ESTIMATING VITAL RATES IN THE DEVELOPING WORLD

ESTIMATING VITAL RATES IN THE DEVELOPING WORLD: A BAYESIAN PROCESS MODELING APPROACH. 2015-2020. K01 Award from NICHD to Tyler McCormick.

My ultimate goal to establish an independent research agenda that develops novel statistical methods for population research in developing nations and other data-constrained environments. I focus specifically on developing estimates for vital indicators, which are especially critical to understanding population dynamics, developing public programs, and implementing or evaluating public health actions. In most parts of the developing world, there is massive uncertainty about even the most basic indicators. Achieving this objective requires an interdisciplinary skill-set that has three components: (i) expertise in statistical modeling, (ii) an understanding the historical, social/cultural and economic underpinnings of core themes in demography and (iii) experience with the complex realities of collecting demographic data in developing countries. After doctoral work in statistics, I am prepared for the first of these three components. My training and development plan proposes a series of activities to address the second two components. My mentoring team consists of Stewart Tolnay (mentor), Sam Clark (co-mentor), Adrian Raftery (advisory committee) and Basia Zaba (advisory committee). First, I will pursue training to understand, and eventually contribute to, substantive questions in demography and ecology. I will work to understand how various social, cultural, and economic factors relate to individuals' demographic outcomes and how these outcomes relate to population dynamics. Though social science questions motivate my study of statistics, I have no formal training in demography and my only formal training in the social sciences is at an undergraduate level. I address this gap in my current training through coursework and directed readings with a highly skilled and experience mentoring team. Second, my statistical training leaves me unprepared to address the complex realities of data collection in developing nations. My statistical training emphasizes analysis tools for data already collected, often under restrictive assumptions. Data used for demographic research in developing nations, however, often violates these assumptions and nonsampling error is rampant. I address this gap through coursework as well as fieldwork experiences. I propose two substantial (consisting of approximately 6-8 weeks each) fieldwork experiences at the Agincourt Health and Demographic Surveillance System in the northeast of South Africa. The Agincourt site, which features prominently in both my development and training plans, includes annual census and special events updates (systematic recording of all births, deaths and migrations), making Agincourt one of the very few places with both high-quality validation data and infrastructure to implement and evaluate new data collection methodologies. During my visits I will, under the supervision of my mentoring team, observe interviews, meet key survey research personnel, and discuss the findings and ideas of my research proposal with Agincourt investigators. My experiences in Agincourt are a tangible link between the research and training components of my proposal. The research proposal focuses on estimating fertility in such situations and understanding the key drivers of changes in national and regional fertility patterns. Fertility is an important determinant of population size and composition. Quality information about fertility is key for formulating national and regional policy, developing public programs, and implementing and evaluating public health actions. I propose a technique for estimating fertility in developing countries that emphasizes the relationship between data collection, model, and outcome. An overarching Bayesian modeling framework incorporates nonsampling error, draws strength from similar respondents, and naturally shares uncertainty between different data sources. The proposed methods would reduce bias by adjusting for variability introduced through nonsampling errors, provide statistically principled measures of uncertainty for national and subnational estimates and generate recommendations for efficient survey design. Using the same modeling framework, I will also evaluate specific hypotheses about observed and projected trends in fertility. Aim 1 develops a model to estimate national and subnational fertility rates in developing nations and evaluates that model using both DHS and Agincourt data. Aim 2 proposes a
A microsimulation environment that facilitates testing hypotheses about fertility patterns and dynamics at an individual or household level. This environment also facilitates testing hypotheses about measurement error, which will again be evaluated extensively using Agincourt and DHS data. Aim 3 develops models to project future fertility rates that incorporate uncertainty in the underlying individual-level covariates that are associated with changes in national and regional rates. I will also make projections using both past Agincourt data (a census has been in place for approximately 20 years) and make actual predictions of future fertility rates in Agincourt that I will evaluate at the end of the project period.