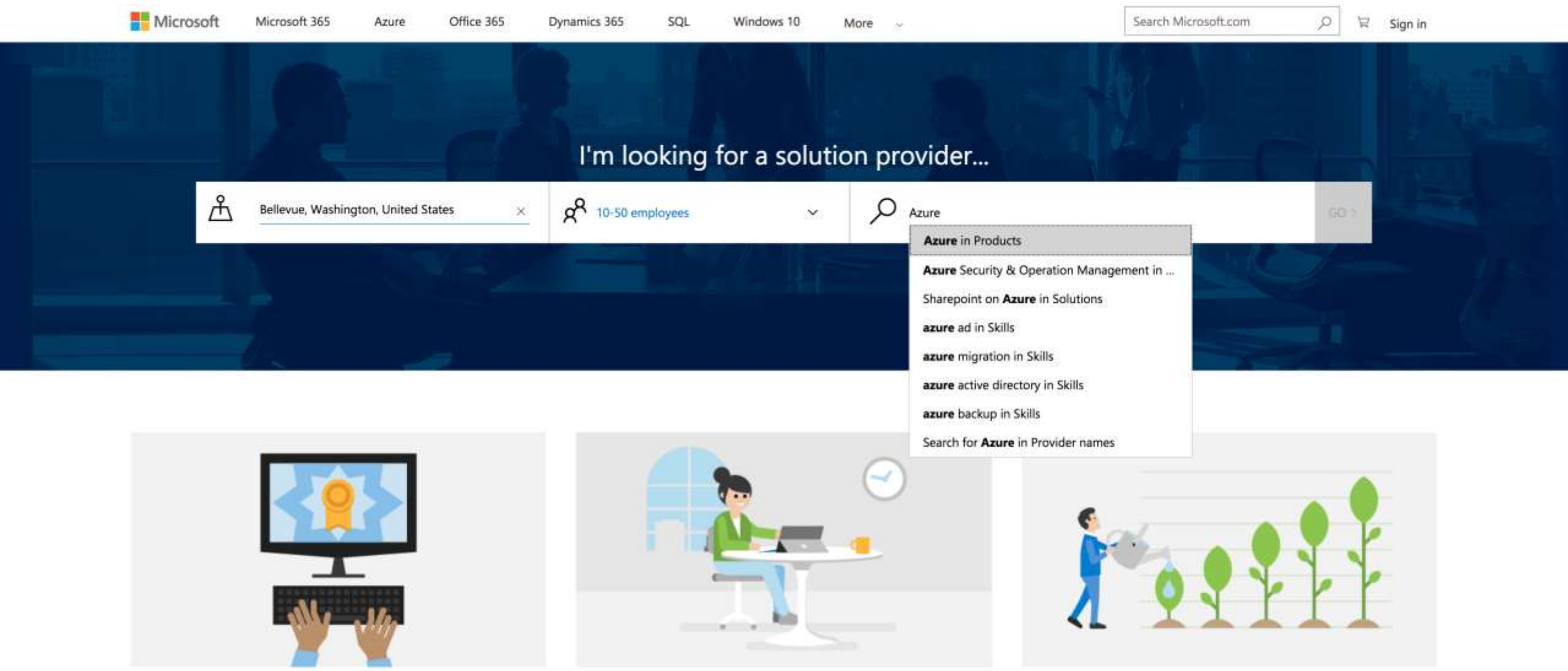
Save time and money

A certified Microsoft solution provider can assess your business goals, identify a solution that meets your business needs and help your business become more agile and efficient.

Grow your business

Microsoft solution providers can help you take full advantage of the cloud, opening a wide array of new opportunities for you to grow your business and your revenue.

Microsoft Partner Center



Microsoft Partner Center

Best match

Logicworks
Seattle, WA, US
SELECT PROVIDER +

CYBAGE SOFTWARE PRIVATE LIMITED
Redmond, WA, US
SELECT PROVIDER +

Softsource Consulting
Portland, OR, US
SELECT PROVIDER +

MAQ Software
Redmond, WA, US
SELECT PROVIDER +

FINE SOLUTIONS LLC
Lynnwood, WA, US
SELECT PROVIDER +

Aldridge
Seattle, WA, US
SELECT PROVIDER +

Most responsive

Databerry LLC.
Seattle, WA, US
SELECT PROVIDER +

Ataira Analytics Inc.
Redmond, WA, US
SELECT PROVIDER +

Actify Media
Helena, MT, US
SELECT PROVIDER +

APEX TECHNOLOGY MANAGEMENT INC
Redding, CA, US
SELECT PROVIDER +

IT GURUS OF ATLANTA LLC
Seattle, WA, US
SELECT PROVIDER +

Intrinium
Spokane, WA, US
SELECT PROVIDER +

Nearest to you

CompuCom Systems Inc.
Bellevue, WA, US
Within 1 mi from you

eMazzanti Technologies
Bellevue, WA, US
Within 1 mi from you

Timextender A/S
Bellevue, WA, US
Within 1 mi from you

VVL Systems & Consulting, LLC
Bellevue, WA, US
Within 1 mi from you

Northwest Cadence
Bellevue, WA, US
Within 1 mi from you

Accordo Group Limited
Bellevue, WA, US
Within 1 mi from you

Microsoft Partner Center

Logicworks

http://www.logicworks.com/ · · 500 Yale Avenue, North, Seattle, WA, US 98109

CONTACT > Search for other providers

This provider has demonstrated competency in the following areas

Gold	Cloud Platform
Silver	Cloud Platform

Microsoft endorses this provider for the following skills and capabilities

Azure, Enterprise Mobility + Security, Office 365, SQL, Windows, Consulting and Professional, Custom Solution, Deployment or Migration, Integration, Licensing, Managed Services (MSP), Project Services, Backup & Disaster Recovery, Cloud Database Migration, Cloud Migration, Containers, DevOps, Hybrid Storage, Networking, Serverless Computing, SQL Server Upgrade

About us

Logicworks is a Gold Cloud Platform Partner and Managed Services Partner that helps customers plan, build, and run Azure cloud infrastructure. For companies that are new to Azure, Logicworks can help develop a cloud strategy, conduct a TCO analysis, and architect a secure cloud environment. Companies with existing Azure environments can leverage Logicworks for 24x7 technical support and access to a proprietary suite of monitoring and automation tools.

More

Expertise

- Azure
  - Azure Backup
  - Azure Migration
  - Azure MSP
- Backup & Disaster Recovery
- Cloud Database Migration
- Cloud Management
- Cloud Migration
- Cloud Security
- Communications & Media
- Consulting and Professional
- Containers
- CSP

Top locations

155 Avenue of the Americas, Fifth Floor, New York, NY, US 10013

1550 Wewatta Street, Denver, CO, US 80202

500 Yale Avenue, North, Seattle, WA, US 98109

2401 Walnut Street, Suite 102, Philadelphia, PA, US 19103

745 Atlantic Avenue, Boston, MA, US 02111

Empirical specification

- We formulate the double index model into a discrete choice problem

- Define a consumer i's utility and search index for product j as

$$u_{ij} = x_j\beta + \omega_j\gamma + \xi_j + v_j + \epsilon_{ij}^U$$

$$s_{ij} = x_j\beta^S + position_j\gamma^S + \xi_j^S + \epsilon_{ij}^S$$

where

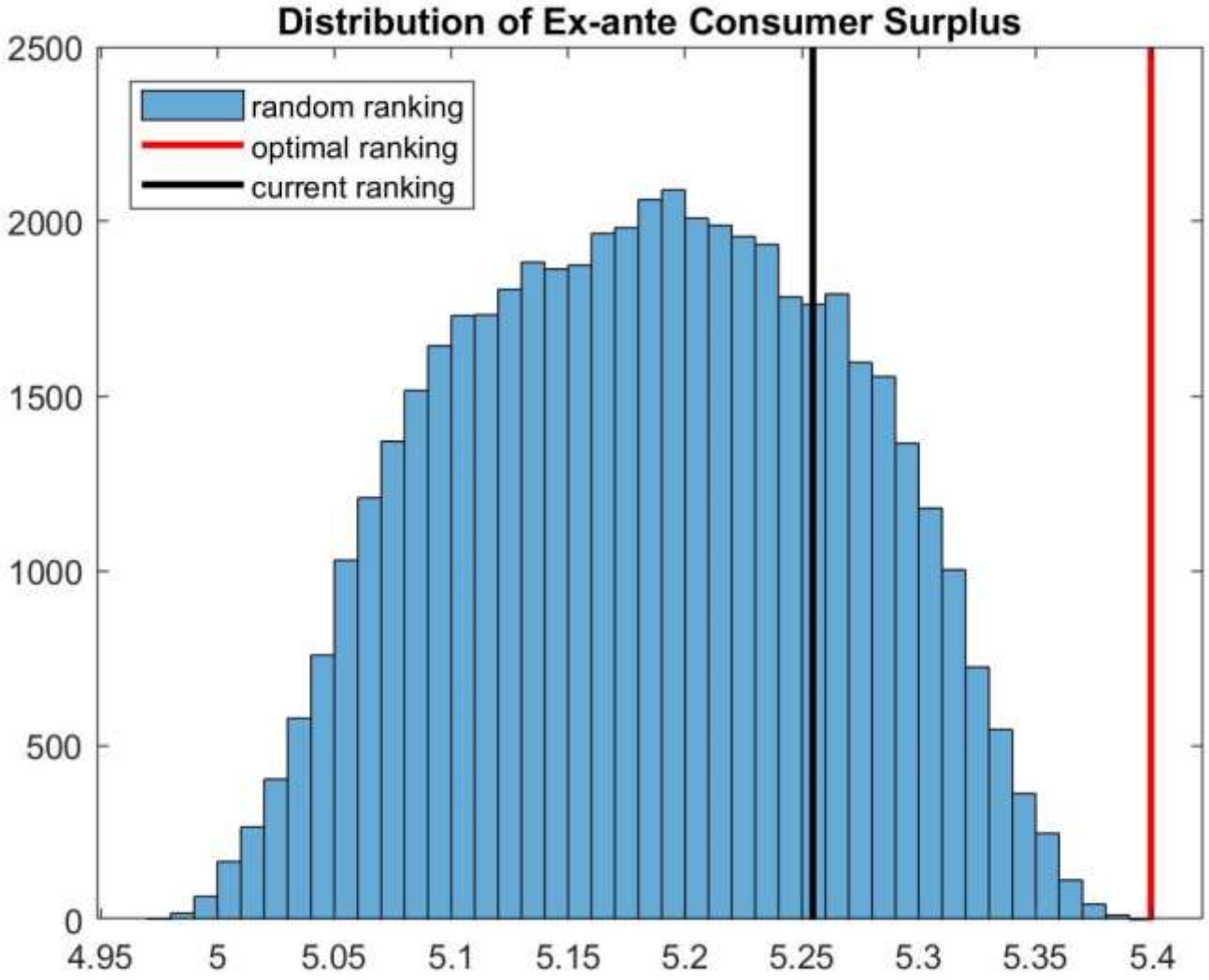
- x_j's are observed characteristics *before* search, e.g. endorsement badge, LinkedIn logo
- ω_j's are observed characteristics *after* search, e.g. endorsement details, partner attributes
- $position_j$ include both the vertical groups and horizontal rankings of product j

- ξ_j's are unobserved characteristics *before* search
- v_j's are unobserved characteristics *after* search
- ξ_j^S's are unobserved search index
- ϵ's are *i.i.d.* distributed Type I EV

Estimates

	Search score			Utility	
	Coef	SE		Coef	SE
Group 1	0.5344***	0.0995	Endorsement: Gold	-1.0515	0.9073
Group 2	-1.2401***	0.1086	Endorsement: Silver	0.5614	0.7568
Rank 1	3.5620***	0.1370	LinkedIn	-0.4866	0.6154
Rank 2	2.4517***	0.1296	length of description	0.3476	2.9137
Rank 3	2.1400***	0.1346	num of locations	0.0788	0.0013
Rank 4	1.7604***	0.1448	Product Endorsement	0.2603	0.6794
Rank 5	1.6875***	0.2024	Service Endorsement	0.4788	0.9005
Rank 6	1.6440***	0.1489	Solution Endorsement	-0.5849	1.4152
Rank 7	0.5501*	0.1904	Specialized in Product	10.0676	1.1063
Rank 8	0.1033	0.1871	Specialized in Service	-0.0873	1.1477
Rank 9	-0.0128	0.2147	Specialized in Solution	0.7035	1.2979
Rank 10	0.1257	0.2073	Industry Focus	-11.8794***	3.8621
Rank 11	-0.1085	0.2340	Expectation Constant	-39.0116***	0.4238
Rank 12	0.3670	0.1860	Constant	40.8532***	0.9427
Endorsement: Gold	-0.0022	0.1328			
Endorsement: Silver	-0.1431	0.1493			
LinkedIn	-0.1236*	0.1044			
Search Constant	-6.5377***	0.1357			

Counterfactual simulations

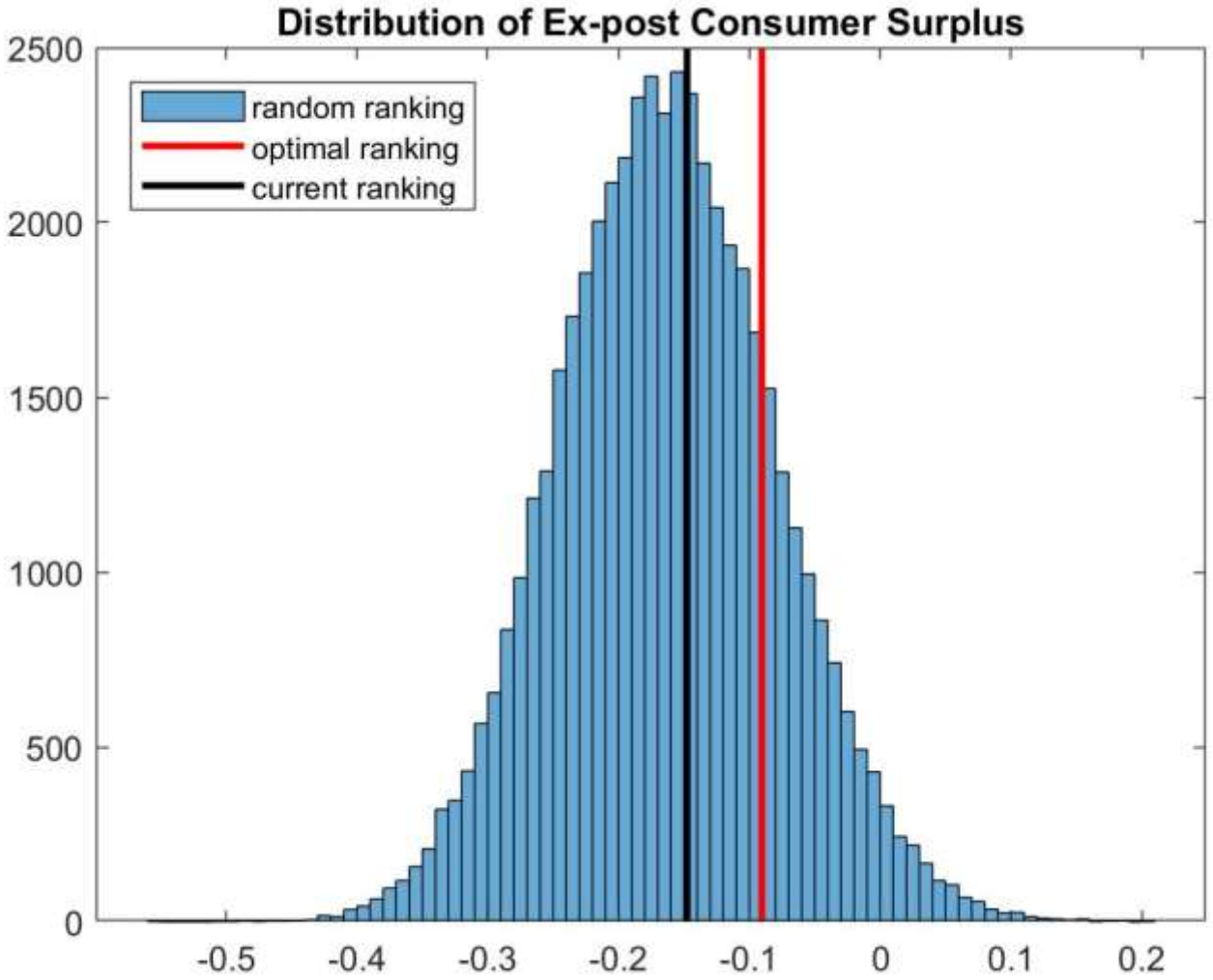


Conclusions

- We propose a double index framework that subsumes several search models
- We propose a rank optimization algorithm to maximize “ex ante” consumer surplus
- We study a real world business platform and improve its ranking algorithm

APPENDIX

“Ex post” consumer surplus



Sequential search with recall

- Following Weitzman (1979), each product delivers an ex-ante unknown payoff $u_j \sim F_j^U$.
- Payoff can be learned by paying a cost c_j

$$c_j = \int_{s_j}^{\infty} (u - s_j) dF_j^U(u)$$

- Let S denotes the set for unsearched products and S^c as those already searched.
- Define $s^* = \max_{\{j \in S\}} s_j$ and $v^* = \max_{j \in S^c} u_j$.
  - If $s^* > v^*$, the consumer keeps on searching and searches the option with search score s^* first.
  - If $s^* \leq v^*$, the consumer stops searching and purchases the option with utility v^*

Simultaneous search

- Ex-ante identical goods with payoffs drawn from F_U
- F_U supported on $[\underline{u}, \overline{u}]$
- Searching n goods entails a cost $c(n)$
- Consumers optimally choose how many goods to sample, n^*
  - $\uparrow n \Rightarrow \uparrow$ search cost
  - $\uparrow n \Rightarrow$ better u draw in expectation
- This strategy is subsumed in the double index model by setting $s_j = \overline{u}$ for n^* randomly chosen goods and $s_j = \underline{u}$ for the rest

Satisficing

- First introduced by Simon (1955)
- Ex-ante identical goods
- Consumers search sequentially until a level of utility s^* is reached
- This strategy is subsumed in the double index model by setting $s_j = s^*$ for all j