E-commerce and Regional Inequality:

A Trade Framework and Evidence from Amazon's Expansion

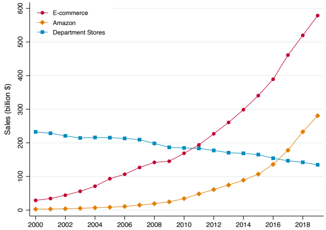
Elmer Zongyang Li

Department of Economics

Cornell University

April, 2023

Brick-and-motor vs. E-commerce



- Secular ↑ online retail sales (**e-commerce**)
- ”Opening to trade” challenges *regional equality*
  - Comparative advantages, worker specializations
  - Pre-existing trade & input-output linkages

This Paper

E-commerce as trade shock $\Rightarrow$ Spatial **GE** and **reallocation** (welfare, empl. dispersion)

Argument: **mobility** of online retailer $\Rightarrow$ $\uparrow$agglomeration, specialization

- **Empirics**: New facts on Amazon sales, retailers, facilities
  - Online retailer concentration, sales & trade
- **Theory**: multi-region & -sector (retail) trade model
  - Location choice of online retailer
  - Search & transport, elastic labor supply
- **Policy**: place-based public finances & online market design

Contribution: new data & extend trade theory $\Rightarrow$ e-commerce on regions

Related Literature

- E-commerce's Impact on Retail Market Structure
  - stores: ↓ demand ↑ product. mark-up. Stanchi(2019), Goldmanis et. al (2010)
  - consumer: ↑ welfare. Dolfen et. al (2021); Fan et. al (2018); Forman et. al (2009)

    This paper: spatial GE related to employment & real GDP
- Inter-national/regional Trade, Urban
  - LM effects under Ricardian trade model: Caliendo et. al (2018); Caliendo, Dvorkin & Parro (2019); Lee (2020); Galle, Rodríguez-Clare & Yi (2022)

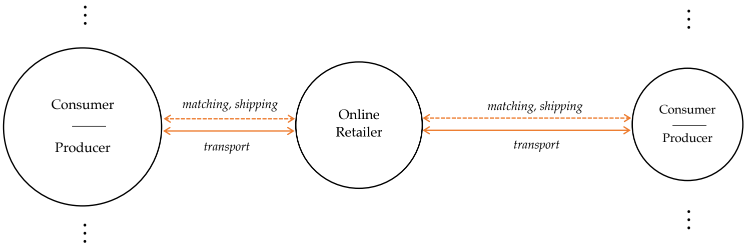
    This paper: apply & extend + new data & ID strategy
- Local LM shocks
  - trade & technology: Autor, Dorn & Hansen (2013), Firpo, Fortin & Lemieux (2013), Pierce & Schott (2016); Autor & Dorn (2013), , Chava et. al (2022)

    This paper: GE regional heterogeneity, welfare (beyond ATE)

Table of Contents

Empirical Patterns

The Online Retail Business Model



Agglomeration: Trade frictions (HME, CA) $\Leftrightarrow$ **Dispersion**: Factor prices

Predictions:

- E-commerce sales should be spatially more concentrated than BM
- Online retailers using FBA should be more concentrated
- Attraction of online retailers → more trade of upstream
- Loss of online retailers → less trade of upstream

Data Sources

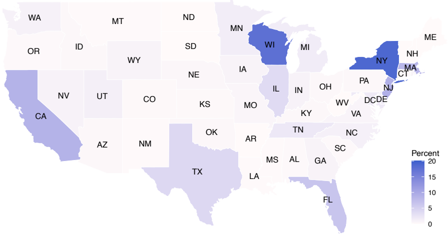
- Amazon Retailers and Products (Keepa.com)
  - A random sample of products on Amazon (36 categories, 2016-2020)
  - Information on prices, and sales ranking, converted to sales
  - Collect sellers' addresses, FBA status
- Amazon Facilities (MWPVL)
  - Addresses, square feet, date, type.[Houde, Newberry & Seim (HNS,2021)]
  - Focus on large fulfill. & distr. centers; drop specialized, small-package
- DOT Commodity Flow Survey (CFS)
  - Origin-destination data on trade value, volume, NAICS category
- Other Datasets
  - Surveys: CBP, BEA, ACS
  - Geography Datasets (topography, climate)

Empirical Patterns

- 1a: Online retail sales is more concentrated than average retail sales...

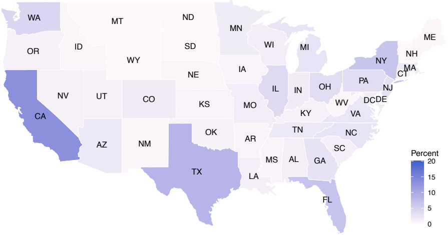
(1) Regional Share of E-commerce Sales

(HHI = 0.16)



(2) Regional Share of Retail Sector Sales

(HHI = 0.05)



Empirical Patterns

- 1a: Online retail sales is more concentrated than average retail sales...
- 1b: ...and is more strongly correlated with manufacturing, and less correlated with population

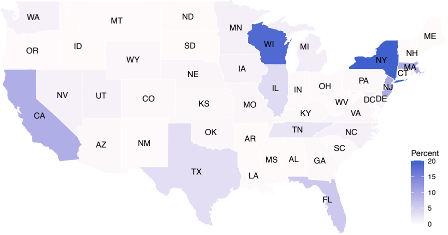
Table 1. E-commerce vs. Retail Sector Sales, Population, and Manufacturing

Dependent Variable:	E-commerce Sales Share (%)			Overall Retail Sales Shares (%)		
% Total population	0.92***		0.61	1.06***		1.08***
	[0.29]		[0.38]	[0.04]		[0.05]
Top manufacturing		4.27**	2.47		3.42***	-0.14
		[1.72]	[2.53]		[0.88]	[0.27]
Observations	255	306	255	255	306	255
R-squared	0.12	0.11	0.14	0.98	0.40	0.98

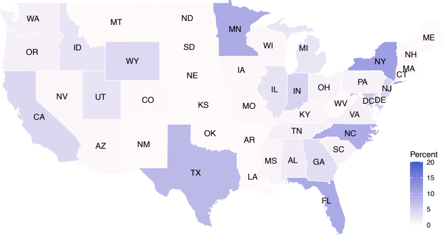
Empirical Patterns

- 2: Online retail sales that are FBA more concentrated than non-FBA

(3) Regional Share of E-commerce Sales with FBA
(HHI = 0.17)



(4) Regional Share of E-commerce Sales without FBA
(HHI = 0.13)



Empirical Patterns

- 3a: Destinations with ↑ online retailers import <u>more</u> WS goods
- 3b: Origins with ↑ online retailers export <u>less</u> WS goods

Dependent Variable:	Δln(Shipment)	ln(Shipment)
Share (%) of online sellers - destination	1.5*	
	[0.8]	
Share (%) of Online sellers - origin	-3.7***	
	[1.0]	
Bilateral distance via Amazon facility		-0.20**
		[0.08]
Origin FE	✓	✓
Destination FE	✓	✓
Year FE		✓
Industry FE	✓	✓
Observations	19,739	43,715
R-squared	0.2	0.4

Empirical Patterns

- 3a: Destinations with ↑ online retailers import more WS goods
- 3b: Origins with ↑ online retailers export less WS goods
- 4: Regions near to fulfillment centers import & export less WS goods

Dependent Variable:	Δln(Shipment)	ln(Shipment)
Share (%) of online sellers - destination	1.5*	
	[0.8]	
Share (%) of Online sellers - origin	-3.7***	
	[1.0]	
Bilateral distance via Amazon facility		-0.20**
		[0.08]
Origin FE	✓	✓
Destination FE	✓	✓
Year FE		✓
Industry FE	✓	✓
Observations	19,739	43,715
R-squared	0.2	0.4

A Spatial Trade Model

Summary

- Environment
  - N regions: n (destination), m (origin)
  - J sectors: j (durable, non-durable)
  - 3 subsectors: M (manufacturer), R (online retailer), B (brick-and-mortar)

1. **Demand**: Consumer search → CES with demand shifter
2. **Intermediate**: comparative advantages → manuf. trade flow
3. **Online Seller**: Location choice → agglomeration, retail trade flow
4. **Worker**: Roy labor supply

Consumer Search

- Set-up
  - A continuum of consumers (n), purchase retail & service (η^j)
  - For retail, pick 1 among $O+1$ sellers ($i=0$ BM; $i \geqq 1$ online)
  - $v_{ni}^j = \ln \eta^j y_n - \ln p_{ni}^j + \varepsilon_{ni}^j$ (ε_{ni}^j the match value)
    - ε_{ni}^j unknown: i.i.d. $E(\varepsilon_{ni}^j) = 0$ for $i = 0$ and $E(\varepsilon_{ni}^j) = ln(\mu)$ for $i \geq 1$
  - To find the seller, they search sequentially (SOM) Weitzman (1979)
    - each step, whether to pay cost k to observe $p_i^j, \varepsilon_{ni}^j$, or continue

1. Any SOM has a discrete choice model (DCM) w/. same demand proof
2. CES demand is a special case of DCM with extreme type I error proof

Theorem

A rep. consumer in n with weights η^j has nest CD-CES demand as below under sequential ordered search and if $\min\{\varepsilon_{ni}, r(k)\}$ is distributed extreme type I

$$C_n = \Pi_{j=1}^J (C_n^j)^{\eta^j}, \quad C_n^j = [(c_{n0})^{\frac{\sigma-1}{\sigma}} + \mu \sum_{i=1}^N (c_{ni})^{\frac{\sigma-1}{\sigma}}]^{\frac{\sigma_j}{\sigma_j - 1}} \text{ for } j \geq 2$$

Production

- Intermediate Varieties (M)
  - A rep. firm in (n, j, M) produces varieties $e^j \in [0,1]$

$$q_n^{j,M}(e^j) = a_n(e^j) l_n(e^j)$$

- Retail Sector (R/B)
  - Collect varieties $e^j \in [0,1]$: $q_n^{j,R/B} = [\int_0^1 q_n^{j,M}(e^j)^{\frac{\alpha^j - 1}{\alpha^j}} d\phi^j(a^n(e^j))]^{\frac{\alpha^j}{\alpha^j - 1}}$

$$Q_n^{j,R/B} = z_n^{j,R/B} \left[(h_n^{j,R/B})^{\beta_n} (l_n^{j,R/B})^{1-\beta_n} \right]^{\gamma_n^j} \left[q_n^{j,R/B} \right]^{1-\gamma_n^j}$$

  - Fréchet $\phi^j(a_n(e^j)) = exp(\sum_n -T_n^j z^{-\theta^j})$, exp. share: $x_{nm}^{j,M} = \frac{(\kappa_{nm}^M c_m^{j,M})^{-\theta^j} T_m^j}{\sum_{g=1}^N (\kappa_{ng}^M c_g^{j,M})^{-\theta^j} T_g^j}$
  - Unit cost: $c_n^{j,R/B} = (\omega_n^{j,R/B})^{\gamma_n^j} (p_n^{j,M})^{1-\gamma_n^j} / z_n^j$. For online: $c_{nm}^{j,R} = c_m^{j,R} \kappa_{nm}^R$

Online Retailer Location

- Optimal Location (R)
  - Online retailer char. by $(z_1^{j,R},...,z_N^{j,R})$, entry cost f_m. Optimal location:

$$m^* = argmin_m\{\frac{\tilde{\sigma}}{z_m^{j,R}} c_m^{j,R} \sum_n (\frac{\kappa_{nm}^R}{P_n^{j,R}})^{\sigma^j - 1} X_n\} \quad (\equiv \frac{\tilde{\sigma}\xi_m^j}{z_m^{j,R}})$$

  - Entry: $\sum_n (\frac{p_{nm}^{j,R}/\mu}{P_n^{j,R}})^{1-\sigma^j} \eta^j X_n \geq \sigma^j w_m^{j,R} f_m$. Thold: $\bar{c}_m^j = \frac{\mu}{\tilde{\sigma}^j}[\frac{\sigma^j}{\eta^j} \frac{w_m^{j,R} f_m}{\sum_n \kappa_{nm}^R P_n^{j,R} X_n}]^{\frac{1}{1-\sigma^j}}$
- Aggregate Retail Trade
  - Multi-variate Pareto : $P(Z_1^j < z_1,...,Z_N^j < z_N) = 1 - (\sum_{m=1}^N [A_{jm} z_m^{-\phi}]^{\frac{1}{1-\rho}})^{1-\rho}$

$$\Psi_m^j = P(m = argmin_m\{\tilde{\sigma}\xi_m^j / z_m^j\} \cap c_m^{j,R} < \bar{c}_m^j) = \psi_m^j (\bar{c}_m^j)^\phi$$

  - where $\psi_m^j = A_{jm}(\xi_m^j)^{\frac{-\phi}{1-\rho}} / \sum_{m=1}^N [A_{jm}(\xi_m^j)^{-\phi}]^{\frac{-\rho}{1-\rho}}$
  - Bilateral online retail exp. share $x_{nm}^{j,R} = \frac{\psi_m^j (\bar{c}_m^j)^\phi (c_{nm}^{j,R}/\mu)^{1-\sigma}}{\sum_h \psi_h^j (\bar{c}_h^j)^\phi (c_{nh}^{j,R}/\mu)^{1-\sigma} + \frac{1}{O}(c_{n0}^{j,R})^{1-\sigma}}$

Labor Supply

- Employment Rate
  - L_n HHs choose sector of employment (includ. home production)
  - Draw $z_n = \{z_n^0, z_n^S, z_n^{2,M}, z_n^{2,R}, z_n^{2,B}, z_n^{3,M}, z_n^{3,R}, z_n^{3,B}\}$ from i.i.d. Fréchet $(v_n, A_n^{j,K})$
  - Probability of working in $\{j, K\}$:

$$\pi_n^{j,K} = \frac{A_n^{j,K}(w_n^{j,K})^{v_n}}{\Phi_n}, \text{ where } \Phi_n = \sum_{j=0}^{J} \sum_{K=M,R,B,\emptyset} A_n^{j,K}(w_n^{j,K})^{v_n}.$$

E-commerce and General Equilibrium Outcomes

- General Equilibrium
  - market clearing details
  - comparative statics w/. *exact hat algebra* details
- E-commerce
  - $\uparrow$ match efficiency μ (Dinerstein et. al 2018; Goldmanis et. al 2010)
  - $\downarrow$ transportation cost of retail κ_{ni}^{R} (Houde, Newberry & Seim 2021)
- Welfare
  - Definition: real income per capita $W_n = \frac{Y_n/L_n}{P_n}$, its change:

$$\hat{W}_n = \underbrace{\hat{w}_n^0 (\hat{\pi}_n^0)^{\frac{-1}{\nu_n}}}_{\text{non-emp. worker special.}} \times \Pi_{j=1}^{J} \underbrace{(\hat{x}_{n0}^{j,R})^{\frac{-\eta_j}{\sigma^j - 1}}}_{\text{industry composition}} \underbrace{(\hat{c}_n^{j,B})}_{\text{input-output local pref.}}$$

Quantification

Estimation Outline

Section	Param.	Description	Estimation/Caliberation
Consumer	η_n^j	Sector share of consumption	CFS 2007
	σ^j	Elasticity of subs. across retailers	Keepa + IV
	$w_n^{j,K} L_n^{j,K}$	Sectoral income	BEA, ACS, CFS 2007
Labor Supply	π_n^j	Share of empployment	CBP, ACS
	ν^n	Fréchet shape of worker product.	Galle, Rodríguez-Clare & Yi (2022)
Production	β_n^j	Share of structures	BEA + Greenwood et. al (1997)
	θ^j	Fréchet shape of sector product.	Caliendo and Parro (2015)
	γ_n^j	Value-added share of retail goods	BEA, CFS
Trade	$x_{ni}^{M,j}$	Interm. expenditure share	CFS 2007
Amazon Shock	$\hat{\kappa}_{nm}^R$	Iceberg cost change	Amazon data + CFS 2007 + IV
	μ	Matching efficiency	E-stats + CES
	Ψ_m^j	Online retailer location probability	Keepa
	O	Measure of online retailers	E-stats
	T_n^j	Fréchet scale of sectoral product.	Assume constant
	A_n^j	Fréchet scale of labor product.	Assume constant

Estimation: Amazon Shock

- Extrapolate Amazon Ice-berg cost shock
  - **Intuition**: Ice-berg is increasing in distance
  - Estimate coefficient of ice-berg cost on shipping distance details

$$ln(\kappa_{nm}^{j,R}) = \delta^{j}\text{Distance}_{nm} + X_{nm}'\theta + \delta_n^j + \delta_m^j + \varepsilon_{nm}^j$$

  - Estimate reduction in shipping distance due to Amazon
    - Build counterfactual facilities based on exog. factors as IV for actual ones
- Back-out online matching efficiency
  - **Intuition**: % online exp. should inform matching, conditional on shipping

$$\sum_{m=1}^{N} x_{nm}^{j,R}/x_{n0}^{j,R} = (\mu)^{\sigma^j - 1} \sum_{m=1}^{N} M_m (p_m^{j,R} \kappa_{nm}^{R} / p_{n0}^{j,R})$$

    - Use Keepa for M_m, above estimated κ_{nm}^R, CES for $p_m^{j,R}, p_{n0}^{j,R}$

δ^{dur}	δ^{nondur}	$\widehat{\kappa}$	μ
1.5	2.1	0.97	1.27
[0.2]	[0.6]	[0.15]	[1.46]

Results

Result: Welfare

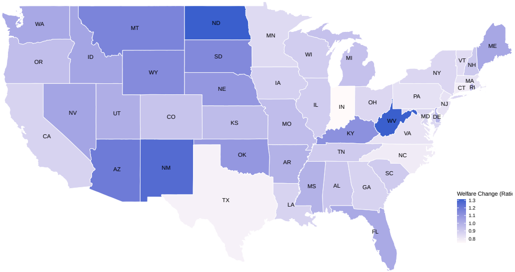


Figure: Total Welfare Change

- ↓ **welfare** overall (avg: 2.3%)
  - driver: price effect +10%, income effect -12%
- ↑ **inequality**
  - Gini: 0.11→0.13 (+18%)

Result: Employment

Sector	Mean	Std. Dev
Manufacturing	0.35	0.04
Online Retail	0.70	0.15
Brick-and-Mortar	0.70	0.15
Service	1.59	0.16
Unemployment	0.91	0.10

Table: Employment Changes by Sector

Conclusion

- E-commerce as trade shock
- Online retailer concentration plays an important role
- Spatial retail trade model
- Amazon $\Rightarrow$ regional inequality (welfare, employment)
- Place-based public finances & online market design

Thank you!

Elmer Zongyang Li
zl685@cornell.edu

Appendix

Weitzman (1979) back

Search is *ordered*: Weitzman (1979) optimal stopping

- Assign thresholds/scores $\bar{v}_i$ st. $E[\max\{\hat{x}_i + \tilde{\varepsilon}_i - \bar{v}_i, 0\}] = 0$, where $\hat{x}_i = \ln y - \ln p_i$
- Therefore, $\bar{v}_i = \hat{x}_i + \gamma_{\varepsilon_i}^{-1}(\ln s_i)$, where $\gamma_{\varepsilon_i}(z) = E[\max\{\varepsilon_i - z, 0\}]$, decreasing function
- Search in decreasing order of the scores
- Stop if find a $\bar{v}_i$ exceeding all remaining

Proof of OSM to DCM back

Proposition: For any OSM, there is a DCM with same demand & payoff.

- Under OSM, consumer's optimal choice is the one for which $v_i^* = \min\{v_i, \bar{v}_i\}$ is largest (Armstrong and Vickers (2015), Armstrong(2017), Choi, Dai and Kim(2018)), where $\bar{v}_i = \hat{x}_i + \gamma_{\varepsilon_i}^{-1}(\ln \mu_i) = \hat{x}_i + r(\ln \mu_i)$, and $\gamma_{\varepsilon_i}(z) = E[\max\{\varepsilon_i - z, 0\}]$, the upside gain function
- Consumer's demand for i, D_i is thus:
  $$P[v_i^* > \max_{j \neq i} v_j^*] = \int_{-\infty}^{\infty} P[z > \max_{j \neq i} v_j^*] f_{v_i^*}(z; x_i, \hat{x}_i) dz = \int_{-\infty}^{\infty} \Pi_{j \neq i} F_{v_j^*}(z; x_j, \hat{x}_j) f_{v_i^*}(z; x_i, \hat{x}_i) dz.$$
- Under advertised price, $x_j = \hat{x}_j, \forall j$. D_i then simplifies to
  $$\int_{-\infty}^{\infty} \Pi_{j \neq i} F_{\omega_j}(\varepsilon_j) f_{\omega_i}(\varepsilon_i) d\varepsilon, \text{ where } \omega_i = \min\{\varepsilon_i, r(\ln \mu_i)\}.$$
  Thus, D_i is equivalent to the demand of a DCM: $v_i = x_i + \varepsilon_i^{DC}$, iff

Proof of DCM to CES back

Proposition: The CES demand is a special case of DCM with extreme type I error.

The following proof follows Anderson, De Palma, and Thisse (1987, 1989) closely

- Consumer's utility $u_i = \ln c_i$, income y. Let price of i: $\tilde{p}_i = \mu_i p_i$
- Random utility/match value ε_i with i, st. net value: $v_i = \ln y - \ln \tilde{p}_i + \varepsilon_i^{DC}$

  Further, re-scale $\varepsilon_i^{DC} = \chi \tilde{\varepsilon}_i$ st. $\tilde{\varepsilon}_i$ mean 0 and unit variance
- The demand for i, D_i is then

$$P[v_i > \max_{j \neq i} v_j] = \int_{-\infty}^{\infty} \Pi_{j \neq i} F_{\varepsilon_j^{DC}}(\varepsilon_j^{DC}) f_{\varepsilon_i^{DC}}(\varepsilon_i^{DC}) d\varepsilon.$$

- And if $\tilde{\varepsilon}_i$ is distributed extreme type I, D_i then simplifies to

$$D_i = \frac{\mu_i p_i^{-1/\chi}}{\sum^n \ldots^{-1/\chi}},$$

Market Clearing Conditions back

- Retail and intermediate goods:

$$X_n^{R,j} = \sum_{i=1}^{N} x_{in}^{R,j}(I_i L_i), \text{where } I_i L_i = \sum_{k=0}^{J} [r_i^{g,k} g_i^{R,k} + \sum_{K=M,R} (r_i^{h,k} h_i^{K,k} + w_i^k l_i^{K,k})] - \Omega_i,$$

$$X_n^{M,j} = \sum_{i=1}^{N} (1-\gamma_i^j) x_{in}^{M,j} X_i^{R,j}.$$

- Trade balance:

$$\sum_{j=0}^{J}\sum_{i=1}^{N}(x_{ni}^{M,j} X_n^{M,j} + x_{ni}^{R,j} X_n^{R,j}) + \Omega_n = \sum_{j=0}^{J}\sum_{i=1}^{N}(x_{in}^{M,j} X_i^{M,j} + x_{in}^{R,j} X_i^{R,j}).$$

- Labor market: $w_n^{M,j} l_n^{M,j} = \beta_n X_n^{M,j}, \; w_n^{R,j} l_n^{R,j} = \gamma_n^j m_n^{R,j} \beta_n X_n^{R,j}$
- Structure: $r_n^h h_n^{M,j} = (1-\beta_n) X_n^{M,j}, \; r_n^h h_n^{R,j} = \gamma_n^j \frac{1}{\rho_n^{R,j}} (1-\beta_n) X_n^{R,j}$
- Capital: $r_n^g g_n^{R,j} = (\frac{\rho_n^j - 1}{1-\beta_n}) w_n^{R,j} \pi_n^{R,j} L_n$

Comparative Statics back

- Employment shares:

$$\hat{\pi}_n^0 = \frac{\hat{A}_n^0 (\hat{w}_n^0)^{\nu_n}}{\hat{\Phi}_n}, \; \hat{\pi}_n^{K,j} = \frac{\hat{A}_n^{K,j} (\hat{w}_n^{K,j})^{\nu_n}}{\hat{\Phi}_n}, \text{ where } \hat{\Phi}_n = \sum_{h=0}^{J} \sum_{K=M,R} \pi_n^{K,h} \hat{A}_n^{K,h} (\hat{w}_n^{K,h})^{\nu_n}.$$

- Input costs: $\hat{c}_n^{M,j} = \hat{\omega}_n^{M,j}, \quad \hat{c}_n^{R,j} = (\hat{\rho}_n^{R,j} \hat{\omega}_n^{R,j})^{\gamma_n^j} (\hat{P}_n^{M,j})^{1-\gamma_n^j}$, where

$$\hat{\omega}_n^{K,j} = \hat{w}_n^{K,j} (\hat{l}_n^{K,j})^{\beta_n} = (\hat{w}_n^{K,j})^{1+\beta_n} (\hat{\pi}_n^{K,j})^{\frac{(\nu_n - 1)\beta_n}{\nu_n}}, \text{and } \hat{P}_n^{M,j} = \Big(\sum_{i=1}^{N} x_{ni}^{M,j} (\hat{\kappa}_{ni}^M \hat{c}_i^{M,j})^{-\theta^j} \hat{T}_i^j \Big)^{\frac{-1}{\theta^j}}.$$

- Trade shares: $x_{ni}^{\prime M,j} = x_{ni}^{M,j} \Big(\frac{\hat{\kappa}_{ni}^M \hat{c}_i^{M,j}}{\hat{P}_n^{R,j}}\Big)^{-\theta_j} \hat{T}_i^j, \quad x_{ni}^{\prime R,j} = x_{ni}^{R,j} \Big(\frac{\hat{\kappa}_{ni}^R \hat{c}_i^{R,j}}{\hat{\mu}_{ni}^j \hat{P}_n^{R,j}}\Big)^{1-\sigma^j}.$
- Market clearing:

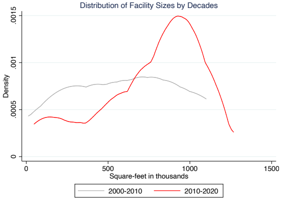
$$X_n^{\prime R,j} = \sum_{i=1}^N x_{in}^{\prime R,j} \eta^j \left[\sum_{k=0}^J \Big(\frac{1}{1-\beta_i}\Big) (\hat{\rho}_i^{R,k} \hat{w}_i^{R,k} \hat{l}_i^{R,k} \rho_i^{R,k} w_i^{R,k} L_i^{R,k} + \hat{w}_i^{M,k} \hat{l}_i^{M,k} w_i^{M,k} L_i^{M,k}) - \Omega_i \right],$$

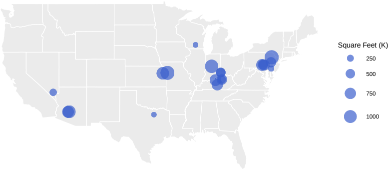
$$X_n^{\prime M,j} = \sum_{i=1}^{N} (1 - \gamma_i^j) x_{ni}^{\prime M,j} X_n^{\prime R,j},$$

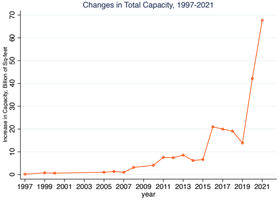
$$\hat{w}_n^{M,j} \hat{l}_n^{M,j} w_n^{M,j} L_n^{M,j} = \beta_n \hat{X}_n^{M,j}, \quad \hat{w}_n^{R,j} \hat{l}_n^{R,j} w_n^{R,j} L_n^{R,j} = \frac{1}{\;} \gamma_n^j \beta_n \hat{X}_n^{R,j}$$

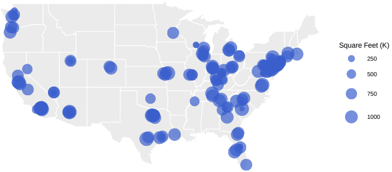
Estimation: Amazon Transportation Shock back

- Data: Amazon's Facility Network
  - address, square feet, date, type.[Houde, Newberry & Seim (HNS,2021)]
  - focus on large fulfill. & distr. centers; drop specialized, small-package









Estimation: Amazon Transportation Shock back

- Need to specify how:

$$\text{origin} \rightarrow \text{facility} \rightarrow \text{destination}$$

- HNS (2021): 90% of orders from 3 closest centers to dest.
- Assume order is processed by among the 3 closest to destination, the closest to origin

Year	Mean	Std. Dev	P25	P75
2007	490.2	376.3	234.9	739
2017	287.9	225.6	124.7	409
Diff.	-202.2	295.6	-249.8	-12.5
Log Diff.	-.5	.6	-.9	0

Estimation: Amazon Transportation Shock back

- Spatial Simulated IV
    - concern: endogeneity of facilities
    - simulate facilities' locations based only on geo. cost factors, to be uses as IV (Duflo et.al, 2007; Lipscomb et.al, 2013; Faber 2014)
    - need orthogonality of geo. factors

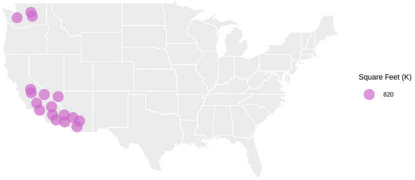
Dependent 1{AMZ Center}		
Temperature (Lag)	Mean	-0.011
		[0.018]
	Minimum	-0.002
		[0.009]
	Maximum	0.046***
		[0.012]
Precipitation (Lag)	Mean	-0.032
		[0.040]
	Minimum	0.043
		[0.044]
	Maximum	-0.015
		[0.013]
Elevation	Mean	-0.001***
		[0.000]
	Minimum	0.000
		[0.000]
	Maximum	0.001***
		[0.000]
Tornado	Magnitude	-0.051
		[0.086]
	Injuries	-0.110
		[0.153]
County, Year FE	X	
Observations	55,259	
Psudo R-squared	0.1663	

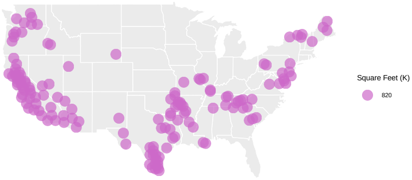
Estimation: Amazon Transportation Shock back

- Spatial Simulated IV
  - concern: endogeneity of facilities
  - simulate facilities' locations based only on geo. cost factors, to be uses as IV (Duflo et.al, 2007; Lipscomb et.al, 2013; Faber 2014)
  - need orthogonality of geo. factors
- Simulation Steps
  - based on observed # of new centers, determine AMZ's budget
  - rank counties by geo. factors
  - highest ranks get new centers

Dependent 1{AMZ Center}		
Temperature (Lag)	Mean	-0.011
		[0.018]
	Minimum	-0.002
		[0.009]
	Maximum	0.046***
		[0.012]
Precipitation (Lag)	Mean	-0.032
		[0.040]
	Minimum	0.043
		[0.044]
	Maximum	-0.015
		[0.013]
Elevation	Mean	-0.001***
		[0.000]
	Minimum	0.000
		[0.000]
	Maximum	0.001***
		[0.000]
Tornado	Magnitude	-0.051
		[0.086]
	Injuries	-0.110
		[0.153]
County, Year FE	X	
Observations	55,259	
Psudo R-squared	0.1663	

Estimation: Amazon Transportation Shock back





Year	Mean	Std. Dev	P25	P75	Corr
2007	623.4	400.3	349.6	897.4	0.10
2017	335.2	278.4	143.9	412.1	0.58
Diff.	-288.2	361.8	-355.9	0	-0.22
Log Diff.	-.7	.8	-1.1	0	-0.02

	Dependent Variables	
	Actual log distance	Counterfactual log distance
First Stage Results		
Counterfactual log distance	0.399***	
	[0.015]	
F-Stats	670	
Robustness		
Avg. lag GDP		0.000
		[0.000]
Avg. GDP growth		-0.004***
		[0.001]