Public and Historical Data in International Macro-Finance

Chenzi Xu, Assistant Professor, Stanford GSB

Historical International Macro-Finance Data Sources

Sunday, August 30, 2020
Reasons to study history

Goods Flows

Capital Flows

Crises, disasters, major events
Reasons to study history
Timeline of international monetary system

1800-1914:
First age of Globalization

1914-1971:
WWI, Interwar, WWII, Bretton Woods

1971-present:
Second age of Globalization

“GOOD DATA”

First globalization: huge increases in goods & capital integration around the world
- Led by UK on pound sterling/gold standard
- Technologies like steam ships, railroads, telegraph helped to integrate information & real markets

WWI–Bretton Woods: collapse in international trade & private capital flows
- Transition from £ to $

Second globalization: broadly very similar to first globalization
- Led by US $ with (mostly) freely floating exchange rates
- ICT revolution, GATT, WTO, regional trade agreements

“Good data” period: 25% of years since First globalization—we have a lot to learn from history!
Opportunities in historical research

1. Contributing facts: the data are not downloadable, so collecting it requires more work. But because data are less readily available, we know less. Generating facts (in the context of answering a question) has a high return

2. Natural experiments: institutional context often simpler so reduced form causal analysis possible in historical contexts that aren’t possible today
Goods Flows
Trade data

One of the most reliable historical indicators of economic activity because tariffs were the main taxable sector for most countries. Much more reliable than historical GDP data

1. Country-level panels: best for early developing countries (Western Europe, the US) and former European colonies (India, Asia, Africa); worst for Latin America
   - Ultimate sources: government reports, either compiled into Statistical Abstracts (done by the UK and US) or found individually. These often contain much more detail than just aggregate trade but have not systematically been digitized (more later)
   - Digitized sources: large-scale efforts to create historical bilateral trade data

2. Within-country panels: product/industry-level variation; almost systematically un-digitized other than exceptions below which are cross-sections or 5-year panels
   - US & UK: digitized but not cleaned at UC Davis
   - Germany: Hungerland Wolf (2018)
3 main sources:

- **TRADHIST database 1827–2014:**
  
  
  - Pros: thorough appendix with sources, contemporary borders/entities, standardized to nominal £, other country characteristics such as GDP and bilateral connections like colonial ties
  - Cons: post-1948 observations account for 84% of data (which are from the IMF)

- **Federico-Tena database 1800–1938:**
  
  https://www.uc3m.es/ss/Satellite/UC3MInstitucional/es/TextoMixta/1371246237481/Federico-Tena_World_Trade_Historical_Database
  
  - Pros: only source with any product composition (split by raw vs manufactured goods) but very coarse, commodity prices (also very coarse)
  - Cons: less well-documented, focus on countries so tend to omit smaller entities where data existed

- **RICardo project 1800–1938:** http://ricardo.medialab.sciences-po.fr
  
  - Pros: best documentation of all sources, contemporary borders/entities (researcher needs to do work of aggregating to consistent borders), appeared to digitize all available data that they found
  - Cons: pre-1840 data very scarce (only France, UK, US)
No single source has comprehensive coverage
- Most datasets work from the same source material
- Different choices were made about which observations to include, currency/exchange rate conversions, other covariates
- For broadest panel, it’s best to merge all datasets together (see next step)

Harmonizing borders
- Depending on your research question, it may be necessary to generate consistent units of observation over time (most likely at the level of “treatment”)
- Simplest solution: aggregate to larger units (but this may not always be suitable)
- Alternative: generate synthetic countries (where values from true countries are split/merged into them based on trade weights)
- Note: this will likely need to be done with each source dataset given the lack of consistent naming conventions (no ISOs)

Intensive margin issues: true zeros not distinguished from statistical zeros; almost always true with trade data but worse with historical data

Extensive margin issues: entry/exit of countries/entities based on data availability
Additional data

Within-country bilateral trade:

- *Lloyd’s List* newspaper is only source for trade flows at port-level on a daily basis historically (18th century–today). Juhasz (2018) & Xu (2020) digitize parts. Quantities are in ships, and no records of goods shipped
- Individual country sources: often provide total value of trade at ports annually or industry composition of trade (need to be standardized using SITC codes)
- No systematic industry-level bilateral trade flows across all countries: best available is UK/US provide industry composition of trade to its own trading partners
- Caveat: sometimes only quantities are recorded without prices so hard to compute values

Major changes in trade costs:

- Sail → Steam: Pascali (2017) estimates sailing times using weather & current patterns; validates with historical log books
- Canals: historical maritime guides print port-to-port matrices of travel times using routes including/excluding canals
Capital Flows
1. Central banks
2. Private capital
3. Public (“official”) flows
Central banks

- Bilateral lending among central banks historically a rare phenomenon until WWI, other than in crisis periods (Horn Reinhart Trebesch 2020b)
- Central bank lending to private entities (banks, corporations, individuals):
  - Mostly through the discount window until WWI (these data are in archives and not systematically digitized)
  - Mostly domestic transactions, but cross-border transactions happened through international banks (Ex: Bank of England lending to Ottoman Bank)
- Available data:
  - Discount policy rates (even daily): easily available from macrohistory databases, GFD, etc.
  - Market rates: NBER macrohistory database has monthly market rates taken from newspapers; higher-frequency observations requires going to the original publications, like The Economist
  - Balance sheets (Bank of England, Banque de France, Reichsbank): daily balance sheets already scanned and online, just need to be digitized. More detailed ledgers (with line-item lending) require in-person visits + photos
Private capital: equities

Main digitized database: Yale’s International Center for Finance Historical Financial Resource Data
https://som.yale.edu/faculty-research/our-centers-initiatives/international-center-finance/data/historical-financial

  - Most relevant from an international perspective: almost all private capital flows went through London
  - Digitized but needs significant amounts of cleaning
- New York (monthly, 1815–1925; daily afterward from CRSP); Shanghai (annual, 1870–1940; abolished afterward); St. Petersburg (monthly, 1865–1914; abolished afterward)

Non-digitized sources: for higher frequencies, need to return to original sources. Subsets have been digitized by different scholars but nothing systematic because of the scale

- Financial publications in each country such as *Moody’s Manuals* (Bernstein 2019), *The Economist* (Xu 2019), *Wall Street Journal* (Picel 2018), etc.
- Scans accessible through Gale newspaper databases; bulk purchases from libraries also possible and not too expensive
(Sovereign) bonds were the primary form of private cross-border capital flows

- Meyer Reinhart Trebesch (2019) has the most comprehensive coverage of these bonds
- Coverage: all foreign-currency bonds traded in London & NY from 1815–2016; 91 sovereigns
- Unique for having monthly prices (note: a large subset of price data come from the IMM); data will be available when published (forthcoming in QJE)

International banks also active, particularly in international trade (issuing capital & deposits domestically to lend abroad)

- Xu (2020) shows cross-section of British lending abroad in 1866
- Kisling Meissner Xu (2019) shows expansion of international banking by all nationalities from 1850–1914 using annual *Banker’s Almanacs*
Public (“official”) flows

Officially defined as lending between sovereigns, which mostly occurred through treaties

- Horn Reinhart Trebesch (2020b) provides first and most definitive dataset (not publicly available yet, but their sources are documented) of this type of lending from 1815–present
- Show that it’s a major form of capital flows, usually substituting for collapses in private capital (due to financial crises, wars, etc.)
Crises, disasters, major events
Major events are rare and only history provides data

Financial crises:

- Reinhart Rogoff (2009) provides most definitive dates of banking, sovereign debt, and currency crises: https://carmenreinhart.com/this-time-is-different/
- Schularick Taylor (2018) have additional macro outcomes (but are limited to advanced countries)
- Many papers study particular crises:
  - UK crises: Xu (2020) for 1866; Paolera Taylor (2001) for 1890

Wars and military conflicts:

- Correlates of War (COW) database: https://correlatesofwar.org/data-sets/COW-war
Thank you!

chenzixu@stanford.edu