Vital Statistics: Birth and Death Files

Overview
The NYC Department of Health and Mental Hygiene’s (DOHMH) Bureau of Vital Statistics makes certain data collected from NYC birth and death certificates publically available. Vital events data is primarily used to inform DOHMH priorities and are widely used in research, internally and externally.

Methodology
The Bureau of Vital Statistics captures demographic and medical information for all births and deaths in NYC. Medical facilities and funeral homes throughout NYC are required to register information on approximately 120,000 births and 55,000 deaths per year.

Publically Available Data
Limited Use Public Data Files: DOHMH provides pre-defined identifiable Limited Use Birth and Death Datasets available from 1996 forward. While direct identifiers (e.g. names, address, etc.), literals, and exact dates are removed, sufficient individual characteristics of births and deaths are included in the dataset that can potentially re-identify individuals. Variables include demographics (age, education, race/ethnicity, etc.) and maternal and birth characteristics (delivery method, gestational age, birth weight, etc.). Data are made available as SAS datasets, with accompanying file formats. A short application and data use agreement is required for access to these datasets.

Birth Aggregate Micro-datasets: These datasets include aggregated counts of births with demographics for three geographical units: community districts, zip code, and census tract. The community district and zip code datasets include univariate counts of maternal and birth data and bivariate counts (e.g. maternal age categories*maternal education categories, maternal race/ethnicity*parity, etc.). The Census Tract datasets include univariate counts of maternal and birth characteristics only. Multiple years are combined for each of the three datasets as described below. Death aggregate micro-datasets will be available in 2016.

|-------------------------------------------------------------|---------------------------------------------------------------|-----------------------------------------------------------------|

Strengths and Limitations
Data characterize all events in the population and there are >15 years of data available; however, data quality varies, and there is time lag between file closure and availability. Due to confidentiality concerns, access to some of the vital event data elements may be restricted or denied.

Sample Research Questions
Possible research could include individual level and multilevel analyses to assess the structural and demographic predictors of premature mortality and infant mortality.

Publications

