General

Merchandise Trade by State:

Estimating interstate trade flows for the United States

UEA 2026, Chicago, IL

Kyle Hood; Gabriel Medeiros; Dirk van Duym



Disclaimer: The views expressed in this presentation are those of the authors and do not necessarily reflect those of the Bureau of Economic Analysis or Department of Commerce

Interstate Trade

- Interstate trade flows are essential for regional IO tables, supply-use systems, spatial equilibrium models, and policy analysis

- But interstate trade flows are **not directly observed**
  - Many approaches **infer flows from production and use totals** rather than measuring actual shipments
  - Alternative: use freight shipment data: this describes **movements**, not **producer-destination flows**

- This paper: build model to link flows in the shipment data, taking into account producer versus passthrough shippers, to recover actual trade flows

Trade Flows

- Location quotients: Infer local supply versus imports from regional specialization relative to national production (Miller & Blair, 2009; Flegg & Webber, 2000).
- Commodity balance / CHARM: Infer imports and exports from regional supply-demand gaps, allowing cross-hauling (Kronenberg, 2009).
- Gravity models: Allocate flows across regions using origin supply, destination demand, and distance/trade costs (Leontief & Strout, 1963).
- RAS / balancing: Adjust estimated trade matrices to match known regional totals and accounting constraints (Polenske, 1978).
- These methods **infer** the geography of interregional trade from assumptions and accounting constraints. They have the benefit of reflecting a complete system of origins, destinations, and goods.
- In contrast, **survey-based trade statistics** allow trade flows to be observed directly.

General

Trade Statistics: The Commodity Flow Survey (CFS) and Freight Analysis Framework

- Since 1993, the CFS has provided information on shipments of some commodities.
- The FAF is a CFS-based statistical framework that integrates multiple data sources to estimate comprehensive U.S. origin to destination freight flows.
- Trade frictions and market integration: Wolf (2000); Hillberry & Hummels (2008); Allen & Arkolakis (2014).
- Infrastructure and urban specialization: Duranton, Morrow & Turner (2014); Dingel (2017).
- Regional production networks and shock propagation: Caliendo et al. (2018); Fajgelbaum et al. (2019).
- Market access and regional outcomes: Knaap (2006); Crafts & Klein (2015).

Source Data: FAF and BEA Comparison, 2017

Merchandise Activity (Billions of Dollars)

	BEA	FAF	Duplicated Value (residual)
Imports	2,370	2,263	
Exports	1,547	1,504	
Output / Domestic origin shipments	6,207	16,637	10,430

- Passthroughs are responsible for the majority of recorded shipments
  - Production-use link is obfuscated
  - Logistics-heavy states would appear as giant producers and/or consumers
- Shipments must be linked on average 3 times to account for FAF shipments

Core Measurement Problem

- Observed shipment flows: $i \rightarrow k;\ k \rightarrow j$
- We want: $\text{Production location } (i) \rightarrow \text{Final destination } (j)$
- Shipment data include: production shipments, but also passthrough shipments (wholesaling, warehousing, retailing)
- Shipment data may also be subject to infeasibility (more on this later)

Sources: Freight Analysis Framework and Commodity Flow Survey

Extension of the CFS to include shipments out of scope of the CFS

FAF Component	Percent of Value 2017 FAF (%)
CFS(domestic)	~71%
Farm, fisheries, logging production	~2%
Petroleum and natural gas	~4%
Foreign Trade	~19%
Retail	~2%
Other	~2%

CFS survey of shipments over the course of 4 weeks (one representative week per quarter)

Shipments scaled to match Econ. Census annual values

Information In Use from CFS/FAF
Domestic Origin (132 geographies)
Domestic Destination (132 geographies)
NAICS at origin (roughly 3 digit)
SCTG Commodity (42 categories)
Trade Type (Domestic, Import, Export)
Value of Shipments
Average Distance of shipments

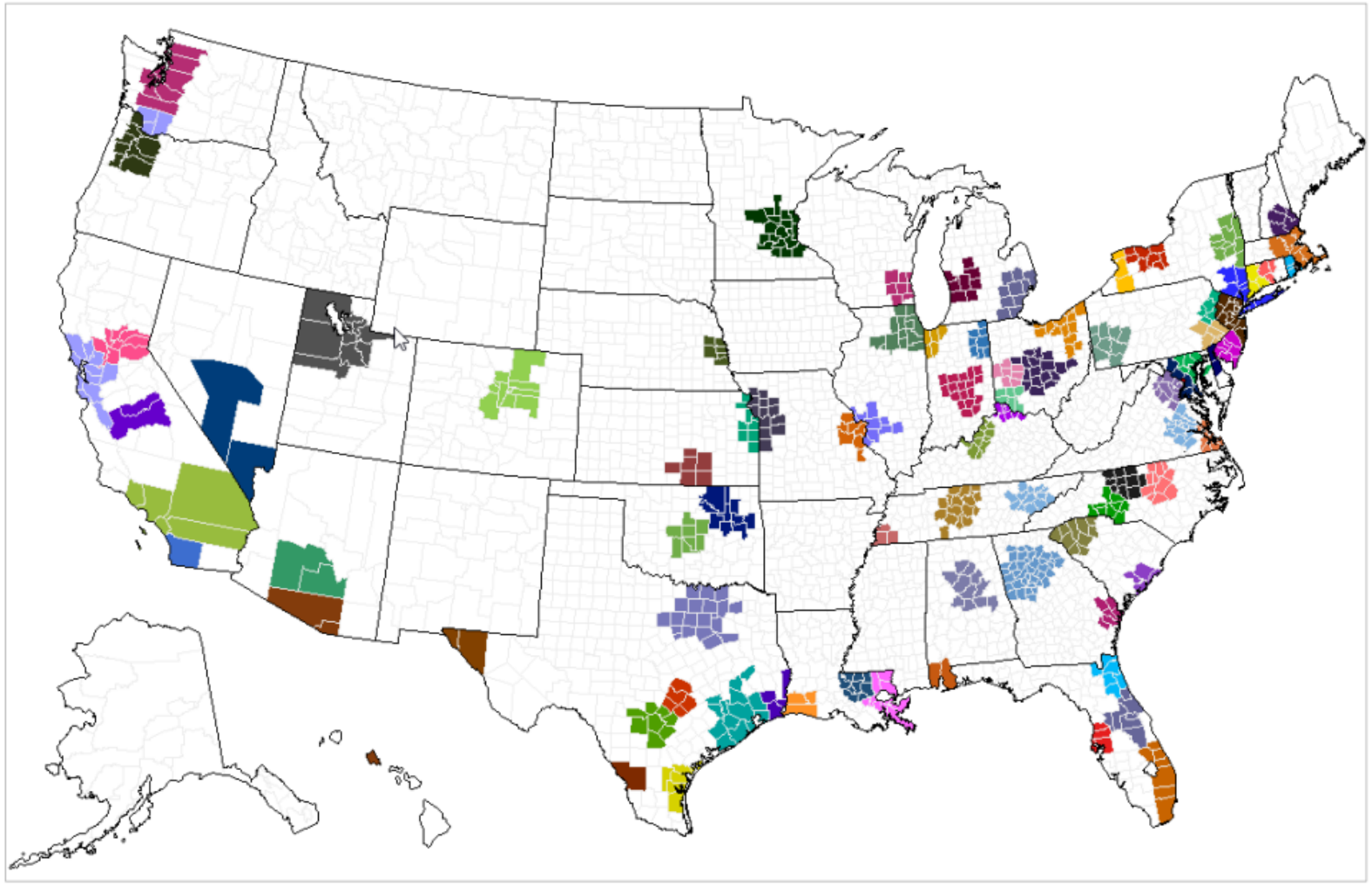
Source Data: SCTG Commodities

Code	Description
01	Live animals/fish
02	Cereal grains
03	Other ag. prods.
04	Animal feed
05	Meat/seafood
06	Milled grain prods.
07	Other foodstuffs
08	Alcoholic beverages
09	Tobacco prods.
10	Building stone
11	Natural sands
12	Gravel
13	Nonmetallic minerals
14	Metallic ores

Code	Description
15	Coal
16	Crude petroleum
17	Gasoline
18	Fuel oils
19	Natural gas and other fossil prods.
20	Basic chemicals
21	Pharmaceuticals
22	Fertilizers
23	Chemical prods.
24	Plastic/rubber
25	Logs
26	Wood prods.
27	Newsprint/paper
28	Paper articles

Code	Description
29	Printed prods.
30	Textiles/leather
31	Nonmetal min. prods.
32	Base metals
33	Articles-base metal
34	Machinery
35	Electronics
36	Motorized vehicles
37	Transport equip.
38	Precision instrument
39	Furniture
40	Misc. mfg. prods.
41	Waste/scrap
43	Mixed freight

Source Data: 2017 CFS/FAF Geographies



Note: Remainder of state (ROS) CFS Areas are shown in white.

132 areas are either:

- Metropolitan statistical areas
  - Broken down into separate FAF regions, one for each state
- Rest of state (part not included in an MSA)
- Whole state

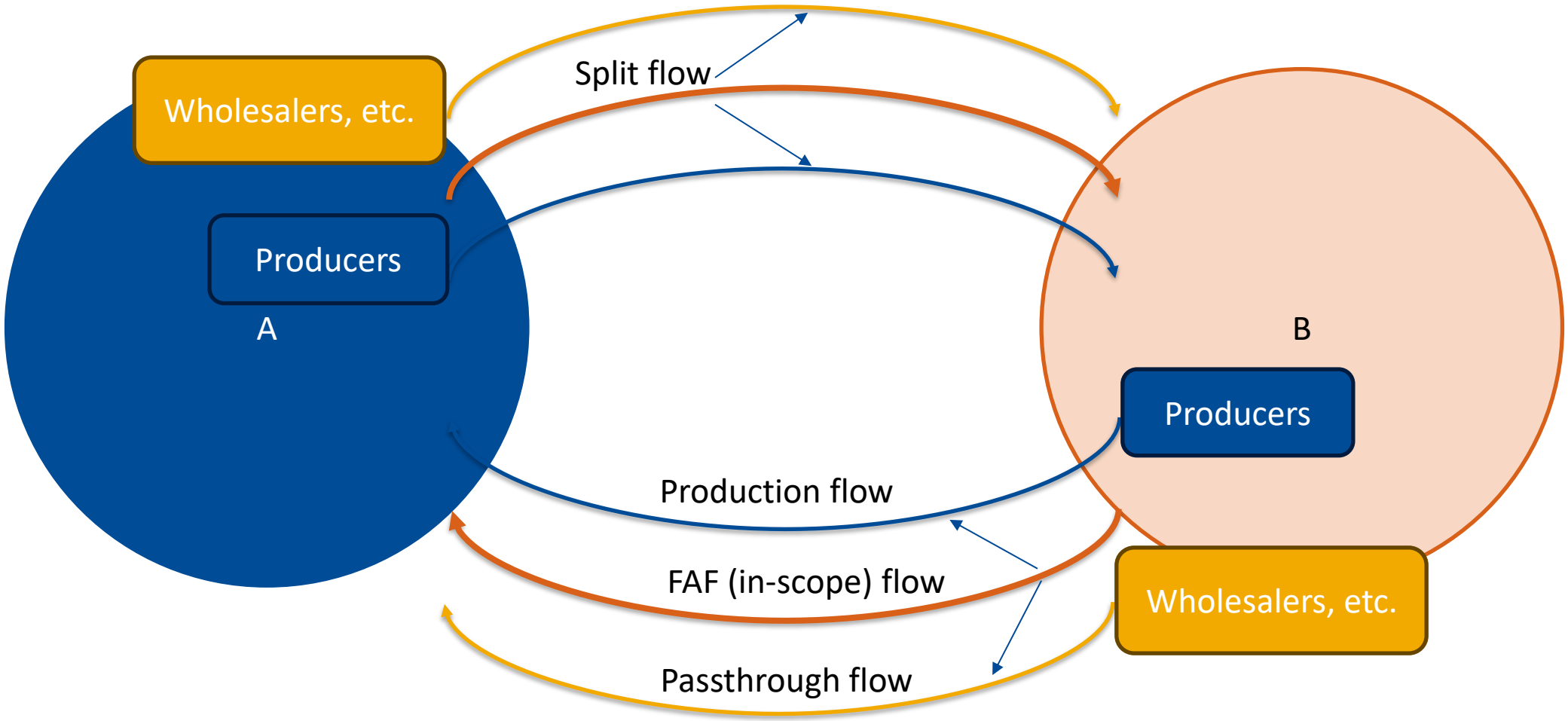
Conceptual Framework: from Shipment Flows to Trade Flows

- Initial decomposition: FAF shipment data and producer/passthrough attribution
  - Observed shipment flow (FAF): X_{ij}
  - Adjust to only include current production
  - We use CFS data on industry of the shipper to decompose into production and passthrough shipments: $X_{ij} = P_{ij} + T_{ij}$
- Flow process (conceptual):
  - Each region produces output Y_i which are shipped to other regions, P_{ij}. A portion of P_{ij} is
    - Absorbed (used as intermediate or final demand)
    - Forwarded to another region k
  - Recursive propagation: commodities may move $i \to j \to k \to l \to \cdots$
    - At each node, some fraction is absorbed and some is forwarded, until all is absorbed
- Key insight: absorption fractions determine implied passthrough flows:
  - We vary absorption shares so that the recursive forwarding process reproduces observed shipments

FAF Shipment Data and Producer/Passthrough Attribution

- FAF provides shipment data for 41 commodities based on the CFS (covering shipping-intensive industries), then expands coverage by adding shipments from non-CFS industries (e.g., farms), crude oil and natural gas flows, moving-company activity, benchmarking to international trade data, and filling suppressions.
- Our objective is to produce data by commodity × origin × destination × producer/passthrough status, but FAF does not publish industry detail, so all FAF additions must be assigned to producer or passthrough categories.
  - Easy cases: e.g., for cereal grains, FAF additions come entirely from farms, so FAF–CFS differences at the OD level are attributed to producers.
  - Harder cases: e.g., for electronics, FAF additions include retailer shipments (passthroughs) and moving-company shipments (out-of-scope), making attribution less direct.

Split Bilateral Flows Based on CFS Information (Origin Industry)



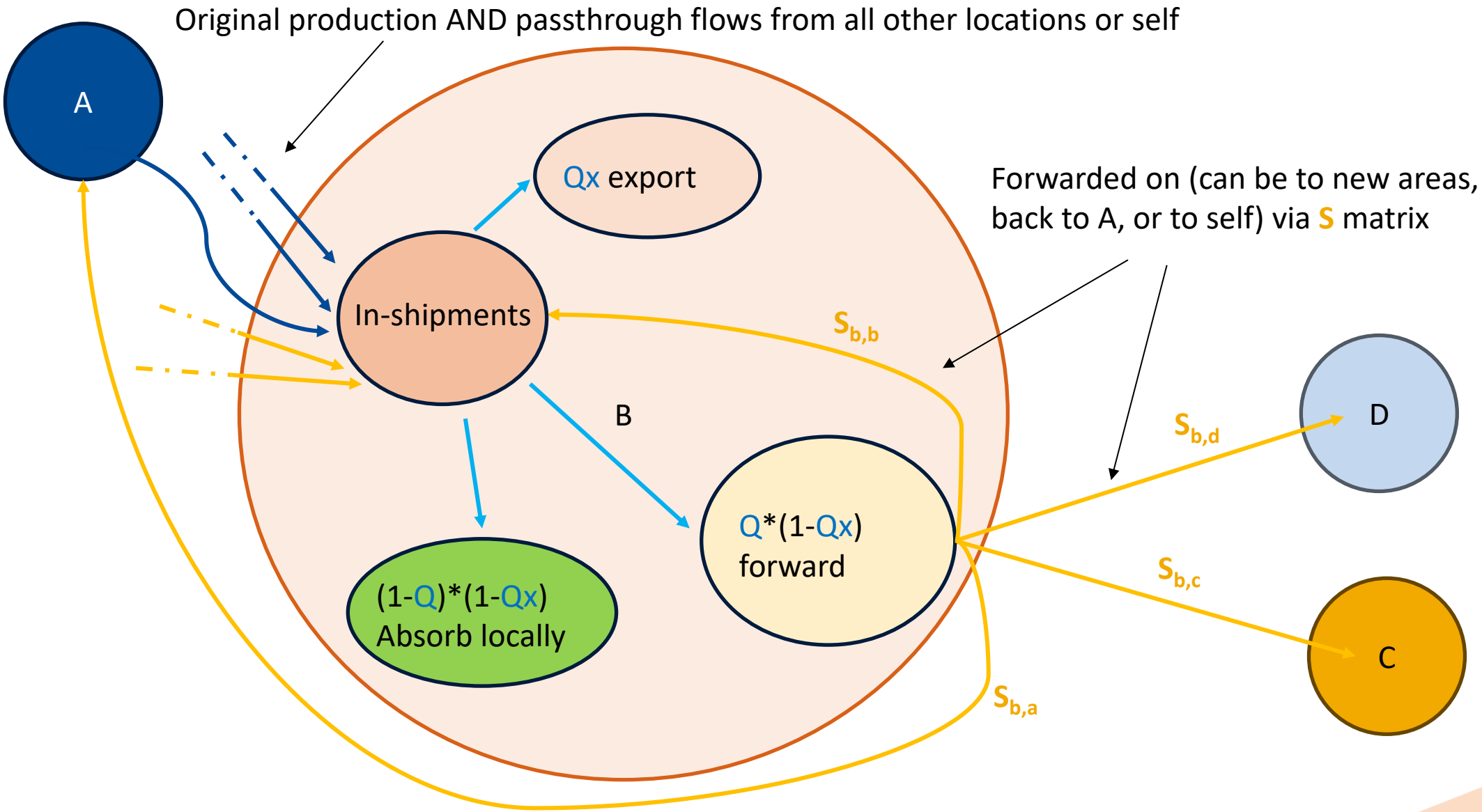
Analytic Impossibility

For some commodities and regions, CFS/FAF implies “impossible” flows

Commodity	Region	Production ($B)	Inflows ($B)	Outflows ($B)	Imbalance ($B)
Electronics	Los Angeles	15.8	50.1	99.0	**−33.1**
Pharmaceuticals	Memphis	4.4	6.0	30.9	**−20.5**
Pharmaceuticals	Sacramento	2.0	8.2	29.8	**−19.5**
Misc. manufactured goods	Seattle	2.4	9.1	29.3	**−17.9**
Pharmaceuticals	Orlando	2.1	7.2	26.6	**−17.2**
Textiles	Los Angeles	5.9	40.4	63.4	**−17.1**
Pharmaceuticals	Chicago	7.0	35.3	59.4	**−17.1**
Pharmaceuticals	Allentown-Bethlehem, PA	1.1	7.6	24.1	**−15.4**
Pharmaceuticals	Atlanta	2.1	13.9	30.4	**−14.4**
Pharmaceuticals	Dallas-Fort Worth	3.3	21.5	38.6	**−13.8**

Imbalance = Production + Inflows − Outflows; negative values indicate outflows exceed available supply.

Model: Tracing Production Flow



Basic Flows Model and Forwarding Rules

Total inflow to a location: $M_i = \sum_j X_{ji}$

- At destination i:
  - Exported: $Q_i^x M_i$
  - Forwarded domestically: $Q_i(1 - Q_i^x)M_i S_{ij}$
  - Absorbed locally: $(1 - Q_i)(1 - Q_i^x)M_i$

Routing rules

- Forwarding shares are **data**:
  - $S_{ij} \propto T_{ij},\ \sum_j S_{i,j} = 1$
  - $S_{ij}^x \propto T_{ij}^x,\ \sum_j S_{i,j}^x = 1$
  - S is T *row normalized* to 1

Parameters

- Unknowns: $\{Q_i, Q_i^x\}$
- Below, we will choose $\{Q_i, Q_i^x\}$ to match T and T^x (with cumulating forwarded flows)
- Allow up to 6 forwardings

Two-Stage Optimization

Stage 1: Solve Necessary Flow Condition

- *Flow feasibility* condition:
  - $\sum_i P_{ij} + \sum_i T_{ij} - \sum_i T_{ji} \geq 0$
- We start by adjusting the *column sums* of T_{ij} to reach flow feasibility, starting with T_{ij}^0
- Solve:
  - $T^* = \text{argmax}_T \left\{ \sum_{i,j} \left(T_{ij}^0 - T_{ij} \right)^2 \right\}$ s.t. flow feasibility
  - Only allow new links within a certain radius
- Backstop: Don't allow total T to change more than 30% (minimize infeasibility s.t. 30% change in T flows)

Stage 2: Adjust Production Flows

- Adjust column sums of $R_j = \sum_i P_{ij}$ to match T^* flows
  - Forwarding rules above give: $S_{ij}^* \propto T_{ij}^*$
  - Call starting point (from data) R^0
- Let the forwarding rules above give us $T(R, Q, Q^x)$, $T^x(R, Q, Q^x)$
- Solve:
  - $R^* = \text{argmin}_{R,Q,Q^x} \{ \alpha \log(R/R^0)^2 \}$ s.t. $T(R, Q, Q^x) = T^*, T^x(R, Q, Q^x) = T^{x*}$
  - Fill out P from R*, with a linear distance-based penalty (γ) on new production flows

$$P^* = argmin_P \left\{ \begin{matrix} \sum_{ij} \left(P_{ij} - P_{ij}^0 \right)^2 \\ +\gamma dist_{ij} 1_{(P_{ij}>0)\times\left(P_{ij}^0=0\right)} \end{matrix} \right\}$$

s.t. P* column sums equal R*

Check: how much does this process change the source data?

- Recall: origins for shipments in the source data are fixed, since they are benchmarked to Economic Census production
- Destination adjustment capped at 30% by commodity for T, uncapped for P

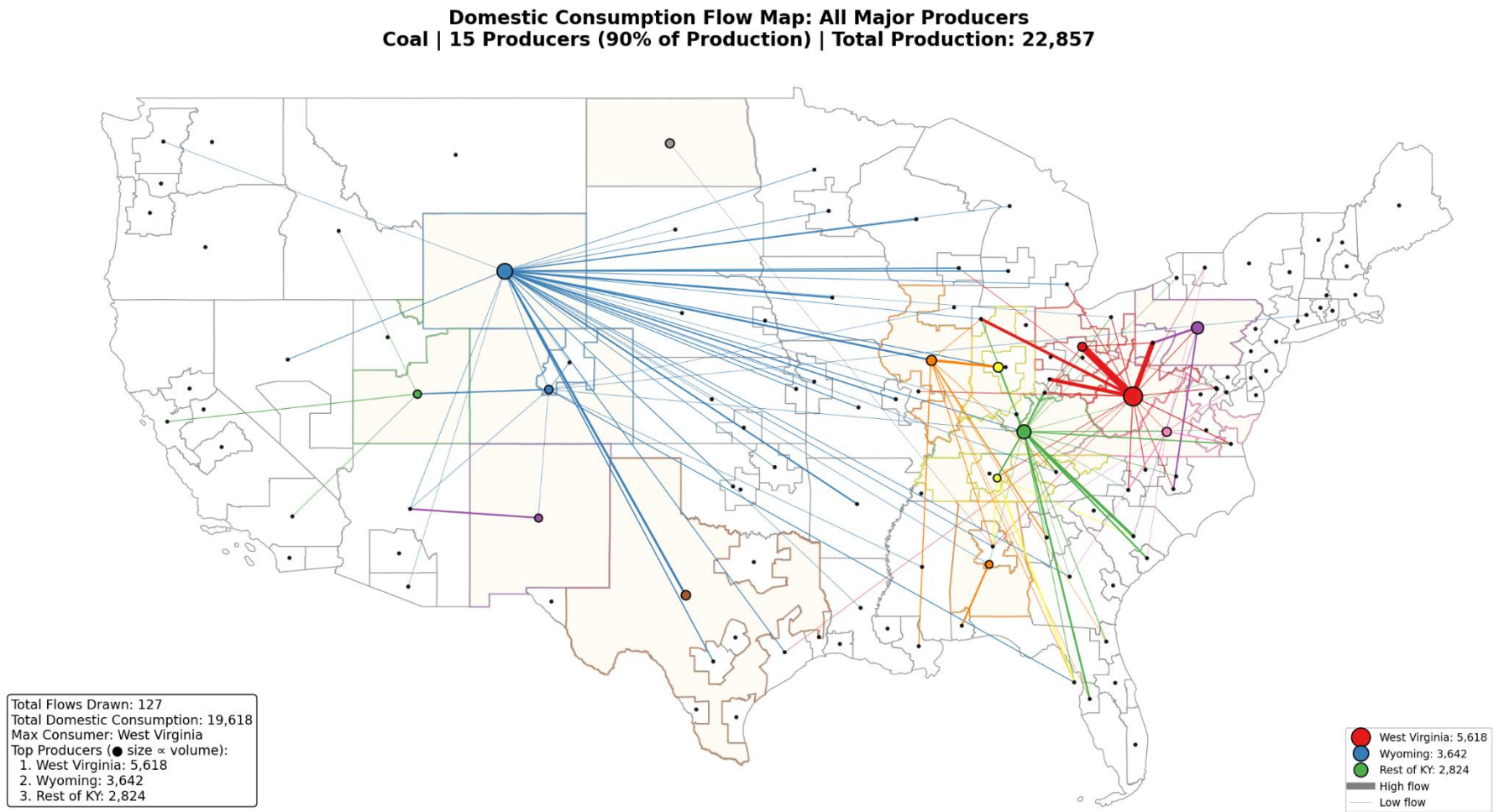
Matrix	**Total Changes ($B)**	**Total Flows ($B)**	**Percentage Change**
T	1,194	8,726	13.7%
P	1,605	5,855	27.4%

Results: Gross Merchandise Trade Values, 2017

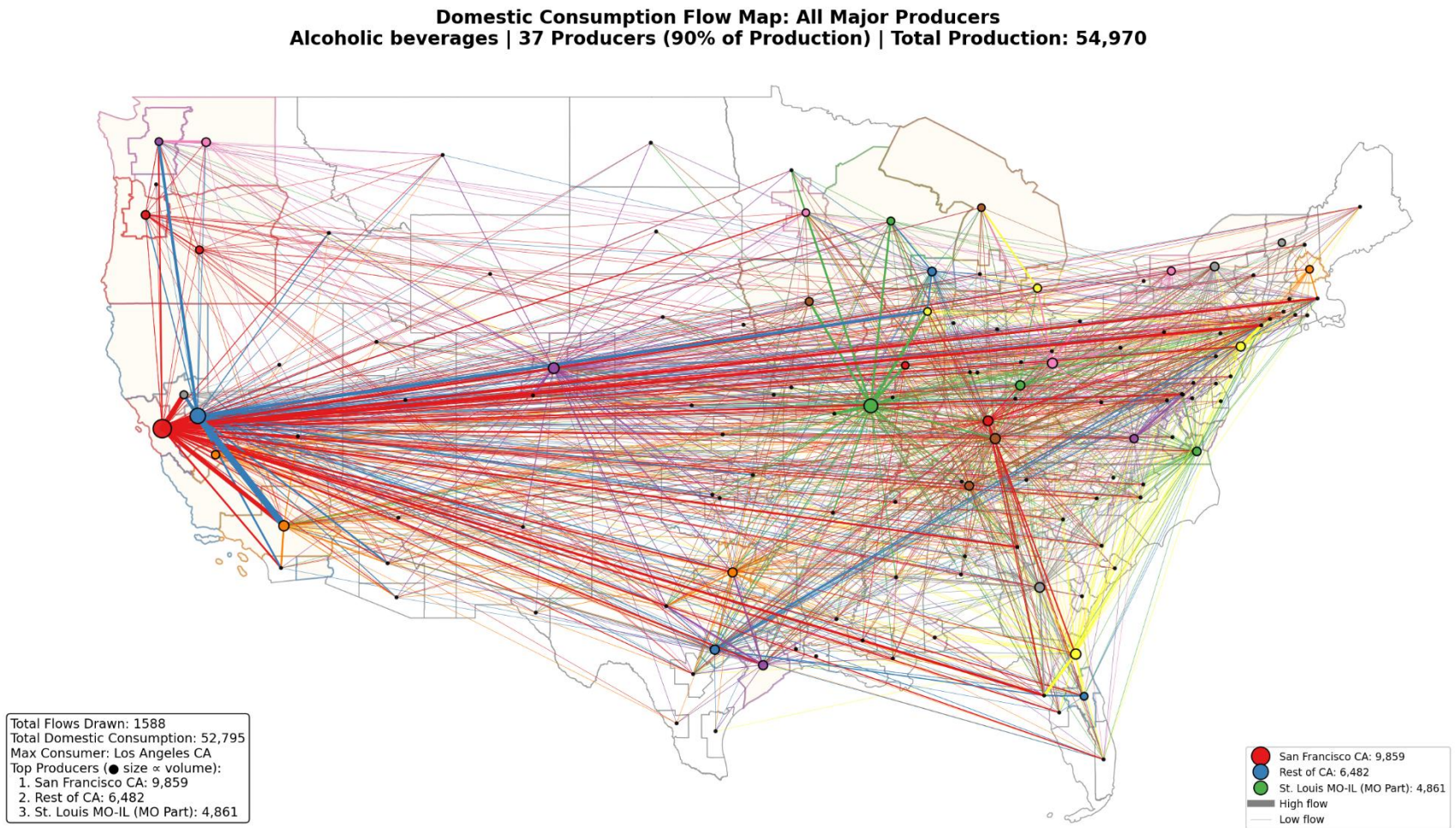
Merchandise	$ Millions
Total Production	6,735,352
Destination:	
Self	1,361,807
Other States	3,869,669
Abroad	1,503,876
Overall Exports	5,373,545

Merchandise	$ Millions
Total Consumption	7,494,438
Origin:	
Self	1,361,807
Other States	3,869,669
Abroad	2,262,961
Overall Imports	6,132,631

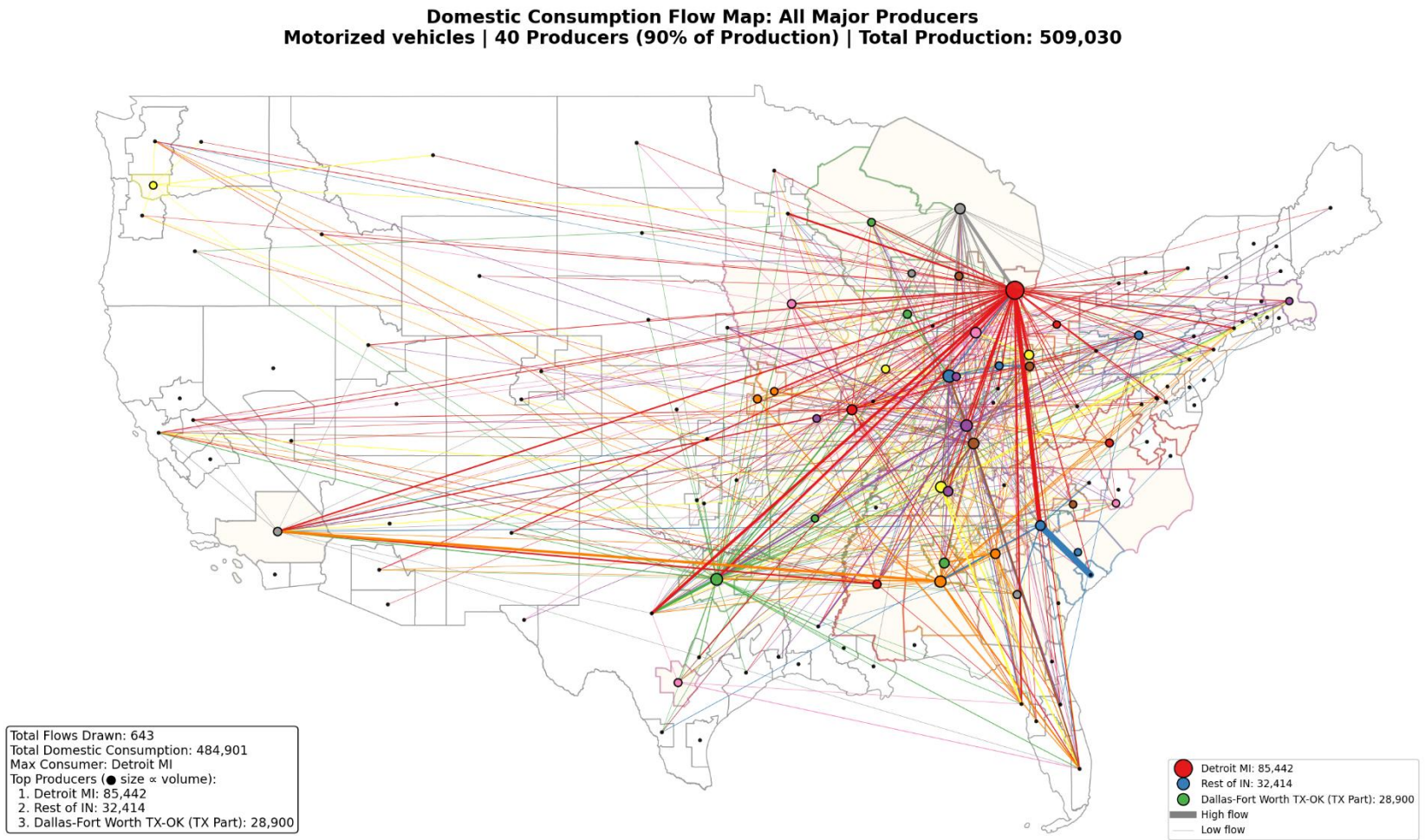
2017 Flows example – coal (domestic)



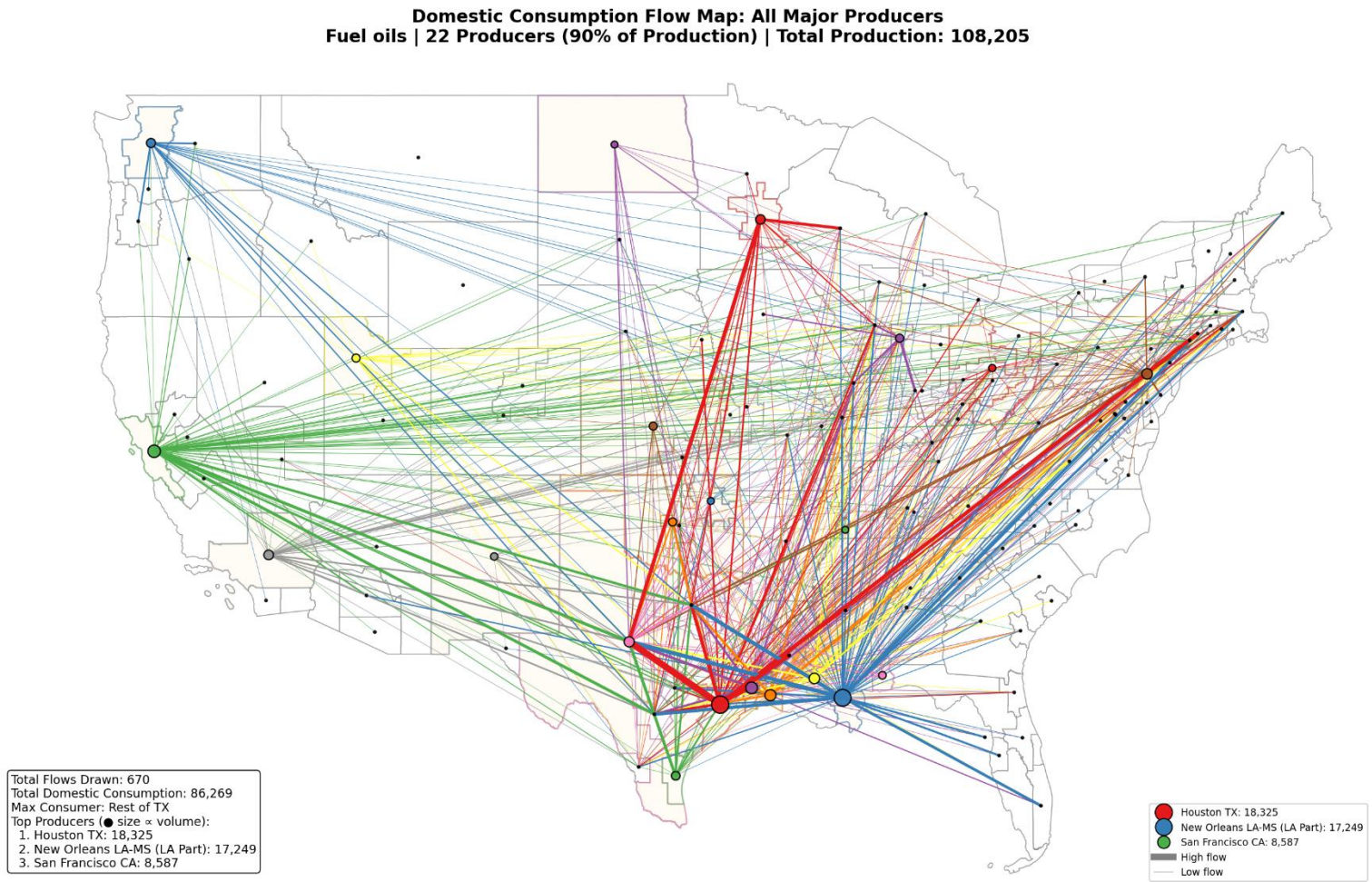
2017 Flows example – alcoholic beverages (domestic)



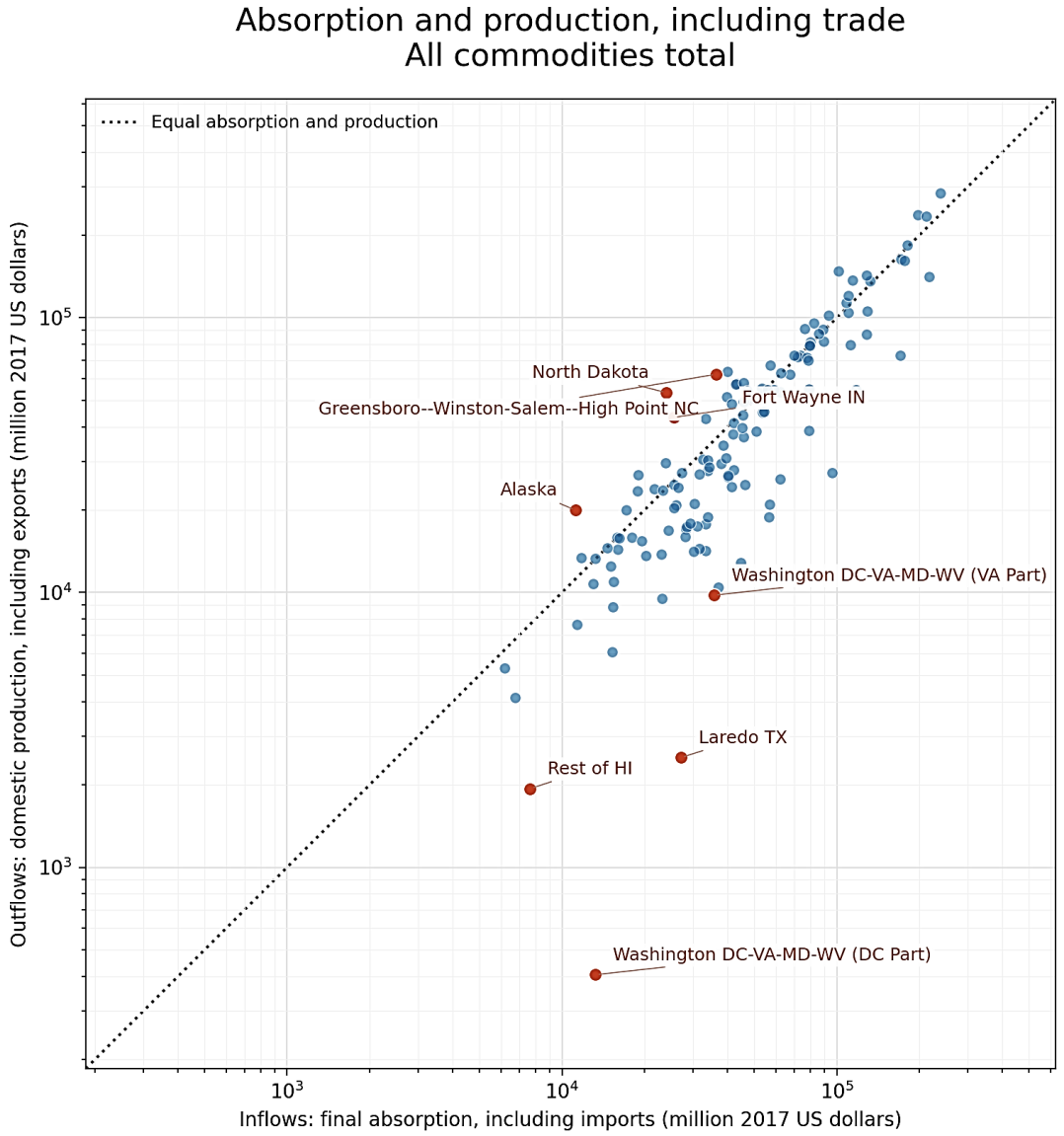
2017 Flows example – motor vehicles (domestic)



2017 Flows example – fuel oils (domestic)



Overall merchandise trade results—all goods total



Conclusions and next steps

- Testing and verification
- Combine trade flows with other production and use information
  - What do the estimated trade flows imply about use of these commodities at the destination?
  - Reconcile with independent estimates of personal consumption expenditures, investment, and intermediate consumption
  - Involves SCTG to NAICS concordance
- Part of larger BEA effort to estimate supply and use tables by state
  - System of Regional Accounts
- Comments and improvements (dirk.vanduym@bea.gov)