

Table 1. Comparing Longitudinal Primary Data and Population-Based Administrative Data

| Characteristics | Longitudinal Primary Data (Panel Study in Income Dynamics) | Population-based Administrative Data (Manitoba and other sites) |
|-----------------------------------|---|---|
| Number of cases | Several thousand or smaller (5,000 households in PSID) | Often over a million |
| Cost | High on a per person basis | Very low on a per person basis |
| Representativeness | Often national | Often from a province or state |
| Population Studied | Subjects sampled and tracked | Built on a registry of an entire population |
| Research Design | Often complex designs need to increase power and control costs | Given a population, complex designs can be imposed retrospectively as needed |
| Record Linkage | Useful in some contexts | Critical to check data quality and expand scope of information sources |
| Individual Follow-up | Before and after an event | Before and after an event |
| Coverage and Loss to Follow-up | Nonresponse and differential attrition possible | Differential attrition possible |
| Updating | New data must be collected and merged with existing data | Multifile information must be cleaned and merged with existing data. Cleaning relies on record linkage |
| Time | Information must be collected (typically annually or at longer intervals) | Information provided at relatively short intervals (from daily to annually) |
| Place | Information at time of study (historical reconstruction possible) | Detailed information usually provided close to date of move |
| Longitudinal | Goes back many years/Corrections for immigration in PSID) | Goes back many years |
| Neighborhoods | Flexible construction from postal code or census area | Flexible construction from postal code or census area; large N may permit flexible assignment to generate nearest 'neighbors' |
| Life Events | Collected as part of design | Possibly available from registry or other sources |
| Family and Intergenerational Data | Collected as part of design (PSID); sibling and intergenerational studies facilitated | Assessing a family composition at any point in time possible; sibling and twin studies facilitated |
| Limits | Important information likely to be collectable for entire sample | Important information may be missing or available only for a subpopulation |
| Measures | Defined by researchers; scaling possible | Defined by others for administrative purposes; creating meaningful variables may be very time-consuming; scaling possible |
| Intellectual History | Scope of data collection often expanded to provide a rich data set | Scope expanded beyond initial health care data by receiving files from other agencies |