

Tax Data for Researching Innovation

Nick Greenia*
Statistics of Income Division
U.S. Internal Revenue Service

* The views expressed in this presentation are the presenter's and not necessarily those of the U.S. Internal Revenue Service.

Two Papers

- *Studying Innovation in Businesses: New Research Possibilities* (N. Greenia, K. Husbands-Fealing, J. Lane)
- *Federal Tax Data: A Case of Under-Utilization for Economic Analysis* (N. Greenia)

Overview

- Conceptual utility of tax data
 - Advantages of directly enlisting tax stat. office
- Content for studying innovation
- Challenges in utilizing tax data
- Overcoming the challenges
- Current innovation study efforts

Utility of Tax Data

- Reliability
- Richness, Variety
- Uniqueness
- Advantages of the tax agency's statistical office

Reliability of Tax Data

- **Universality:** In the interest of tax-incentivized businesses (and individuals) to be captured by the system; tax data also may be useful or mandatory for attracting investors/lenders
- **Accuracy:** Compliance program ensures a certain level of quality and regularity for the population of filers – especially entrepreneurs seeking to attract capital or even go public
- **Electronic filing (both mandatory & voluntary)** – eliminating transcription & even optical scanning

Richness, Variety

- Both employer & non-employer firms (50 million US commercial businesses -- corporations, sole-props, pass-through's); employees & contract labor; pension plans; non-profit organizations
- Linkages : employers & employees, corporate families (parent/subsidiary corporations); household families (parents & dependents); estates & beneficiaries; “passthrough” entities & shareholders/partners

Uniqueness of Tax Data

- Effective tax rates -- post-return transactions, including amended returns; audit results; and especially carrybacks of net operating/capital losses & certain credits; e.g, foreign tax credit
- For statistical purposes, indefinite retention of data is possible (in perpetuity) – enables linked population time series data sets at the record level
- Captures small start-up enterprises not considered conventional “businesses” by some data producers, but may be ideal for first identifying innovative activity

Advantages of Tax Agency's Statistical Office

- Confidentiality – data need not be transferred to another agency's stewardship; e.g., another statistical office
- All tax data accessible, not some subset
- Statistical tax data edited for analytical purposes, (not subject to FOIA & Privacy Act) preserve data "purity", unaffected by audit or other administrative actions
- Direct access as
 - Sampling frame for innovation surveys
 - Followup performance data to validate & even complement survey responses

Content for Studying Innovation/Entrepreneurship

- Business Births
 - Non-employer (sole proprietorships)
 - Employer (application for EIN), tax returns
- Legal form of ownership (LFO): sole prop, corporate, passthroughs (taxed at the partner or shareholder level – can be individuals and firms)
- Track changes in LFO with transactions documenting business life-cycle; e.g., Bill Gates: sole prop to Microsoft corp status; even to non-profit status: Bill & Melinda Gates Foundation

Content (cont.)

- Complete Balance sheets (econ. stocks)
- Complete Income statements (econ. flows)
 - Research tax credit
 - Qualified research detail (see Form 6765)
 - Capital infusions might show up in several items
- Capital flows; e.g., passthroughs to investors; estates to beneficiaries; carrybacks freeing up past tax year refunds for current usage
- Demographic data; e.g., industry/occupation classifications
- Physical location of firm, track firm migrations
- Transactions capturing: mergers, acquisitions, bankruptcies, final & part-year return transactions

Challenges to Utilizing Tax Data

- Data Architecture Holes
- Data Inaccessibility

Holes in the Data Architecture, Including Unknown Quality

- Employee Plans & Sponsors' Tax Returns
 - Form 5500 Series, Business Tax Returns
- Business Ownership Hierarchy
 - Subsidiaries, Parent, Ultimate Parent
- Business Evolution; Mergers & Acquisitions
 - Business expansion
 - M&A
- Effective Tax Rates
 - Details re Amended returns, Carrybacks, Audit Actions
- Professional Employer Organizations
 - Who's filing returns at firm level, for workers?

Data Architecture Needs

- Fundamental Principle:
 - Tax Administration needs complement Economic Analytical Needs
 - Principle related to improving data accessibility

Barriers to Data Access

Importance of --

- Statute (all accesses)
- Voluntary Tax Compliance System
- Protection of Taxpayer Confidentiality, including perceptions
- Fundamental tension between right to privacy and confidentiality vs. policy analysis needs, respondent burden, cost

Access Mechanisms, Enhancements

- Statute/Regulations
 - Statistics production mandate
 - Use of agents/contractors
 - Outside requests for special tabulations or studies
- Policy Agreement
 - Census/IRS Criteria Agreement (Sept 2000)
- National Economic Imperative
 - Example: US Survey of Consumer Finances
 - Federal Reserve Board as US Treasury contractor/agent,
 - NORC as subcontractor
- Organizational will, driven/supported by statistical research community

Current Initiatives to Study Innovation using Tax Data

- Innovation Survey for large businesses (NSF via Census Bureau):
 - Limitations: Very few tax items available at Census
- Innovation Survey for small businesses (NSF via IRS/SOI):
 - Limitations: Unknown, has not been done before
 - Advantages
 - Few limitations on tax items accessible – mostly a function of what is already processed to electronic medium
 - NSF is SOI's contractor, directly participates in not only survey design (theirs) but also research analysis of response data in conjunction with any further tax data needed to validate and even complement survey data.

Current Initiatives (cont.)

- American Recovery and Reinvestment Act (stimulus funds) -- pilot project under development via Inter-Agency Task Group to establish an infrastructure for studying outcomes from funds spent on research & development
 - Potential use of federal tax data to identify, track, and analyze recipients of such R&D funds