Estimating models of plant location choice

Keith Head
UBC

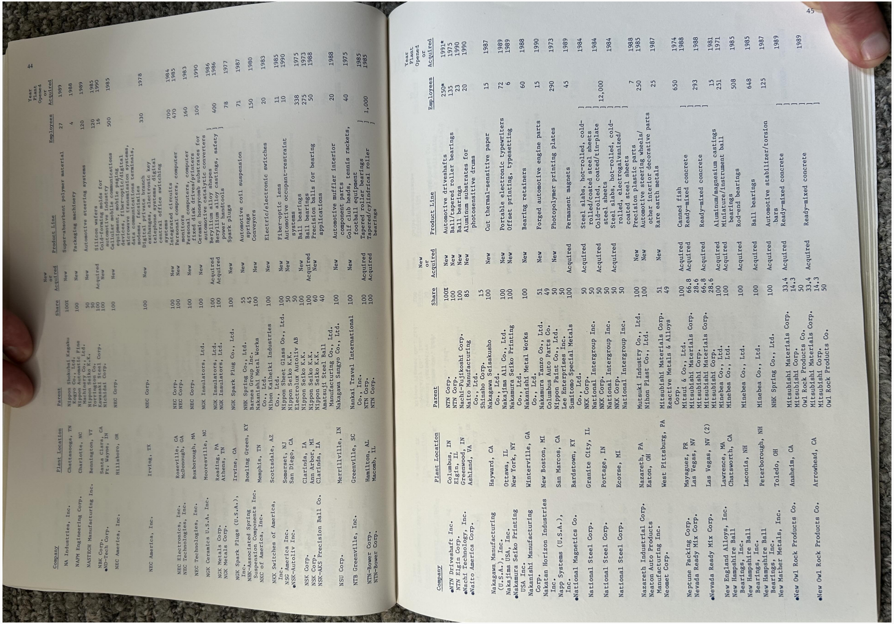
Thierry Mayer
Sciences Po

Marc Melitz
Harvard

Chenying Yang
Singapore Management U.

UEA Montreal
October 4, 2025

1992: My starting point for location choice estimation



Zooming in on an example

Plant Location	Parent	Share	New or Acquired	Product Line	Employees	Opened or Acquired
Chattanooga, TN	Nippon Shokubai Kagaku Kogyo Co., Ltd.	100%	New	Super-absorbent polymer material	27	1989
Charlotte, NC	Nippon Automatic Fine Machinery Co., Ltd.	100	New	Packaging machinery	4	1988
Bennington, VT	Nippon Seiko K.K.	50	New	Automotive steering systems	120	1989
	Torrington Co.	50				
		100	Acquired	Silicon wafers	120	1985
	Kawasaki Steel Corp.				16	1990

Plant Location	Parent	Share	New or Acquired	Product Line	Employees	Opened or Acquired
San Diego, CA	Nippon Seiko K.K.	50	New	Automotive [illegible] systems		
	Electrolux Autoliv AB	50				
Clarinda, IA	Nippon Seiko K.K.	100	New	Ball bearings	338	1975
Ann Arbor, MI	Nippon Seiko K.K.	100	Acquired	Ball bearings	275	1973
Clarinda, IA	Nippon Seiko K.K.	60	New	Precision balls for bearing applications	50	1988
	Amatsuji Steel Ball Manufacturing Co., Ltd.	40				
				Automotive muffler interior	20	1988

Year	Product	Location
1973	Ball bearings	Ann Arbor, MI
1975	Ball bearings	Clarinda, IA
1988	Precision balls	Clarinda, IA
1989	Steering systems	Bennington, VT

Motivation for estimating location choice models

Government policies aim to shape location decisions

Does it work? We need estimated models of facility location choices to conduct counterfactuals.

1. China's special economic zones
2. "Million dollar plants"
3. 2022 Inflation Reduction Act (industrial policy for green energy)

China's FDI special economic zones

- Starting in 1980, China designated a series of incentive zones to attract foreign investment
- Reductions in income taxes, tariffs on imported inputs, land user fees.
- Other privileges include ease of establishment, autonomy of operations, preferential access to inputs, streamlined bureaucracy.
- Head and Ries (1996, JUE), estimated that these zones increased investments by 30%, comprising a static increase of 13%, and a 17% additional gain via agglomeration.
- Possibly an underestimate.

Shenzhen (the first SEZ), then and now

pop. 30,000

pop. 18mn (2024)

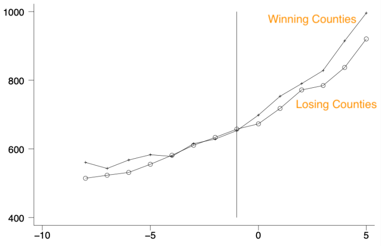
Bidding for Industrial Plants: Does Winning a 'Million Dollar Plant' Increase Welfare?

Michael Greenstone & **Enrico Moretti**

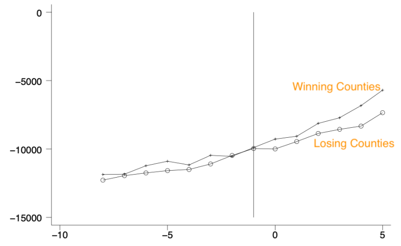
WORKING PAPER 9844 **DOI** 10.3386/w9844 **ISSUE DATE** July 2003

Increasingly, local governments compete by offering substantial subsidies to industrial plants to locate within their jurisdictions. This paper uses a novel research design to examine the consequences of successfully bidding for a plant on county-level labor earnings, property values, and public finances. Each issue of the corporate real estate journal Site Selection includes an article titled The Million Dollar Plant that describes how a large plant decided where to locate. These articles report the county where the plant chose to locate (i.e., the 'winner'), as well as the one or two runner-up counties (i.e., the 'losers'). The losers are counties that have survived a long selection process, but narrowly lost the competition. We use these revealed rankings of profit-maximizing firms to form a counterfactual for what would have happened in the winner counties in the absence of the plant opening. We find that a plant opening is associated with a 1.5% trend break in labor earnings in the new plant's industry in winning counties (relative to losing ones) after the opening of the plant (relative to the period before the opening). Property values may provide a summary measure of the net change in welfare, because the costs and benefits of attracting a plant should be capitalized into the price of land. We find a positive, relative trend break of 1.1% in property values. Further, we fail to find any deterioration in local governments' financial position. Overall, the results undermine the popular view that the provision of local subsidies to attract large industrial plants reduces local residents' welfare.

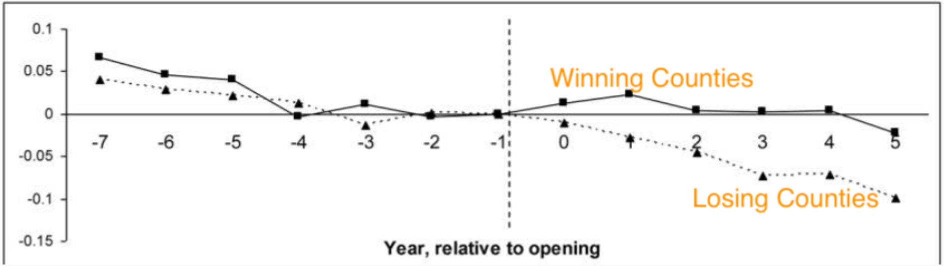
Winning counties trend better than losers



Wagebill



Property Values



Total factor productivity (2010, JPE)

Policies promoting EV production in the US (2022–2024)

- Consumer subsidy for passenger vehicles up to \$7,500 ($\approx$ 15%–20%)
    1. no production contingencies (before IRA, most of EU, Canada)
    2. require vehicle **assembly** in North America (after IRA)
    3. require battery **cells** manufactured in North America (by full IRA)
- Tax credit for US battery production: 20–30% of battery cost
- Tariffs on China-made EVs (25% $\rightarrow$ 100%)
- Dept. of Energy loans (e.g. \$6.6bn for Rivian in Georgia)
- 13 states have zero-emission vehicle (ZEV) mandates

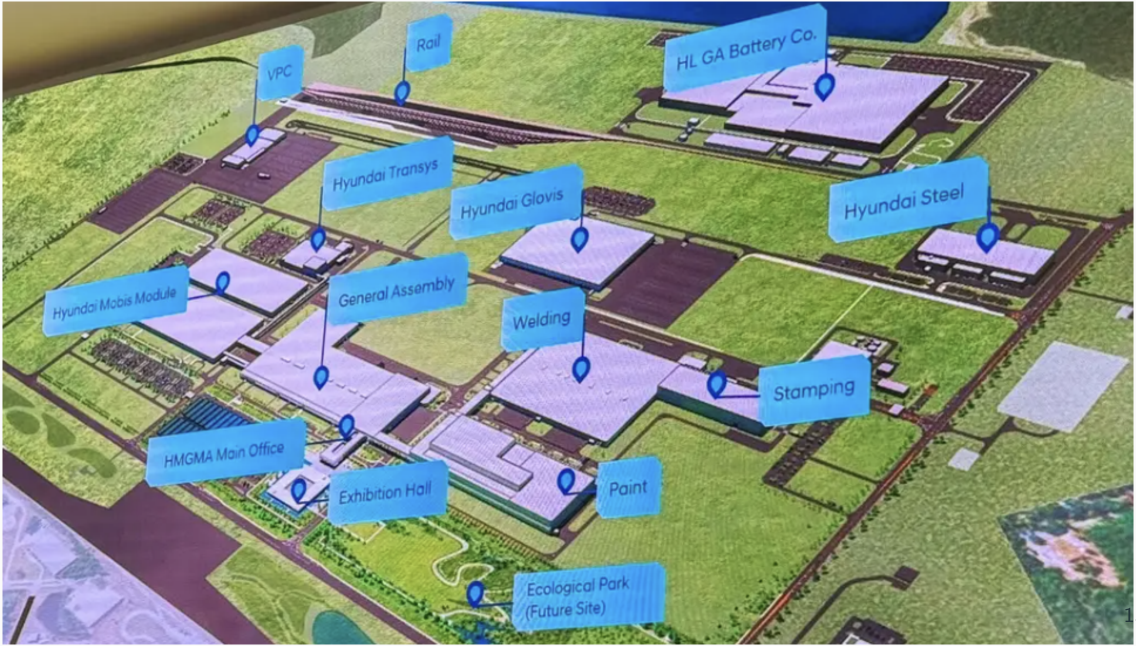
Top 10 models excluded by Rule 1 from the IRA buyer credit

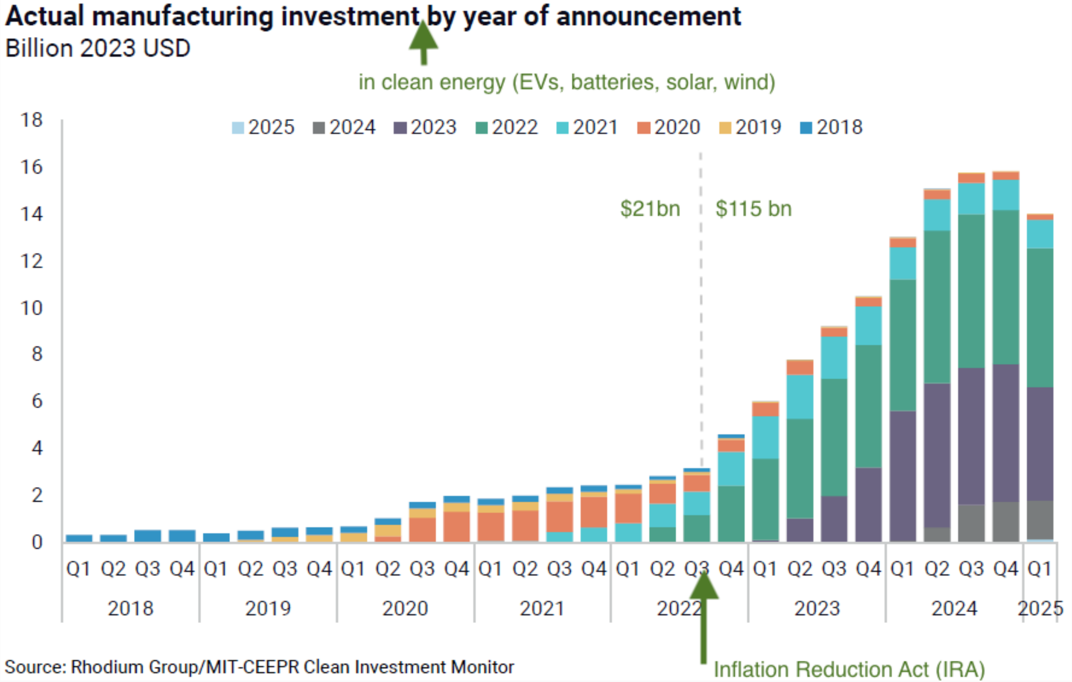
Rank	Brand	Model	2022 Sales (US)	Assembly location announcements
1	Hyundai	Ioniq 5	23741	Adds from Korea to Savannah, GA
2	Kia	EV6	21978	Adds EV9 from Korea to West Point, GA
3	BMW	i4	11462	Stays in Germany
4	Polestar	2	8758	Polestar 3 to Ridgeville, SC
5	Porsche	Taycan	8425	Stays in Germany
6	BMW	iX	7394	Stays in Germany
7	Kia	Niro	7262	Stays in Korea
8	Audi	e-tron	7233	Stays in Belgium
9	Hyundai	Kona	4719	Stays in Korea
10	Volvo	C40	4693	Stays in Belgium

Hyundai Ioniq Savannah, GA plant (December 2022)

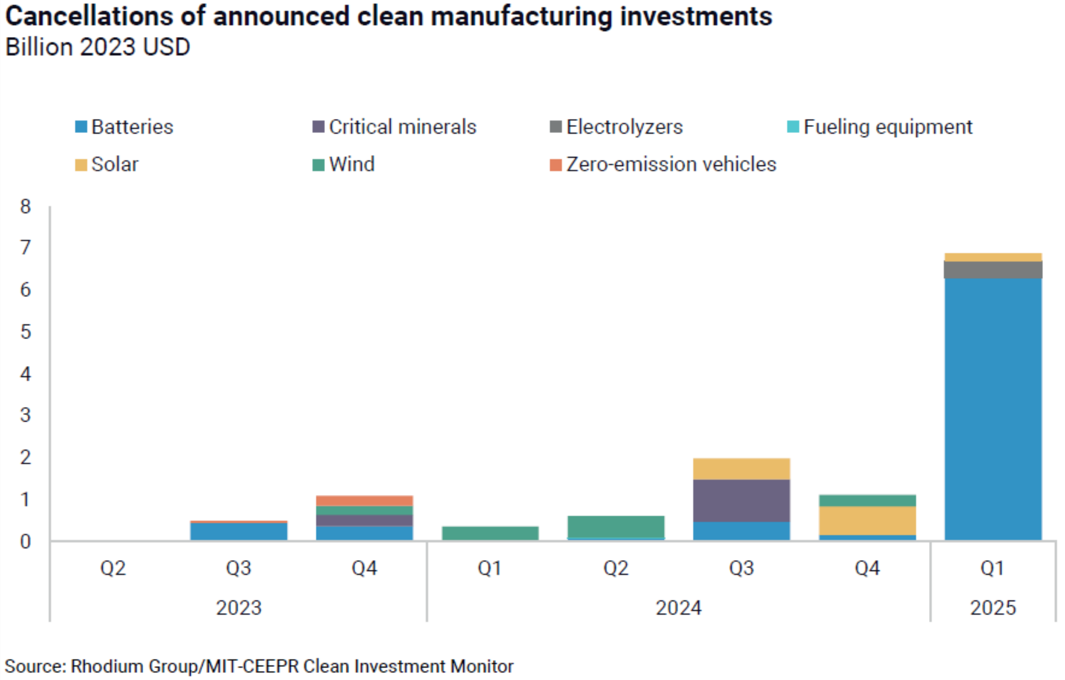
Hyundai Ioniq near Savannah, GA plant (October 2024)

Hyundai Metaplant has multiple production stages





Actual manufacturing investment by year of announcement
Billion 2023 USD
in clean energy (EVs, batteries, solar, wind)
2025 2024 2023 2022 2021 2020 2019 2018
18
16
14
12
10
8
6
4
2
0
$21bn
$115 bn
Q1 Q2 Q3 Q4 Q1 Q2 Q3 Q4 Q1 Q2 Q3 Q4 Q1 Q2 Q3 Q4 Q1 Q2 Q3 Q4 Q1 Q2 Q3 Q4 Q1 Q2 Q3 Q4 Q1
2018 2019 2020 2021 2022 2023 2024 2025
Source: Rhodium Group/MIT-CEEPR Clean Investment Monitor
Inflation Reduction Act (IRA)


Cancellations of announced clean manufacturing investments
Billion 2023 USD
Batteries
Critical minerals
Electrolyzers
Fueling equipment
Solar
Wind
Zero-emission vehicles
8
7
6
5
4
3
2
1
0
Q2
Q3
Q4
Q1
Q2
Q3
Q4
Q1
2023
2024
2025
Source: Rhodium Group/MIT-CEEPR Clean Investment Monitor

A brief history of location choice estimation

An abbreviated—and very selective—review of methods

Models of static, granular plant location choice:

1. Multinomial choice (conditional logit)
2. Binary choice (logit/probit)
3. Combinatorial discrete choice for interdependent decisions.
   (a) Jia–AFT-AES, solve pb via super/sub-modularity
   (b) Multi-stage and multi-product problems solved by MILP
   (c) SMM estimation

Conditional logit

- When ?
  (a) Each investment i is an independent multinomial choice, $y_{i\ell}$, between L_i options. $y_{i\ell^*} = 1$ and $y_{i\ell} = 0$ for all $\ell \neq \ell^*$
  (b) With single sourcing, conditional on "activated" supplier facilities, input sourcing is also a conditional logit problem.
- How?
  (a) Estimated via "clogit" in Stata
  (b) Guimarães, Figueirdo, Woodward (2003, REStat), pointed out that likelihood of Poisson is the same as clogit, except for a constant.
  (c) With chooser fixed effects, we can estimate the multinomial choice model in R using fixest::fepois (very fast)
- Issues:
  1. The smaller problem: non-independence of choices (IIA) $\rightarrow$ Nested logit
  2. The bigger problem: interdependence (substitutes and complements)

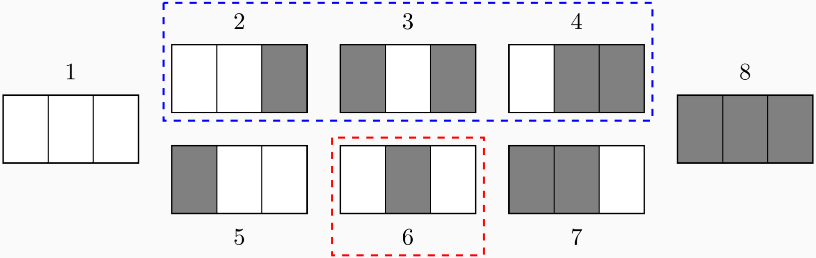
Binary choice

- When?
  (a) Each investment i is an independent binary choice, $y_{i\ell}$ is evaluated for each of the L options, multiple 1s are allowed.
  (b) This approach makes sense for investments serving autarkic markets, non-traded services, such as retailers...
  (c) ...assuming no chain effects
- How?
  (a) Head and Mayer 2019 used binary choice for decision of each car brand to open dealership network in a country (No Renaults in Canada)
- Issues:
  1. The smaller problem: if many FEs, incidental parameter bias (BIFE)
  2. The bigger problem: interdependence (substitutes and complements)

Combinatorial discrete choice: Supermodularity

- Multinomial / binomial choice: evaluate profits L times for each i.
- Combinatorial problem: evaluate 2^L possible configurations (a lot).
- Jia (2008) solves the interdependence problem by adapting Tarski's (1955) fixed point theorem and Topkis's (1978) monotonicity theorem
- Her method allows her to solve the problem of Wal-mart choosing store locations in the presence of "chain effects": Nearby stores split the costs of operation, delivery, and advertising to achieve scale economies.
- The complementarity between nearby stores leads to super-modular profits.
- Supermodularity implies far fewer than 2^L evaluations.

Squeezing: How supermodularity can reduce evaluations



1. Configurations $i = 1 \ldots 8$ of $[y_i^W, y_i^C, y_i^E]$
2. Evaluate 7 and 8: If $\pi_8 < \pi_7$ then position $y_i^E = 0$ for all i.
3. No need to evaluate $i = 2, 3, 4$ (blue dashed rectangle)
4. Evaluate 1 and 5: If $\pi_1 < \pi_5$ then position $y_i^W = 1$ for all i.
5. $i = 6$ is out; select larger of π_7 and π_5 (evals $= 4 \ll 8 =$ BF)

Combinatorial discrete choice: Submodularity

- The one-stage facility location problem is one in which a plant in one location substitutes for plants in nearby locations, because it offers an alternative way to serve the same markets.
- When both fixed costs and transport costs are important, we have the concentration vs proximity tradeoff.
- Arkolakis, Eckert, and Shi develop an algorithm that serves submodular profit functions as well as supermodular ones.
- As with Jia's algo, the AES squeezing algo dramatically cuts the number of configurations to evaluate.

UFLP is a limit case of AES, allows comparison to MILP

- AES cost function for variety ω (general case)

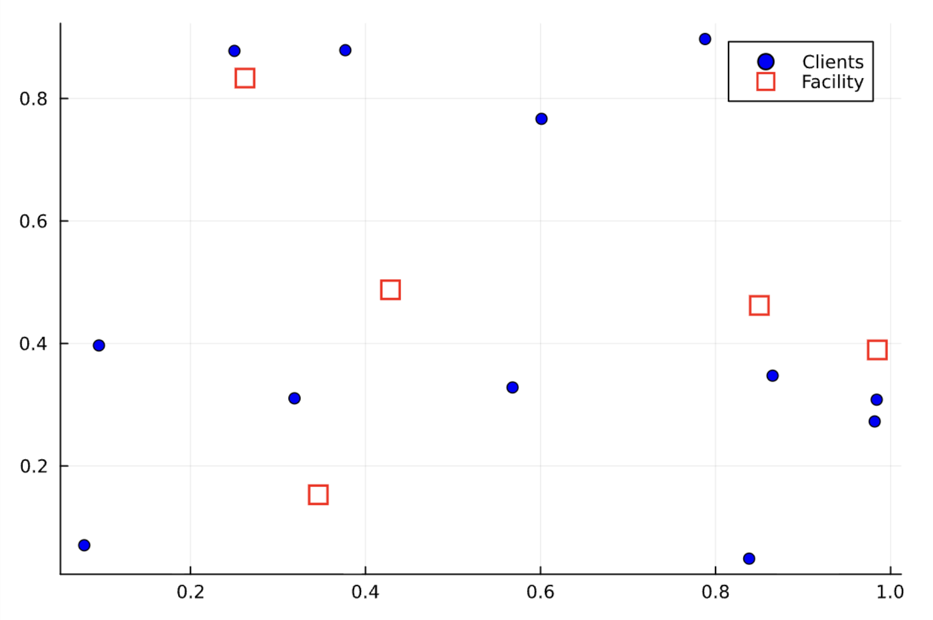
$$c_n(\mathcal{L},\omega) = \left[\sum_{\ell\in\mathcal{L}} c_{\ell n}(\omega)^{-\theta}\right]^{-1/\theta}, \quad \text{where} \quad c_{\ell n} = w_\ell \tau_{\ell n}(\omega)/z_\ell(\omega).$$

- Limit case as $\theta \to \infty$: $\quad c_n(\mathcal{L},\omega) = \min_{\mathcal{L}} c_{\ell n}(\omega).$
- Perfect substitution implies single-sourcing (least cost plant)
- Plug minimized cost into the profit function (taking P_n as given).

$$\pi(c_{\ell n}) \propto c_\ell^{1-\sigma}$$

- Firm chooses $\mathcal{L}$ to maximize the sum of all destination-specific profit
- Randomness: cost shocks (w_ℓ) and locations on grid ($\tau_{\ell n}$)
- Simulate L potential locations, N consumer markets

Facility location problem: setup



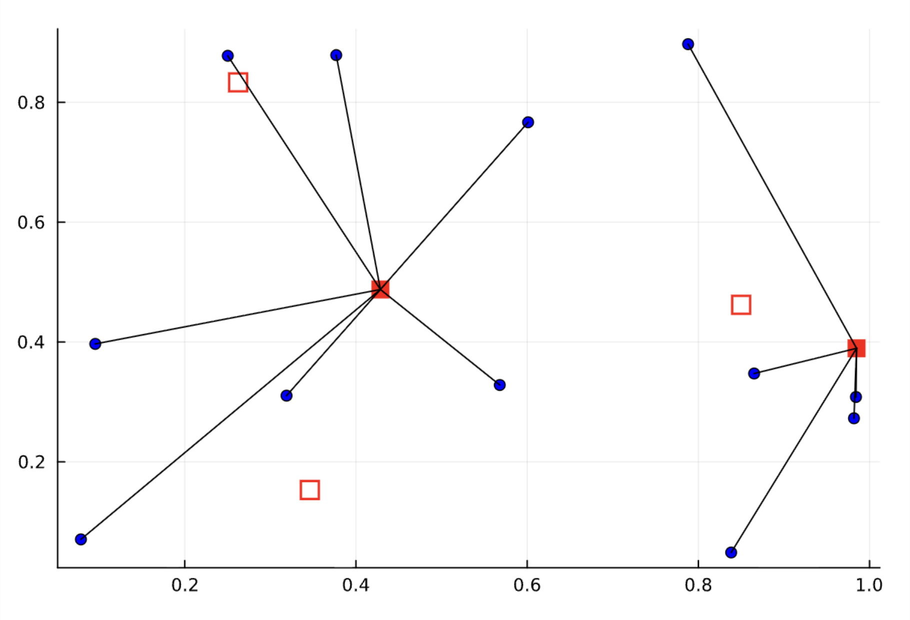
Single stage FLP can also be formulated as a MILP

$$\max \sum_{n \in N} \sum_{\ell \in L} \pi(c_{\ell n}) x_{\ell n} - \sum_{\ell \in L} \phi_\ell y_\ell \quad \text{subject to}$$

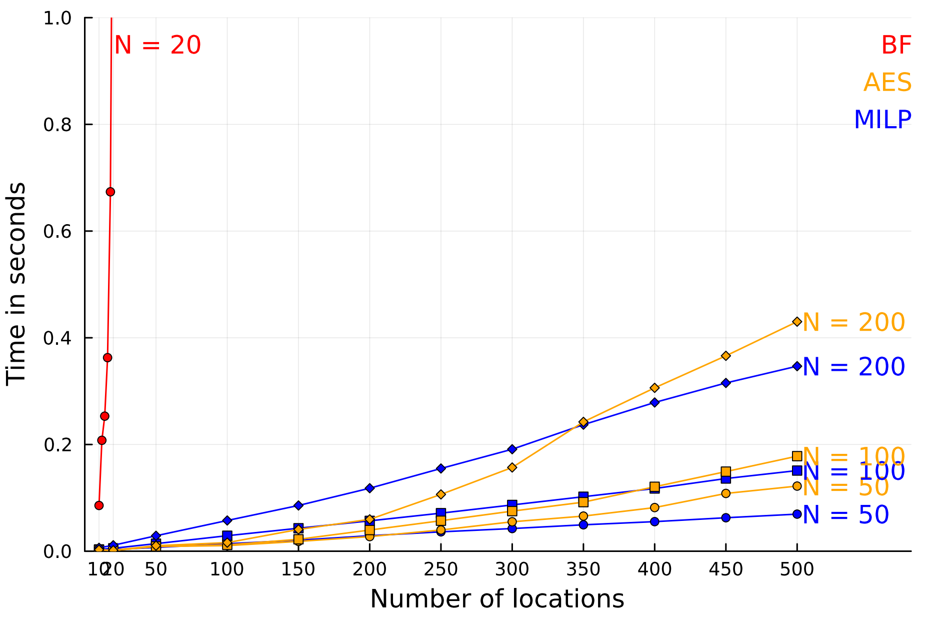
$$\begin{aligned} \sum_{\ell \in L} x_{\ell n} &= 1, \quad n \in N \\ x_{\ell n} &\leq y_\ell, \quad n \in N, \ell \in L \\ x_{\ell n} &\geq 0, \quad n \in N, \ell \in L \\ y_\ell &\in \{0, 1\}, \quad \ell \in L \end{aligned}$$

- Mixed integer linear programming (MILP) problem
- Consumers, production sites located randomly on the unit square
- Costs depend on distance from ℓ to n and a random ℓn shock

Example of solved facility location problem



Comparing AES, MILP, and brute force

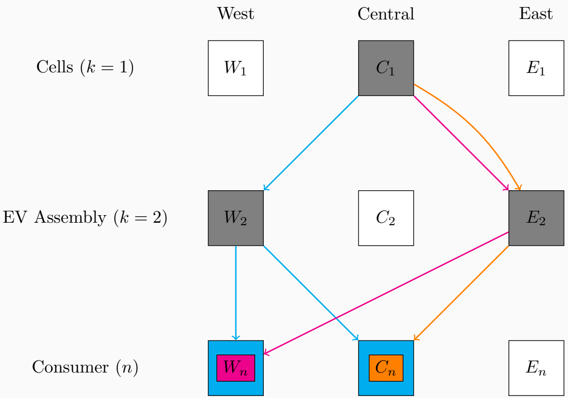


Multi-stage production calls for MILP

- One mechanism for submodularity: plants at same stage k are substitutes
- Two mechanisms for supermodularity:
    1. plants at different stages $k \leq K$ complement each other
    2. distribution $(K+1)$ facilities complement production plants

  $\Rightarrow$ Parametric restrictions to ensure global sub- or super- may not exist for our model and we prefer not to rely on them.
- Multi-stage profit maximization problems are easy to formulate as MILP.

Optimization over Paths with Constraints



Firm chooses

- models m to sell in n s.t. market entry cost $\phi_{nm} \rightarrow z_{nm} = \{0, 1\}$,
- plants to open at each k s.t. plant fixed cost $\phi_{\ell k} \longrightarrow y_{\ell k} = \{0, 1\}$,
- the **optimal path** ℓ_{nm} $\forall n, m$ through open facilities

The cyan, magenta, orange lines: paths chosen by 3 car models

Gray squares are **activated** facilities. Colour squares are entered markets.

From Paths to Profits

- Key requirement for integer programming to work: variable profit can be written as a function of paths and a market aggregator
- Delivered MC $c(\ell_{nm})$ depends on the path chosen ℓ_{nm} and variable cost parameters (to be estimated)
- Quantity demanded, q, is determined by firm's $c(\ell_{mn})$, its appeal ξ_{mn}, market size, and an index of costs and appeal of competitors
- Variable profits for tuple (m, n) if path ℓ_{mn} is chosen: $\pi(c(\ell_{mn}), A_n)$
- A_n is a function of path costs of all models in market n, including gas vehicles.

MMM UFLP: objective, variables (x, y, z), & constraints

$$\max_{\mathbf{x},\mathbf{y},\mathbf{z}} \quad \sum_{m\in M_f}\sum_{n\in N}\sum_{\ell_1\in L_1}\sum_{\ell_2\in L_2} \pi(c(\boldsymbol{\ell}_{mn}), A_n) x_{mn\ell_1\ell_2}$$
$$- \sum_{g_1\in G_1}\sum_{\ell_1\in L_1} \phi_{fg_1\ell_1} y_{fg_1\ell_1} - \sum_{g_2\in G_2}\sum_{\ell_2\in L_2} \phi_{fg_2\ell_2} y_{fg_2\ell_2} - \sum_{m\in M_f}\sum_{n} \phi_{mn} z_{mn}$$

subject to

$$\sum_{\ell_1\in L_1}\sum_{\ell_2\in L_2} x_{mn\ell_1\ell_2} \leq z_{mn}, \quad n \in N, m \in M_f \tag{1}$$

$$\sum_{\ell_1\in L_1} x_{mn\ell_1\ell_2} \leq y_{fg_2(m)\ell_2}, \quad n \in N, m \in \mathcal{M}_f, \ell_2 \in L_2 \tag{2}$$

$$\sum_{\ell_2\in L_2} x_{mn\ell_1\ell_2} \leq y_{fg_1(m)\ell_1}, \quad n \in N, m \in \mathcal{M}_f, \ell_1 \in L_1 \tag{3}$$

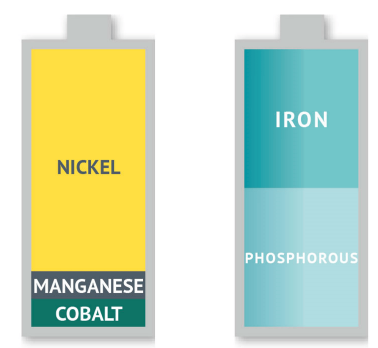
$$x_{mn\ell_1\ell_2} \geq 0, \quad y_{fg_1\ell_1} \in \{0,1\}, \quad y_{fg_2\ell_2} \in \{0,1\} \quad z_{mn} \in \{0,1\}. \tag{4}$$

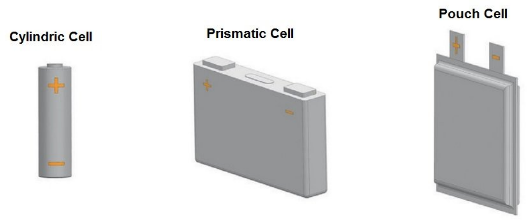
What does activate ($y_{fg_k\ell_k} = 1$) mean in the BEV industry?

- A plant is active if it has paid the indivisible fixed cost required to produce positive amounts for a client f (car maker)
- Paid for each group g_k of potential outputs
  - For cells, $g_1(m)$ maps models to a combination of cell material categories and shapes, e.g. Tesla Model S uses "Nickel-Cobalt-Manganese/Cylinder"
  - For vehicles, $g_2(m)$ gives the platform, e.g. "GEN III" for Tesla Model 3 and Model Y, "MEB" for VWs such as ID.4, ID.3, Audi Q4, and Skoda Enyaq.

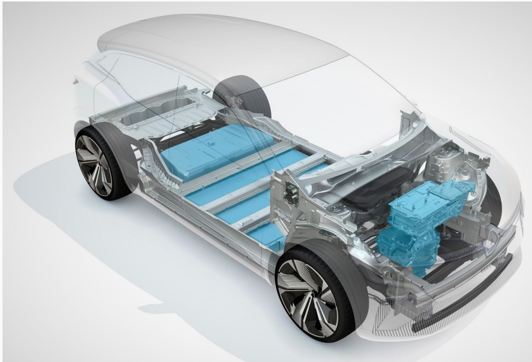
Fixed costs pertain to location-firm-groups

Battery cells grouped by
material category ↓ & shape →



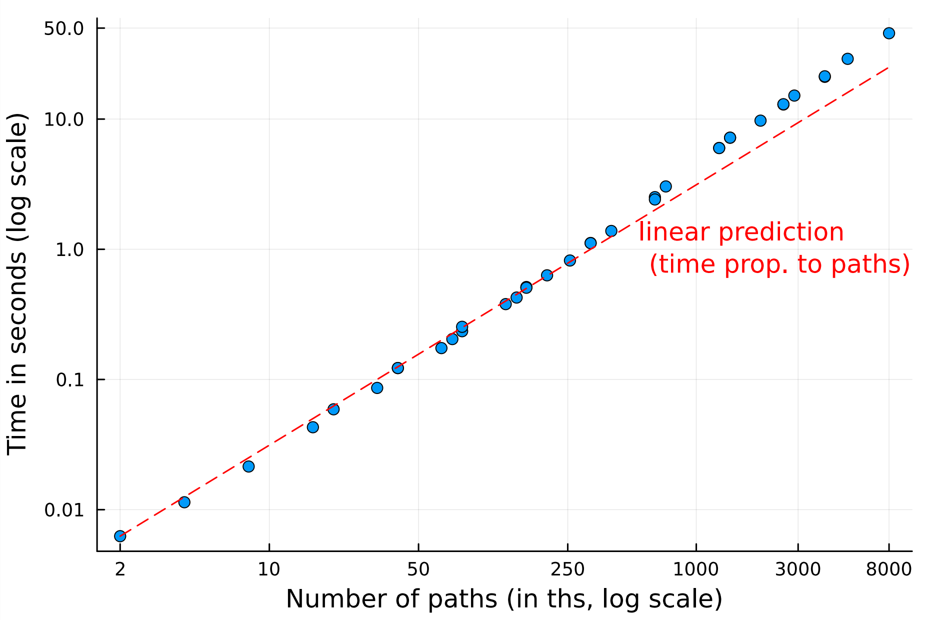


Vehicle assembly grouped by platform ↓

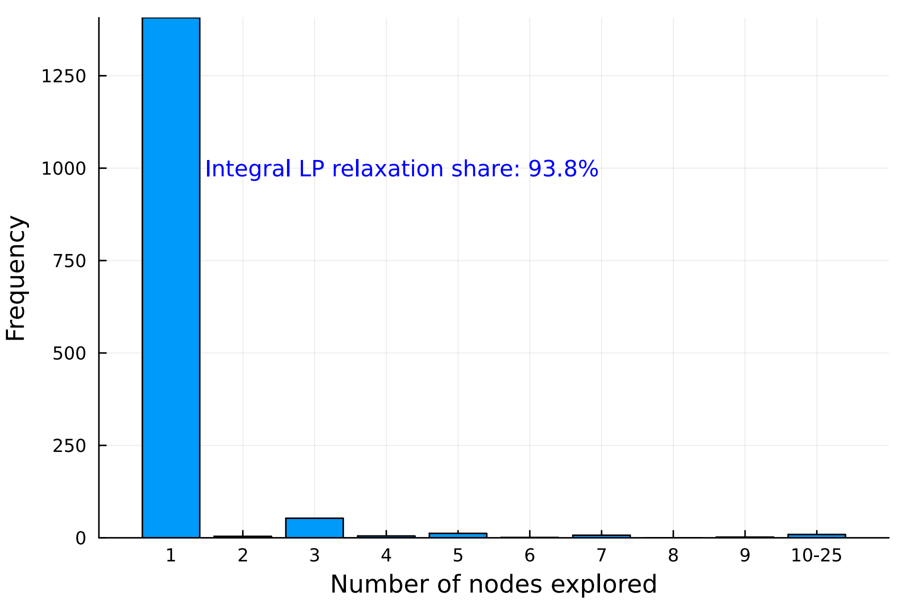


$g_1(m)$ assigns models to cell groups, $g_2(m)$ assigns models to assembly groups

Computation time $\approx$ linear in number of paths ($L_1 \times L_2 \times N \times M$)



High success rate of linear relaxation



What range of problems does our method handle?

1. No restrictions on complementarity or substitution between facilities (super vs sub modularity)
2. Firms with multiple products, grouped by characteristics
3. Endogenous market entry to multiple markets
4. Market interdependence between firms (via price index)
5. Because our focus is the GVC supply side, we simplify other aspects via CES demand, constant markups, efficient bargaining across firms.
6. The method permits many generalizations (e.g. more stages, multiple inputs at a given stage)

Necessary industry characteristics for this framework

1. Inputs from different plants are perfect substitutes (no love of variety) if all dimensions of the product are specified $\Longrightarrow$ disaggregated sourcing data needed (ours comes from IHS Markit)
2. Constant marginal costs; plants are "uncapacitated": no long run capacity constraints
3. Together 1+2 $\Longrightarrow$ single sourcing from the least-delivered-cost plant

Estimation of the model applied to BEV value chain

Six EV features that fit with the MUFLP framework

1. Cells account for over 1/4 of final cost (3/4 of battery cost, which are 1/3 of vehicle)
2. Cells production is usually close (median distance of about 500km) to vehicle assembly
3. Cell factory investment cost are huge: avg. $2.5bn (based on 83 news articles), assembly plants average $0.8bn (198 articles).
4. Cell plants serve few clients (often VI or JV); large fixed costs to add clients
5. Multi-sourcing declines with product detail to a few percent for both cells and assembly.

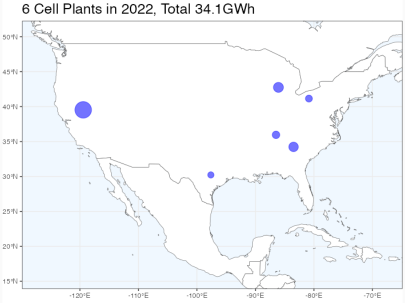
Challenges in estimating the model

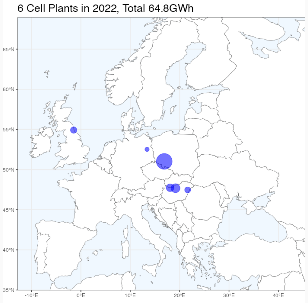
- Our estimation requires solving the MUFLP millions of times
  - Many sets of draws of path and fixed cost shocks (100)
  - Global optimization requires many starting parameter sets (1000),
  - Many iterations to fit the parameters (4700)

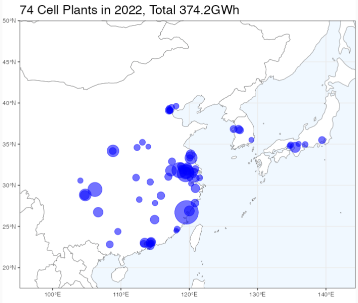
$\Rightarrow$ need some dimension reduction: lower number of locations and parameters

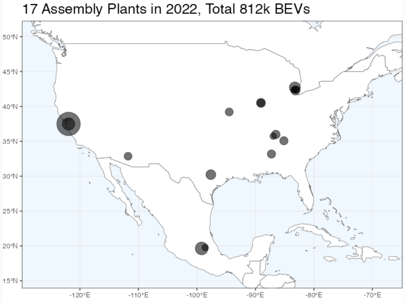
- We cannot estimate 1260 fixed costs fixed costs and 591 thousand path costs, so we parameterize them in terms of observables.

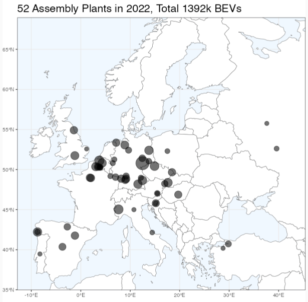
Mapping the cell and vehicle plants in 2022

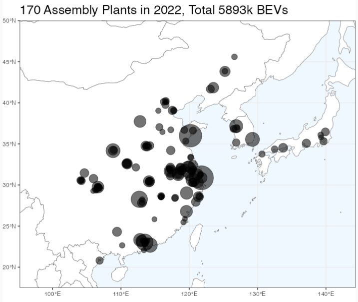












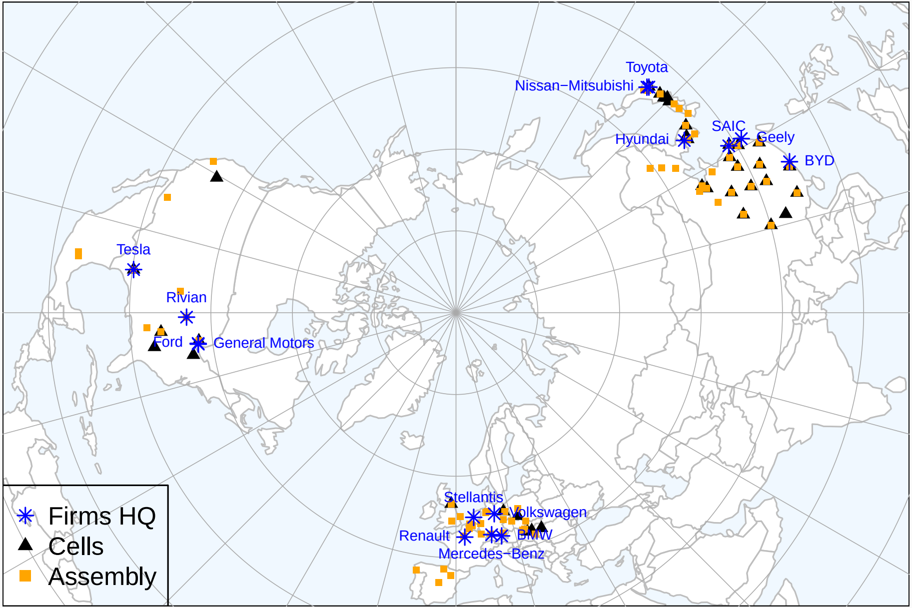
Data

Data (from S&P) contains universe of cell, assembly locations and flows for all EV models since 2015. We calibrate the model to 2022.

We reduce the estimation data to

- **Top 15 EV makers** and their **137 BEV models**, together 77% of world sales and 99% outside China
- Locations are across **24 countries**, together 97% world sales, 99.9% world battery production and 99% world EV production
- Entry is defined at geographic aggregates and product group level, **12 potential locations** for a median of **2 cell groups** per firm, and **15 potential locations** for a median of **4 assembly platforms** per firm $\longrightarrow 2^{24+60}$ total configurations

The top 15 EV makers and their location alternatives



Estimation roadmap

Step 1: Estimate bilateral trade costs for cells and vehicles
Conditional on active facilities ($\mathbf{y}$,$\mathbf{z}$) for each firm, estimate discrete choice sourcing ($\mathbf{x}$) via sequential (cells, assembly) nested logit for Cobb-Douglas production

Step 2: Use CES demand structure to map variable profits in levels to function of path observables and remaining parameters

Step 3: Estimate marginal cost of production, fixed costs, and endogenous market demand shifters by matching moments of the data to the simulated model (SMM) subject to market equilibrium as constraints (MPEC)

Step 1: Variable costs of paths

- Variable costs

$$\text{Cells costs:} \quad c^1_{m\ell_1\ell_2} = w^1_{\ell_1}\tau^1_{\ell_1\ell_2}\varepsilon^1_{m\ell_1\ell_2},$$
$$\text{Vehicle path costs:} \quad c(\boldsymbol{\ell}_{mn}) = \left(w^2_{\ell_2}\right)^{\alpha_{22}}\left(c^1_{m\ell_1\ell_2}\right)^{\alpha_{12}}\tau^2_{\ell_2 n}\varepsilon^2_{m\ell_2 n}.$$

- Conditional on active facilities ($\mathbf{y}$,$\mathbf{z}$), each firm sources cells and vehicles to minimize:

$$\ln c(\boldsymbol{\ell}_{mn}) = \alpha_{22}\ln w^2_{\ell_2} + \ln\tau^2_{\ell_2 n} + \alpha_{12}(\ln w^1_{\ell_1} + \ln\tau^1_{\ell_1\ell_2}) + u(\boldsymbol{\ell}_{mn})$$

Sequential Nested Choice estimation of variable costs

- **Stage** $k = 1$**:** Battery cell sourcing
  - **Chooser:** Assembly plant in ℓ_2 for model m
  - **Choice:** Battery cell plant in ℓ_1
  - **Choice set:** Plants $L_1(m)$ making the model's cell type for that maker
  - **Determinants:** trade costs (β_τ^1), fixed effects of supplier countries
- **Stage** $k = 2$**:** Vehicle sourcing
  - **Chooser:** dealership network in country n sourcing model m
  - **Choice:** EV assembly plant in country ℓ_2
  - **Choice set:** Plants that assembles the model's platform
  - **Determinants:** trade costs (β_τ^2), fixed effects of supplier countries, inclusive cost from cell stage $\rightarrow -\beta_\Phi^2$

Nested Logit Choice probabilities

Stage 1 (cells) estimating equation is

$$\mathbb{P}^1_{\ell_1|\ell_2} = \exp\left[\mathrm{FE}^1_{\ell_1} + \mathrm{FE}^1_{\ell_2}(m) + \beta^1_D \ln D_{\ell_1\ell_2} + \beta^1_t \ln\left(1 + t^1_{\ell_1\ell_2}\right)\right],$$

Stage 2 (vehicle assembly) estimating equation is

$$\mathbb{P}^2_{\ell_2|n} = \exp\left[\mathrm{FE}^2_{\ell_2} + \mathrm{FE}^2_n + \beta^2_D \ln D_{\ell_2 n} + \beta^2_t \ln\left(1 + t^2_{\ell_2 n}\right) + \beta^2_\Phi \mathrm{FE}^1_{\ell_2}(m)\right],$$

Inclusive cost: $\mathrm{FE}^1_{\ell_2}(m) = -\ln \sum_{\ell \in L_1(m)} \exp[\mathrm{FE}^1_\ell + \beta^1_D \ln D_{\ell_1\ell_2} + \beta^1_t \ln\left(1 + t^1_{\ell_1\ell_2}\right)]$

Nested Logit Sourcing Results

	Cells	Vehicles
Border	-0.953[a]	-1.04[a]
	(0.319)	(0.254)
log distance	-0.382[a]	-0.112[c]
	(0.021)	(0.062)
RTA	0.458	0.869[a]
	(0.320)	(0.214)
Inc. cost of cells		-0.234[a]
		(0.084)
log GDP per capita	0.213[c]	0.206[b]
	(0.118)	(0.087)
log(1+tariff)	-8.49[a]	-8.56[a]
	(2.34)	(1.74)
Observations	7,945	15,793
Squared Correlation	0.322	0.265

- Home plants (border =0) are $\approx$ 3 times more likely to be chosen
- Trade agreements also important
- Tariff elasticities are large ($\theta_1 \approx \theta_2 = 8.5$)
- Coef on inclusive cost $\rightarrow$ cell cost share of about 23%
- GDP/cap effects based on within-country variation since regression includes country fixed effects

Path cost changes $\implies$ level of variable profits

Log path costs (Logit, SMM):

$$\ln c(\boldsymbol{\ell}_{mn}) = \frac{1}{\beta_t^2}\Big\{\mathrm{FE}^2_{\ell_2} + \beta_D^2 \ln D_{\ell_2 n} + \beta_t^2 \ln\left(1 + t^2_{\ell_2 n}\right) - \underbrace{\beta_\Phi^2}_{\text{IC coef.}} \left[\mathrm{FE}^1_{\ell_1} + \beta_D^1 \ln D_{\ell_1\ell_2} + \beta_t^1 \ln\left(1 + t^1_{\ell_1\ell_2}\right)\right]\Big\} + \underbrace{u(\boldsymbol{\ell}_{mn})}_{-\text{Gumbel}(0,-1/\beta_t^2)}.$$

Expected variable profits levels using observed market shares s°_{mn} (for observed paths $\boldsymbol{\ell}^\circ_{mn}$) and cost change of potential path costs relative to the observed path, $\hat{c}_{mn\boldsymbol{\ell}} = c(\boldsymbol{\ell}_{mn})/c(\boldsymbol{\ell}^\circ_{mn})$:

$$E[\pi_{mn\boldsymbol{\ell}}] = \frac{1}{\eta} s^\circ_{mn} R_n (\hat{c}_{mn\boldsymbol{\ell}})^{1-\eta} A_n$$

where $A_n = E\left[(\hat{P}_n)^{\eta-1}\right]$ captures endogenous change in market demand

Endogenous market demand

- CES demand for all cars (EV and ICE) with demand elasticity $\eta = 4$
- Firms only know their own cost differentials $\hat{c}_{mn\ell^*}$ but not other firms
- Assume
  - Constant total car expenditures R_n
  - Monopolistic competition and constant markup
  - Paths of EVs do not affect ICE prices: $\hat{P}_n^{ICE} = 1$
- Change in the price index satisfies:

$$\hat{P}_n^{1-\eta} = s_n^{\circ,ICE} + \sum_m s_{mn}^{\circ} \hat{c}_{mn\ell^*}^{1-\eta}$$

- Minimize SMM objective subject to solving for the expected market demand $\hat{A}_n$ using sample mean across J simulations *as constraints*:

$$\hat{A}_n = \frac{1}{J}\sum_{j=1}^{J}(\hat{P}_{n,j})^{\eta-1} = \frac{1}{J}\sum_{j=1}^{J}\left(s_n^{\circ,ICE} + \sum_m s_{mn}^{\circ}\hat{c}_{mn\ell^*,j}^{1-\eta}\right)^{-1}$$

Completing the parameter set

- Last parameters needed are the fixed costs $\phi_{fg_k\ell_k}$, of activation $(y_{fg_k\ell_k})$.
- Distribution of fixed costs draws by location-stage:

$$\phi_{fg_k\ell_k} \sim \text{LogNormal}\left(\ln\left[\mu_{fg_k\ell_k} \times R_W^{\text{EV}}\right], \sigma_k\right),$$
$$\text{with } \mu_{fg_k\ell_k} = \exp\left[\ln \rho^k_{\mathcal{N}(\ell_k)} + \rho^{\text{HQ},k} \ln D_{h(f)\ell_k} + \rho^{\xi,k} \ln \tilde{\xi}_{fg_k}\right],$$

- Means expressed as a fraction of worldwide EV revenues, R_w^{EV}.
- The expectation of the fixed cost draws, $\rho^k_{\ell_k}$, depends on
  - distance to headquarter (HQ) w/ elasticity $\rho^{\text{HQ},k}$,
  - platform inferred quality w/ elasticity $\rho^{\xi,k}$.
  - continent-stage means, $\rho^k_{\mathcal{N}}$, for $\mathcal{N} \in \{\text{As}, \text{Eu}, \text{Am}\}$
- To avoid selection bias in the sourcing equations, the SMM also estimates continental variable cost differences: $\text{FE}^k_{\mathcal{N}}$ for $\mathcal{N} \in \{\text{As}, \text{Eu}\}$

Simulated method of moments (SMM)

Target 47 moments formed by stage $k = 1, 2$ & continent $\mathcal{N} \in \text{Am}, \text{As}, \text{Eu}$

1. Number of production lines $\left(\sum_{\ell_k \in \mathcal{N}} y_{fg_k \ell_k}\right)$ [6]
2. Share of continent $\mathcal{N}$ spending on EVs from continent $\mathcal{N}'$ [18]
3. Number of models offered in continent $\mathcal{N}$ [3]
4. Shares of production lines (firm HQ continent by production continent) [18]
5. Standard deviation of realized fixed cost from news articles [2]

Let $v(\cdot)$ be the difference between simulated moments and data. SMM solves

$$\min_{\mathbf{FE}, \boldsymbol{\rho}, \boldsymbol{\sigma}} v\left(\mathbf{FE}, \boldsymbol{\rho}, \boldsymbol{\sigma}\right)' W v\left(\mathbf{FE}, \boldsymbol{\rho}, \boldsymbol{\sigma}\right),$$

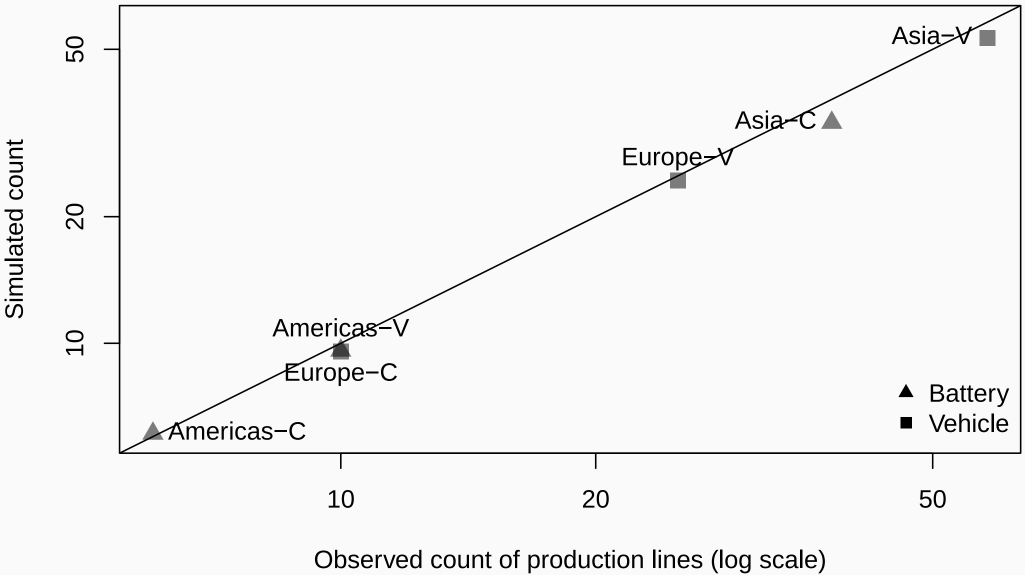
where W is a weighting matrix. The simulated moments average over $J = 100$ sets of draws of path cost and fixed cost shocks.

Simulated Method of Moments Estimates

Par.	$\mathcal{N}$	Description	Est.	SE
FE^1	As	Var. cost adv. (C)	4.91	(1.16)
FE^1	Eu	(by region, Am = 0)	5.94	(0.94)
FE^2	As	Var. cost adv. (V)	-0.21	(0.4)
FE^2	Eu	(by region, Am = 0)	-0.25	(0.48)
ρ^1	Am	Fixed cost	0.30	(0.21)
ρ^1	As	of cell plant	0.12	(0.05)
ρ^1	Eu	(by region)	0.54	(0.2)
ρ^2	Am	Fixed cost	0.16	(0.12)
ρ^2	As	of assembly	0.03	(0.05)
ρ^2	Eu	(by region)	0.18	(0.09)
$\rho^{\text{HQ, dist},1}$		FC HQ dist. elas. (C)	0.47	(0.08)
$\rho^{\text{HQ, dist},2}$		FC HQ dist. elas. (V)	0.78	(0.07)
$\rho^{\xi_{fg_1}}$		FC quality elas. (C)	3.40	(0.39)
$\rho^{\xi_{fg_2}}$		FC quality elas. (V)	2.76	(0.44)
σ_1		FC dispersion (C)	2.12	(0.31)
σ_2		FC dispersion (V)	1.99	(0.27)

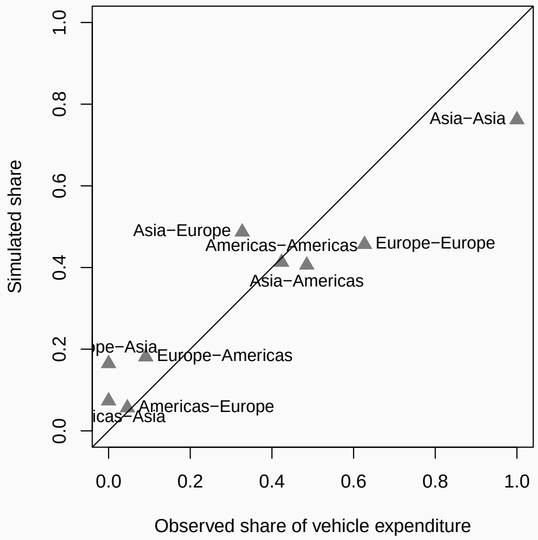
- Asia and EU have lower cell variable costs than US
- but their assembly costs are higher.
- US and EU fixed costs (FC) of assembly are 5–6 times those in Asia (0.16/0.03, 0.18/0.03)
- Proximity to V HQ lowers FC.
- Plants making higher quality have higher FC.
- Fixed costs of open facilities ≈ 7% of world revenue.

Calibrated Fit to Data: Production Lines by Continent

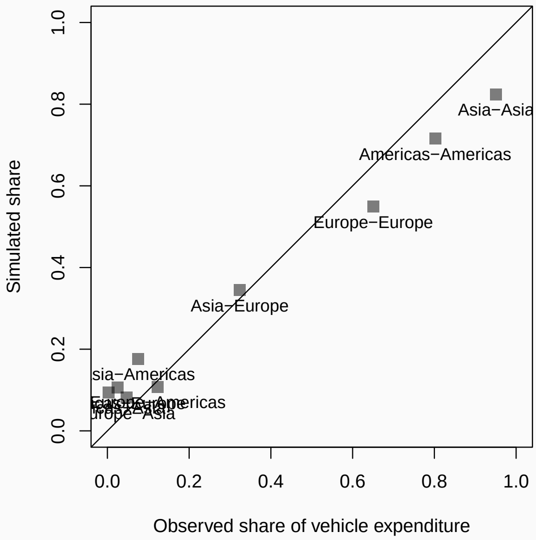


Calibrated Fit to Data: Inter-Continental Market share (revs)

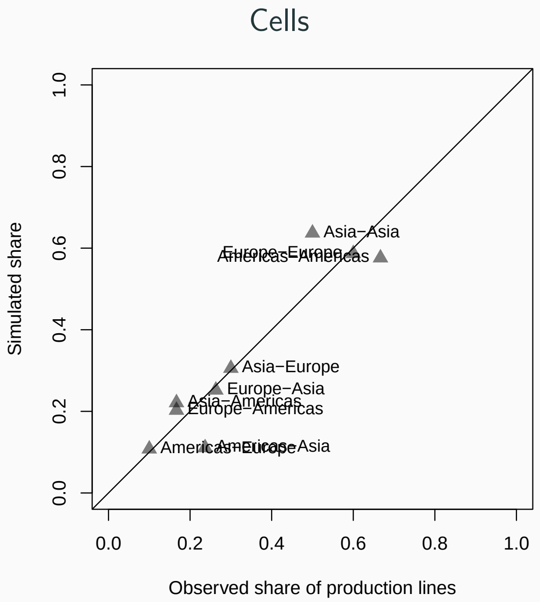
Cells

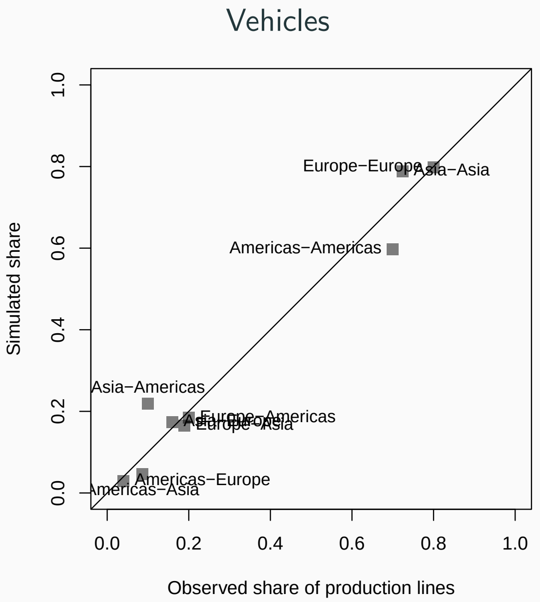


Vehicles



Calibrated Fit to Data: Production lines by headquarters





Reflections of an SMM newbie

- Irony: the "hard" problem of combinatorial discrete choice was the easy part; SMM was brutal.
- The computer resources for SMM are massive due to high number of sims (J) and starts (TikTak)
- It only gets worse when you try to bootstrap ($B \times J = 400$ cores)
- Selection and weighting of moments allows uncomfortably high "researcher degrees of freedom"
- Nostalgia for simple methods of my youth

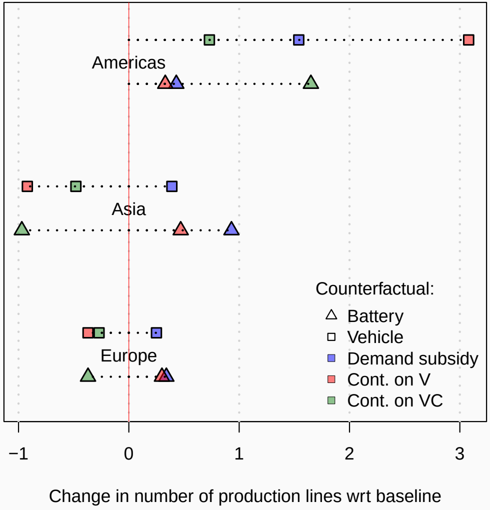
Counterfactual: BEV Policies

Policies we evaluate

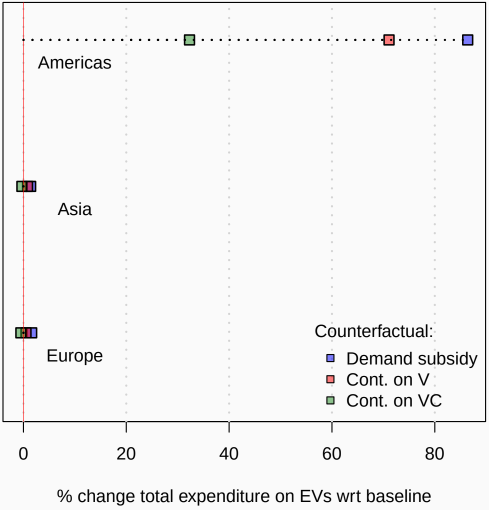
1. **Subsidy to EV buyers** 20% subsidy to all consumers from a given continent for all vehicles.
2. **Subsidy contingent on V** Only cars assembled in the same continent as the buyer are eligible for the 20% subsidy.
3. **Subsidy contingent on VC** Both cell production and vehicle assembly must be in the same continent as the buyer to be eligible for 20% subsidy.
4. Production subsidies of 20% for continental cells.
5. Tariffs of 20% on imported cells and vehicles.

Predicted Impact of North American BEV Subsidies

Production Lines ($\sum_f y_{g\ell}$)

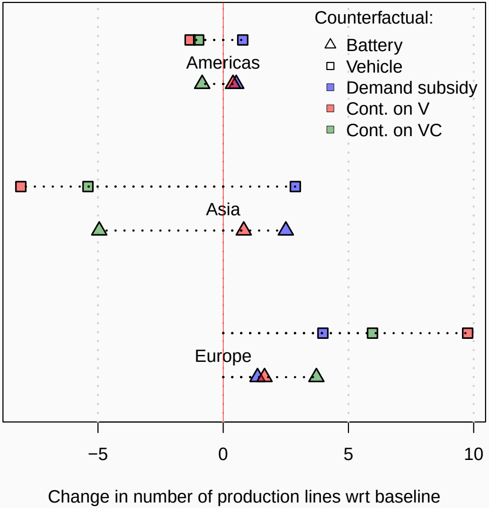


BEV Expenditures ($\sum_m p_{mn} q_{mn}$)

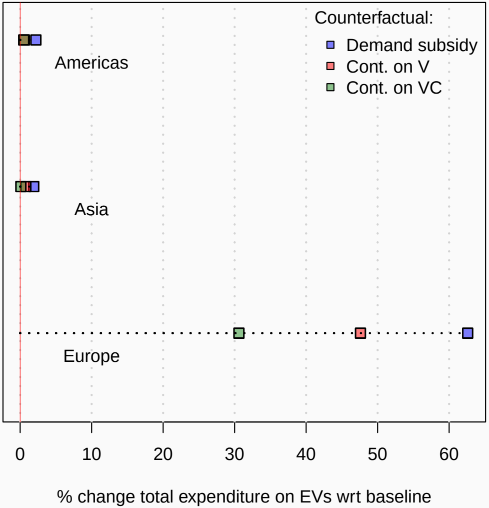


Predicted Impact of European BEV Subsidies

Production Lines ($\sum_f y_{g\ell}$)



BEV Expenditures ($\sum_m p_{mn} q_{mn}$)



What drives increased EV spending?

- The calibrated demand elasticity converts a uniform subsidy s into an expenditure increase of $(1-s)^{1-\eta}-1=(1-0.2)^{1-4}-1=$95%.
- But since EVs compete with each other their actual increases are smaller (Price index adjusts)
- New facilities permit new paths and closures cut off paths: delivered marginal costs c_{mn} change.
- Many models not eligible under policies 2 (cont. on V) and 3 (cont. on VC).
- We decompose the implied contributions of subsidy and path cost changes.

Contributions of subsidies and cost reductions to EV sales

Policy	Eligible share		Cost index change		Shifters		EV Exp.
	path	revenue	subsidy	costs	variety	demand	change
			$\hat{\tilde{t}}_{\mathcal{N}}$	$\hat{\tilde{c}}_{\mathcal{N}}$	$\bar{s}_{\mathcal{N}}^{\mathrm{EV}}$	$\hat{A}_{\mathcal{N}}$	$\hat{\bar{R}}_{\mathcal{N}}^{\mathrm{EV}}$
1: Unconditional	100.0	100.0	-20.0	-4.5	2.5	-16.5	86.4
2: Continental V	43.4	90.5	-17.4	-3.4	1.9	-13.6	71.1
3: Continental V+C	21.5	68.3	-14.7	2.2	0.8	-6.2	32.3
4: Subsidy C	18.3	46.0	-2.7	0.5	0.9	-1.0	5.1
3+4	28.1	77.9	-19.2	3.2	1.2	-8.8	46.2
5V: Tariff V	63.1	12.5	3.1	1.1	-0.9	1.9	-9.8
5: Tariff V+C	85.5	53.9	4.8	1.7	-1.2	3.1	-16.1

The unconditional buyer subsidy yields the largest reduction in path costs and biggest increase in EV takeup

What we've learned so far

- Increasing returns (fixed costs) and multiple production stages generate complex, jumpy policy responses
- MILP can be effective for solving optimization problems that combine substitution and complement patterns across locations
- With IRS, protectionist policies potentially align with emissions goals.
- Precise estimation of the core parameters presents challenges. But we have moderate confidence that
  - A clean consumer credit would have done more to promote EV adoption
  - Non-enviromental objectives (e.g. national security benefit of domestic battery capacity) needed to justify upstream restrictions.