...DBOOK ON LINKING TRADE ...BUSINESS STATISTICS

Diana DOYLE
Trade and Competitiveness Statistics, OECD

The OECD Handbook on Linking Trade and Business Statistics

- Linking Trade and Business Statistics is an increasingly common practice among OECD countries to develop **new, policy relevant statistics** on economic globalization

- However: **limited** international statistical **guidance** is available, even if important compilation **challenges remain**

- The OECD Handbook aims to:

  - Identify the main **statistical challenges** that are encountered in this data linking process and

  - Provide **suggestions** and concrete **examples** for overcoming these, considering the variability in national practices and resources

  - Outline the required **statistical infrastructure** necessary for moving from individual, 'ad-hoc' microdata linking projects to a fully integrated system necessary for integrated international economic accounting – **with the SBR at its core**.
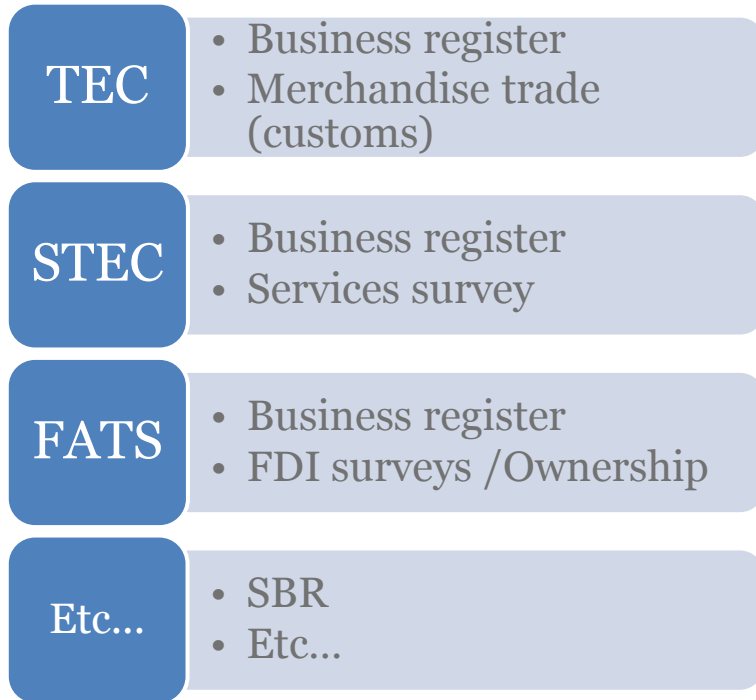
Achievements and next steps

- Linking trade and business statistics has long been on the agenda of the Working Party on Trade in Goods and Services Statistics (WPTGS) and builds on longstanding experience in EU countries supported by Eurostat

  - Important progress has been made to include an increased number of different datasets to produce wide spectrum of globalisation indicators

  - OECD WPTGS 2015 - Crea tion of Informal Reflection Group on this topic; OECD WPTGS 2016 - Held WPTGS Workshop on Linking Trade and Business Statistics to underline and address challenges and plan way forward

  - Conclusion: OECD should initiate the writing of a "Handbook on Linking Trade and Business", in collaboration with WPTGS members and other interested international organisations

- ***Organisations/Countries providing contributions to-date include:*** Eurostat, United States, Canada, Mexico, France, Spain, Netherlands, Denmark, Turkey

- ***Next steps:*** Begin dissemination of Chapters to Reflection Group, **integrate further country contributions** and Present draft chapters to WPTGS March 2018
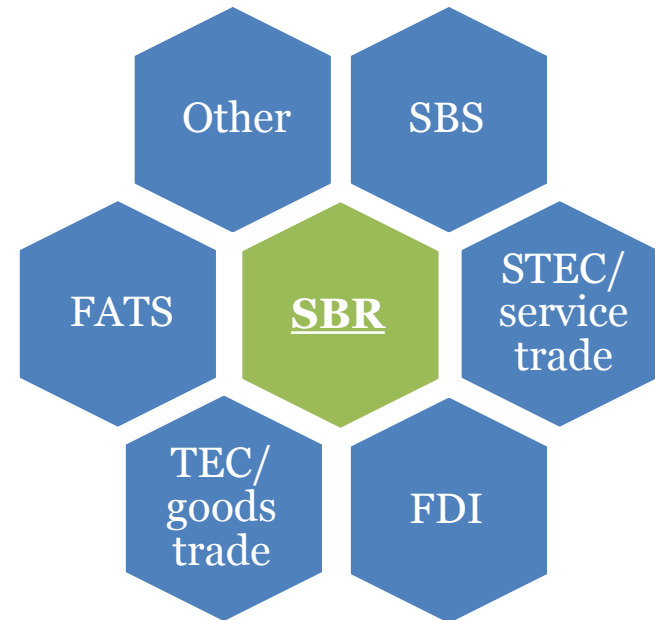
Linking trade and business statistics

From individual projects...

TEC
- Business register
- Merchandise trade (customs)

STEC
- Business register
- Services survey

FATS
- Business register
- FDI surveys /Ownership

Etc...
- SBR
- Etc...

...to an integrated approach

Other

SBS

FATS

SBR

STEC/ service trade

TEC/ goods trade

FDI

Key features of the SBR for Linking Trade and Business Statistics

- The SBR identifies the **population of economically active enterprises for statistical purposes** and **their relationship with legal units for administrative purposes**

- SBR provides coverage of the **firm universe** through **profiling firms** and **updating profiles**

- Establishes and maintains links between **legal/administrative AND statistical identifiers (and identification variables)**

- Maintains both a **live register** and **frozen frames (**snapshots taken periodically and made available for survey frames and for linking trade and business statistics)

- Assigns **statistical units** to each entity within a **firm hierarchy, a parent-child structure which** (ideally) includes up to the ultimate foreign or domestic parent, in the case of **multinational enterprises**, which tend to be the most complex firms

- Facilitates coordination and **harmonisation (**without all economic areas needing to maintain awareness of its efforts

Data linking challenges addressed in the OECD Handbook

- ***Statistical infrastructure*** challenges
  - The organization of the SBR
  - Institutional arrangements and legal systems
- ***Compilation*** challenges
  - The practice of data linking:
  - Grossing up to published totals
  - Differences in statistical units
  - Conceptual/definitional differences
  - ...etc!
- ***Dissemination*** challenges
  - What policy questions can be addressed
  - Dealing with confidentiality

Statistical infrastructure challenges: the SBR

- Pressures to **move from direct data collection at enterprises to the use of data in administrative registrations**, leading to complex **mixed mode designs** in data collection

- Increased use of administrative data in statistics production further enhances integrative and central role of SBR, but also imposes further challenges with implications for data linking

- Handbook: What are the challenges related to a **shared stage of linking trade and business statistics** AND **survey frame production**, with the **legal/administrative enterprise structure** a more important part of the business register than before?

- How will **SBR adapt** toward facilitating linkage in data domains, such as international trade, and how will those **data domains adapt** toward SBR norms?

Statistical infrastructure challenges: the SBR

- Within linkage, which is the **dominant statistical domain**: SBR or other domain?
    - E.g. apply SBR characteristics to trade in goods OR identifying traders within SBR universe?
- **Profiling coverage adequate** for current and future linking?
    - Updating – is timeliness timely enough? Does it accommodate changes over time requirements? New data sources?
- **Frozen frame content evaluated** for surveys and linking projects?
    - Decisions taken regarding mergers and acquisitions (both in profiling and in timing of profiling), tools to facilitate matching when no common identifier
- **Feedback integration from linkage programs** on firm structure, etc in live registers and frozen frames? Same mechanisms or new?
- **SBR definitions** relative to definitions in new data sources

Statistical infrastructure challenges: legal/institutional arrangements

- **Legal/Institutional challenges may be more severe in decentralized statistical system**

  – Legal constraints mean difficulties in obtaining access to SBR, access to administrative data, sharing of microdata and can prevent linking data

  – Some decentralized statistical systems have made progress with innovative arrangements; however, they highlight that **MOUs not embedded in system**; not permanent; setbacks likely

  – While challenges exist in centralized agencies (access protocols, including anonymising data shared, etc), as well, there are many **additional steps and resources to navigate data sharing, linkage strategy and publication** in decentralized legal framework (options more limited)

  – In decentralized systems, key data domains are often located in distinct agencies

    - Trade in goods and Trade in services

  – Multiple existing business registers

Compilation challenges (examples)

- Important practical questions remain regarding the best techniques regarding how to link (**common identifiers, propensity score matching**) and on what data sources would be most useful to link

- Despite SBR serving as the common frame for all surveys, each survey may have a distinct survey frame, with distinct weighting

- **Differences in sample coverage** across data sources
  - dealing with different **grossing up** factors across sources
  - the treatment (possible imputation) of missing data
  - and the possibility of positive coordination across sample surveys to facilitate subsequent data linking (including strategies for dealing with changes in respondent burden)
  - calculation of sample error

- Dealing with **large, complex firms**

- Compilation challenges for internal production processes vs external dissemination

Compilation challenges (examples)

- **Absence of one clear data source** for a key data domain (trade, employment, etc)

- Allocating **indirect exports** to producer (ie exports flowing from a wholesaler/distribution agent that from production statistics clearly did not produce good)

- Very **different concepts and approaches** to measuring trade in goods and trade in services in most statistical agencies (in decentralized systems, the two are often in distinct agencies), meaning that the creation of a trader universe a key challenge despite high demand

- **Differences in statistical units** used across datasets (or across national statistical agencies) create large inefficiencies and make data linking problematic

- Definitions of key actors: SMEs, MNEs

Dissemination challenges

Confidentiality

- Increased granularity due to linking raises confidentiality concerns, especially for **smaller countries**:
  - How to disseminate lower and higher quality estimates for experimental and established linkage programs, **communicating the quality to data users while providing maximum data points**
  - How to identify and reflect the confidentiality level of the original data source
  - How to **evaluate confidentiality for derived indicators**

Addressing policy questions

- How to identify successful linkages (fewest resources for highest quality result), leveraging strengths of system
- How to **formulate indicators and policy messages** that can be derived from linking (firm size, ownership, trading status)

Proposed table of contents (1)

Proposed table of contents (2)

Chapter 1. Introduction

Motivation behind OECD Handbook on Linking Trade and Business Statistics and outline of Chapters

Chapter 2. Identifying the Challenges

Enumerating the challenges relating to statistical infrastructure, compilation and dissemination challenges that will be discussed at length (complete with proposed solutions and national examples) in later chapters

Chapter 3. Statistical infrastructure Challenges

Business registers; main globalization-related datasets involved; units and classifications

Chapter 4. Compilation challenges: data linking practices

Record linking; linkage practices in absence of common identifiers

Proposed table of contents (3)

Chapter 5. Compilation challenges: Statistical units and classifications

Dealing with large/complex businesses (e.g. consolidation); exploiting information on group structures; assigning indirect trade to enterprises; 'creative' classifications

Chapter 6. Compilation challenges: Imputations and grossing up

Post-hoc recalibration of grossing up factors; employing logistics and other regressions for imputations; shrinking the universe (e.g. only produce statistics for large firms)

Chapter 7. Compilation challenges: Designing statistical systems with data linking in mind

From data linking project to developing an enriched statistical system: new stratification variables for other official sources; new survey frames; positive survey coordination.

Proposed table of contents (4)

Chapter 8. Data dissemination: confidentiality concerns

Strategies for dealing with confidentiality: obtaining explicit or implicit consent from respondents, coefficients of variation; 'smart' aggregations; using sample frame size.

Chapter 9. Data dissemination: Policy relevant indicators derived from linked trade and business statistics

Examples of indicators and policy messages that can be derived from these indicators (focus on firm size, ownership, trading status).

Chapter 10. Data dissemination: Using linked trade and business data to enhance statistics on firm heterogeneity in Global Value Chains

Use of linked trade and business data to refine GVCs analysis: extended SUTs and examples of ongoing empirical projects.

Thank you for your attention!

Statistical business register experts are kindly invited to comment on the outline of the Handbook on Linking Trade and Business Statistics.

In addition, they are invited to contribute relevant national expertise and examples regarding the compilation issues that are addressed.

Contributions as well as requests to be added to the distribution list to receive chapters of the Handbook on Linking Trade and Business Statistics can be directed to: Diana DOYLE, Diana.Doyle@oecd.org.