California Longitudinal Administrative Database

The poverty rate for California, 21.0 percent, is the highest in the country (when poverty is measured with the Supplemental Poverty Measure). The cost of running a high-poverty state is very high whether measured in terms of human suffering, gross domestic product foregone, tax revenues foregone, higher program costs, or such externalities as increased crime and health costs. As it stands, California runs its poverty and labor market policy largely in the blind, indeed it has mainly watched from the sidelines even as other states (e.g., Florida, Tennessee, North Carolina, New York) have come to rely heavily on administrative data to monitor their labor markets and evaluate and develop policy. The state has had no choice but to cobble together analyses of proposed policy changes (e.g., the EITC supplement) on the basis of incomplete and inadequate data; it is forced to track trends in poverty with data that are collected only infrequently; and the state’s counties have been obliged to guess about how best to target social service spending because national surveys are not large enough to provide small-area or small-group estimates. Moreover, given that individual-level data are not linked over time, the state does not know who is moving into and out of poverty, employment, and other labor market statuses. The state is obliged, in short, to attempt to run one of the world’s biggest economies and labor markets without the data needed to make informed decisions.

This problem can be solved. The California Longitudinal Administrative Database (CLAD) is a new administrative-data infrastructure for California that will make it possible to monitor real-time trends in labor market outcomes as well as evaluate how new and existing programs are delivering on their stated objectives. Although one could of course set up one-off evaluations of state programs, it is more cost-effective to carry out the requisite evaluations by building an administrative database in California that adds value by linking existing data. This new infrastructure entails linking administrative tax data, administrative program data, and survey data.