The Local Area Incidence of Exporting

Christoph E. Boehm
UT Austin & NBER

Aaron Flaaen
Federal Reserve Board

Nitya Pandalai-Nayar
UT Austin & NBER

Jan Schlupp
Amazon

April 2025

Any views expressed are those of the authors and not those of the U.S. Census Bureau. The Census Bureau has reviewed this data product to ensure appropriate access, use, and disclosure avoidance protection of the confidential source data used to produce this product. This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 2439. (CBDRB-FY25-0247). The views expressed are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Board of Governors of the Federal Reserve System or UT Austin. Schlupp concluded work on this project before joining Amazon.

Motivation

Question: What do we know about the geography of U.S. exports?

Motivation

Question: What do we know about the geography of U.S. exports?

Answer: Not Much

- Common US trade datasets do not identify the origin of export (or destination of imports)
- Firm-level data does not resolve issue — most trade is by large, multi-unit firms
- State-level data based on shipment's origin of movement, not necessarily production

Motivation

Question: What do we know about the geography of U.S. exports?

Answer: Not Much

- Common US trade datasets do not identify the origin of export (or destination of imports)
- Firm-level data does not resolve issue — most trade is by large, multi-unit firms
- State-level data based on shipment's origin of movement, not necessarily production

Why it Matters

- Studies on local-area effects of exporting must impute local exports from employment shares
- Many research questions require establishment-level exports
  (because other datasets are available only at establishment level)

This Paper

- New algorithm to build microdata on the geography of U.S. exports
  - Establishment-level (with zipcodes), currently implemented for 1993-2021
  - Destination, product and date information
- Stylized facts
  - Geography plays a role in export destination specialization patterns
  - High concentration of exports across space, more than employment or sales
- Application: Local-Area effects of the Great Trade Collapse
  - Employment, payroll and wages declined in more exposed counties
  - Contrast using publicly available data
    - Noisy estimates, weak instrument
- Application: Which MSAs are most exposed to retaliatory tariffs?

Related Literature

- **Studies on sub-national or regional trade:**
  Allen and Arkolakis (2014), Atalay et al (2014) , Caliendo, Parro, Rossi-Hansberg and Sarte (2018), (Caliendo, Dvorkin and Parro (2019) [Commodity Flows Survey, US], Santamaria et al (2020) [Spain], Boehm, Flaaen and Pandalai-Nayar (2024)

- **Facts on US national exports:** Bernard and Jensen (1995), Bernard, Jensen, Redding and Schott (2007a,b,c, 2010), Bernard, Jensen and Schott (2009), ...

- **Labor market effects of trade shocks:**
  Autor, Dorn and Hanson (2013), Pierce and Schott (2016), Feenstra, Ma and Xu (2019), Topalova (2007), Kovak (2010), Dix-Carneiro and Kovak (2017), Hakobyan and McLaren (2016), Benguria and Saffie (2020), etc.....

- **Great Trade Collapse:** Levchenko et al (2010), Bems (2010), Alessandria et al (2010), Bown and Crowley (2014), etc....

Constructing A New Dataset of U.S. Export Transactions at the Plant Level

Key Ingredients

- Longitudinal Foreign Trade Transactions Database (LFTTD):
  - Universe of transaction-level trade in goods, with links to firm (not establishment)
  - Includes date, value, quantity, product code (detailed HS-10), destination country, among other variables

Key Ingredients

- Longitudinal Foreign Trade Transactions Database (LFTTD):
  - Universe of transaction-level trade in goods, with links to firm (not establishment)
  - Includes date, value, quantity, product code (detailed HS-10), destination country, among other variables
- Longitudinal Business Database (LBD):
  - Universe of U.S. private nonfarm establishments
  - Annual employment and payroll
  - Establishment identifiers linked to firm identifiers
  - Industry (NAICS), as well as information to identify location

Key Ingredients

- Longitudinal Foreign Trade Transactions Database (LFTTD):
  - Universe of transaction-level trade in goods, with links to firm (not establishment)
  - Includes date, value, quantity, product code (detailed HS-10), destination country, among other variables
- Longitudinal Business Database (LBD):
  - Universe of U.S. private nonfarm establishments
  - Annual employment and payroll
  - Establishment identifiers linked to firm identifiers
  - Industry (NAICS), as well as information to identify location
- Set of product-industry concordances from Pierce and Schott (2012)

Summary of Establishment Export Allocation Algorithm

- The LBD provides the set of possible establishments associated with firm of transaction
- For single-unit firms, allocation is trivial...

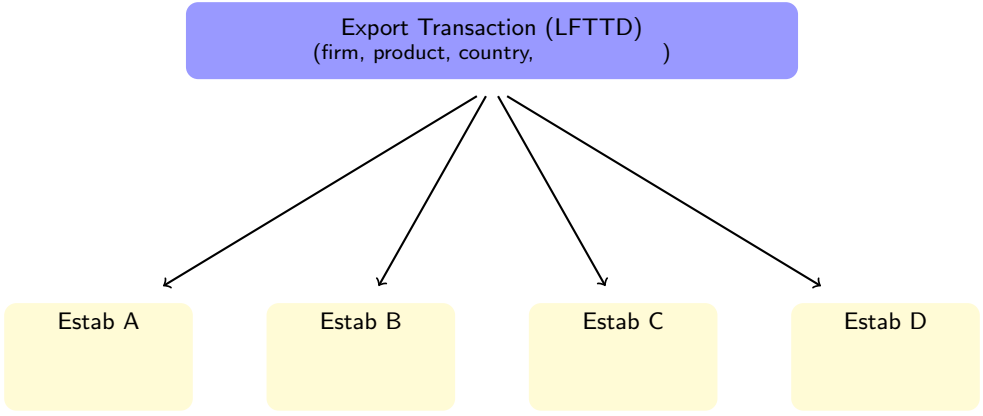
For Multi-unit Firms, Allocate to Establishments Using 3 Pieces of Information

1. Geographic information
   - Variable in LFTTD identify origin of *movement*, can be different than origin of *production*
2. Production Associated Industry
   - Find industries that are principal producers of each product – use Census of Manufacturers Products Trailer File and concordance in Pierce and Schott (2012)
3. Survey information in ASM/CMF

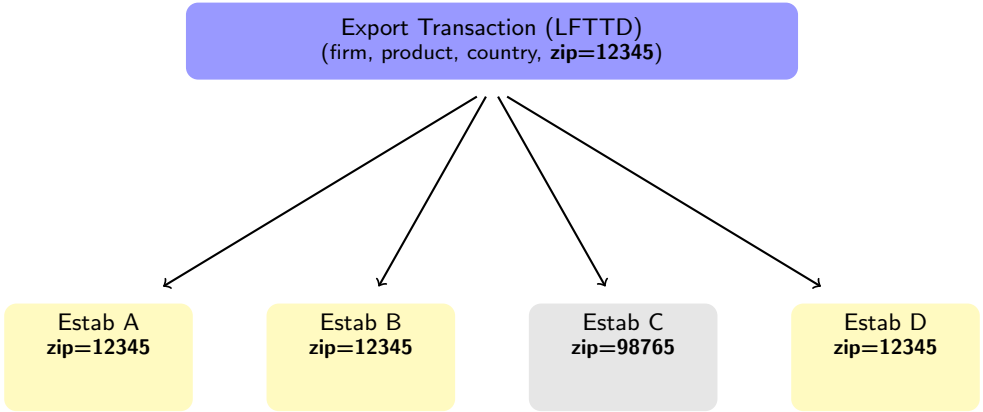
Export Transaction (LFTTD)
(firm, product, country,)

Export Transaction (LFTTD)
(firm, product, country,)

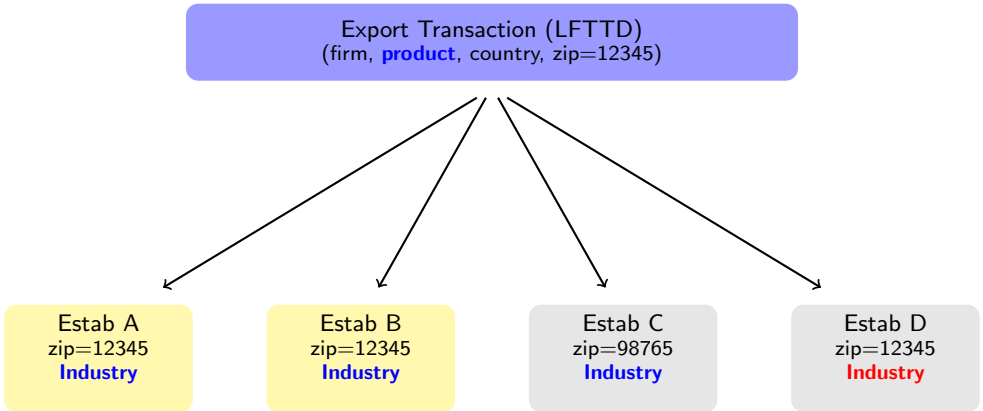
firm identifier – via LBD – links to set of establishments of firm



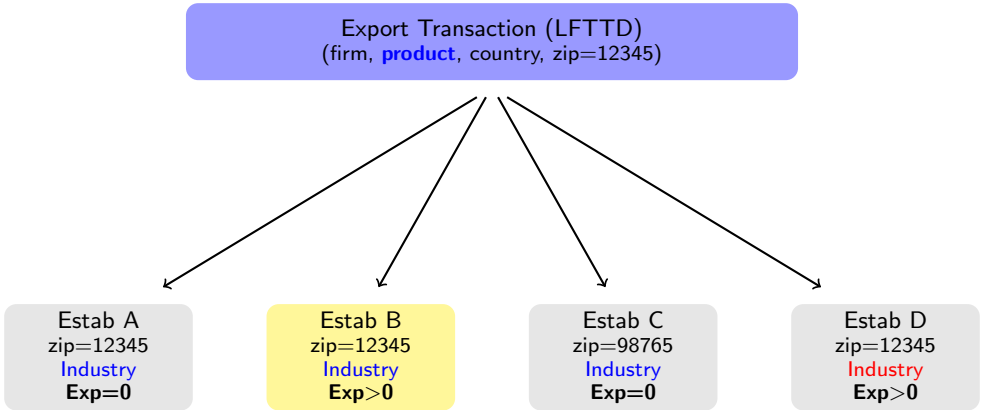
firm identifier – via LBD – links to set of establishments of firm



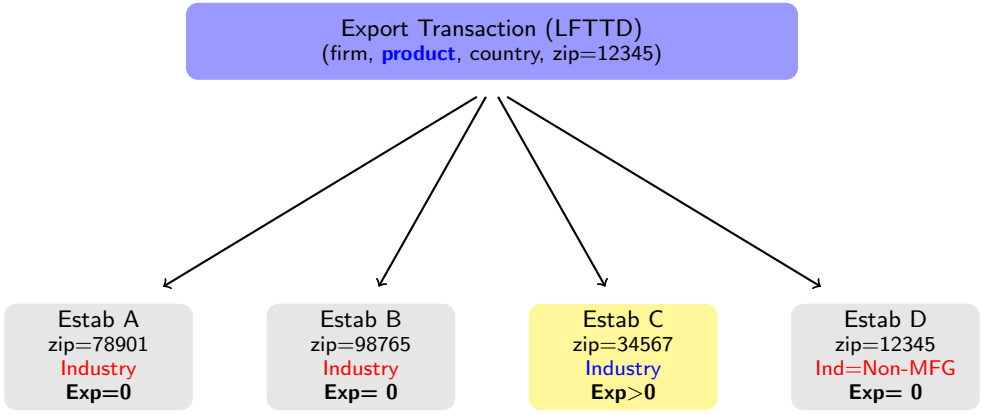
1. Zipcode: LFTTD identifies *origin of movement* of transaction



2. Production Associated Industry:
   - CMF: Map products into likely industries of production

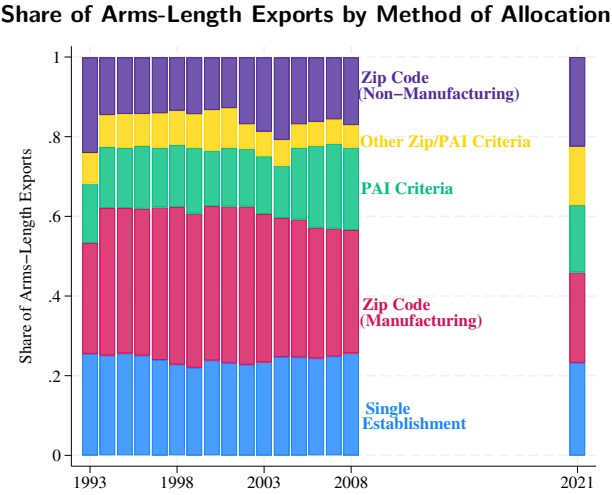


3. Survey-based export indicator (discontinued from CMF 2017)
   - Positive exports in ASM or CMF



- **This is an illustration! In reality we have over 15 iterative steps over which we perform the allocation**

Composition of Export Allocation



- Allocation remarkably stable over nearly 30 years: high quality matches 65-75%
- Share of each type of allocation varies due to census years, changes in underlying datasets, etc

Comparison to Origin-of-Movement Data

States	Shares (in percent) Public Data	This paper	Percentage point Difference	Percent Difference
	Top 5 states: Public Data > Data in this paper			
Washington	2.4	1.7	0.7	50.7
Tennessee	2.4	1.7	0.7	45.2
New York	6.2	5.6	0.6	15.9
Illinois	4.5	4.0	0.4	16.8
Kentucky	2.1	1.7	0.3	25.6
	Bottom 5 states: Public Data < Data in this paper			
California	11.1	11.6	-0.5	-0.9
Indiana	2.9	3.4	-0.5	-10.0
Louisiana	2.4	2.9	-0.5	-14.6
Massachusetts	2.2	2.8	-0.6	-16.1
Connecticut	1.0	1.7	-0.7	-37.3

Sources: Author's calculations using LFTTD, CMF as explained in the text, and Census Bureau State-Level Trade for 2021.

- Overall correlation between sources high at 0.99, but important differences

Stylized Facts

New Stylized Facts

This new data reveals new facts on U.S. exporting:

1. Geography plays a role in destination of export specialization patterns

2. High concentration of exports across space

New Stylized Facts

This new data reveals three new facts on U.S. exporting:

1. **Geography plays a role in destination of export specialization**
2. High concentration of exports across space

Fact 1: Geography of Export Destinations

For the first time, we see the destination of exports at a granular level:

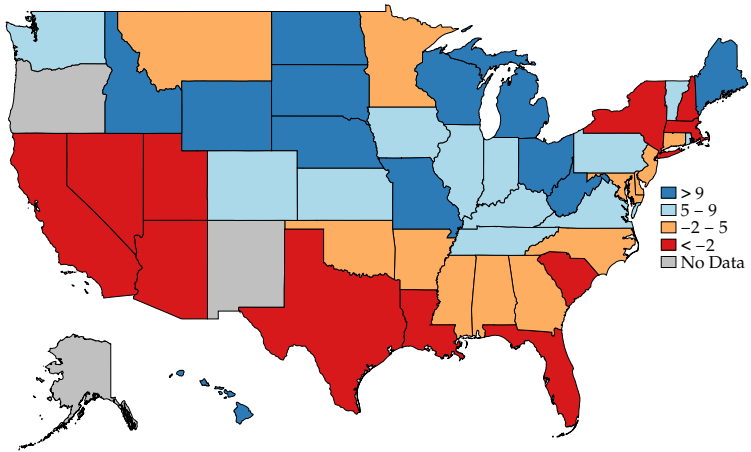
For a first look at these patterns we:

- Aggregate to the state-level

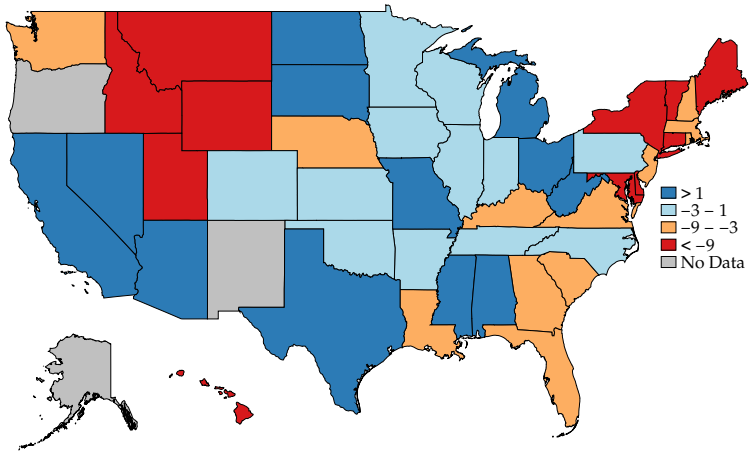
- Calculate the *relative* destination composition of state exports: for exports of state i to destination k

$$\frac{x_{ik}}{\sum_j x_{ij}} - \frac{X_k^{US}}{\sum_j X_j^{US}}. \tag{1}$$

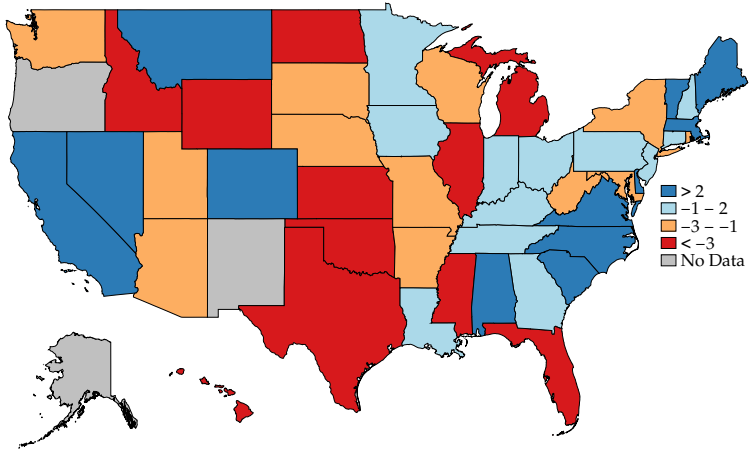
Relative Export Composition: Canada



Relative Export Composition: Mexico



Relative Export Composition: China



▸ Europe ▸ Other Asia ▸ Rest of World

New Stylized Facts

This new data reveals new facts on U.S. exporting:

1. Geography plays a role in destination of export specialization
2. **High concentration of exports across space**

Exports are Highly Concentrated

- Economic activity is highly concentrated across space [Axtell (2001), Gabaix (1999, 2011)]
- Exporting is highly concentrated across firms [Bernard et al (2007)]

Exports are Highly Concentrated

- Economic activity is highly concentrated across space [Axtell (2001), Gabaix (1999, 2011)]
- Exporting is highly concentrated across firms [Bernard et al (2007)]

To highlight export activity across space we:

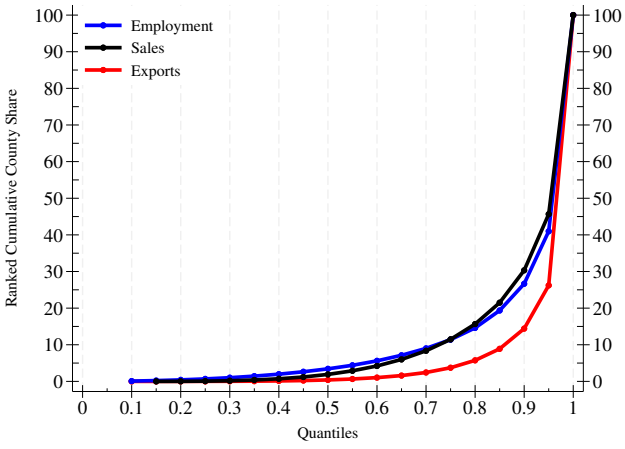
- Aggregate employment/sales/exports by county
- Rank counties by each measure
- Calculate cumulative shares, by quantile

Exports are Highly Concentrated

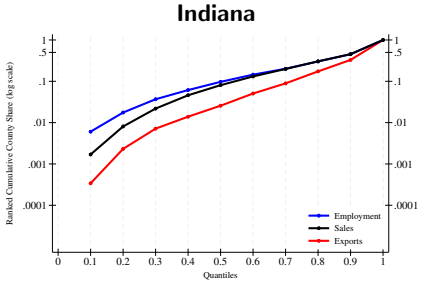
- Economic activity is highly concentrated across space [Axtell (2001), Gabaix (1999, 2011)]
- Exporting is highly concentrated across firms [Bernard et al (2007)]

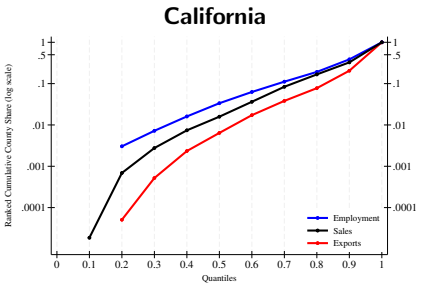
To highlight export activity across space we:

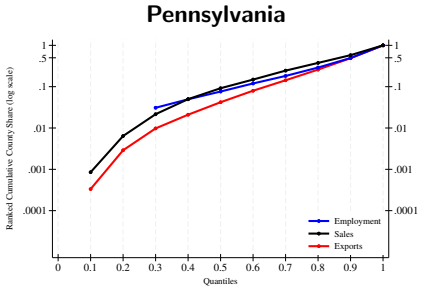
- Aggregate employment/sales/exports by county
- Rank counties by each measure
- Calculate cumulative shares, by quantile

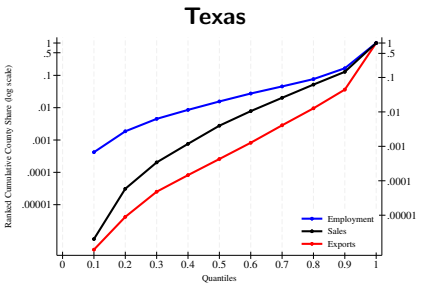


Relative Concentration Holds Within States, but Degree Varies Across States









Exports are more Highly Concentrated than Employment / Sales

Proportion in Top County Decile, Selected States

	Employment	Sales	Exports
Pennsylvania	50.7%	42.1%	51.1%
Indiana	54.7%	54.7%	67.1%
California	61.4%	67.3%	79.5%
Texas	81.5%	85.4%	96.4%
Georgia	67.8%	58.1%	80.8%

- Findings indicate that practice of imputing exports to local areas using employment shares could be problematic

Local-Area Effects of the Great Trade Collapse

Estimating Equation

$$\frac{y_{c,2009} - y_{c,2007}}{\text{emp}_{c,2007}} = \alpha + \beta \frac{\text{exp}_{c,2009} - \text{exp}_{c,2007}}{\text{emp}_{c,2007}} + \text{controls} + \varepsilon_c$$

- Start with data (employment, exports, etc) aggregated to the county-level
- $y_{c,t}$ is an outcome of interest (e.g. employment in county c at time t)
- $\text{exp}_{c,t}$ is the value of exports
- $\text{emp}_{c,t}$ employment

Notes

- Wages as outcome measured in percent changes
- Controls include county exports/worker, lagged dependent variable

Instrument

- Endogeneity concern – exports and employment simultaneously determined
- Instrument: *modified* World Import Demand (Hummels et al 2014)

$$\text{inst}_c^{\text{WID}} = \frac{\text{exp}_{c,2007}}{\text{emp}_{c,2007}} \cdot \sum_{d,p} s_{c,d,p,2007}^{\text{exp}} \left(\Delta \ln \text{imp}_{d,p} - \overline{\Delta \ln \text{imp}}_p \right).$$

  - $s_{c,d,p,2007}^{\text{exp}} = \frac{\text{exp}_{c,d,p,2007}}{\text{exp}_{c,2007}}$ is c's export share of product p to destination d
  - $\overline{\Delta \ln \text{imp}}_p = \frac{1}{D} \sum_d \Delta \ln \text{imp}_{d,p}$
- De-meaning with $\overline{\Delta \ln \text{imp}}_p$ purges common global demand declines—potentially correlated domestic demand (Boehm and Pandalai-Nayar 2021)
- Shift-share design: standard errors and first stage following Borusyak et al 2021

Results: Employment, Payroll, Wages

Dependent variable:	$\frac{emp_{c,2009}-emp_{c,2007}}{emp_{c,2007}}$		$\frac{pay_{c,2009}-pay_{c,2007}}{emp_{c,2007}}$		$\frac{wage_{c,2009}-wage_{c,2007}}{wage_{c,2007}}$	
	(1)	(2)	(3)	(4)	(5)	(6)
$\frac{exp_{c,2009}-exp_{c,2007}}{emp_{c,2007}}$	16.4***	17.7***	0.96***	0.96***	1.92***	1.56**
	(4.65)	(6.51)	(0.34)	(0.44)	(0.77)	(0.75)
	[4.77]	[6.47]	[0.34]	[0.45]	[0.77]	[0.75]
Other controls	yes	yes	yes	yes	yes	yes
State fixed effects	no	yes	no	yes	no	yes
First stage F	37.42	31.27	36.89	31.28	37.37	31.29
Shock-level first stage F	25.18	21.84	24.62	21.81	25.06	21.82

- Large employment effects, one-for-one passthrough to payroll
- Instrument is uniformly strong
- OLS

Manufacturing and Non-Manufacturing

Dependent variable	Employment		Payroll		Wages	
	Mfg. (1)	Non-Mfg. (2)	Mfg. (3)	Non-Mfg. (4)	Mfg. (5)	Non-Mfg. (6)
$\frac{exp_{c,2009} - exp_{c,2007}}{emp_{c,2007}}$	9.72***	7.95***	0.52***	0.87	2.08***	0.90
	(2.76)	(4.49)	(0.12)	(0.60)	(0.79)	(0.88)
	[3.90]	[3.70]	[0.19]	[0.62]	[0.86]	[0.92]
Other controls	yes	yes	yes	yes	yes	yes
State fixed effects	no	no	no	no	no	no
Shock-level first stage F	25.02	25.26	24.65	25.12	25.27	24.89

- Larger effects on manufacturing
- Some spillovers to non-manufacturing

Comparison with Publicly Available Data

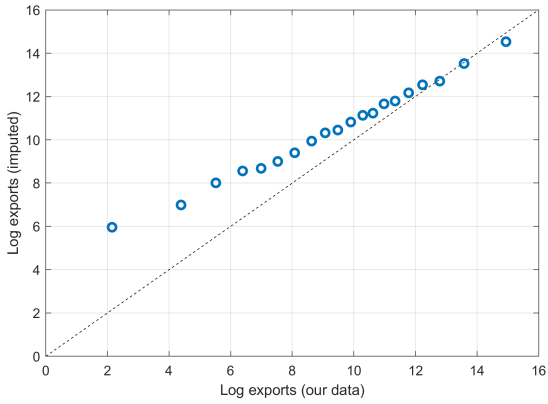
Imputation Approach with Public Data

- State trade data from USA Trade + CBP employment
- Imputation procedure

$$\widehat{\exp}_{c,d,p,t} = \frac{\text{emp}^{\text{CBP}}_{c,p,t}}{\sum_{c' \in s} \text{emp}^{\text{CBP}}_{c',p,t}} \cdot \exp_{s(c),d,p,t},$$

  - $s(c)$ is the state s of county c
- Alternatively: impute in growth rates (results very similar)

Imputed vs Allocated County-Level Exports



Sources: Author's calculations using LFTTD, CMF

- Imputed exports larger than measured exports for most counties
- Consistent with stylized fact on export concentration

Estimates using Public Data

	Dependent variable: $\frac{\text{emp}_{c,2009}-\text{emp}_{c,2007}}{\text{emp}_{c,2007}}$					
	(1)	(2)	(3)	(4)	(5)	(6)
$\frac{\widehat{\text{exp}}_{c,2009}-\widehat{\text{exp}}_{c,2007}}{\text{emp}_{c,2007}}$	21.9	22.4	29.9			
	(27.7)	(28.6)	(28.8)			
	[21.4]	[22.1]	[27.7]			
Other controls	yes	yes	yes			
State fixed effects	no	no	yes			
Instrument	Public	Public	Public			
Shock-Level First stage F	0.11	0.11	0.02			

- Estimates noisy, instrument weak

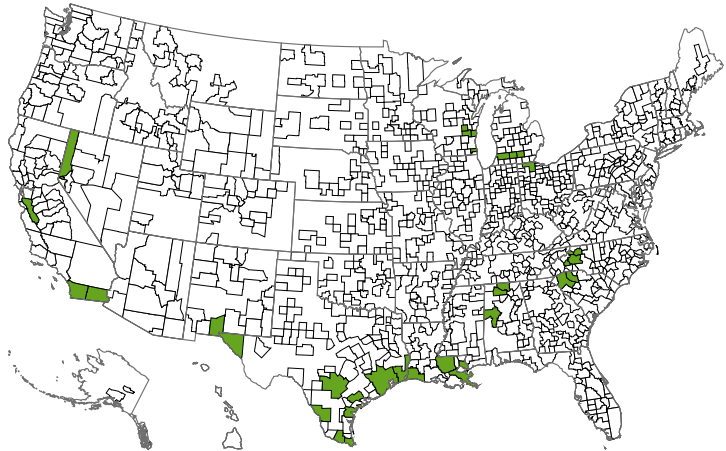
Estimates using Public Data

	Dependent variable: $\frac{emp_{c,2009}-emp_{c,2007}}{emp_{c,2007}}$					
	(1)	(2)	(3)	(4)	(5)	(6)
$\frac{\widehat{exp}_{c,2009}-\widehat{exp}_{c,2007}}{emp_{c,2007}}$	21.9	22.4	29.9	39.9***	40.9***	76.4**
	(27.7)	(28.6)	(28.8)	(12.7)	(13.0)	(43.1)
	[21.4]	[22.1]	[27.7]	[9.63]	[10.2]	[36.9]
Other controls	yes	yes	yes	yes	yes	yes
State fixed effects	no	no	yes	no	no	yes
Instrument	Public	Public	Public	Baseline	Baseline	Baseline
Shock-Level First stage F	0.11	0.11	0.02	8.33	7.68	3.23

- Estimates noisy, instrument weak

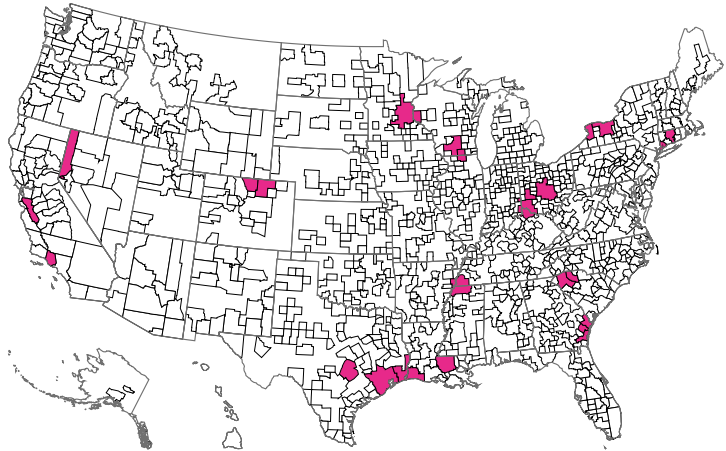
MSA Export-Destination Exposure

MSA export exposure: Mexico



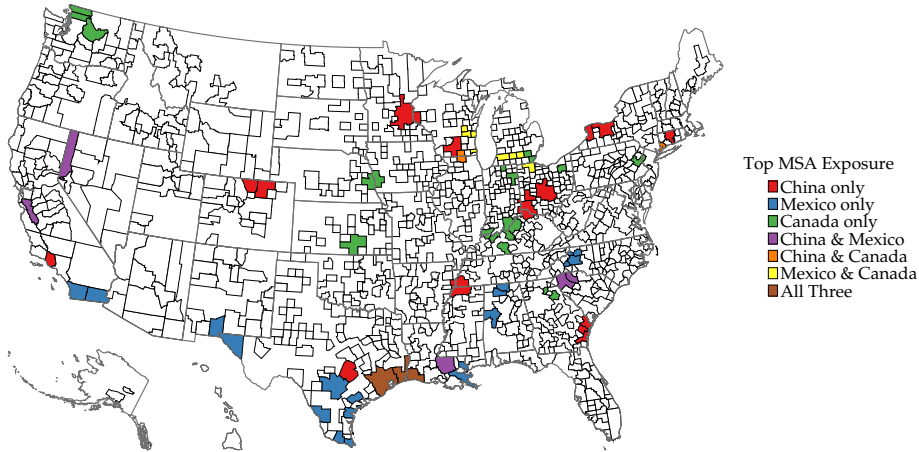
- For each export-destination, top 20 MSAs ranked by exports/worker
- Some areas near the border exposed, but also Detroit MSA, Tuscaloosa AL, Greenville SC

MSA export exposure: China



- Among most exposed: Austin, Madison, Columbus, Rochester, San Jose, CA
- High-education high/tech areas

MSA export exposure varies

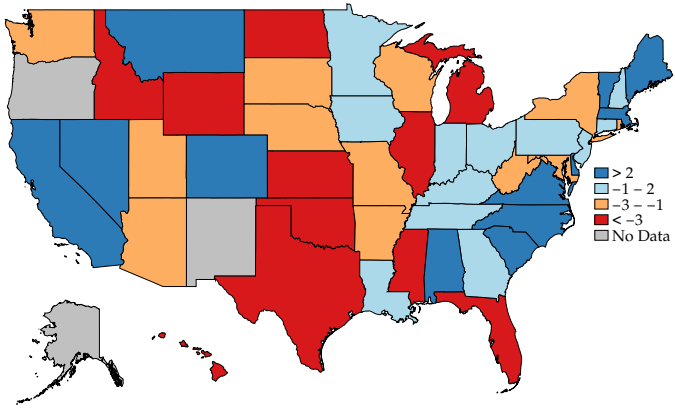


- For each export-destination, top 20 MSAs ranked by exports/worker
- Highlights regions most exposed to trade-war retaliation (several from multiple destinations)

Conclusion

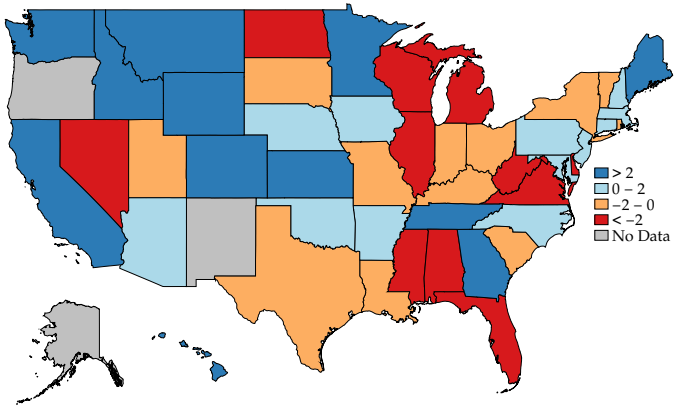
- New algorithm to construct dataset of exports with much more geographic detail within US Census data architecture
- New stylized facts: exports more concentrated than employment/sales even within states, heterogeneity across states
- Local area effects of exporting: counties more exposed to foreign demand shocks lost 17 additional jobs per million dollars of exports
- "Imputing" local-area trade from publicly available data might be problematic for exports
- New measurement of MSAs most heavily exposed to top trading partners' trade policies

Relative Export Composition: Europe



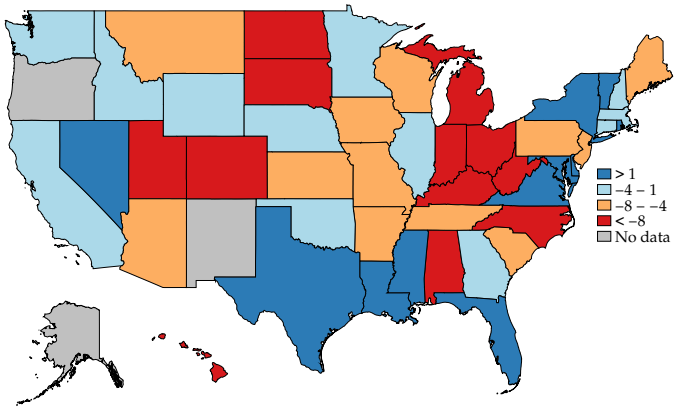
Back

Relative Export Composition: Other Asia



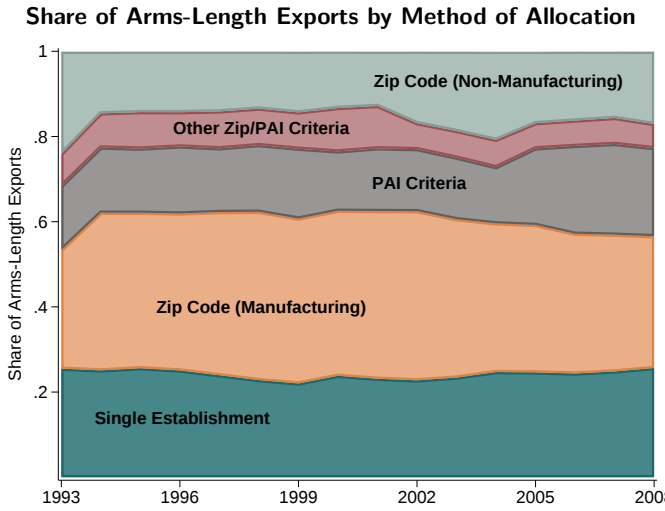
Back

Relative Export Composition: Rest of World



Back

Composition of Export Allocation



Back

OLS

Dependent variable:	$\frac{emp_{c,2009}-emp_{c,2007}}{emp_{c,2007}}$		$\frac{pay_{c,2009}-pay_{c,2007}}{emp_{c,2007}}$		$\frac{wage_{c,2009}-wage_{c,2007}}{wage_{c,2007}}$	
	(1)	(2)	(3)	(4)	(5)	(6)
$\frac{exp_{c,2009}-exp_{c,2007}}{emp_{c,2007}}$	4.38***	3.80***	0.18***	0.15***	0.15**	0.08
	(0.68)	(0.71)	(0.03)	(0.03)	(0.09)	(0.07)
$exp_{c,2007}/emp_{c,2007}$	0.00	0.00***	-0.04***	-0.03*	-0.20***	-0.18***
	(0.00)	(0.00)	(0.02)	(0.02)	(0.04)	(0.04)
Lagged dep. variable	0.01	-0.01	0.05***	0.02	-0.20***	-0.21***
	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)
State fixed effects	no	yes	no	yes	no	yes
R^2	0.01	0.13	0.05	0.19	0.09	0.17

Back