

CDER: An Overview and Recent Development on Business Microdata

Jiang Beryl Li
Economic Analysis Division
Statistics Canada

100

STATISTICS CANADA

ONE HUNDRED YEARS AND COUNTING

Prepared for Canadian Economics Association 2018 Conference



Statistics
Canada

Statistique
Canada

Canada

Roadmap

- ❑ What is CDER?
- ❑ What microdata are available?
- ❑ What is new?
- ❑ How to access CDER?



100

Canadian Centre for Data Development and Economic Research (CDER)

□ Background

- Created to allow Statistics Canada to make better use of its business data holdings without compromising security
- Launched in Jun 2011 and expanded in Oct 2012

□ Activities

- Provides external researchers with secure access to business microdata for analytical projects
- Serves as a repository for business microdata
- Leads the development of new business microdata

□ Key information

- Accessible at Statistics Canada's head office in Ottawa
- Application and approval process required for access
- Researchers must cover the full cost of their project



Business Microdata at CDER

- ❑ Stand-alone, research-ready data already in use
 - E.g., Trade by Characteristics of Enterprises; Survey of Financing and Growth of SMEs linked to tax data; and Annual Survey of Manufactures
- ❑ Linkable File Environment (LFE)
 - Specific variables from a set of files where linkages have been done, but files are so large that extractions are made upon request
- ❑ Developmental datasets and other linkage environments
 - Analytical databases containing derived variables for specific analyses (e.g., National Accounts Longitudinal Microdata File); additions to LFE; other linkage environments (e.g., Canadian Employer-Employee Dynamic Database); new stand-alone data



100

Trade by Enterprise Characteristics

- ❑ Enterprise level, 2010-2016
- ❑ By Industry, province, CMA, and by employment of enterprise
- ❑ Trade by Exporter Characteristics (TEC)
 - Value of exports by HS8 commodity
 - By country of destination, including U.S. States breakdown
- ❑ Trade by Importer Characteristics (TIC)
 - Value of imports by HS10 commodity
 - By country of origin, country of export, and by related-parties
- ❑ Uses
 - Trade dynamics: barriers to trade, firms' life cycle in the export/import market
 - Firm dynamics: trade orientation, owner demographics and firm characteristics (if linked to administrative data)
 - Industry or product concentration: Herfindahl index

100

Survey of Financing and Growth of SMEs (SFGSME)

- ❑ Cross-sectional survey
 - 2000, 2001, 2004, 2007, 2011, and 2014
- ❑ Survey information
 - Use of financing during start up
 - Requests for financing (term loans, mortgages, lines of credit, credit cards, government loans, equity) and outcomes (approved/rejected, collateral, term, interest rate, amounts requested/received)
 - Business information (exports, R&D, innovation, IP use, plans for growth, public procurement participation)
 - Owner information (age, education, experience, country of birth, language of primary decision maker; %female, %aboriginal, %visible minority)
- ❑ Linked to admin. data on firm performance, 2000-2015
- ❑ Uses
 - Financing cost; lending relations; and financial constraints on firm dynamics



100

Annual Survey of Manufactures (ASM)

- ❑ ASM (1961-2012) – a series of longitudinal datasets
 - Cross-sectional, establishment level survey of manufacturing
 - Principal industrial statistics and commodity file... 100s of variables in the latest database
 - Essentially a census up to 2012
- ❑ ASM has been linked to other sources
 - National Pollutant Release Inventory (NPRI) and Greenhouse Gas Reporting Protocol (GHGRP), plant level, 2000 to 2012
 - General Index of Financial Information (GIFI), ASM-enterprise level, 2000 to 2012
 - Research and Development in Canadian Industries, ASM-enterprise level, 2000 to 2009
- ❑ Uses
 - Productivity, trade, and innovation



100

National Accounts Longitudinal Microdata File (NALMF)

- ❑ Longitudinal database of Canadian enterprises, 2000-2015
 - Successor to T2-Longitudinal Employment Analysis Program
 - Tracks a richer set of firm characteristics over time (e.g., employment, payroll, revenue, profit, assets, tangible assets, R&D, investment, and productivity)
 - Updated longitudinal structure
- ❑ Main data sources
 - Statistics Canada's Business Register; Corporation Income Tax (T2); Payroll Account Deductions (PD7) and Statements of Remuneration Paid (T4 slip); and Goods and Services Tax (GST/HST)
- ❑ Uses
 - GDP and employment across firm-size categories
 - Activities at sub-provincial levels (e.g., CMAs and ERs)
 - Firm dynamics, survival, and growth



100

Canadian Employee-Employer Dynamics Database (CEEDD)



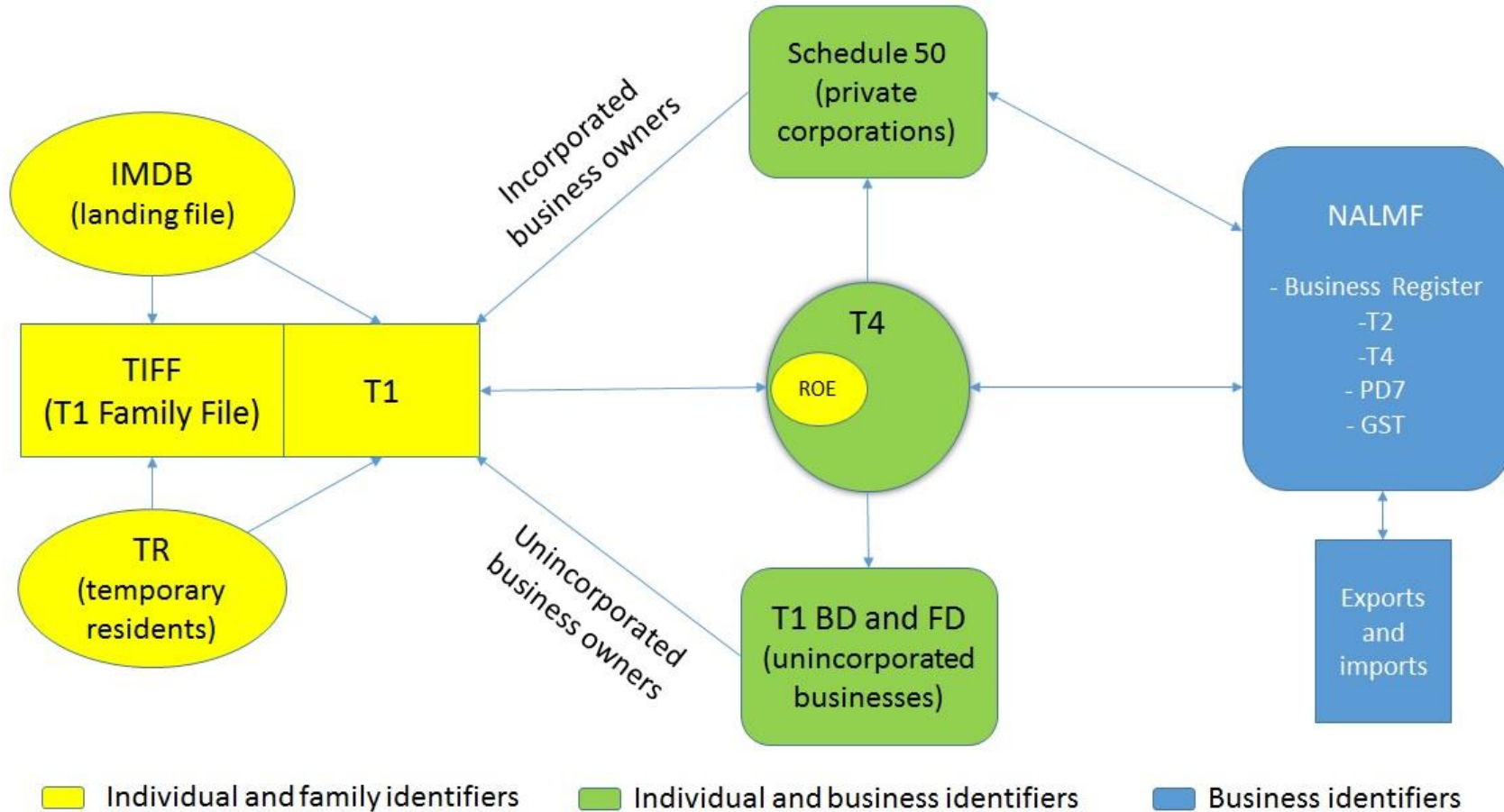
100

- ❑ Matched employer-employee database, 2001 to 2015
- ❑ Covers the universe of
 - Individual tax filers and their families
 - Employees receiving T4 slips
 - Unincorporated businesses and corporations
 - Unincorporated business owners reporting self-employment income
 - Owners of Canadian-controlled Private Corporations
- ❑ Links across various administrative tax files
 - T1: Personal Master File (T1PMF); Historical Personal Master File (T1H); Family File (T1FF)
 - Longitudinal Immigration Database (IMDB); Temporary Residents File (TR)
 - T1: Business Declaration File (T1BD); Financial Declaration File (T1FD)
 - National Accounts Longitudinal Microdata File (NALMF)
 - T2 Schedule 50; T4 Statement of Remuneration Paid Files (T4); Record of Employment (ROE)
 - Trade by Exporter Characteristics (TEC); Trade by Importer Characteristics (TIC)





CEEDD – Linkage environment



Current Use of CEEDD

- ❑ Labour market and firm dynamics
 - E.g., productivity and wage dispersion; the impact of merger and acquisition; spillovers and entrepreneurship on a spatial scale; hiring decision; and wage setting policy, etc.
- ❑ Immigrants
 - E.g., careers and network; initial firm allocation and earning growth; business ownership; and trade orientation, etc.
- ❑ Government programs: evaluation and policy development
 - E.g., employment at the economic regional level; business ownership by gender/age/immigration status/marital status



100



CEEDD

❑ Business Owner Module (2005-2015)

- Immigrant
- Gender
- Age
- Family-owned (future direction)
- Indigenous group (future direction)

❑ Geographic information

- Census metropolitan areas (CMAs)
- Self-contained labour areas (SLAs): ongoing



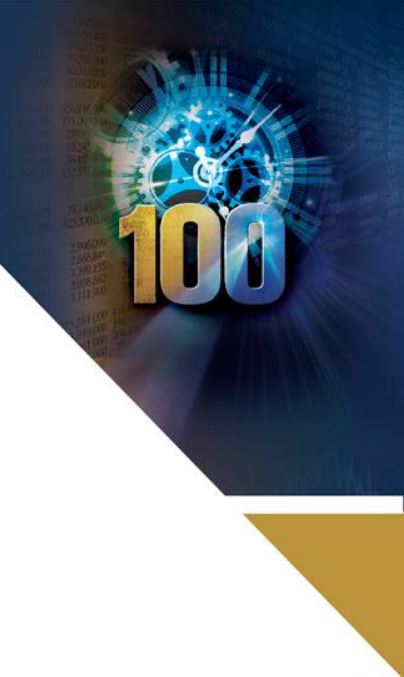
100

Treasury Board Secretariat Horizontal Review

- ❑ Collects data on Government of Canada programs supporting business innovation and clean technology
 - Financing data from 98 program streams of 22 federal organizations (incl. departments, agencies and crown corporations) over 2007-2016
 - By organization, program, year, and funding/award type
- ❑ Matched records linked to Statistics Canada's business microdata in the LFE
- ❑ Allows for research and analysis in federal government funding of innovation and clean technology



Survey Redesign on Manufacturing & Service Producing Sector



□ Manufacturing and wholesale industries

- Generic Modules

- ✓ Revenue and expenses; e-commerce; sales for secondary activities; purchased services; origin and destination of goods manufactured/sold (top 4 commodities)

- Recent Development

- ✓ Publication now in December following the end of reference period
- ✓ Employment question back on Annual Survey of Manufacturing and Logging (ASML)
- ✓ Commodity data for Annual Wholesale Trade Survey
- ✓ Increased sample for food industries in ASML

□ Retail and service industries

- Generic Modules

- ✓ Revenue and expenses; e-commerce; client type; client location; purchased services; and international transactions

- Survey-specific module (characteristics)

- ✓ Revenue source breakdown; industry-specific information (e.g., publishing, film, accommodation, and property)

New Business Microdata at CDER



100

- ❑ Cluster mapping and site selection databases (ongoing)
 - 67 industries by geographic region
 - Key economic indicators (employment, business location counts, wage and salary indicators) as a variety of social-economic and demographic indicators
- ❑ Indices of remoteness and accessibility
 - An index of remoteness, complemented by a set of accessibility measures to selected services
 - Computed for all CSDs of Canada that recorded some population in 2011 or that were connected to the road network
- ❑ Crowdsourcing
 - E.g., collect cannabis prices, map buildings across Canada, pinpoint disasters and road closures

New Business Microdata at CDER (cont'd)

- ❑ International Accounts and Trade
 - Canada and the World Statistics Hub (13-609X)
 - Trade by Characteristics of Enterprises: 2010-2016
 - Inward and Outward Foreign Affiliates Statistics (FATS)
 - Enhanced details on Securities Statistics
- ❑ Industrial Organization and Finance
 - Financial and Taxation Statistics for enterprises
 - Quarterly Survey of Financial Statements
 - Biannual Survey of Suppliers of Business Financing
 - Corporate Returns Act



New Business Microdata at CDER (cont'd)

□ Investment, Science, and Technology

- Canadian Survey of Cyber Security and Cybercrime (RY2017 in fall 2018)
- Canadian Internet Use Survey
- Survey of Innovation and Business Strategy (RY2017 in fall 2018-spring 2019)
 - ✓ Innovation
 - ✓ Advanced technology use
 - ✓ Business strategies and practices
 - ✓ Global value chains
- Canadian Housing Statistics Program (next release in June 2018)
 - ✓ Link numerous administrative files (e.g., tax files, property assessment files, etc.)
 - ✓ Used to understand and address foreign-home-ownership-related issues
 - ✓ By geography(CMAs & CSDs in Vancouver and Toronto), property/owner characteristics, property use/financing



New Business Microdata at CDER (cont'd)

❑ Farm-level data

- Census of Agriculture: 1986, 1991, 1996, 2001, 2006, and 2011
- Farm Financial Survey, 2013 and 2015
- Livestock, crops and horticulture (potential)

❑ Price Indices

- Cover a range of goods, services and construction outputs
- By type: transaction price; unit value; margin price; list price; and model price
- Unpublished indices are available in many cases (incl. geography/product/industry) and with a more recent starting date



Application Process

- ❑ By type of researcher: government (federal/provincial) and non-government
- ❑ For non-government researchers, the process is as follows
 - 1) Contact CDER and draft a proposal
 - 2) Submit final proposal
 - 3) Proposal evaluation: peer review and approved by an internal Statistics Canada committee
 - 4) Undergo a security screening
 - 5) Take the Oath of the Office of the Statistics Act to become deemed employee
 - 6) Sign a Microdata Research Contract
- ❑ Contact:
 - Website: <http://www.statcan.gc.ca/cder>
 - Email: statcan.cder-cdre.statcan@canada.ca

100

Improving Access to Business Microdata

1 - Synthetic Data Projects

- More sophisticated synthetic data developed to facilitate access outside of headquarters
 - Synthetic Longitudinal Business Database
 - ✓ Wholly synthetic Longitudinal Employment Analysis Program (LEAP) database, 2001 to 2015
 - ✓ Used to study business and employment dynamics
 - ✓ Final runs on master file will allow the use of more detailed industry and geography
 - Pilot project
 - ✓ Calls for proposal
 - ✓ Approved research projects that focus on multivariate analysis will be carried out in the RDCs in 2018-2019
 - ✓ Researchers will develop models using the synthetic/treated data at the RDCs
 - ✓ Programs will be sent to CDER by RDC analysts to be run on the actual data
 - ✓ Final results are released from CDER by e-mail
 - Disclosure-limited Survey of Financing and Growth of Small and Medium Sized Enterprises (SFGSME) linked to CRA tax data



100

Stay Tuned this Summer: News from Statistics Canada!



Access to More Business Microdata in the Research Data Centres

The Canadian Centre for Data Development and Economic Research will launch a call for proposals for economic research using the Survey on Financing and Growth of Small and Medium Enterprises 2014 in the Research Data Centres. This survey is designed to find out what kinds of financing small and medium enterprises are using, and to collect information on recent attempts to obtain new financing. The survey also contains information on barriers to growth, the financial position of small and medium enterprises, the characteristics of ownership and the extent to which the enterprise is involved in innovation and intellectual property.



Data Challenge for Students

The Economic Analysis Division at Statistics Canada will announce a student data challenge for the 2018/2019 school year. The challenge will be based on a set of data that tracks quarterly firm dynamics in Canada. The goal of the challenge is to provide a forum for students to demonstrate their capabilities for analysis, synthesis and/or data visualization.



Research Fellowship for PhD candidates

The Statistics Canada Analytical Studies Research Fellowship-Research Affiliate Program will provide PhD candidates, students currently enrolled in a doctoral programme who are working on their theses, an opportunity to access microdata holdings that are only available for use at Statistics Canada's Ottawa headquarters.

To receive a notification when the above announcements are made,
please send an e-mail to:

statcan.cder-cdre.statcan@canada.ca



100



Improving Access to Business Microdata

2) Productivity Partnership

- ❑ Team of experts from the academic, private, and public sectors
- ❑ Lowers barriers to access business data for academics and students by providing funding for productivity research projects
- ❑ Statistics Canada liaison researcher: Beryl Li
- ❑ Important partner in the development of synthetic datasets
- ❑ Collaboration and support to clarify the CDER application process steps

<https://productivitypartnership.ca>

THANK YOU!

For more information,
please visit

www.statcan.gc.ca/eng/cder/index



#StatCan100



Statistics
Canada Statistique
Canada

www.statcan.gc.ca

Canada

100