Market Counterfactuals with Nonparametric Supply: An ML/AI Approach

Harold D. Chiang Jack Collison Lorenzo Magnolfi Christopher Sullivan

July 7, 2025 — Alpine IO Symposium

Motivation

- The outcomes of government policy are shaped by the response of markets
- Thus, we seek answers to a wide range of policy questions from empirical models of demand and supply, used to perform market counterfactuals
  - E.g., effect of taxes and subsidies, product regulations, competition policy, and more
- Credible empirical models of markets require balance between data and structure
- Key achievement of IO approach: enable flexible estimation of demand, rich subst. patterns
  - Otherwise, ans. to counterfactual questions driven by assumption/parametric restrictions
- What about supply?

Motivation

- Market counterfactuals crucially depend on the supply (i.e., conduct and cost) specification
- Examples:
  - Price response to selective tax/tariff depends on strategic complements/substitutes
  - Economies of scale/scope matter in response to industry consolidation/mergers
- Standard approach imposes strong assumptions on supply
  - E.g., Bertrand-Nash oligopoly with constant marginal cost
- Can we provide useful market counterfactuals without restrictive assumptions on supply?

What We Do I

- Consider a nonparametric model of supply
- Supply fct. that captures markup + costs, depends on endogenous prices and quantities
- Reflects notion that prices are set in market equilibrium, but no structure on conduct or cost
- Show identification with appropriate supply instruments
  - Can leverage both variation in (own and rival) demand shifters, and rival cost shifters

What We Do II

- Estimate model with ML/AI: deep learning + objective function with instruments
- Adapt Variational Method of Moments (VMM) (Bennett and Kallus, 2023)
- Better performance with high-dimensional data than standard nonpar-IV (e.g., series)
- We develop an inference procedure to quantify uncertainty in prediction

What We Do III

- Simulations show that nonparametric model is practical with moderate sample sizes/variation, outperforms misspecified models
- Simulations across many counterfactuals showcase performance
- Application: mergers in airline markets
- Portable method, computationally manageable

Literature

- We build on existing nonparametric approaches to markets for differentiated products
  - Nonparametric identification: Berry and Haile (2014)
  - Nonparametric approaches to demand: Compiani (2022), Tebaldi, Torgovistky, and Yang (2023), Brand and Smith (2025), ...
  - Flexible approaches to supply: Gandhi and Houde (2020), Otsu and Pesendorfer (2024)
- Complementary to testing/parametric estimation approaches
  - Modern testing approaches (e.g., Backus, Conlon, and Sinkinson 2021, Duarte et al. 2023) also let data shape supply by selecting conduct model within menu
  - Trade-off: flexibility vs. data/variation requirements
- Part of broader trend of using ML/AI to enhance structural modeling
  - E.g., Kaji, Manresa, and Pouliot (2023)
  - We use Bennet and Kallus (2023) for ML/AI approach to nonparametric IV

Roadmap

Market Equilibria and Counterfactuals

Nonparametric Model of Supply

Estimation and Inference

Monte Carlo Simulations

Empirical Application

Conclusion

A Model of Market Equilibrium

- We observe data on a set of differentiated products $\mathcal{J}$ across $\mathcal{T}$ markets:
  - Consumers and firms' behavior results in outcomes p_{jt} (prices) and s_{jt} (market shares)
  - Exogenous observables include characteristics x_{jt}, cost shifters w_{jt}
  - Exogenous unobservables are unobserved quality ξ_{jt} and unobserved cost shifter ω_t
  - Useful transformations of endogenous variables include demand derivatives D_t and quantities q_t
- Market equilibrium is determined by demand and supply:

$$s_t = \underbrace{\mathscr{s}(p_t, x_t, \xi_t)}_{\text{Demand}}, \qquad p_t = \underbrace{\Delta(p_t, s_t, D_t; \cdot)}_{\text{Markups}} + \underbrace{c(q_t, w_t, \omega_t)}_{\text{Marginal costs}}$$

- Markups can depend on other exogenous variables, e.g., ownership matrix $\mathcal{H}_t$
- Allow for conduct and cost to depend on firm identity, denote Δ_j, c_j scalar valued functions
- General static setting, can be extended to other endogenous non-price variables (not today)

Assumptions: DGP and Observables

1. *(Equilibrium Selection)* There exists a unique equilibrium, or the equilibrium selection rule is such that the same p_t arises whenever the vector $(\mathrm{w}_t, x_t, \omega_t, \xi_t)$ is the same.
2. *(Separability of Cost)* The cost function is separable in unobservable shocks:
$$c(q_t, \mathrm{w}_t, \omega_t) = \bar{c}(q_t, \mathrm{w}_t) + \omega_t.$$
3. *(Known Demand)* The matrix of demand derivatives is known, so that D_t is observed.
4. *(Markup Dependence)* The markup function Δ depends only on endogenous market shares s_t and the matrix of demand derivatives D_t.

- Assumption 4 general (includes most standard static oligopoly models) but not without loss
  - e.g., Bertrand/profit weights: $\Delta = (\mathcal{H}_t \odot D_t)^{-1} s_t$ Cournot: $\Delta = \left(\mathcal{H}_t \odot (D_t^{-1})'\right) s_t$
- Demand is known, supply (Δ and $\bar{c}$) is not

Market Counterfactuals

- Policy changes of interest exogenously change primitive object $a \to \tilde{a}$
- New market outcomes can be computed by solving fixed point:

$$\tilde{p}_t = \tilde{\Delta}(\tilde{p}_t, \tilde{\mathfrak{z}}(\tilde{p}_t, \tilde{x}_t, \tilde{\xi}_t)) + \tilde{c}(\tilde{q}_t, \tilde{w}_t) + \tilde{\omega}_t$$

- We can express counterfactuals as a map $F(\tilde{p}, \tilde{s}, \cdot)$ from structural objects and exogenous variables to outcomes of interest, e.g., prices, shares, consumer welfare, etc.

Estimating Counterfactuals

- Evaluating the map F requires knowledge of (counterfactual) primitives, exogenous observables, and unobservables

- Researchers use a combination of data and theory/assumptions
  - We typically estimate functions $\mathcal{s}$ and c, and assume Δ with a model of conduct
  - Allows us to specify $\tilde{\xi}, \tilde{\omega}$ and either fix or deterministically change $\tilde{\mathcal{s}}, \tilde{c}, \tilde{\Delta}$

- Trade off practicality and data limitations with the dangers of misspecification
  - Estimation of nonparametric models has a curse of dimensionality and requires rich data
  - Misspecification can result in misleading results and subsequent counterfactuals

- Next: a feasible nonparametric model of supply

Roadmap

Market Equilibria and Counterfactuals

Nonparametric Model of Supply

Estimation and Inference

Monte Carlo Simulations

Empirical Application

Conclusion

Nonparametric Model of Supply

- Recall DGP, under assumptions 1-4:

$$p_{jt} = \Delta_j(s_t, D_t, \mathscr{H}_t) + c_j(q_t, w_{jt}) + \omega_{jt}$$

- Without further restrictions, as $q_{jt} = s_{jt} M_t$ for scalar market size M_t:

$$p_{jt} = h_j(s_t, D_t, w_{jt}; \mathscr{H}_t, M_t) + \omega_{jt}$$

for supply function $h(\cdot) \equiv \Delta(\cdot) + \bar{c}(\cdot)$

Remarks

- Notice that formulation of h does not enforce separability of cost and markup
  - Extension I: we can assume a model of conduct, and recover costs flexibly (not today)
  - Extension II: we can assume a cost function, and recover markup functions flexibly (not today)
  - Extension III: we can allow separability with more instruments and variation (not today)
- For counterfactuals, find the prices $\tilde{p}_t$ that solve:

$$\tilde{p}_t - \hat{h}(s(\tilde{x}_t, \tilde{p}_t), D(\tilde{x}_t, \tilde{p}_t), \tilde{w}_t; \tilde{\mathscr{H}}_t) - \hat{\omega}_t = 0$$

- We can define a map F to counterfactuals of interest using estimated objects

Which Counterfactuals?

- This nonparametric supply structure enables a wide range of counterfactuals
  - Changes in ownership resulting from mergers, firm/product exit, and product divestment
  - Regulations that alter product characteristics or cost shifters, e.g., fuel economy standards
  - Unit and ad valorem taxes with variation across products and/or markets
- We can measure equilibrium prices, quantities, and changes in consumer welfare (and, in the case of a tax, government revenue and incidence)
- Important limitations of our approach:
  - Cannot measure markups and cost levels separately
  - Cannot alter cost or markups separately
  - (Can be addressed w/ extensions of the method)

Identification of h

- Key identification challenge: the s_t and D_t arguments of h are endogenous
- We rely on a moment condition with supply instruments z_{jt} for identification. Assume:

5. *(Instrument Exogeneity and Exclusion)* The vector of instruments z_{jt} that satisfies $\mathbb{E}[\omega_{jt} \mid z_{jt}, \mathrm{w}_{jt}] = 0$ contains demand shifter(s) $x_{jt}^{(e)}$ that are excluded from the vector w_{jt}.
6. *(Completeness)* For all functions $B(s_t, D_t, \mathrm{w}_t; \mathscr{H}_t)$ with finite expectation, if $\mathbb{E}[B(s_t, D_t, \mathrm{w}_t; \mathscr{H}_t) \mid z_{jt}, \mathrm{w}_{jt}] = 0$ almost surely, then $B(s_t, D_t, \mathrm{w}_t) = 0$ almost surely.

- Result: under 1.-6., h_j is identified for each j
    - Proof follows arguments akin to Berry and Haile (2014)

Instruments and Data Requirements

- Candidate supply instruments - need $2J + J(J-1)/2$, many candidates available
  - Intuitively: rival cost shifters move s_t; (own and rival) prod. characteristics move D_t
  - Other instruments (e.g., variation in exogenous tax rates) may be available
- Must include demand shifters *excluded* from cost
  - If not, e.g. w/ logit demand, may just recover inverse demand $h = \mathscr{s}^{-1}$
- Concern: will variation in the data suffice?
- We show simulation evidence later on that standard datasets may have enough variation
- Conduct/cost testing toolkit allows some flexibility in supply with discipline from theory
  - E.g., procedures in Backus, Conlon, Sinkinson (2021), Duarte et al. (2023)

Roadmap

Market Equilibria and Counterfactuals

Nonparametric Model of Supply

Estimation and Inference

Monte Carlo Simulations

Empirical Application

Conclusion

Estimation and Inference VMM Inference

- Classic nonparametric estimators (e.g., series) are well studied for GMM-type setups
    - For nonparametric IV problem, Ai and Chen (2003); Newey and Powell (2023)
    - See reviews by Carrasco et al. (2007); Chen (2007)
- But, curse of dimensionality and instability in classical nonparametric estimation
    - Documented in, e.g., Bennett et al. (2019); Bennett and Kallus (2020)
    - (We make no general claim about the usefulness of classical methods)
- We employ the Variational Method of Moments (VMM)
    - VMM accounts for endogeneity via moment conditions that standard neural networks ignore
    - We develop a method for our nonparametric supply, adapting VMM + DNN
    - Derive uniform prediction bands for prices, shares, consumer surplus, tax revenue
- Method in a nutshell:
    - Deploy two DNNs to learn both optimal instruments and structural supply fcn h

Why Neural Networks and VMM?

- Why neural network structures?
  - Learn complex structures and achieve faster convergence rates than nonparametric benchmarks
  - See Bauer and Kohler (2019); Schmidt-Hieber (2020); Kurisu et al. (2025)
- *Example*: Bertrand-Nash with sparse profit weights
  - Prices (below) can be written in a sparse tensor decomp that depends on latent dim $\ell < J$

$$\omega_j(p, c, \mathcal{H}, D) = p_j - c_j - [(\mathcal{H} \odot D')^{-1} s]_j,$$

  - Results in Schmidt-Hieber (2020) then implies that DNNs achieve faster convergence rates
- Why the Variational Method of Moments?
  - In fixed-dimensional parametric settings, VMM coincides with OWGMM
  - Inference properties are known; we develop inference for a complex functional of parameters

Roadmap

Simulations Roadmap

How does our method perform?

I Does it handle realistic, high-dimensional environments well with moderate sample sizes?

II But, what's inside the black box?

III What range of counterfactuals can it handle?

Simulations Setup I

- Simple parametric simulations to evaluate absolute and relative performance
- For $T = 100; 1,000; 10,000$, market t has either
  - A: $J_t = 2; 3$ with equal probability (small)
  - B: $J = 30$, owned by 5 firms, as US beer market in Miller and Weinberg (2017)
- Training data is random 80% of the dataset (sampled by market)
- Demand is
  - A: simple logit, scalar unobservable ξ_{jt}, three observable x_{jt}
  - B: RCNL as in Miller and Weinberg (2017)

Simulations Setup II: Supply

- We simulate data under two different assumptions on conduct...
  - *Bertrand-Nash*: Identity ownership matrix
  - *Profit-Weight*: Off-diagonal weights of $\tau = 0.5$
- ...cost specifications...
  - *Linear*: Linear costs with two independent cost shifters w_{jt}
  - *Economies of Scale*: Marginal costs are quadratic in quantities
- ...and policy instruments
  - *Unit Taxes*: Variation in unit taxes across market
  - *Ad Valorem Taxes*: Variation in ad valorem taxes across markets

Comparison of Models

- We recover ω^B, ω^M, and ω^P under Bertrand, Monopoly, and perfect competition

- In the nonparametric supply model, we estimate h and recover ω_{jt}^V:

$$p_{jt} = \hat{h}_j\left(s_t, D_t, w_{jt}, \mathscr{H}_t\right) + \omega_{jt}^V$$

- Supply instruments: own x_{jt}, rival w_{-jt}, sum of rival x_{jt}, w_{jt}

- We run two types of simulation exercises:

1. For A, B: predict in test sample (20% of data) w/ estimated function $\hat{h}$ and setting $\omega_{jt}^m = 0$
2. For A: Simulate counterfactual interventions in mkt t, predict residual ω_{jt}^m under model m

- For both exercises, used demand and estimated supply to compute fixed point

Test Sample Price Prediction Performance in A

Table 1: MSE Across Models, Bertrand DGP (Small Network), environment A

T	True Model	Standard Models			Flexible	D_t included
		B	M	P		
100	0.90	0.90	578.77	9.41	1.91	No
					1.71	Yes
$1,000$	0.89	0.89	1022.81	8.17	2.21	No
					1.07	Yes

- Small network has 3×3 hidden layer

Test Sample Price Prediction Performance, Profit-Weight DGP in A

Table 2: MSE Across Models, Profit-Weight DGP (Large Network), environment A

T	True Model	Standard Models			Flexible	D_t included
		B	M	P		
1,000	0.89	3.69	66.42	8.22	1.73	No
					1.74	Yes
10,000	0.96	4.02	77.08	8.79	1.23	No
					1.05	Yes

- Large network has 100×100 hidden layer
- (When we repeat this exercise 100 times w/ different random draws, initialization of NN, we find tight MSE ranges)

Test Sample Price Prediction Performance, Profit-Weight DGP in B

Table 3: MSE Across Models, Profit-Weight DGP, environment B

T	True Model	Standard Models			Flexible ($\#h = 10$)	Flexible ($\#h = 200$)
		B	M	P		
$1,000$	1.19	3.86	3.76	3.31	2.96	1.58
$10,000$	1.07	3.96	3.86	3.41	2.74	1.28

- $\#h$ is number of layers

Key Takeaways

- Performance is
    - in A, already reasonable with 100 markets, with 1,000 may be already close to match truth
    - in B, ok with 1,000 markets, probably need closer to 10,000
- Adding derivatives helps, especially in larger samples
- Larger network structure useful to capture complex models of supply, but need more data

Peeking Inside the Black Box: Pass-through

- Key question: How do we interpret the flexible $\hat{h}$ we recover?
- A useful object for comparison is the pass-through matrix implied by $\hat{h}$
- To compute pass-through:
  - Pick median post-merger market by inside share from simulations
  - Increase costs c by 10%, loading increases on the residual ω^V
  - Solve for equilibrium prices under different models of conduct
  - Compare price before and after cost change, report price change/cost change

Pass-through Comparison

Table 4: Simulated Pass-through Matrices

Panel A: Bertrand DGP				Panel B: Profit-Weight DGP			
True Model		Flexible Supply		True Model ($\kappa = 0.5$)		Flexible Supply	
0.77	0.17	0.69	0.12	0.35	-0.31	0.44	-0.21
0.10	0.60	0.12	0.63	0.02	0.98	0.01	0.91

- Flexible model learns economics of supply side, implies pass-throughs close to the true ones
- (Holds beyond this one market)

Market Counterfactuals

- Results thus far show test sample performance
  - Predict for markets out of the training sample, but from same DGP
- Key aspect of counterfactual prediction: (somewhat) out-of-sample
- Caveat when using our method:
  - As with any nonparametric approach, will struggle too far from the support of the data
- Next set of simulations shows "how far is too far"
- (Throughout, $T = 1,000$, env A, flex model estimated with small network structure)

What Market Counterfactuals?

- Predict prices after product regulations alter cost shifters simulations
  - E.g., environmental regulations could increase production costs
- Predict market shares after product regulation on product characteristics simulations
  - E.g., bans on menthol in cigarettes or caps on sugar content could change consumption
- Predict welfare changes due to product entry or exit simulations
  - E.g., introduction of a new vehicle or merging firms drop products
- Predict welfare changes due to mergers simulations
  - E.g., mergers in which existing products have new ownership
- Predict revenues after changes in unit and ad valorem taxes simulations
  - Governments could impose taxes on goods

Computational Cost

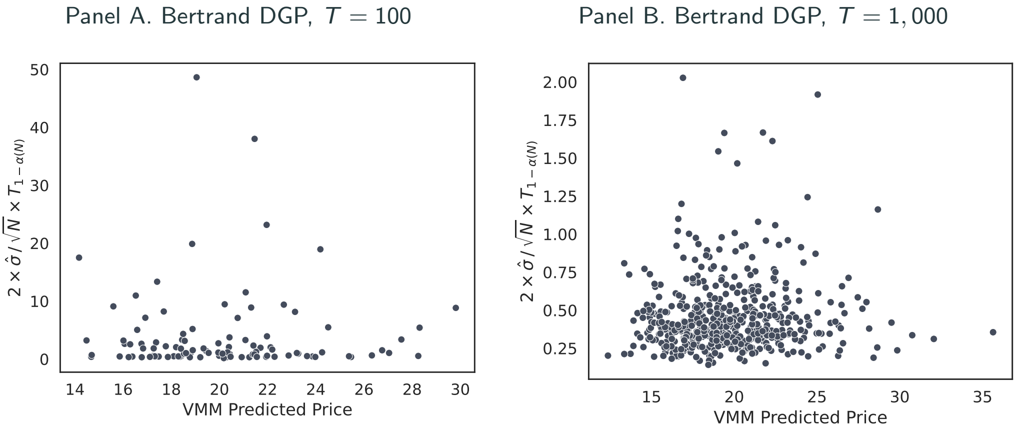
- Implementation and computation is manageable
- We use the Python package `torch` for all models
- Model fit takes minutes with $T = 100$, an hour for $T = 1,000$, and $\sim$a day for $T = 10,000$

Uncertainty in Counterfactuals

- What about uncertainty in prediction?
- With our VMM estimator, we can compute standard errors for counterfactual outcomes
- Two aspects:
    - Computing standard errors is computationally tractable
    - Uncertainty in prediction seems reasonable in simulations
- Exercise: show market-by-market prediction errors for product exit counterfactuals

Inference on Counterfactual Prices

Figure 1: Inference on Counterfactual Product Exit Prices



- With $T = 1,000$, tight prediction intervals

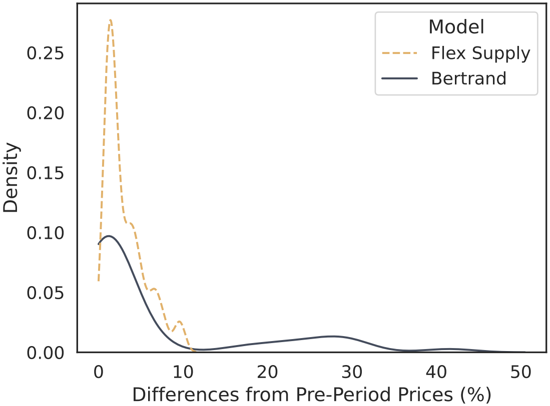
Roadmap

Application to Airline Mergers

- Good environment to test our method: airline markets in the US have rich data from DB1B
  - Fares, passenger counts, distances, carrier identifiers, etc.
  - Origin and destinations of trips
  - Several large mergers in sample
  - 1mln+ obs pooling quarterly data 2005-2019, we use $\sim$10,000 pre-merger markets for VMM
  - Previous merger retrospectives (Peters, 2006)
- Estimate simple nested logit demand model
- Goal: predict unilateral price effects of American-US Airways merger in Q4 2013
  - Zoom in on markets that move from 3 $\rightarrow$ 2 firms post-merger
  - Treated markets are markets in which both merging firms are present
- (We abstract from many interesting aspects of the industry here...)

Merger Simulation: Predicted Price Changes Inference

Figure 2: Predicted Price Change Distribution

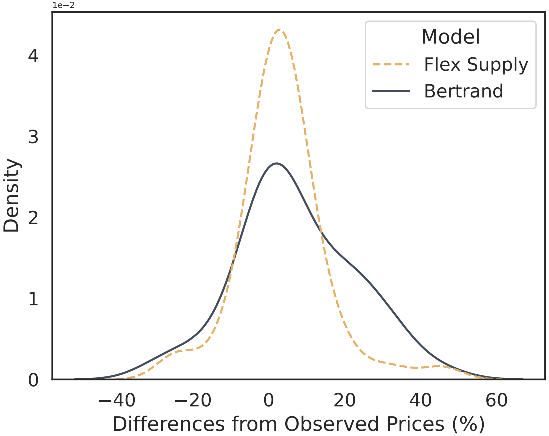


	Price Increases (%)	
Model	Median	Mean
Bertrand	1.45	6.66
VMM	2.05	2.16

- In theory, our flexible supply model could predict price decreases but it doesn't here

Merger Simulation: Comparing Predicted and Observed Post-merger Prices

Figure 3: Merger Simulation Comparison



Model	MSE
Bertrand	365.71
VMM	66.93

Roadmap

Conclusion

- Market counterfactuals crucially depend on the supply model
  - We develop a nonparametric model of supply for a range of counterfactuals
  - Estimation technique uses deep learning + objective function with instruments
  - Inference procedure quantifies uncertainty of a complex functional
- Simulation exercises and an empirical application showcase the method
  - Outperform misspecified models across a host of counterfactuals
  - Merger simulation in the airline industry outperforms the standard merger simulation toolkit

Thank You!

Variational Method of Moments (VMM) Back

- Our model and assumptions imply a moment condition for the structural supply function:

$$\mathbb{E}[p_{jt} - h_j(s_t, D_t, w_t, \mathcal{H}_t) \mid z_t, w_t] = 0$$

- The VMM estimator (Bennett and Kallus, 2023) for our setting is:

$$\hat{\theta}_N = \text{argmin}_{\theta \in \Theta} \text{sup}_{f \in \mathcal{F}_N} \frac{1}{TJ} \sum_{j,t} f(z_{jt})^T \omega_{jt}(\theta) - \frac{1}{4TJ} \sum_{j,t} (f(z_{jt})^T \omega_{jt}(\tilde{\theta}_N))^2 - R_N(f, h)$$

$$\text{s.t.} \quad \omega_{jt}(\theta) = p_{jt} - h_j(s_t, D_t, w_t, \mathcal{H}_t; \theta,) \qquad \forall j \in J$$

  - $\tilde{\theta}_N$ is preliminary estimate
  - Both f and h are neural networks, allowing flexible controls of model complexity
  - Quadratic term motivated by optimal weighting of making each moment condition zero
  - $R_N(\cdot)$ is regularizer that penalizes complexity

Quantifying Uncertainty Details + Algorithm Back

- We establish simultaneous confidence intervals for d predicted or counterfactual prices $\hat{h}$:

$$\sqrt{N}(\hat{h} - h_0) \xrightarrow{d} N(0, \nabla_{\theta'} h_0 \Omega_0^{-1} \nabla_{\theta'} h_0^T)$$

- Prices are not the only counterfactual of interest in economics research
  - E.g., quantities, consumer surplus, government revenue

- Assuming smoothness of the counterfactual map F in prices, we establish:

$$\sqrt{N}(F(\hat{h}) - F(h_0)) \xrightarrow{d} N(0, \nabla_h F(h_0) \nabla_\theta h_0 \Omega_0^{-1} \nabla_\theta h_0^T \nabla_h F(h_0)^T)$$

- We can quantify uncertainty on other economic objects of interest
  - Inference is possible for product-level, market-level, and aggregate objects

Inference: Simplest Case ($d = 1$) Back

- Note that $\nabla_{\theta'} h(\theta_0)$ is $d \times b$; in the simplest case, suppose that $d = 1$
- Lemma 9 in Bennett and Kallus (2023) states that for any $\beta \in \mathbb{R}^b$, we have:

$$\beta^T \Omega_0^{-1} \beta = -\frac{1}{4} \inf_{\gamma \in \mathbb{R}^b} \sup_{f \in \mathcal{F}} \left\{ \mathbb{E}[f(Z)^T \nabla_\theta \omega(X; \theta_0) \gamma] - \frac{1}{4} \mathbb{E}[(f(Z)^T \omega(X; \theta_0))^2] - 4\gamma^T \beta - R_N(f, h) \right\} \quad (1)$$

- Take $\beta = \nabla_\theta h_x(\theta_0)$ and the above solution to the optimization problem becomes:

$$\sigma_x^2 = \nabla_\theta h_x(\theta_0) \Omega_0^{-1} \nabla_\theta h_x(\theta_0)^T$$

- This is the asymptotic variance for $\sqrt{N}(h_x(\hat{\theta}_N) - h_x(\theta_0))$
  - $\nabla_\theta h_x(\theta_0)$ can be difficult to compute analytically
  - Numerical differentiation can be employed (e.g., Hong et al. (2015))
  - Expectations can be replaced by sample means, $\hat{\theta}_N$ can be used in place of θ_0
  - These together yield a feasible version of Equation (1) which provides an estimator $\hat{\sigma}_x^2$ for σ_x^2

Inference: Extending to $d \geq 2$ Back

- The approach above cannot obtain a covariance matrix when $d \geq 2$
- Holm's Step-Down procedure using the estimates for $\hat{\sigma}^2_{x_j}$ and $h(\hat{\theta})$ for each $j = 1, ..., d$
- The set of critical values T_α is known for significance levels $\frac{\alpha}{d+1-k}$ and $k = 1, ..., d$
  - We can use a folded normal distribution with $t = 1$ to account for bias
- For any ordering of x and fixed ordering T_α, we can compute the confidence interval:

$$h_x(\hat{\theta}) \pm N^{-\frac{1}{2}} \hat{\sigma}_x T_\alpha$$

- We compute this for all permutations of $j = 1, ..., d$, resulting in $d!$ permutations of x
- This is because we must consider any possible ordering of the p-values of $x_1, ..., x_d$

Inference Algorithm

Back

1. Estimate $\hat{\sigma}^2_{x_j}$ for $\sigma^2_{x_j}$ for $j \in \{1, ..., d\} \equiv J$ by solving the feasible version of Equation (1)
2. Fix values $T_\alpha = \{T_{\alpha_k} : k = 1, ..., d\}$ where $\alpha_k = \frac{\alpha}{d+1-k}$
3. For each permutation $\tilde{J}$ of J:
   3.1 Arrange values $\tilde{x}$ and $\hat{\sigma}_{\tilde{x}}$ with permuted indices $\tilde{J}$
   3.2 Construct bounds as $h_{\tilde{x}}(\hat{\theta}) \pm n^{-\frac{1}{2}} \hat{\sigma}_{\tilde{x}} T_\alpha$ with fixed T_α
4. Simultaneous confidence interval as the union of $2 \times d \times d!$ linear constraints from Step (3)

Regulation of Cost Shifters

- For all counterfactuals, use RMPSE $\simeq$ avg magnitude of percentage errors
- Implementation: add 1 to $w_{jt}^{(1)} \sim U(0,1)$

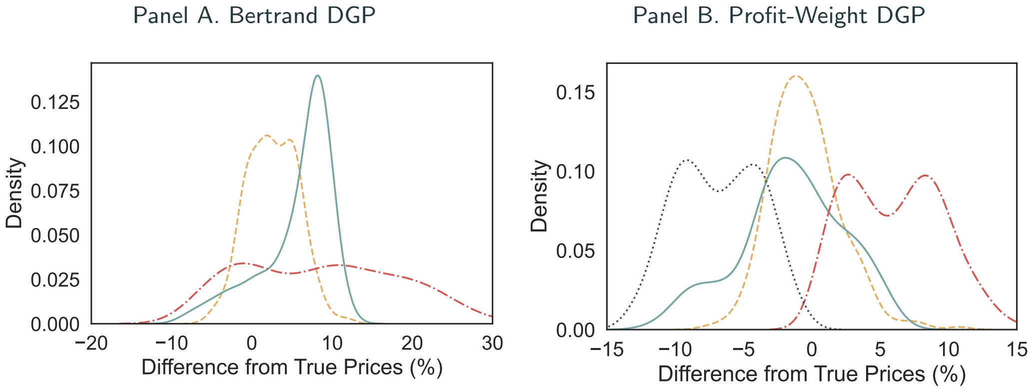
Table 5: RMPSE in Prices for Cost Shifter Regulation

Fitted Model	Panel DGPs			
	A. Bertrand	B. Profit-Weight	C. Bertrand (Scale)	D. Profit-Weight (Scale)
Bertrand (Scale)	--	--	--	5.0
Bertrand (Const.)	--	5.1	3.5	6.0
Monopoly	10.2	5.1	11.6	6.5
Perf Comp	5.0	6.3	5.0	3.8
Flex Supply	2.8	2.5	3.5	3.0

- Good absolute and relative performance for a fairly out-of-sample counterfactual

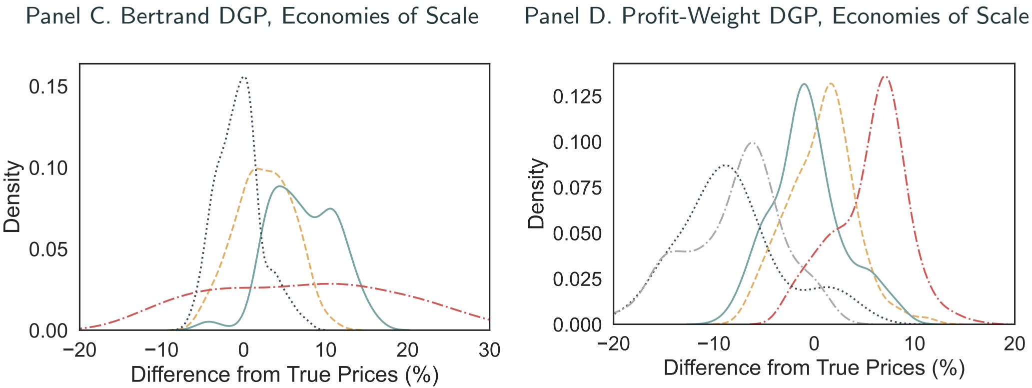
Regulation of Cost Shifters

Figure 4: Regulation of Cost Shifters



Regulation of Cost Shifters with Economies of Scale

Figure 5: Regulation of Cost Shifters



Back

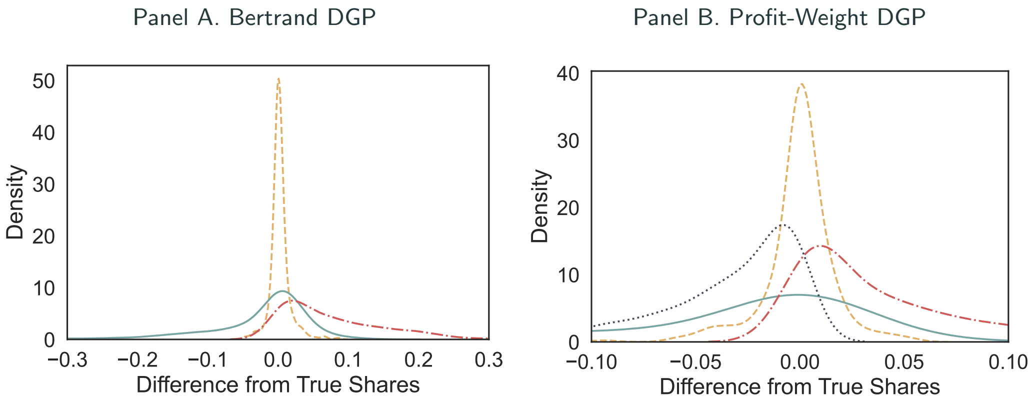
Regulation of Product Characteristics

- Implementation: add 1 to $x_{jt}^{(1)} \sim U(0,1)$

Table 6: RMPSE in Shares for Regulation of Product Characteristics

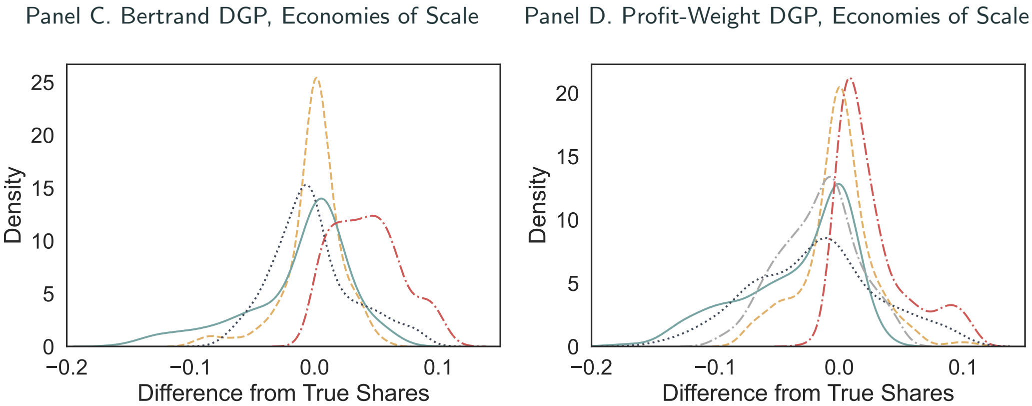
Fitted Model	Panel DGPs			
	A. Bertrand	B. Profit-Weight	C. Bertrand (Scale)	D. Profit-Weight (Scale)
Bertrand (Scale)	--	--	--	20.3
Bertrand (Const.)	--	21.0	10.2	28.3
Monopoly	31.9	22.3	15.9	11.1
Perf Comp	32.9	57.7	12.7	16.4
Flex Supply	5.0	6.6	6.6	9.2

Figure 6: Regulation of Product Characteristics



Regulation of Product Characteristics with Economies of Scale

Figure 6: Regulation of Product Characteristics



- Predicting shares seems somewhat harder

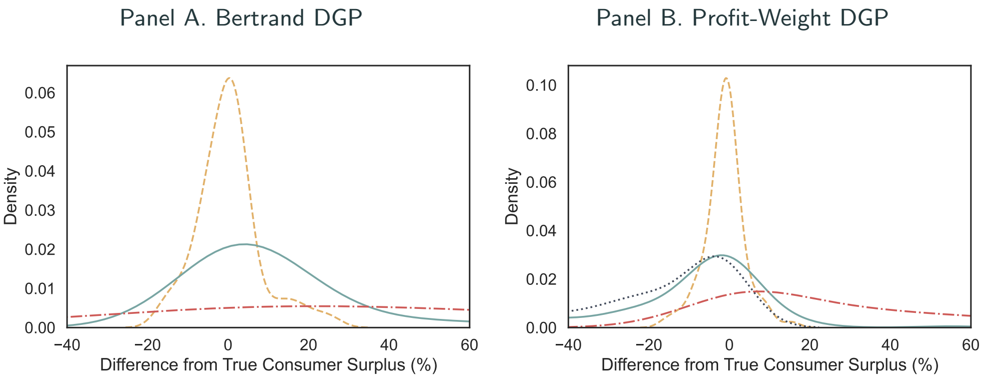
Product Exit

- Implementation: drop a product

Table 7: RMPSE in Consumer Surplus for Product Exit

Fitted Model	Panel DGPs			
	A. Bertrand	B. Profit-Weight	C. Bertrand (Scale)	D. Profit-Weight (Scale)
Bertrand (Scale)	--	--	--	25.0
Bertrand (Const.)	--	23.9	7.7	23.2
Monopoly	135.4	47.8	239.4	79.6
Perf Comp	40.1	22.1	19.1	19.6
Flex Supply	7.8	5.0	5.4	5.3

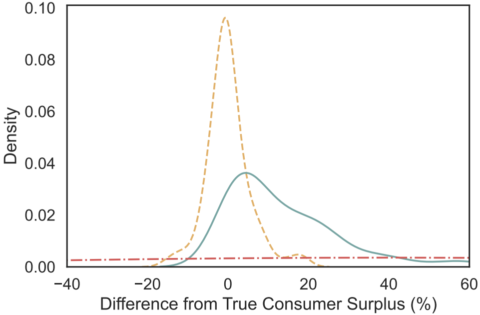
Figure 7: Product Exit



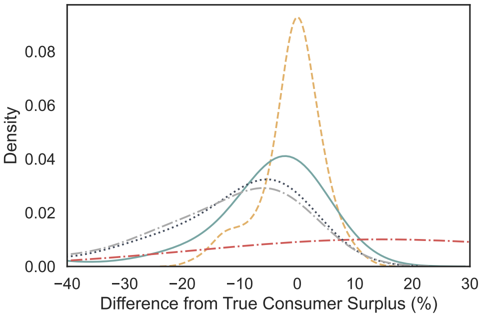
Product Exit with Economies of Scale

Figure 7: Product Exit

Panel C. Bertrand DGP, Economies of Scale



Panel D. Profit-Weight DGP, Econ. of Scale



Back

Multi-product Merger Simulation

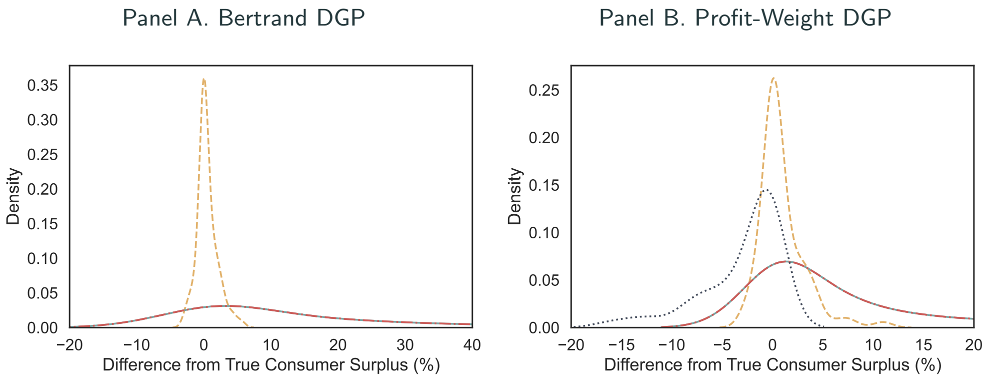
- Implementation: change ownership of one product

Table 8: RMPSE in Consumer Surplus for Mergers

Fitted Model	Panel DGPs			
	A. Bertrand	B. Profit-Weight	C. Bertrand (Scale)	D. Profit-Weight (Scale)
Bertrand (Scale)	--	--	--	4.8
Bertrand (Const.)	--	5.0	0.8	4.9
Monopoly	23.1	10.4	20.8	9.3
Perf Comp	23.1	10.4	20.8	9.3
Flex Supply	1.5	2.4	4.9	3.9

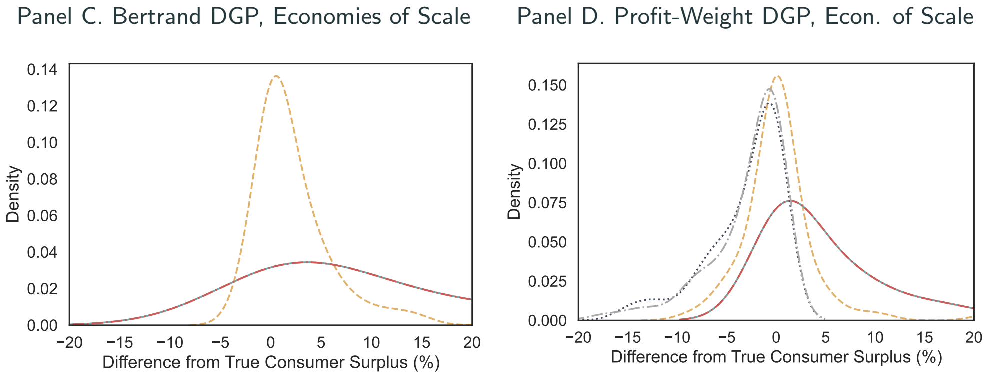
Multi-product Merger Simulation

Figure 8: Merger Simulation



Multi-product Merger Simulation with Economies of Scale

Figure 9: Merger Simulation



Back

Laffer Curves

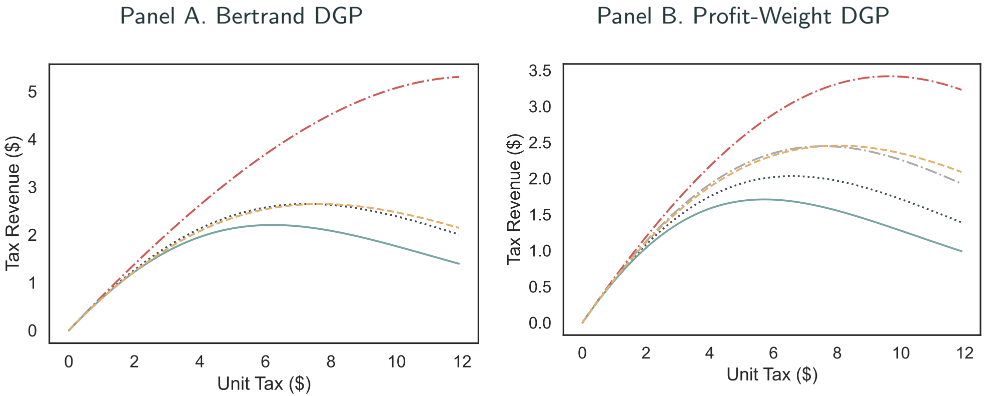
- Implementation: augment DGP with unit and ad valorem taxes, generate some variation in tax rate in training data

Table 9: MSE in Government Revenue for Laffer Curves

Fitted Model	Panel DGPs			
	A. Bertrand (Unit)	B. Profit-Weight (Unit)	C. Bertrand (AV)	D. Profit-Weight (AV)
Bertrand	--	0.13	--	0.27
Monopoly	2.88	0.54	16.67	2.04
Perf Comp	0.17	0.45	0.96	1.34
Flex Supply	0.00	0.00	0.03	0.05

Laffer Curves for Unit Taxes

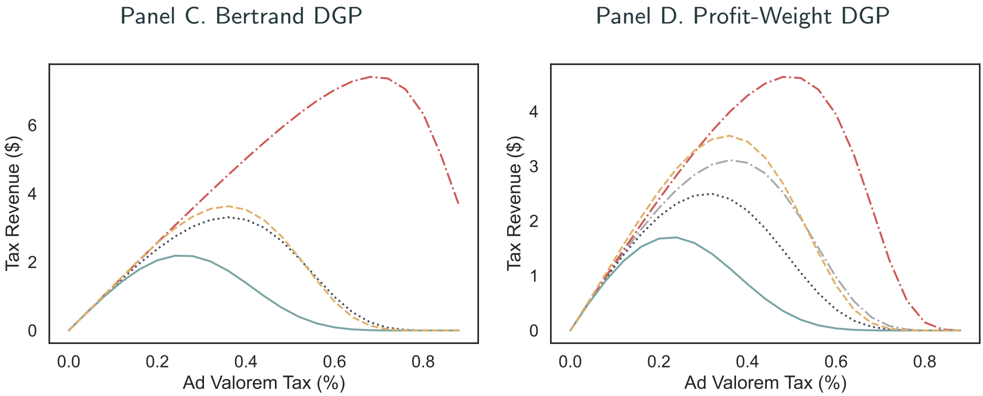
Figure 10: Laffer Curves



- In training data unit tax is $U[4, 8]$

Laffer Curves for Ad Valorem Taxes

Figure 11: Laffer Curves

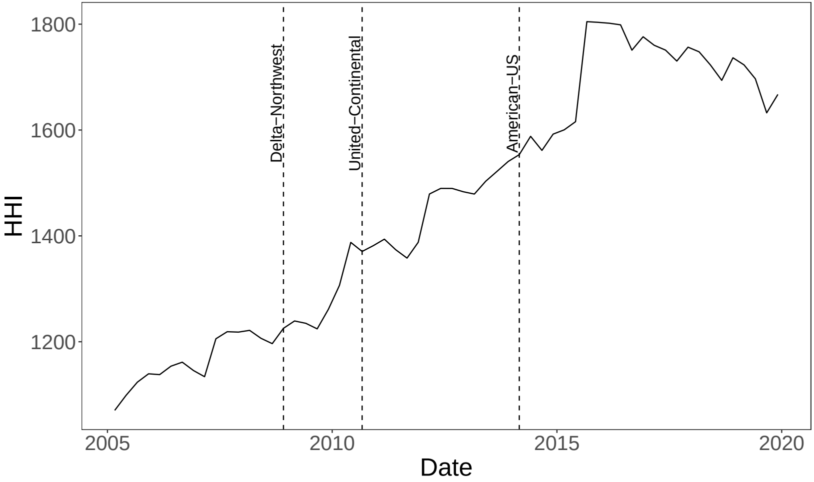


- In training data ad valorem tax is $U[0, 0.8]$

Back

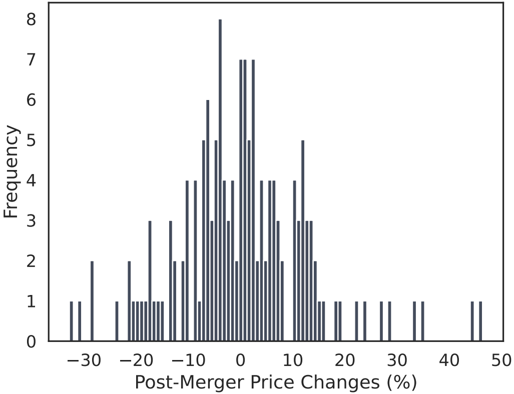
Airline Concentration Back

Figure 12: HHI in the Airline Industry



Observed Price Changes after AA-US Merger Back

Figure 13: Price Change Distribution



- Price changes after the AA-US merger in $3 \rightarrow 2$ markets

Demand Back

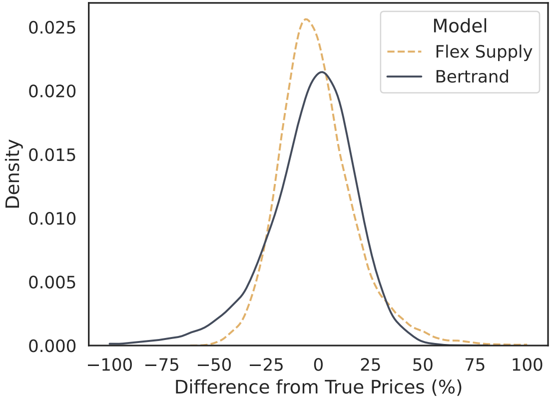
Table 10: Demand Estimates

	$\log(s_{jt})$ - $\log(s_{0t})$
Average Fare	-0.0048***
	(0.0004)
$\log(S_t)$	0.8356***
	(0.0133)
Share Nonstop	0.4030***
	(0.0282)
Average Distance (1,000's)	-0.4881***
	(0.0498)
Average Distance^2 (1,000's)	0.0485***
	(0.0045)
$\log(1 + \text{Num. Fringe})$	-0.2642***
	(0.0057)
R^2	0.94238
Observations	1,283,472
Own-price elasticity	-5.1652
Origin-destination fixed effects	✓

- Elasticities broadly in line with literature (e.g., Berry and Jia, 2010)

Fit: Pooled In-Sample and Out-of-Sample Results Back

Figure 14: Model Comparison



Model	Sample	MSE
Bertrand	All	1949.59
VMM	All	1242.95
Bertrand	Train	1932.70
VMM	Train	1235.39
Bertrand	Test	2016.33
VMM	Test	1272.82

- Reduction of $\sim 40\%$ in passenger-weighted MSE relative to Bertrand

Merger Simulation: Inference

Back

Figure 15: Width of Confidence Intervals