Economies of Density in E-Commerce: A Study of Amazon's Fulfillment Center Network

Jean-François Houde[1] Peter Newberry[2] Katja Seim[3]

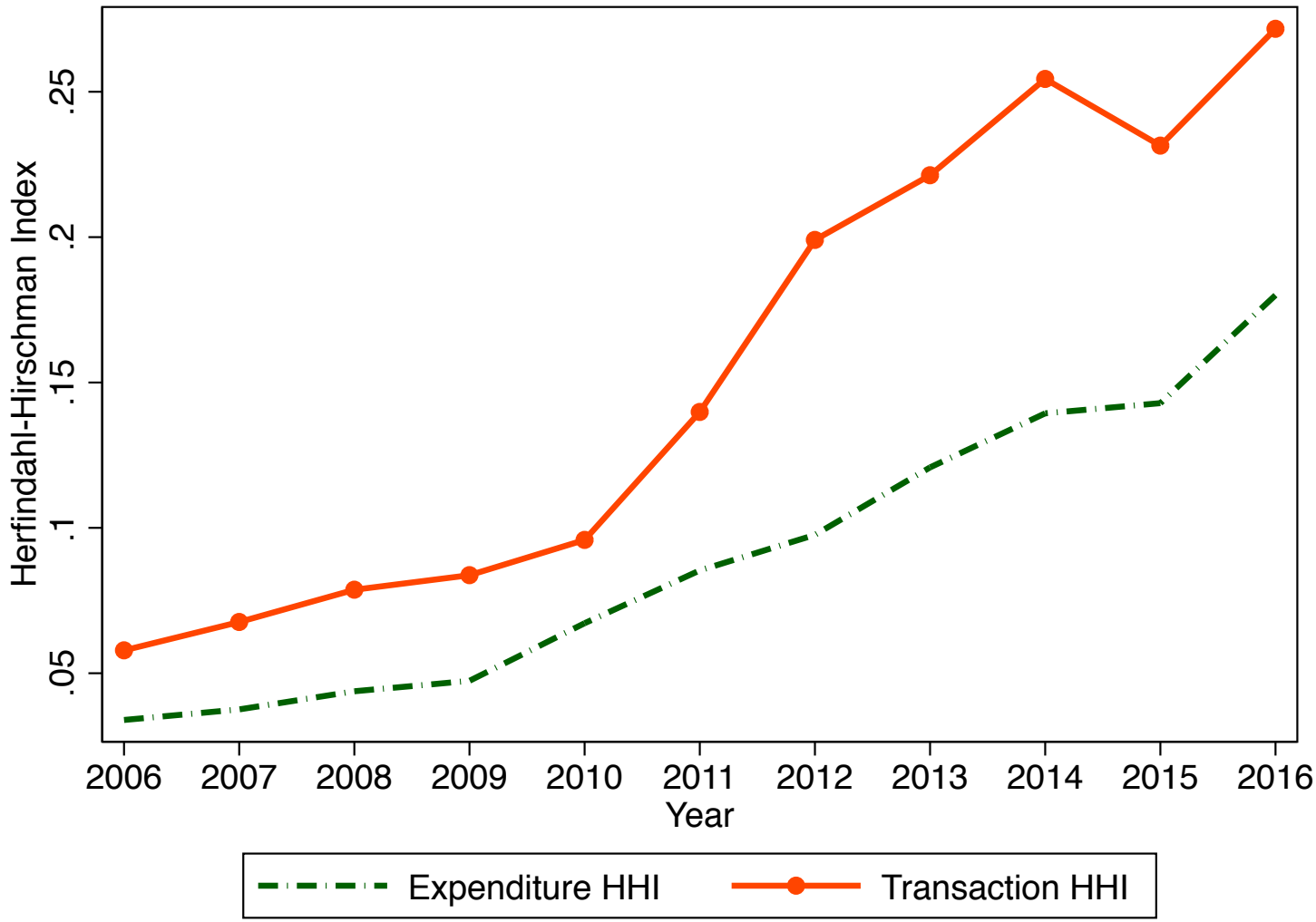
[1]Department of Economics
Cornell University & NBER

[2]Department of Economics
Pennsylvania State University

[3]The Wharton School
University of Pennsylvania & NBER

Postal Economics Conference March 29, 2018

Concentration in Online Retail



Source: ComScore Web Behavior Database.

Research Question

- Expenditure concentration: HHI of 300 to 1,300.
  - Amazon's sales growth: US revenue of $5bn to $80bn.
  - Growth and decentralization in Amazon's distribution network: 8 fulfillment centers (FCs) in 6 states to over 100 FCs in 28 states.

Research Question

- Expenditure concentration: HHI of 300 to 1,300.
  - Amazon's sales growth: US revenue of $5bn to $80bn.
  - Growth and decentralization in Amazon's distribution network: 8 fulfillment centers (FCs) in 6 states to over 100 FCs in 28 states.
- What is the source of Amazon's scale advantage?
  - Platform effects: product variety + reputation + one-stop shopping.
  - Economies of density in distribution:
    - ★ Shipping times: willingness-to-pay for convenience.
    - ★ Shipping costs: declining distribution costs.

Cost/Benefit of a Dense Network

- Benefits of a centralized network:
  - Low fixed-costs + Return to scale + Save on sales taxes

Cost/Benefit of a Dense Network

- Benefits of a centralized network:
  - Low fixed-costs + Return to scale + Save on sales taxes
- Cost of a centralized network:
  - Heavy reliance on third-party suppliers: Transportation and sorting
  - **Shared asset:** Steep premium/delays during congested periods

Cost/Benefit of a Dense Network

- Benefits of a centralized network:
  - Low fixed-costs + Return to scale + Save on sales taxes
- Cost of a centralized network:
  - Heavy reliance on third-party suppliers: Transportation and sorting
  - **Shared asset:** Steep premium/delays during congested periods

Cost/Benefit of a Dense Network

- Benefits of a centralized network:
  - Low fixed-costs + Return to scale + Save on sales taxes
- Cost of a centralized network:
  - Heavy reliance on third-party suppliers: Transportation and sorting
  - **Shared asset:** Steep premium/delays during congested periods
- **Bottom line:** Building a dense network reduces reliance on suppliers for the first leg (especially planes, less competitive) + permits (partial) vertical integration into sortation segment
  - Reduce risk of delays and lower shipping cost

This Paper

- **Objective:** Quantify the demand (convenience) and cost (shipping) effects of network expansion

This Paper

- **Objective:** Quantify the demand (convenience) and cost (shipping) effects of network expansion
- **Step 1:** Estimate demand for Amazon
  - Price/tax elasticity: entry into new state causes loss in revenue due to new sales tax liability for in-state customers.
  - Convenience elasticity: marginal disutility of shipping speed (proxied by distance and other measures).
  - Controls: value of online channel, platform "quality", relative prices of different channels, offline competition.

This Paper

- **Objective:** Quantify the demand (convenience) and cost (shipping) effects of network expansion
- **Step 1:** Estimate demand for Amazon
  - Price/tax elasticity: entry into new state causes loss in revenue due to new sales tax liability for in-state customers.
  - Convenience elasticity: marginal disutility of shipping speed (proxied by distance and other measures).
  - Controls: value of online channel, platform "quality", relative prices of different channels, offline competition.
- **Step 2:** Estimate cost saving from decentralized network
  - Revealed preference approach
  - Specify profit as a function of variable shipping and local fixed costs
  - Quantify cost savings that rationalize observed FC network without explicitly solving optimal roll-out problem (à la Holmes 2011).

Data: FC Network, Distances and Controls

- Fulfillment center information from MWPVL, International:
  - Observed and planned FCs between 2002 to 2018
  - Information: location, opening date, type, size, and employees
- Shipping distances and delivery time:
  - *Shipping distance:* Straight-line distance from county centroid to closest FC location
    - ⋆ **Assumption:** Shipments come from *nearest* FC (relaxed somewhat in robustness specs.)
    - ⋆ Also proxies for delivery time.
  - *Expected delivery time:* Shortest delivery time between county centroid to FC location for USPS 4 mail classes (# days)
    - ⋆ **Assumption:** Shipments come from FC with shortest time.
  - *(NEW) Prime Same/Next day:* Indicator variable for availability.
- Other controls:
  - County-level demographics, offline competition, and wages [Census]; warehouse rental rate [SNL, Inc].

Sales Taxes and Online Spending

Sales tax:

- Source: County/year average taxes from TDS
- Date of change in Amazon's tax status for each state.
  - Tax Nexus:
    - ★ *Legal definition:* Retailers with "sufficient physical presence" in a state must collect and pay tax on sales in that state
    - ★ **Implication:** Give competitive advantage to online retailers

Household Expenditures:

- *ComScore Web Behavior Database:* online purchasing behavior for 50-100k households each year from 2006-2013 (NEW: 2014-2016).
- *Forrester Technographics Survey:* Conditional probability of buying online from 2006-2013 (except 2008-2009, NEW: 2014-2016)).
- *CEX:* Average retail spending (offline)

Demand Model

- Estimate CES demand model across 4 shopping modes: offline, Amazon, taxed online retailer, and non-taxed online retailer.

Demand Model

- Estimate CES demand model across 4 shopping modes: offline, Amazon, taxed online retailer, and non-taxed online retailer.
- Key Parameters: price (tax) elasticity and convenience effect.

Demand Model

- Estimate CES demand model across 4 shopping modes: offline, Amazon, taxed online retailer, and non-taxed online retailer.
- Key Parameters: price (tax) elasticity and convenience effect.
- Control for mode-year effects (prices, national changes in convenience, variety, etc.)

Demand Model

- Estimate CES demand model across 4 shopping modes: offline, Amazon, taxed online retailer, and non-taxed online retailer.
- Key Parameters: price (tax) elasticity and convenience effect.
- Control for mode-year effects (prices, national changes in convenience, variety, etc.)
- Tax elasticity around -1.4:
  - Compared to Einav et al (2014) $[-1.7]$, and Baugh et al (2015) $[-1.2, -1.4]$.
  - Going from 0% to 6.5% tax rate $\rightarrow$ reduction in demand by 9.1%.
  - Robust to allowing Amazon's elasticity to differ from other modes; alternative ways of constructing representative consumer's spending measure; county-year fixed effects; time-varying tax effects. ▸ Robustness.

Demand Model

- Estimate CES demand model across 4 shopping modes: offline, Amazon, taxed online retailer, and non-taxed online retailer.
- Key Parameters: price (tax) elasticity and convenience effect.
- Control for mode-year effects (prices, national changes in convenience, variety, etc.)
- Tax elasticity around -1.4:
  - Compared to Einav et al (2014) $[-1.7]$, and Baugh et al (2015) $[-1.2, -1.4]$.
  - Going from 0% to 6.5% tax rate $\rightarrow$ reduction in demand by 9.1%.
  - Robust to allowing Amazon's elasticity to differ from other modes; alternative ways of constructing representative consumer's spending measure; county-year fixed effects; time-varying tax effects. ▸ Robustness.
- Convenience effect:
  - Distance to FC does not impact demand.
  - Not a good measure of shipping times?
  - Sameday/nextday dummy variables not significant.
  - **Takeaway:** Expansion of FCs increased convenience at a national level.

Quantifying Cost Savings

- **Goal:** Use estimated demand and observed distribution network over time to infer cost savings realized from densification of FC network

Quantifying Cost Savings

- **Goal:** Use estimated demand and observed distribution network over time to infer cost savings realized from densification of FC network
- **Challenge:** Solving fully specified, dynamic model of network roll-out computationally challenging **and** requires knowing the *end game* (still playing..)

Quantifying Cost Savings

- **Goal:** Use estimated demand and observed distribution network over time to infer cost savings realized from densification of FC network
- **Challenge:** Solving fully specified, dynamic model of network roll-out computationally challenging **and** requires knowing the *end game* (still playing..)
- **Approach:** Revealed-Preference (Holmes, 2011)
  - Compare discounted profit stream under observed and alternative roll-outs

Quantifying Cost Savings

- **Goal:** Use estimated demand and observed distribution network over time to infer cost savings realized from densification of FC network
- **Challenge:** Solving fully specified, dynamic model of network roll-out computationally challenging **and** requires knowing the *end game* (still playing..)
- **Approach:** Revealed-Preference (Holmes, 2011)
  - Compare discounted profit stream under observed and alternative roll-outs
  - Perturbed roll-out = Swap opening dates of two FCs.
  - **Advantage:** Profit comparison does not rely on post-sample continuation values

Revealed Preference Tradeoff

- Expanding FC network:
  - Revenue change: Taxes
  - Cost change: Fixed operating cost and shipping cost

Revealed Preference Tradeoff

- Expanding FC network:
  - Revenue change: Taxes
  - Cost change: Fixed operating cost and shipping cost
- **Scenario 1:** Build in low-tax/low-pop county; rel. to high density
  - Net gain in revenue, wages and rents ($R^* - R > 0$)
  - Shipping distance increase ($D^* - D' > 0$)

  $\Rightarrow$ Upper bound on $\$x$ incurred per unit of distance

$$R^* - \theta D^* > R' - \theta D' \Rightarrow \theta \leq (R^* - R')/(D^* - D')$$

Revealed Preference Tradeoff

- Expanding FC network:
  - Revenue change: Taxes
  - Cost change: Fixed operating cost and shipping cost
- **Scenario 1:** Build in low-tax/low-pop county; rel. to high density
  - Net gain in revenue, wages and rents $(R^* - R > 0)$
  - Shipping distance increase $(D^* - D' > 0)$

  $\Rightarrow$ Upper bound on $\$x$ incurred per unit of distance

$$R^* - \theta D^* > R' - \theta D' \Rightarrow \theta \leq (R^* - R')/(D^* - D')$$

- **Scenario 2:** Build in high-tax/high-pop county; rel. to low density
  - Net losses in revenue, wages and rents $(R^* - R' < 0)$
  - Shipping distance reduction $(D^* - D' < 0)$

  $\Rightarrow$ Lower bound on $\$x$ saved per unit of distance

$$R^* - \theta D^* > R' - \theta D' \Rightarrow \theta \geq (R' - R^*)/(D' - D^*)$$

Revealed Preference Tradeoff

- Expanding FC network:
  - Revenue change: Taxes
  - Cost change: Fixed operating cost and shipping cost
- **Scenario 1:** Build in low-tax/low-pop county; rel. to high density
  - Net gain in revenue, wages and rents ($R^* - R > 0$)
  - Shipping distance increase ($D^* - D' > 0$)

  $\Rightarrow$ Upper bound on $\$x$ incurred per unit of distance

$$R^* - \theta D^* > R' - \theta D' \Rightarrow \theta \leq (R^* - R')/(D^* - D')$$

- **Scenario 2:** Build in high-tax/high-pop county; rel. to low density
  - Net losses in revenue, wages and rents ($R^* - R' < 0$)
  - Shipping distance reduction ($D^* - D' < 0$)

  $\Rightarrow$ Lower bound on $\$x$ saved per unit of distance

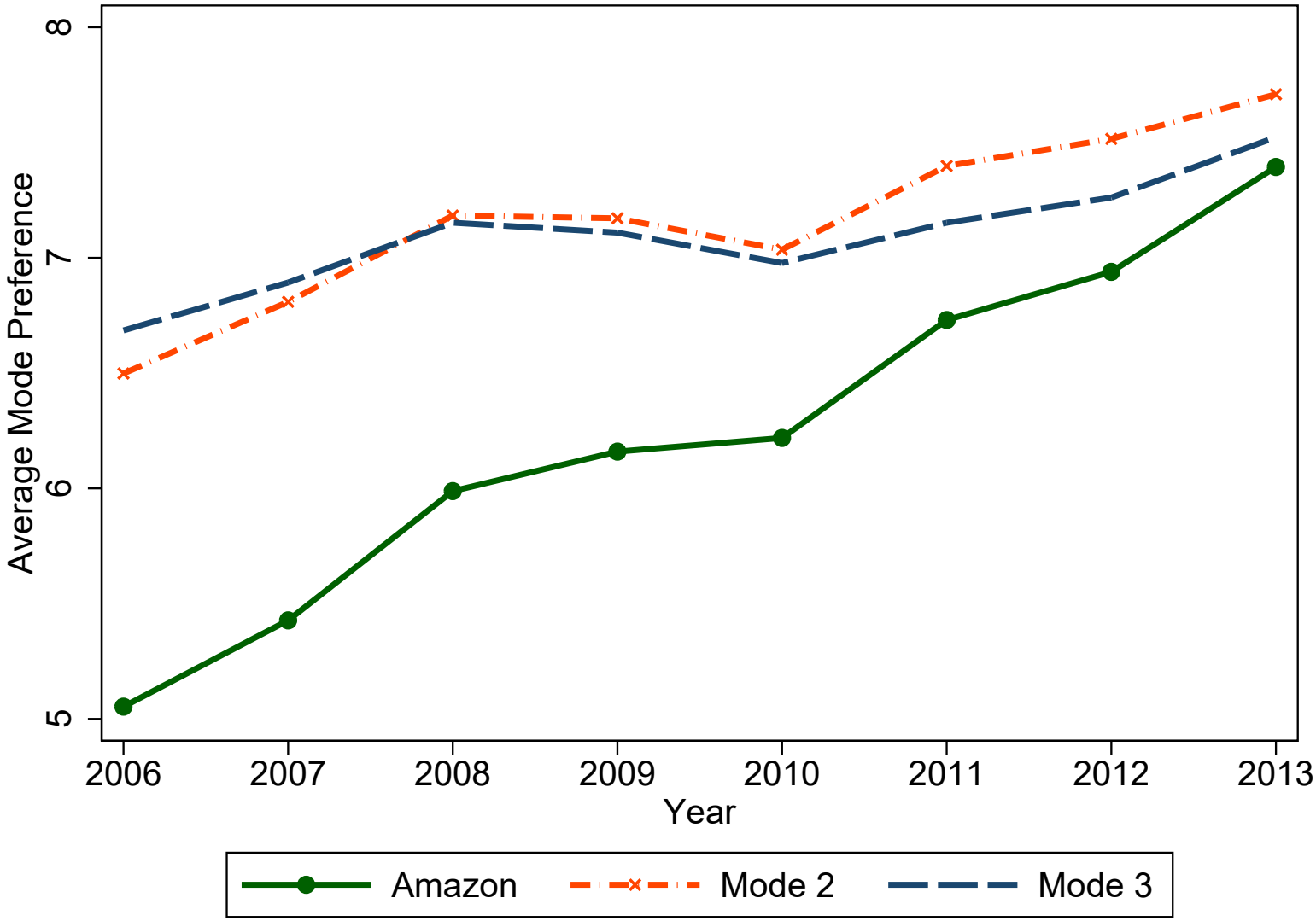
$$R^* - \theta D^* > R' - \theta D' \Rightarrow \theta \geq (R' - R^*)/(D' - D^*)$$

- **Takeaway:** Tax elasticity identifies implicit economies of density.

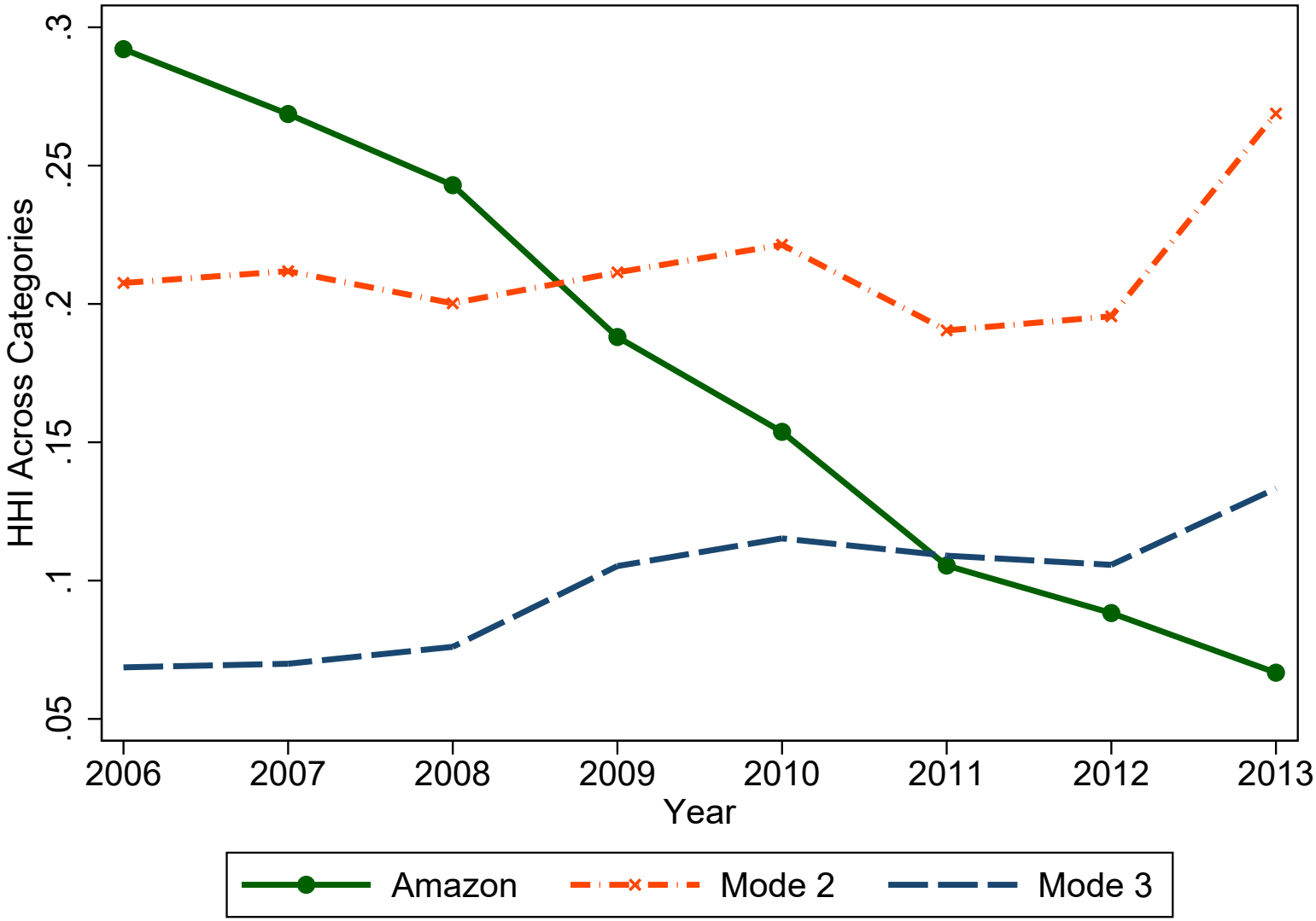
Shipping Cost Estimates: Results

- Net shipping cost for $30 of goods from $0.17 to $0.47 per 100 miles.
- Similar estimates:
  - Assuming shipments come from the lowest cost FC (determined using data from Commodity Flow Survey).
  - Different assumptions about when tax rules are implemented.

Amazon's "quality" growth is fueled by a combination of lower prices, faster delivery, and enhanced variety



Platform Quality and Variety: HHI Across Product Categories



Decomposition: Density *versus* Variety

- We cannot directly infer cost passthrough from prices because of the increased in variety
- Similarly, we can only measure the net effect of cost passthrough and shipping-time reduction on Amazon quality

Decomposition: Density *versus* Variety

- We cannot directly infer cost passthrough from prices because of the increased in variety
- Similarly, we can only measure the net effect of cost passthrough and shipping-time reduction on Amazon quality
- Back of the envelope decomposition from 8 observations...

$$\text{Amazon FE}_t = 7.16^* - 1.02^* \times \log \text{Avg. shipping cost}_t - 6.64^* \cdot HHI_t$$

- Growth in Amazon platform WTP 2006-2013
  - *Dense network:* 46%
  - *Centralized network:* 29%

Decomposition: Density *versus* Variety

- We cannot directly infer cost passthrough from prices because of the increased in variety
- Similarly, we can only measure the net effect of cost passthrough and shipping-time reduction on Amazon quality
- Back of the envelope decomposition from 8 observations...

$$\text{Amazon FE}_t = 7.16^* - 1.02^* \times \log \text{Avg. shipping cost}_t - 6.64^* \cdot HHI_t$$

- Growth in Amazon platform WTP 2006-2013
  - *Dense network:* 46%
  - *Centralized network:* 29%

- **Bottom line:** Roughly 35% of Amazon's growth in average WTP is associated with denser network (pass-through + shipping time)

Conclusion

- Quantified the trade-off associated with the expansion of FC network:
  - Consumers sensitive to sales tax.
  - No demand side benefit to expansion.
  - Cost saving significant.

  $\rightarrow$ Suggestive evidence that (as in brick-and-mortar retail) economies of density significant drivers of concentration and market position.
- Extensions:
  - Distortions from taxes: Would cost savings have been more important if *tax nexus* didn't exist?
  - Convenience effect of same-day shipping likely more pronounced, but still concentrated in urban areas, post-sample.
  - Missing piece: Sortation facility network (since 2014)

Initial Evidence: Transaction-Level Regression Model

Estimate the following linear probability model of Amazon purchase:

$$Pr(A_{ohijt} = 1) = \beta_0 + \alpha \ln(1 + \tau_{it} \mathbf{1}_{it}^{taxable}) + \gamma d_{it} + \beta_1 C_h + \beta_2 Z_{it} + \lambda_{ijt} + \epsilon_{ohijt}$$

- A_{ohijt}: indicator of Amazon transaction on purchase occasion o from household h in county i in year t.
- τ_{it}: sales tax rate in county i and year t.
- $\mathbf{1}_{it}^{taxable}$: tax status for Amazon purchases for county i and year t.
- d_{it}: measure of shipping speed from Amazon (distance or shipping time from FC to county centroid).
- Controls: Z_{it}, local competition, C_h, household demographics and county, product category j, and year FEs in λ_{ijt}.

Initial Evidence:
Propensity of Buying from Amazon

Variable name	(1)	(2)	(3)	(4)
Tax Elasticity	-0.156**	-0.142*	-0.158**	-0.152**
	(0.074)	(0.074)	(0.075)	(0.074)
Local Express Delivery		-0.019*		
		(0.010)		
Log Distance			0.001	
			(0.002)	
1 or 2 Day Priority				0.019*
				(0.011)
Obs	2,291,291	2,291,291	2,291,291	2,291,291
R-Sq	0.355	0.355	0.355	0.355

*** 1% ** 5% * 10%.

Initial Evidence: County-Level Expenditure Model

Diff-in-diff model of effect of "Amazon Tax":

$$ExpAm_{it} = \beta_0 + \sigma \mathbf{1}_{it}^{taxable} + \gamma d_{it} + \beta_1 C_{it} + \beta_2 Z_{it} + \lambda_{it} + \epsilon_{it}$$

- $ExpAm_{it}$: log of average household expenditures on Amazon from county i in year t.
- $\mathbf{1}_{it}^{taxable}$: sales tax status.
- d_{it}: measure of shipping speed.
- Controls: Z_{it}, local competition, C_{it}, consumer demographics, and county and year FEs in λ_{it}

DiD: Effect of taxes on Amazon Expenditures

Variable name	(1)	(2)	(3)	(4)
Amazon Purchase Taxed	-0.105*	-0.105*	-0.104*	-0.108*
	(0.060)	(0.060)	(0.061)	(0.061)
Local Express Delivery		0.147		
		(0.185)		
Log Distance			-0.002	
			(0.035)	
1 or 2 Day Priority				-0.057
				(0.107)
Obs	12,486	12,486	12,486	12,486
R-Sq	0.448	0.448	0.448	0.448

*** 1% ** 5% * 10%.

- Avg tax rate of 6.5% $\rightarrow$ 1pp increase in tax reduces spending by 1.6%.

▸ Return

Household Purchasing: comScore vs Forrester

Year	Online Expenditure	Online Transactions	% Zero Expenditure	Adjusted Expenditure	Adjusted Transactions	% Offline Shoppers Only
2006	$239	2.4	51.8%	$318	3.1	55.5%
2007	$254	2.5	52.0%	$318	2.9	60.8%
2008	$196	2.0	60.0%	$333	3.2	-
2009	$141	1.4	67.9%	$355	3.4	-
2010	$125	1.4	68.6%	$369	3.5	32.1%
2011	$131	1.4	69.7%	$424	4.2	23.0%
2012	$152	1.8	64.0%	$434	4.6	23.9%
2013	$120	1.7	65.3%	$377	4.7	15.5%

▸ Return

Modes

Sales Rank	Taxed	Non-Taxed
1	walmart.com	dell.com
2	jcpenney.com	qvc.com
3	staples.com	yahoo.net
4	victoriassecret.com	hsn.com
5	officedepot.com	yahoo.com
6	bestbuy.com	quillcorp.com
7	apple.com	overstock.com
8	target.com	ebay.com
9	sears.com	orientaltrading.com
10	costco.com	zappos.com
Total (%)	192 (34)	375 (66)

Return

Robustness (1)

	(1)	(2)	(3)	(4)	(5)	(6)
Tax Elasticity	-1.307***	-1.687***	-1.867***		-1.203*	-1.199***
	(0.286)	(0.585)	(0.575)		(0.625)	(0.401)
Tax Elasticity (Amazon)				-1.166**		
				(0.515)		
Tax Elasticity (Mode 3)				-2.900***		
				(0.847)		
Obs	42,399	52,617	29,053	42,399	43,811	42,400
R-Sq	0.315	0.162	0.135	0.185	0.137	0.199
Regression	A-Weights	Zeros	2008-2013	Individual Tax Effect	No Forrester Adjustment	No Population Weights

Robustness (2)

▸ Return.

Variable name	(1)	(2)	(3)	(4)
Tax Elasticity	-1.325***	-1.215***	-1.401***	-1.644***
	(0.498)	(0.443)	(0.481)	(0.525)
Entry Year 0			-0.034	
			(0.050)	
Entry Year -1			-0.015	
			(0.061)	
Entry Year -2			0.093	
			(0.072)	
Entry Year -3			0.046	
			(0.073)	
Entry Year $¡$ -3			-0.078	
			(0.058)	
Tax*(Entry Year 0)				-0.620
				(0.759)
Tax*(Entry Year -1)				-0.045
				(0.845)
Tax*(Entry Year -2)				-0.719
				(1.061)
Tax*(Entry Year -3)				-1.104
				(1.082)
Tax*(Entry Year $¡$ -3)				0.551
				(0.406)
Obs	42,399	42,399	42,399	42,399
R-Sq	0.195	0.240	0.186	0.186
Fixed Effects	Year-State, County	Year-County	County	County

Fulfillment Center Roll-out and Amazon's Wholesale Cost

- **Assumption:** Only distribution, but not wholesale, cost affected by FC network

Return

Fulfillment Center Roll-out and Amazon's Wholesale Cost

- **Assumption:** Only distribution, but not wholesale, cost affected by FC network
  - $\Rightarrow$ Pre-shipping mark-up remains constant, regardless of FC layout.

Return

Fulfillment Center Roll-out and Amazon's Wholesale Cost

- **Assumption:** Only distribution, but not wholesale, cost affected by FC network
  - $\Rightarrow$ Pre-shipping mark-up remains constant, regardless of FC layout.
- **Motivation:** Suppliers distribute products to wide range of retailers, not just Amazon.
  - Economies of scale in distribution $\rightarrow$ combine shipments to different retailers when possible.
    - ★ Books: Barnes & Noble requires publishers to ship direct to store.
    - ★ Other products: Walmart requires shipment to distribution centers. Avg distance from Amazon FC to closest Walmart FC falls from 92.2m (2006) to 65.4m (2013).

▸ Return

Fulfillment Center Roll-out and Amazon's Wholesale Cost

- **Assumption:** Only distribution, but not wholesale, cost affected by FC network
  - $\Rightarrow$ Pre-shipping mark-up remains constant, regardless of FC layout.
- **Motivation:** Suppliers distribute products to wide range of retailers, not just Amazon.
  - Economies of scale in distribution $\rightarrow$ combine shipments to different retailers when possible.
    - ★ Books: Barnes & Noble requires publishers to ship direct to store.
    - ★ Other products: Walmart requires shipment to distribution centers. Avg distance from Amazon FC to closest Walmart FC falls from 92.2m (2006) to 65.4m (2013).
  - $\Rightarrow$ When delivering to host of retailers, suppliers unlikely to incur higher cost from expansion of FC network.

▸ Return