

World History Conference, Sydney, 3-9 July 2005

*Abstract: Statistical Confidentiality and the Construction of Anonymized Restricted-Access Census Microdata Samples: Kenya, 1969-1999*

Agnes A. Odinga and Robert McCaa, Minnesota Population Center July 9, 2004

Kenya has one of the richest collections of census microdata in the world, with four datasets at ten year intervals beginning in 1969. This valuable trove is little used by scholars or public policy-makers. The first step in providing broader, if still restricted access--and reaping the benefits to be gleaned from these valuable sources--is to anonymize the microdata to attain the highest levels of statistical confidentiality. The IPUMS International project, in cooperation with a group of National Statistical Agencies in Europe, the Americas, Asia and Africa, is developing uniform standards for anonymizing census samples of individuals and households. This paper summarizes research on statistical confidentiality and, then as a test case, applies emerging international practices to samples for the 1969, 1979, 1989 and 1999 censuses of Kenya. The results are promising. For example, of the thirty-six person variables in the 1989 census microdata, four are suppressed entirely (because they report finely detailed information on place of residence), and another six undergo some degree of aggregation. While this will disappoint purists who demand total access to the original data, this approach strikes a balance between access and statistical confidentiality, sacrificing some degree of detail to safeguard statistical confidentiality to a maximum, yet still make it possible for researchers to gain access to the Kenyan census microdata.