BLS Import & Export Price Index Expansion Using Census Administrative Trade Data

TAC

Dominic Smith and the Administrative Data Team

June 17, 2022

Challenges in MXPI

- BLS publishes Import and Export Price Indexes based on survey data and secondary sources

- Challenges
  - Declining response rates
  - Intermittent trade
  - Small samples in certain product areas

Can Administrative Data Help?

- Administrative trade data maintained by the Census Bureau used for sampling
- Where can administrative data be used in place of directly collected data?
- **Results:** Indexes accounting for 40% of trade can be replaced with UVIs

The Project in Three Steps

1. What information in the data should be used to create homogeneous items? (Do we have enough info?)

2. What index methodology should be used to aggregate items to Unit Value Indexes (UVIs)?

3. When is a UVI good enough to replace an existing price index?

Administrative Trade Data

- Maintained by the U.S. Census Bureau
- Almost all goods that enter/exit the U.S.
- Information about trading parties, location, and other transaction characteristics
- Value and quantity => Unit Values
- Sampling frame for MXPI
- Previously used for unit value indexes, but we're using more characteristics to create items

What information in the data should be used to create homogeneous items?

Which characteristics are price determining?

Color	Upper Material	Sole Material	Size	Country of Origin	Unit Value
Brown	Canvas	Rubber	11	Vietnam	100
Black	Canvas	Rubber	10	Vietnam	100
Black	Leather	Rubber	12	China	105

- Color/size probably don't matter. Materials and origin do
- In administrative data have variables collected at border
  - Which do we use?
  - Do we have enough info?

Tradeoff

- Given characteristics (W) adding characteristic x to item key

1. Increases how much of the variation in transaction UVs explained by items
2. Decreases the probability that items are traded in consecutive periods

Item/Variety = Set of transactions that share values for set of characteristics

Match Adjusted R-Squared (Chessa, 2021)

- For each set of characteristics (W) take the geometric mean of

1. Fraction of items traded in consecutive periods

$$\mu_t^W = \frac{\sum_{k \in K_{0,t}^W} q_{k,t}^W}{\sum_{i \in G_t} q_{i,t}} = \frac{Quantity\ of\ items\ traded\ in\ consecutive\ periods}{Total\ quantity\ of\ goods\ traded}$$

2. Fraction of UV variation explained by item UVs

$$R_t^W = \frac{\sum_{k \in K^W} q_{k,t}^W \left(\vec{p}_{k,t}^W - \bar{p}_t\right)^2}{\sum_{i \in G_t} q_{i,t} \left(p_{i,t} - \bar{p}_t\right)^2} = \frac{Item\ Price\ Variation}{Transaction\ price\ variation}$$

Implementation Details

- Require item keys include
  - Harmonized System Code
  - Country of Destination/Country of Origin
  - Unit of Quantity
- Consider 1028 (2^{10}) import keys and 64 (2^{6}) export keys
- Calculate statistics for three years and average over time

- Result: Longer keys in more heterogeneous areas
  - 70+ item keys used

What index methodology should be used to aggregate items to Unit Value Indexes (UVIs)?

Four Stages of Aggregation

1. Transactions (unit values) => Items (Weighted Mean)
2. Items => Classification Group (Base-year Tornqvist)
3. CG => Upper-Level Strata (Modified Laspeyres)
4. Upper-Level Strata => Headline (Modified Laspeyres)

Step 1: Unit Value (s) -> Item (k)

- Quantity weighted mean of transaction unit values
  - i.e. Unit value = revenue/quantity

$$p_{k,t} = \frac{\sum_{s \in S_k} q_{s,t} p_{s,t}}{\sum_{s \in S_k} q_{s,t}}$$

Step 2: Base-Year Strategy for CG Indexes

- Tornqvist formula to aggregate item Uvs [Item (k) -> CG (c)]
- Base-year strategy – Use UV in current month compared to average in previous year

$$r_{c,t} = \left(\prod_{k \in K_c} \left[\frac{p_{k,t}}{p_{k,y(t)-1}} \right]^{\frac{v_{k,y(t)-1}/V^{t}_{y(t)-1} + v_{k,t}/V_t}{2}} \right)$$

- Great for goods only traded part of the year! (Seasonal goods)
- Almost no chain drift (Multi-period Identity Test)

More Details: Suopera, Vartia, Nieminen, Montonen (2021)

Base-Year Strategy: More Details

- Number that is published

$$p_{c,t} = p_{c,y(t)-1} r_{c,t}$$

- Only chain once per year
  - MPIT looked significantly better than monthly chaining
  - Bi/Multilateral methods deemed unfeasible

Step 3: CG (c) -> Upper-Level Strata (e)

- Laspeyres formula using values in the base period
  - 12-24 months prior to t, updates in January

$$r_{e,t} = \sum_{c \in C_e} \frac{v_{c,b}}{\sum_{i \in C_e} v_{i,b}} \frac{p_{c,t}}{p_{c,t-1}}$$

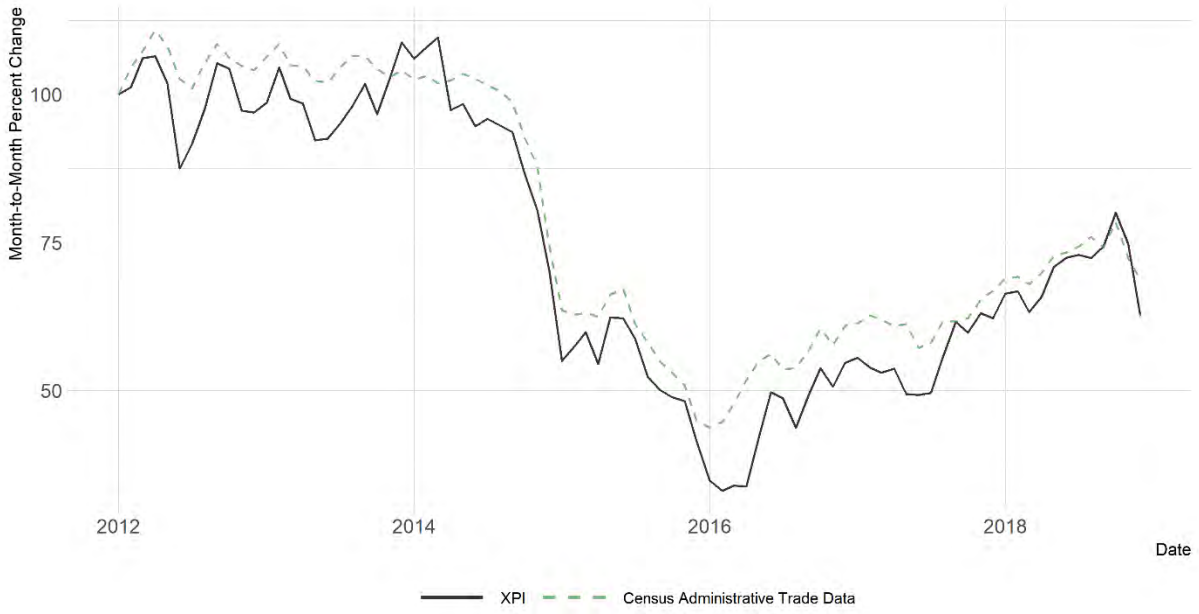
- Same strategy for step 4.
- **Both consistent with current MXPI methods**

When is a UVI good enough to replace existing price index?

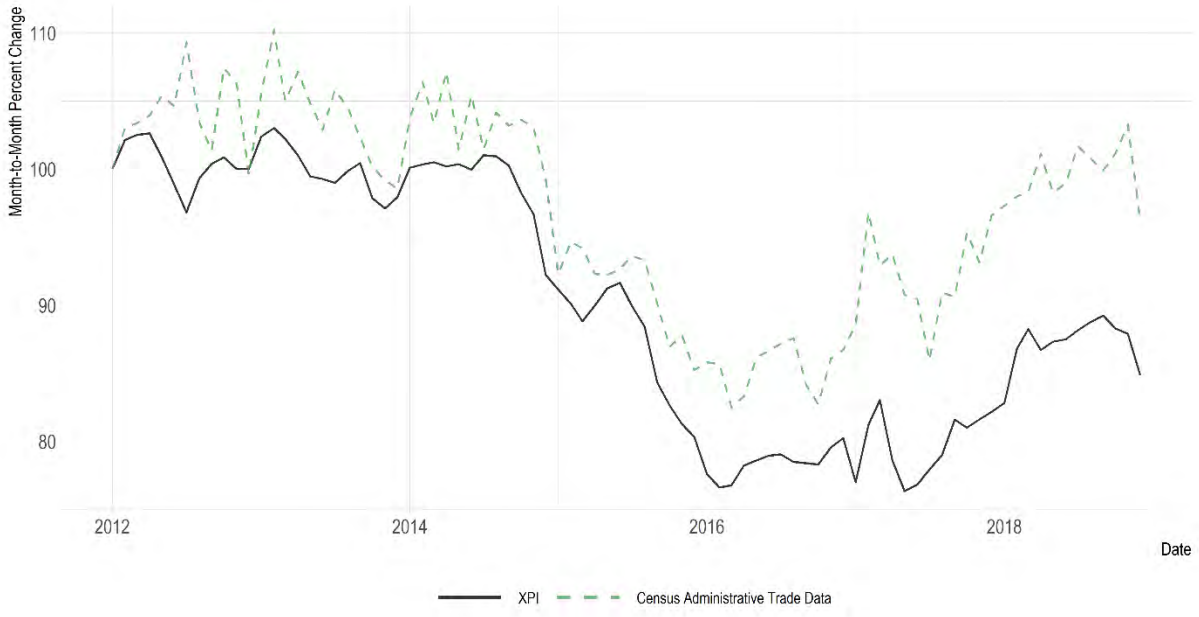
Rating Indexes

- MXPI Benchmark
  - Statistics
  - Visual Analysis
- Variability statistics
- Comparison to MXPI microdata

Fuel Oils


Month-to-Month Percent Change
100
75
50
2012
2014
2016
2018
Date
XPI
Census Administrative Trade Data

Industrial Organic Chemicals



Miscellaneous Household Goods

Trucks Example

- UVI for XX almost matches MXPI
- Heterogeneous product area, but are items homogeneous given MARS?
- Item Key:
- Explains X percent of price variation
- Coefficient of variation of items within month

Sources of Differences between MXPI and UVI

- Unit Value Bias
  - Change in composition of items. Quality increasing => UVIs biased up
- Substitution Bias
  - Tornqvist updates weights quicker than Laspeyres
  - MXPI biased up
- Sampling
  - Administrative data has everything
  - For detailed BEA categories MXPI may be based on <10 prices
  - MXPI has sampling variability and more months with no price change

Index Rating

	Imports		Exports	
UVI Rating	MPI Benchmark	No Benchmark	XPI Benchmark	No Benchmark
High Quality	46	10	22	6
Marginal Quality	12	7	17	11
Poor Quality	45	14	37	32

Notes: Numbers represent the number of unit value indexes at the BEA end-use level that were determined to be a given quality level.

High+marginal account for ~40% of index

Top Level UVI Effect is Small!

Index Classification	Change in Constant all-goods official price index
High+Marginal Export	0.6
High+Marginal Import	1.0

Notes: The first two column represent the performance of high quality and high+marginal indexes against their benchmarks. The third and fourth columns show the impact of these indexes on the top-level numbers. The final column shows the fraction of import or export value accounted for by indexes of a given type in 2015.

Change in the Number of Published Indexes

- Increase the number of published indexes by:
  - 28 (Exports), 36 (Imports)
- Replace:
  - 39 Published Import Indexes
  - 28 Published Export Indexes

Conclusion

- BLS planning to use Administrative data for ~35% of MXPI starting January 2024
- Using recent research for item definition and index calculations

Discussion Questions

1. Are there any concerns with MARS or base-year strategy?
2. Do you agree that replacing MXPI with UVIs is not a break in series?
3. Any additional considerations regarding whether to replace survey data with administrative in certain areas?

Timeline to Use Trade Data in MXPIs

- September 2022 –
  - Publish historic time series of MPIs using administrative trade data
  - Revise XPI's release with new methodology
  - Finalize selection of 5-digit BEA End Use indexes that meet quality criteria using administrative trade data
- **May 2023**
  - **Begin parallel processing of MXPIs using administrative trade data**
- January 2024
  - Release MXPIs using administrative trade data for homogeneous product areas.