

It was . . . necessary to be sure the African chosen would undertake his work efficiently and successfully, as with a period of only a few days to be employed, he might be tempted to sit under a banana tree and write the first figures which came into his head [on the census forms]. —C. J. Martin, “The East African Population Census, 1948”

INTRODUCTION

AN ANTHROPOLOGIST AMONG THE DEMOGRAPHERS

Assembling Data in Survey
Research Worlds

In 1948, C. J. Martin (1949, 315), director of the East African Statistical Department, speculated that African data collectors for the census in Uganda, Kenya, and Tanzania might invent the data they were meant to record. In mid-June 2008—sixty years later—I sat with a group of Malawian data collectors in a minibus parked in a village in central Malawi where they were administering household-level surveys for an American-led longitudinal cohort study, the Longitudinal Study of AIDS in Malawi (LSAM).¹ They had finished their work for the day and were conversing about one of their colleagues as he sat under a tree nearby, pencil in hand and head bent over a survey questionnaire. As he checked the questionnaire to ensure that each question had been answered by the respondent, those in the van jokingly accused him of “cooking data” (*kuphika madata*). Soon after the conversation, the minibus hurried back to the LSAM field office nearby, where the team’s completed questionnaires were deposited in cardboard boxes until the information they contained would be carefully entered into a growing database by a data entry team.

| | |
|---------------------------|--|
| DATE OF INTERVIEW | <input type="text"/> [<input type="text"/>][<input type="text"/>] (Day, Month) |
| TIME STARTED | <input type="text"/> [<input type="text"/>][<input type="text"/>] (24 HOUR TIME) |
| INTERVIEWER NAME | <input type="text"/> |
| INTERVIEWER NUMBER | <input type="text"/> [<input type="text"/>] |

| | |
|---|--|
| RESPONDENT'S IDENTIFICATION | |
| Village name and number | <input type="text"/> [<input type="text"/>] |
| Headman's name | |
| Head of compound | |
| Respondent's name and Respondent ID | <input type="text"/> [<input type="text"/>] |
| Respondent's other names/nicknames | #living children |
| Respondent's level of education (circle and fill in level): (0) No school (1) Primary-Level | <input type="radio"/> |
| (2) Secondary-Level | <input type="radio"/> |
| (3) Higher | <input type="radio"/> |
| Respondent's birthplace (District and Village) | |
| Respondent's father's name | |
| Respondent's age (estimate if respondent doesn't know) [<input type="text"/>] | Check if age was estimated by interviewer [<input type="checkbox"/>] |
| Respondent's marital status 1....MARRIED 33....NEVER MARRIED 44....SEPARATED 55....DIVORCED 66....WIDOWED | |
| Husband's name | |
| Husband's other names/nicknames | |
| Husband's birthplace (District and village) | |
| Number of other wives that husband has | |
| Husband's level of education (circle and fill in level): (0) No school (1) Primary-level | <input type="radio"/> |
| (2) Secondary-level | <input type="radio"/> |
| (3) Higher | <input type="radio"/> |

| SUPERVISOR | LOGGED BY | CHECKED BY | ENTERED BY |
|------------|----------------------|----------------------|----------------------|
| INITIALS | <input type="text"/> | <input type="text"/> | <input type="text"/> |
| DATE | <input type="text"/> | <input type="text"/> | <input type="text"/> |

FIGURE I.1. LSAM questionnaire, 2008.

Cooking data refers to fabricating, falsifying, or fudging the information one is meant to collect from survey respondents in a standardized and accurate manner. Martin’s fears that enumerators might “write the first figures which came into [their] heads” on their forms reflect his stakes in the first endeavor to accurately map African populations in the territories his office oversaw, express racialized hierarchies of suspicion, and illustrate how data collectors’ practices in the field might spoil census data that would later be analyzed in the office. Meanwhile, in 2008, the phrase “cooking data” operated among Malawian fieldworkers as playful commentary on colleagues’ work performance, indicating that they had come to articulate and embody the habits, investments, and standards central to the collection of high-quality data, as imparted to them by American demographers during intensive pre-fieldwork training sessions. These two accounts point to the tensions between standardization and improvisation, and concerns about data quality that are at the core of this book and continue to preoccupy those who administer surveys in sub-Saharan Africa today. Amid demographers’ interest in measuring and quantifying population-based phenomena—such as HIV/AIDS and other health issues—surveys like the ones administered by LSAM’s fieldworkers are a major source of health-related evidence in sub-Saharan Africa. They act as localized sensors of a global system by feeding the demand for numbers on which to base evidence-based policy and practice (Cartwright and Hardie 2012; Adams 2013; Geissler 2015a, 15).

Cooked data are a specter that has long haunted survey projects by invoking ways in which data’s future certainty and value as evidence might be unraveled by human error or deviations from the standards or recipes governing their collection. Adjectives such as “cooked” versus “raw” and “dirty” versus “clean” figure across multiple scales of data talk in survey research worlds: fieldworkers, demographers, data entry clerks, policy makers, and statisticians alike employ such terms to comment on the quality of quantitative data at various stages of their collection, analysis, and storage. While we tend to think of data as abstract and intangible, these vivid descriptors draw attention to their materiality and life course. Numbers, of course, come from somewhere. A careful consideration of the social lives of numbers, rather than viewing them as stable and objective measures of reality, provides crucial context for interpreting quantitative evidence that we often deem too big or too technical to wrap our heads around. As an ethnography of the production of quantitative data, this book encourages its readers to be a little bit less in awe of numbers by understanding them as “creatures that threaten to become

corrupted, lost, or meaningless if not properly cared for" (Ribes and Jackson 2013, 147). It also considers how the activities of data collection not only produce numbers but shape personhood, sociality, and truth claims.

Cooked data conjure their culinary opposite: raw data. Data are units of information (such as a number, response, or code written into a box on a survey page by a data collector) that, in aggregate form such as LSAM's public-use database of survey data collected since 1998, might become evidence for policy making, public health interventions, academic analysis, or medical practice by government, nongovernmental organizations (NGOs), scholars, and other institutions in Malawi. Whereas actors in survey research worlds take raw data to be transparent or naked—that is, prior to analysis or interpretation—cooked data have been subjected to processes that shape or transform them in two main ways. In the first sense—the "cooking data" mentioned by the fieldworkers and Martin above—raw data become deformed, dirty, or useless through bad data practices and human error or other contingencies in the field. The most egregious—and mythologized—form of cooking data in the field occurs when a fieldworker fabricates numbers or fills out a survey willy-nilly.² In the second sense, cooked data are raw data that have been processed, organized, and analyzed according to demographic standards and norms; this form of cooking is codified and validated by experts and mostly takes place in the office once data arrive from the field.³ Talk of raw and cooked data recalls Lévi-Strauss's (1969) classic study *The Raw and the Cooked*. He argues that the interplay between the categories raw and cooked is the building block of hundreds of myths found across many cultures and therefore forms the basic structure of human thought. Raw and cooked are heuristics that allow humans to differentiate what comes from nature and what is produced in and by human culture, including data.

An extensive literature authored by statisticians and survey researchers has aimed to diagnose, document, and mitigate instances of cooking or data fabrication by data collectors, both during and after collection (Crespi 1946; Finn and Ranchhod 2013; Waller 2013; Kennickell 2015), with a more recent contribution suggesting that data fabrication by fieldworkers might function as critical commentary on inequalities inherent to research projects in low-income countries or as an expression of low morale (Kingori and Gerrets 2016). However, accounts of data practices in the field take for granted a fundamental difference between raw and cooked data, a binary that I hope this book destabilizes. In titling this book *Cooking Data*, my intention is not to suggest that the data produced by survey projects are fabricated or falsified, nor is it to provide advice to researchers about how to mitigate cooking among

fieldworkers. This book shows how all data—even that verified as clean by demographers—are cooked by the processes and practices of production.

I view survey research worlds as embedded in a heterogeneous social field inhabited by people whose practices, rhetoric, and relations are informed by epistemic conventions that underlie what the collection of good, clean data is supposed to be. I suggest that it is in the field where surveys are administered—rather than in researchers’ offices—that we can gain insight into what research means for the people who are tasked with collecting data by asking respondents questions and for those who have to answer the questions, as well as what kinds of worlds and persons it brings into being. In Malawi, this book shows, the effort to render the AIDS epidemic and its context visible and knowable to a demographic or global health gaze is constitutive of, and entangled with, attempts by fieldworkers and research subjects to achieve their own interests as members of a research world.

As an explicit expression and validation of underlying disciplinary norms or virtues, data talk and the units of information it comments on are not unlike Lévi-Strauss’s myths. Data and myths are both anonymized artifacts of collective labor and seem to “come from nowhere” (Lévi-Strauss 1969, 18); consider how the wide circulation of statistics as the collective currency of policy makers and statisticians reinforces a kind of mythology that takes information as objective, free-floating, abstract, and universal (Poovey 1998, xii; Bowker 2005, 73). Take, for example, the claim made by two demographers who analyzed LSAM’s survey data in an article published in a major HIV/AIDS research journal that “only 15.6 % of women and 8.1 % of men did not share their HIV test result with their spouse” (Anglewicz and Chintsanya 2011). This statement paints a particular picture of Malawian social life and garners legitimacy not only from the numbers it cites, but also from the respectable and long-standing data set from which the numbers are extracted. How did these numbers get all the way from the field in Malawi into the pages of a journal? What is their life story? This book demystifies data by tracing their life course and travels amid and with human and nonhuman actors whose heterogeneous work constitutes caring for data. *Cooking Data* foregrounds the social transactions that characterize survey research worlds all the way from the collection of raw data to the presentation of evidence in policy.

I borrow the phrase “cooking data” from my informants—both Malawian fieldworkers and survey researchers—to open an analytical space for the central questions of this book: How do raw units of information—numbers written onto a questionnaire by data collectors—acquire value as statistics that inform national AIDS policy and interventions? How do on-the-ground

dynamics and practices of survey research cultures mediate the production of numbers? Finally, how are quantitative health data and their social worlds coproduced and with what consequences for local economies, formulations of expertise, and lived experience? In attempting to answer these questions, I draw theoretical inspiration from science and technology studies and critical medical anthropology to illustrate how the lives of data and the lives of those who produce it in one of the poorest countries of the world are impossible to disentangle; data reflect and cohere new social relations, persons, practices, forms of expertise, and expectations. Following recent scholarship in postcolonial science studies, in this book I track how the survey project—a particular kind of socioscientific assemblage—travels; I also consider what matters to whom about research conducted in resource poor contexts. Finally, I show how survey projects, following a long legacy of scientific and development projects dating from the colonial period, are inevitably messier and less comprehensive endeavors than we might expect (Tilley 2007, 2). The blank first page of LSAM’s 2008 questionnaire that precedes this introduction invites future respondents’ answers; likewise, I invite the reader to join me as I track the travels of data in survey research worlds.

Demographers’ Dreams: The Assembly Line of Data

The chapters that follow explore the everyday relations between persons, data, technologies, and infrastructures that temporarily transform parts of Malawi into a field of demographic health research. Foreign survey researchers—demographers, economists, and sociologists such as those affiliated with LSAM—working in Malawi necessarily share responsibility for the quality of data collected with many collaborators, all with different interests in research: Malawian research partners, fieldworkers, HIV testing and counseling teams, data entry clerks, and research participants, for example. As is elaborated in chapter 4, raw information collected by workers in the field may be edited to remove assumptions and ambiguity as it is assembled, making data seem better or more certain than it actually is and enhancing its performative capacity and citability (Latour and Woolgar 1979; Bledsoe 2002, 130; Espeland and Stevens 2008, 421–422; Sana and Weinreb 2008, Tichenor 2017). In their polished form, data reflect the capacity and expertise of all of their handlers, even if epistemic rhetoric and metrics for good data tend to obscure the degree of uncertainty absorbed by data in their travels (March and Simon 1958, 165).

Survey research entails long periods of data collection in the field and confronts epistemic threats from start (survey design) to finish (good numbers

ensconced in a database): mistranslation of questionnaires, poorly trained interviewers, respondents who lie, respondents who refuse to participate or who cannot be found, poor weather conditions, inaccurate data entry, and lost data. Making quantitative data demands designing and implementing a material and human infrastructure—a machinery of knowledge production—that requires managing the unruly people, places, and things that characterize fieldwork, a messy outdoor scientific activity (Kuklick and Kohler 1996; Knorr-Cetina 1999; Ribes and Jackson 2013). These efforts are costly in time and money; in 2008, for example, data collection activities, including data entry, took 70 percent of the survey-based Marriage and Youth Project's (MAYP, discussed below) total project budget. Efforts at standardization and harmonization symbolized by the creation of a streamlined survey script to be administered by fieldworkers serve the goal of clean and high-quality data: data that are accurate, reliable, efficiently and ethically collected, and representative of sufficiently large and bounded samples over time. Indeed, survey researchers employ the term “quality assurance” to consider ways that data processes align or depart from pre-defined operational standards (Usten et al. 2005; Lyberg and Biemer 2008). As we will see in chapter 1, survey researchers endorse a shared set of epistemic virtues that ensure the data they collect will be deemed objective, clean, and consumable (Daston and Galison 2010).

Throughout, the book foregrounds data's materiality and social lives as they move along what demographers imagine to be an assembly line of human and nonhuman actors. Survey researchers themselves take interest in the many stages of a survey, typically bookended by establishing the structure of the study at its birth (usually in a proposal for funding) and ending with the dissemination of findings drawn from the data (Pennell, Levenstein, and Lee 2010). They determine how to best document the production of data at all stages to help data users assess data quality, defined as the degree to which data conforms to requirements agreed upon by producers and users. While demographers may idealize data activities as a kind of assembly-line process that produces identical widgets or units of information, this book shows that survey research activities and data production look more like a life course in practice—where any individual datum results from an unfolding series of transactions, experiences, and relations.

The assembly line—associated with Henry Ford's introduction of the continuously moving technology to mass produce standardized goods—is a compelling image for thinking through and tracing data's travels. First, the Fordist assembly-line process subordinated human skill or creativity by training workers at one station to do the same repetitive task over and over

again; the prefieldwork training sessions for Malawian fieldworkers discussed in chapters 2 and 4 likewise aim to harmonize the practices and procedures that constitute the data collection phase of research, characterized by administration of the same survey in the same manner to different respondents over and over again. Indeed, demographers and survey researchers in other disciplines generally view fieldworkers as a liability, harboring suspicions about the ability of the fieldworkers to do the work well and their potential to mess up data collection by cooking or fabricating data (True, Alexander, and Richman 2011). Fieldworkers across time and space are consistently framed by survey researchers as unreliable, as prone to cheating or cutting corners when collecting data, and as suspicious, thus requiring close surveillance to prevent unwanted edits to data in the field (Crespi 1946; AAPOR 2003; Biemer and Lyberg 2003; Sana and Weinreb 2008; Spagat 2010; Finn and Ranchhod 2013).

Yet, even as project design tools and survey instruments predetermine and limit the actions of fieldworkers, these individuals improvise, reinvent, and improve upon standards as they implement them in the field, far from the researchers' eyes and ears. As this book shows, making good data requires creativity and tinkering as much as it does harmonization and consistency. One major interest in writing this book is to present fieldworkers—often cast as unskilled laborers—as central actors in the story of the production of data. Intermediary local actors such as these have long been eclipsed in accounts of (post)colonial science that cast heroic scientists and Western experts as drivers of knowledge production, though anthropologists and historians have aimed to foreground the maneuverings, knowledge practices, and experiences of a wide variety of middle men and invisible technicians, including fieldworkers (Shapin 1989; Schumaker 2001; Raj 2007; Watkins and Swidler 2012; Bank and Bank 2013; Kingori 2013; Molyneux et al. 2013; Graboyes 2015; Maes 2015; du Plessis and Lorway 2016; Jacobs 2016).

Data collectors have long been portrayed as interchangeable with one another, and often do the grunt work or dirty work of survey research, including trudging from house to house in the field, collecting information, stool, urine, or blood samples, and so on. This book demonstrates that it is the creative and innovative tactics of fieldworkers that ensure that data collection proceeds smoothly, and their artful negotiation between top-down standards and bottom-up particularities—a kind of cooking data—that produces clean data as arbitrated by survey research standards. For this reason, three of the book's five empirical chapters center on the practices and interests of fieldworkers, taking the knowledge work they perform on a daily basis seriously as a form of expertise that emerges from their interactions with data and experience in

the field—the spatial anchor from which much global health knowledge today emerges. Not unlike its construction in anthropology, the demographic field is the practical basis of analytical discourse (Fabian [1983] 2002, 21).

Researching Research in Malawi

This book is an ethnography of survey research projects that were collecting household-level data in Malawi in 2005 and 2007–2008. Driven by demand for current and detailed demographic and socioeconomic data on households in developing countries, and on the characteristics of those who live in those households, the data collected by these projects are a key source of evidence for economic and social policy analysis, development planning, program management, and decision making. The household survey has become the predominant mechanism for collecting information on populations in such contexts. I spent time with four projects working across the southern and central portions of the country in five districts. The book draws principally on fieldwork conducted with two projects: **LSAM** and **MAYP**. Both were collecting survey data and **HIV** tests in multiple waves from samples of thousands of Malawians, most of whom live in rural areas. Data from **LSAM** have tracked demographic, socioeconomic, and health conditions in rural Malawi, and **MAYP** data track a sample of young adults as they transition to marriage. As the longest-standing cohort study in the sub-Saharan African context, **LSAM**'s data set begins in 1998, the first year it undertook field research in-country; since then, there have been six more survey waves, the last in 2012. From 2007–2009, **MAYP** collected data in three waves.

The book also incorporates ethnographic insights drawn from my field-work with two other projects: the Girls Schooling Intervention Project (**GSIP**) and the Religion and Malawi (**RAM**) project. A cash-incentives experimental study targeting girls of school-going age, **GSIP** also collected survey data and conducted **HIV** tests. The other, **RAM**, was a snowball-sampling-driven project collecting qualitative (interview and focus groups) and quantitative (questionnaire) data from religious leaders and church and mosque members in periurban southern Malawi that sheds light on the role of religious leaders in educating members of their churches and mosques about **HIV/AIDS**. While the bulk of ethnographic data in the book draws from the time I spent with **LSAM** and **MAYP**, some anecdotes and insights, as noted in the text, come from my time with **GSIP** and **RAM**. (See table I.1.)

In 2005, I first spent three months as a graduate student research assistant to **LSAM**, where my primary work was aiding with everyday research tasks

TABLE 1.1 Survey Project Information

| Project | Sample Size and Characteristics | Data Collected |
|---------|---|--|
| LSAM | 4,036 (2,361 women, 1,675 men) | Survey data, HIV tests, anthropometric data (height, weight, BMI) |
| MAYP | 1,185 (598 women, 587 men) | Survey data, interview data, HIV tests |
| GSIP | 3,810 young women | Survey data, interview data, HIV tests, health facility assessments, school and market surveys |
| RAM | 620 men and women (80 religious leaders, 508 members of religious groups, 32 people living with HIV/AIDS, and 24 focus groups with religious leaders and religious group members) | Survey data, interview data, focus group discussion data |

Source: Compiled by the author.

and overseeing a side project headed by a Malawian demographer that aimed to inventory cultural practices and their relationship to HIV risk across three districts in Malawi. I began to take interest in the culture and politics of survey research worlds and to formulate the research questions that animate this book. My relationships with LSAM principal investigators and Malawian researchers then led me to the other projects that agreed to host me in 2007–2008. American and European demographers headed LSAM, MAYP, and GSIP in collaboration with Malawian coprincipal investigators. All aimed to collect data that would shed light on social and economic trends over time relevant for understanding the trajectory of Malawi’s AIDS epidemic, one of the most severe in the world. The fourth study, RAM, was led by two American researchers with PhDs in nursing whose work and institutional affiliations were aligned with global health nursing and who sought to understand what kinds of information religious leaders disseminated to their congregations about HIV/AIDS. All four projects employed Malawian fieldwork supervisors, data collectors, and data entry clerks for the duration of their fieldwork periods.

I participated in all aspects of fieldwork including survey design meetings, the recruitment and training of project staff, everyday fieldwork practices such as checking questionnaires with data collection teams, evening social events, trips to the airport to collect shipments of HIV test kits or other

equipment, mapping exercises, data entry, and transcribing interviews. During data collection for each project, I lived alongside or with members of research teams. I spent the most consecutive time in the field with and around LSAM (three months in 2005 and then five months in 2008) and MAYP (three months in 2008). In addition to being a participant-observer during data collection, I also spent an extra month living in LSAM's and MAYP's sampling areas (Balaka and Salima, respectively) after the projects had departed in order to interview people living in recently surveyed households with my research assistant. While I initially planned to spend time only with LSAM, my broad interest in the politics of collaborative research and data collection led me to include the other projects in my research design so as to provide comparative context and to capitalize on the different tempos and data collection schedules of each project, all of which spent at least a few months engaged in fieldwork during the time span I was in Malawi. When I was not in the field with survey teams, I attended AIDS conferences and workshops where AIDS policy was discussed as well as interviewing a wide range of people involved in the world of AIDS research in Malawi, including research participants, chiefs and other traditional authorities, researchers, policy makers, government ministers, institutional review board (IRB) members, NGO staff, and district officials. Finally, I spent time in the Malawi National Archives reading documents, correspondence, and papers related to survey projects implemented in colonial Nyasaland. These censuses, surveys, and other enumerative efforts administered since the 1930s in Malawi provide useful historical context for my discussions of present-day surveys.

Throughout the book, I use the term "demographer" to refer to the core American, European, and Malawian researchers who were involved with LSAM, MAYP, and GSIP. Of those I interacted with most (thirteen), six held or were pursuing MAs or PhDs in economics, two in sociology, and five in demography. What unifies these researchers is their investment in the survey as a key tool in collecting data that will shed light on population dynamics, economic trends in rural Malawi, health issues, and the effects of the HIV epidemic on each of these. The questionnaire—in its imperative to collect standardized information that can be converted into numbers—is the base of these researchers' future analysis of a clean quantitative data set, to be followed by the dissemination of their results through journal articles, books, conferences, and other venues.

Demographers who were in academia at the time of this research were based at population studies or global health centers at the University of Malawi or American universities or, since few universities give degrees in de-

mography, in other social science departments, primarily economics and sociology (Riley and McCarthy 2003; Cordell 2010). Three of the demographers were based at the World Bank at the time of my research in 2007–2008. Chapter 1 elaborates on how demographers render the statistical household to communicate differences in populations across time and space, an agenda I suggest is at the core of the discipline and unifies the researchers and others who produce and utilize the data sets discussed in this book (see appendix 2 for a sample household roster page from LSAM’s 2008 survey that is representative of the same tool as implemented by MAYP and GSIP, as well). In the section that follows—and in chapter 1—I sketch an ideal-type demography that fails to capture the complexity and diversity of persons trained in this discipline, but nonetheless provides a heuristic sense of the general commitments of demographers for the reader; in this endeavor, I find Susan Watkins’s (1993) term “the culture of [demography]” and Saul Halfon’s (2006) term “population-based epistemic community” useful entry points. While culture(s) are unstable and dynamic, one can nonetheless extract patterns via ethnographic study of a discipline’s thought, practices, and products.⁴

Demography and HIV/AIDS in Southern Africa

By 1998, more than two-thirds of the people living with HIV resided in sub-Saharan Africa, and by 2002, HIV/AIDS had become the leading cause of death for both men and women aged fifteen to fifty-nine globally (Carael and Glynn 2008, vii). Once it was realized that there was an AIDS epidemic and that it was worst in southern Africa, where Malawi is located, international organizations flooded into the region to attempt to stem the tide of the epidemic. Researchers contributed to these efforts by producing and disseminating knowledge of the ways that HIV can be prevented, treated, and contained. Rural Malawians widely associate the term “AIDS” with the Chichewa term for research (*kafukufuku*, notably used also to mean survey), pointing to the history of efforts since the 1990s, usually by outsiders, to document and thus contain the HIV virus through the collection of information, anthropomorphic data, and bodily fluids.

Approximately 10 percent of Malawi’s population of 16.9 million is HIV positive, and it is ranked 173 of 188 countries on the Human Development Index (UNDP 2015). The mostly rural population engages in small-scale farming and depends heavily on rain-fed agriculture to grow maize to prepare the staple food dish, *nsima*. Subsistence agriculture is complemented by growing

small cash crops (mostly tobacco and cotton), casual agricultural labor, and selling vegetables and secondhand clothing.

The projects discussed in this book all take up HIV/AIDS as a central indicator in the data they collect. Zuberi et al. (2003, 472) suggest that the rise in AIDS mortality is the most important feature of African population since the early 1990s, particularly in southern and eastern Africa, making population-based surveys and HIV testing important tools through which to know and measure the significant impact of HIV on rural Malawians' lives (Garenne 2011). Although Malawi's "silent epidemic" probably began before 1980—the first case was diagnosed in 1985—a strict ban imposed by postindependence life president Dr. Kamuzu Banda on discussing (or researching) family planning (until 1982) or social problems that would challenge his discourse of Malawi as his land of milk and honey prevented the topic from becoming a point of public discussion until much later (Kerr and Mapanje 2002; GoM 2003; Lwanda 2005; Illife 2006). Pushed by the Global Program on AIDS in Geneva and by Western donors, Banda did establish a short-term plan to contain AIDS by mid-1987 and set up the National AIDS Control Programme in 1989, but its mandate and objectives were impeded by political stagnancy (Wangel 1995). It was only after democratization in 1994, when Banda lost the election and newly elected president Bakili Muluzi publicly prioritized AIDS, that international organizations began unimpeded and intensive work in this arena, eventually complemented by an enhanced governmental response led by the National AIDS Commission (NAC), established in 2001 as a condition for receiving World Bank funding for AIDS (Putzel 2004). The NAC has since overseen AIDS prevention and care initiatives and coordinated the country's AIDS response.⁵ Today, Malawi's AIDS budget continues to rely on international sources, with funds flowing from the World Bank, Global Fund, WHO/UNAIDS, and the President's Emergency Plan for AIDS Relief (PEPFAR), among many others.

The social sciences have played a central role in formulating policy and interventions into the AIDS epidemic in Malawi. Since the early 1990s, research has focused on assessing, among other things, AIDS-related beliefs, attitudes, and practices; determining the economic effects of HIV on the population; documenting support networks' care strategies for infected individuals; identifying a wide variety of ever-shifting risk groups (adolescent girls, truck drivers, sex workers, migrant laborers, and today's key populations, such as men who have sex with men); understanding low rates of condom use and/or family planning; and determining the feasibility and impact of HIV

prevention and treatment efforts, lately male circumcision and distribution of antiretroviral therapy (McAuliffe 1994; Bisika and Kakhongwe 1995; Chirwa 1997; Illife 2006; GoM 2015).

The HIV virus interacts maliciously with tuberculosis, malaria, and bacterial infections and has significantly affected social and economic life in Africa. The impact of AIDS on social institutions in southern and eastern Africa has triggered interest in infectious disease, as manifest in the Global Fund to Fight AIDS, Tuberculosis and Malaria and several other global health initiatives. Anthropologists have documented the burgeoning projectification of the African landscape, with exceptional focus on AIDS (Nguyen 2010; Crane 2013; Dionne, Gerland, and Watkins 2013; Meinert and Whyte 2014; Prince 2014; Benton 2015; Moyer 2015); a body of excellent work in critical global health studies has examined how resource-poor settings become central sites for the rise of global health science that unfolds in clinics, trials, laboratories, and hospitals, particularly amid what Watkins and Swidler (2012) term “the AIDS enterprise.”

This book builds on this scholarship but takes readers outside the wards, laboratories, and offices of global health and into the field that is the site of survey research. Understanding the population impact and dynamics of infectious disease is crucial to global health efforts to reduce morbidity and mortality and for decisions on where to best direct resources; data collected, cleaned, and analyzed by demographers plays a key role in untangling these variables and is vital to the measurement and practice of development in Africa. Indeed, many of the cooperative formations and partnerships between states, parastatals, and other organizations that fall broadly under the headings “development” and “global health” in Africa take as their main goal the achievement of indicators or targets that evaluate severity of health or economic conditions in a population over time, with AIDS as a central concern. Close scrutiny of the everyday socioscientific practices of survey research worlds can thus shed useful light on the politics of making numbers amid the rise of data-driven global health research in Africa.

An Ethnographer in Demographyland

I met with Richard Castells, a WHO epidemiologist, at Giraffe Lodge, a twenty-minute kabaza journey from LSAM’s field headquarters in Balaka District. With another American epidemiologist, he has been commissioned by NAC to develop a new AIDS prevention strategy in

collaboration with a local consultant. He is in Malawi for a short time to gather data from reports, interest groups, and interviews. . . . I noticed that Richard prefaced a lot of his sentences with “One thing I’ve noticed just from looking at the data . . .” I think this works to give him a kind of numerical authority that helps to obscure the fact that he has spent little to no time in Malawi, but lots of time amid numbers and statistics from “Malawi.”⁶

A close reading of this excerpt from my field notes illuminates the enduring chasm between anthropology and the more quantitative applied and practical sciences. Richard, by virtue of his disciplinary training as an epidemiologist, holds intellectual interests and commitments very different from my own. Even in the semiprivate genre of field notes, I perform a boundary between Richard and myself: I have been in Malawi for a long time, Richard for a short time. I make clear that I took a local form of transportation to the lodge (*kabaza*, bicycle taxi), and leave unmarked that Richard likely traveled there in an air-conditioned SUV from the capital. I view our meeting as potential data (e.g., “I noticed that . . .,” the act of recording field notes soon after the meeting), whereas Richard likely did not write up field notes after we parted. Richard, too, nods to our difference when he implies a contrast between his “*just . . . looking at the data*” and the kind of things I have been up to for over a year at this point in Malawi. Finally, my prose emphasizes the difference between an anthropological approach to Malawi (spending time in-country) and a demographic, expedient one (spending “lots of time amid numbers and statistics” that, in my view, will only ever capture a partial and scare-quoted “Malawi”).

In 2007–2008, as an ethnographer of survey research worlds in Malawi, I came to identify as an anthropologist among the demographers, playing on Bernard Cohn’s (1987) elaboration of the differences between the culture, forms of expertise, and even modes of dress of “Anthropologyland” and “Historyland.” Like Cohn, I recognized myself as a sympathetic outsider to practitioners and thinkers from a discipline whose goals were at odds with the tenets anthropologists hold dear. I did not become a demographer, even if I did learn better how to see and think like one. I “played the stranger” to the culture of demography by “adopting a calculated and informed suspension of [my] taken-for-granted perceptions” of demographic practice and its products (Shapin and Schaffer 1985, 6). My own distrust of numbers aligned unexpectedly with some (certainly not all) demographers’ explicit recognition that their data are fraught with limitations. The acknowledgment

of uncertainty built into demographic methods and epistemology works to grant numbers a provisional certainty within the discipline. This book is decidedly not an effort to reclaim the power of numbers—they have enough power already—but rather seeks to present a fine-grained answer to the deceptively simple question: What's in a number? This project resonates with Caroline Bledsoe's (2002) brilliant study of fertility practices in the Gambia, but whereas she seeks to understand the vital events in women's lives that numbers claim to represent (the “lives behind the numbers”), this book aims to trace the lives of numbers themselves, and the social worlds and persons they produce as they come into being. The book illustrates how producing numbers is a technoscientific endeavor that generates new kinds of knowledge, persons, and politics along the way.

So what kinds of things did I get up to as an anthropologist among the demographers? In the following vignette and ensuing discussion, I aim to demonstrate my own position in the larger infrastructure of survey research.

I sat in the LSAM minibus, red pen in hand, checking a survey that had recently been handed in by Ephraim, a fieldworker. Upon finishing, I called Ephraim over from where he was playing bao with an elderly man. He took a quick look at the red marks on his survey and headed back to the household for his callback. From the seat behind me, Esau, a supervisor who had been looking over my shoulder, said, “Crystal, you are not strict enough with them [the fieldworkers]. You need to reprimand them more strongly when they make mistakes . . . or they will just ‘cheat’ you.”⁷

Esau not only chastises me for being too easy on fieldworkers, but manifests a reversal of the anthropological gaze as he, one of my informants, “strains to read over [my] shoulder” a survey that will soon enter into a larger “ensemble of texts” destined to gain meaning as data for demographers (Geertz 1977, 452). His gaze embodies a question I was asked again and again, usually playfully or with a wink, by informants ranging from district health officials to researchers to survey respondents after I explained my research on research to them: “But who will research you?” For many months, I joined fieldworkers making numbers in the field. Yet even as I participated in the daily peregrinations of fieldwork—searching for sample households, checking surveys, filling in log books, and commiserating over bad weather—I was recognizably different from my fellows. Aside from my obvious status as a white person (*mzungu*), I was an anthropologist. My intentions were not purely to collect clean data for survey projects, but to study them along the way.

In the scene above, Esau solidifies a boundary between us. Though I am trained as an anthropologist, he sees me as a novice fieldworker who has not yet absorbed the skills necessary to making good data. He thinks I am more easily cheated than Malawian supervisors. As an honorary fieldworker, I have different investments and lower material stakes than he and other fieldworkers do in living from project to project (for them, a livelihood; for me, fieldwork funded by grants). Finally, in marking up a survey, I play a role in assembling data. I am complicit as I critique, in other words. My corrections to Ephraim's survey alter and affect the quality of the data that will eventually become evidence. Somewhere in Malawi, perhaps, the surveys covered in my pen marks many years ago still sit in a dusty storage room, material traces of data now transferred into databases.

It was my complicity in the larger infrastructure of survey research worlds that afforded me a deeper understanding of where and how quantitative health data come to be facts. Along the way, I learned, as well, that my critical gaze was shared by the people I was studying: some demographers, too, are well aware of the shortcomings of their numbers, but keep making them for the sake of policy, journal articles, and a faint sense that they might somehow improve the lives of rural Malawians. Like their informants who complained about the “too small” gift of soap they received after participating in a survey (see chapter 3), demographers recognized that soap is an inadequate gift for data, but kept giving it because it fit best into the ethical guidelines for human subjects research that govern their activities. Fieldworkers did not need me to tell them that their project-to-project lifestyle exploits them and articulated fine analyses of the structural effects of global health and the AIDS industry on their livelihoods and the well-being of the villagers they encountered. Rural research respondents made clear their critiques of extractive logics undergirding survey research, even if they did not dress them up in the jargon familiar to scholars, but talked about bloodsuckers instead. Policy makers told me they knew that policy was not as evidence-based as we might think and explicitly theorized the gap between themselves and researchers in their ivory towers. Tracing data’s life course from survey design meetings to downstream sites reveals a diversity of actors whose practices and rhetoric reflect their position relative to the other actors in survey research worlds and to the data they are meant to collect and protect.

I took up a temporary position as an honorary fieldworker on the demographers’ assembly line, likewise training my mind and body to absorb guidelines and standards for clean data. I wore a *chitenje*, proper field attire, when interacting with research subjects. I grew faster and more efficient at checking

surveys as time went on. I surveilled fieldworkers to ensure they were doing what they were supposed to. With supervisors, I kept meticulous log books that tracked the outcomes of research encounters. While *Cooking Data* stands in as the primary material artifact of my fieldwork, the numbers I helped produce are delinked from me and float anonymously in databases. The I-witnessing of the anthropologist, so evident in ethnographic representations, is contrasted with the collective and anonymized labor of survey research (Geertz 1988).

As anthropologists of global health and science, it is important to consider our own role in reproducing the logics, intentions, and data of the institutions we study, even as we position ourselves as critics of them. Discussions with my disciplinary fellows indicates that we check surveys, fill out bureaucratic forms, check pulses, file papers, lead trainings, create leaflets, author grant proposals for NGOs, and so on in the field. This, too, is the labor of fieldwork today, but often remains obscured by normative definitions of critique that still require us to present ourselves as somehow “float[ing] above” our subjects and seeing what they don’t, even if we long ago exchanged Malinowski’s “white canvas tent . . . on a beach” for clinical wards, minibuses, and air-conditioned NGO offices (Taussig 2009, 120–121). Critique seems to rely on preserving a kind of god’s-eye view whereby the objects of global health and other enumerative projects can only be seen from the outside (Haraway 1988), covering over how anthropologists make global health in the process of studying it, and continue to be as “doubly ambivalent,” perhaps, as our colonial predecessors—in quiet collaboration with power and institutions even as we critique them (James 1973, 42).

Bad Numbers: Anthropologies and Histories of (Postcolonial) Quantification

Despite efforts from both sides, anthropology and demography have largely maintained their distance. In this section, I hope to elaborate this divide without valorizing anthropology (my own discipline), instead emphasizing that what are considered good data—trustworthy, valuable, and usable—in each field can help us see why the two disciplines often do not see eye to eye. This divide parallels the broader critical position that anthropology adopts relative to disciplines and projects that rely on quantitative evidence. Surveys, censuses, and other enumerative projects are key sites of biopower where vital aspects of life are enlisted into political calculation, governance, and management (Foucault [1978] 2007, 333–361). Anthropologists have shown

how numbers—rather than stable or objective stand-ins for reality—are provisional and malleable entities that reflect their political and epistemological contexts (Andreas and Greenhill 2010; Lampland 2010; Erikson 2012; Hodzic 2013; Adams 2016a).

Demography, glossed as the quantitative study of human populations, with central interest in size, growth, density, migration, and vital statistics, is a positivist science rooted in the assumption that reality can be observed, measured, and counted accurately. Surveys such as those discussed in this book are at the core of the discipline's effort to successfully count, describe, and monitor people and events; as a methodological instrument, the survey claims to collect “identical data from . . . varied settings” that can be easily analyzed by statisticians who may never set foot in the geographic places—the field—where the data originated (Riley and McCarthy 2003, 55). Inevitable progress toward low fertility (which implies also progress toward modernity vis-à-vis normative interpretations of the demographic transition) is at the core of demographic thought, and a search for universal explanations for trends in population finds expression in the numerical data demographers collect and the methods they use to make knowledge (Bledsoe 2002, 19–56).

In this brief overview of demography's interests and pursuits—which are elaborated in chapter 1—we observe how far afield they seem from those of the anthropologist. Demographic approaches to human population, in general, stress the individual rational actor *Homo economicus*, neglect the historical and political context of demographic variables, and rely on quantitative data and methods that masquerade as objective and value neutral (Riley and McCarthy 2003, 40; Sreter, Sholkamy, and Dharmalingam 2004). Further, amid demographers' growing interest in enlisting quasi-anthropological methods into their work since the 1970s, anthropologists have been dissatisfied with their treatment and definitions of culture, viewing them as too simplistic, dated, or unreflexive (Greenhalgh 1990, 1995, 4, 13; Hammel 1990; Kertzer 1995; Kertzer and Fricke 1997; Coast 2003).⁸ Leading demographers of Africa Caldwell and Caldwell's (1987) important article on the cultural context of high fertility in sub-Saharan Africa—cited 803 times at this writing—identifies the need to place fertility in a broader context than surveys can capture, yet still falls into many of the above traps and describes culture as a “seamless whole” to boot (410). Demography has looked to anthropology as a quick fix in response to critiques of its “culture blindness” from outside the discipline. From the anthropologist's perspective, meanwhile, anthropology's totem—culture—has been made profane in the course of its travels to Demographyland.

The gulf between anthropology and demography is reflected, as well, in the different orientations that the respective disciplines have toward numbers, and particularly toward the history of numbers as tools of imperial and state power. Whereas anthropologists are number averse and harbor suspicions of quantification as a mode of knowing, demographers are happily awash in numbers and consider well-collected quantitative data to accurately represent reality. Indeed, a main point of controversy between anthropologists and demographers is how they might answer the question, What is the relationship between data and the social reality it claims to represent or count? Whereas demographers invest much time and money in revealing or discovering reality, anthropologists contend that classificatory exercises such as counting or surveying create reality or “make up” people (Hacking 1986; Greenhalgh 2004).⁹ Whereas the former seek to control the field even from afar, the latter remain open to the many surprises it holds; both approaches, it is important to note, carry with them different costs and benefits that underscore their investments in collecting a particular kind of good data. To oversimplify, demographers deem description and interpretation to be autonomous endeavors, while anthropologists have, since at least the 1980s, made much of their labor debunking that separation. Anthropologists, as we will see in detail in chapter 1, have thoroughly critiqued the categories, variables, and taxonomies at the heart of survey design for failing to acknowledge the diversity and dynamism of cultural contexts and definitions (Hirschman 1987; Bledsoe, Houle, and Sow 2007; Johnson-Hanks 2007; Loveman 2007; Bledsoe 2010).

Anthropologists and other scholars have shown that there has long been a link between those who measure or count population-based phenomena and those who seek to govern or control populations. Demography, in its focus on the very aspects of a population—birth, death, health, longevity, and so on—that Foucault places at the heart of governmentalized societies, is profoundly implicated in biopolitical projects (Foucault 1978 [2007]). Statistics are the “science of the state” (Foucault 1991, 96), a major tool through which the state sees and knows its citizens (Anderson 1991; Appadurai 1996; Scott 1998). The census—and its technologies, including the survey—shape the way states and other actors imagine their dominion, and its categories are key tools of power and empire in their ability to exoticize and classify citizens into moralized groupings and to affect the distribution of goods, allocations of social power, and services (Cohn 1987, 224–254; Kertzer and Arel 2002; Greenhalgh 2004; Cordell 2010; Mamdani 2012). As I show elsewhere, following James Scott (1998), demographers of Africa engage in a kind of “seeing like

a research project” (Biruk 2012), and such optics produce the kinds of indicators and numbers that are at the core of global governance regimes today (Davis, Kingsbury, and Merry 2012; Gerrets 2015a).

Here it is worth noting, however, that the modern state at the core of Foucault’s theorizations of biopower is an analytical category that might fail to capture the nuance of the colonial state’s imagining and management of its subjugated populations and, moreover, obscure the racism and racialization constitutive of biopolitical projects in Africa today. Megan Vaughan (1991) usefully suggests that even as colonial subjects were “unitized” by enumerative practices such as censuses or taxation efforts or, for example, weighing and measuring Nyasa migrant laborers, these processes were merely preliminary to the colonial state’s agenda of aggregation, producing a collective Otherness invested in the overriding difference of race. Colonial medical discourse denied the possibility that Africans were self-aware subjects, throwing a wrench into Foucault’s fulcrum of biopower: the subjectified “speaking subject” (Vaughan 1991, 8–13). Indeed, we might better consider how “racializing assemblages,” where sociopolitical processes—here, counting—that parsed populations into human (colonizer) and not-quite-human (colonized) were the pivot of colonial governance (Weheliye 2014); such taxonomies hinged on “cultural difference” and were the alibi of racialized violence enacted in the name of civilization or hygiene projects (Pierre 2013). Agnes Riedmann (1993) documents African demography’s role, in particular, as an agent of cultural imperialism. Global governance regimes, including human rights and global health, likewise stake claims to a form of suffering predicated on racialized bodies whose difference is often depoliticized by benevolent universalizing language. As will become especially clear in chapter 4, the legacies of racialized colonial imaginings of African others persist in some of the survey tools implemented today in the name of health and development.

King (2002) suggests that the conversion logics that undergirded colonial health projects invested in replacing traditional knowledge and practices with modern biomedical and scientific thought have shifted. He argues that the defining feature of postcolonial global health is integrating local places into global networks of information exchange, an endeavor undergirded by modern projects of total surveillance (782). Demography today has inherited its slot among the human sciences as a “policy-implicated discipline” (Szczerba, Sholkamy, and Dharmalingam 2004, 20). Demand for demography’s products remains high, even if the focus of research is often limited by the strings attached to funding flows to policy-relevant topics. For demography’s products to remain saleable, they must be quantitative, standardized, and

replicable (Demeny 1988; Riley and McCarthy 2003, 77); as we will see in chapter 5, data carries with it a brand that imbues it with quality and makes some data more in demand than others.

African Demography

Demography came of age during the mid-twentieth century, largely through its institutionalization in the United States. Amid rising concern about population growth, population became central to American definitions of development, and funding from both private and government organizations for population research increased (Riley and McCarthy 2003, 61–67). The 1960s and 1970s saw the founding of a number of population studies centers based at major American universities (Michigan, North Carolina, Brown, Johns Hopkins, Penn, and Columbia) funded by institutions such as the Hewlett Foundation, Mellon, National Institute of Health and Childhood Development, and the National Institutes of Health, amid the rise of the international family planning movement and Rockefeller and Ford Foundation funding that fostered the field of demography (Demeny 1988; Cleland and Watkins 2006).

Field sites in Africa are a major source of data for the long-term demographic projects based at such population studies centers; the University of Pennsylvania's center, for example, has “always been heavily weighted toward international population research . . . with a strong ameliorative component” (UPPSC 2017). Some suggest that the relatively secure funding available for demographic research has enabled demographers to avoid critically examining their premises; as Greenhalgh (1995, 10) contends, postmodernism did not enter demography as it did the other social sciences (Riedmann 1993, 96–110). In general, these critiques suggest that demography is a field weak or thin on theory and the most matter-of-fact discipline (Desrosières 1998). Its main investments are methodological: improving data collection and analysis processes to collect more and better data (McNicoll 1990).

The surveys discussed herein, as legacies of technoscientific projects in the service of colonial interests, raise the specter of the exploitation, extractive logics, racism, and ethnocentrism that have underlain science in Africa, and global demography's presumed “right to invade” in the name of knowledge production (Riedmann 1993). As can be seen in chapter 3, impoverished survey participants in 2007–2008 drew on extensive past experience with research projects to evaluate whether or not to participate in a survey headed by researchers from wealthy countries that might bring them no returns; subjects

were highly research conscious and expressed their suspicions or wariness of the means and ends of projects by employing resistive tactics that threatened to influence data quality. Residents across sub-Saharan Africa have by now become accustomed to projects in their midst. Diverse actors were interested in counting and enumerating Africa's population(s) even before the first official or modern census efforts. Owusu (1968) notes that precolonial head counts carried out by chiefs saw the heads of families drop articles such as grains of cereal, beads, or cowrie shells that stood in for the number of a chief's dependents, for example.

Early colonial counting practices largely entailed unscientific walking tours by district officials, estimating local populations with the help of word-of-mouth information from local people, or via simple head counts. These ad hoc techniques were likely adopted by the earliest census takers in Nyasaland in 1901 (Deane 1953, 143; Zuberi and Bangha 2006; Gervais and Mandé 2010). The first systematic attempt to describe the population dynamics of sub-Saharan Africa was Kuczynski's (1949) *Demographic Survey of the British Colonial Empire*, meant to be useful evidence to help in implementation of the Colonial Development and Welfare Act (1940), which provided for large investment in development, agricultural, and health research (Havinden and Meredith 1993). In colonial Malawi, the late 1930s saw the implementation of an ambitious nutrition survey project whose commitments and implications are elaborated in the course of this book and which was symptomatic of a mid-1930s rising colonial interest in coordinating and funding health and agricultural research initiatives in Nyasaland and the Rhodesias (CAA 1935; 1936). As Tilley (2011) documents, from the mid-1930s, the ambitious African Survey led by Lord Hailey shaped research priorities in Britain and colonial Africa, solidifying its role as a living laboratory increasingly dotted by scientific field stations. The migration of the loose discipline of population studies to Africa was somewhat coterminous with the rise of international health as a field of practice and the rise of development as a central concern (Packard 2016, 181–186).

Scholars and policy makers have paid close attention to population in Africa and the global South since World War II; the first world population conference that drew institutes, researchers, and implementers from around the globe was in 1954, and the first round of the African Census Program was initiated in the mid-1960s (Ghana held the first modern census on the continent in 1960; Malawi's first census was in 1966). Access to populations increased by the 1970s and 1980s via censuses; knowledge, attitudes, and practices surveys of fertility in the 1960s; the World Fertility Survey; and the Demographic and Health Surveys (DHS) program (Tarver 1996, 7–8).¹⁰ In 1984, the Union for

African Population Studies—whose 2007 conference in Arusha, Tanzania, figures in chapter 5—was founded through a UN initiative to promote the scientific study of population and application of research evidence in Africa. Headquartered in Accra, Ghana, the association has convened a general conference on African population every four years since 1988 in an African country (UAPS 2017).

In the mid-1990s, the institutionalization of population studies and demography on the continent continued with the establishment of the African Population and Health Research Center (APHRC) in Nairobi, Kenya, and the Africa Centre for Health and Population Studies in South Africa, both of which play a key role in collecting field-based survey data in Africa and as collaborators with foreign researchers engaged in data collection. Likewise, increasing opportunities for training of African demographers and statisticians—some of which have been included in proposals as capacity-building activities by the projects discussed in this book—has made a dent in the huge volume of statistics and publications produced exclusively by researchers from other countries (Oucho and Ayiemba 1995, 73).¹¹ Nonetheless, as will become evident in chapters 1 and 5, disparities in access to data, graduate training, and statistical software and asymmetries in the material conditions of foreign and Malawian researchers poke holes in global health's dominant rhetoric of "partnership" and "collaboration" (Crane 2010b).

Rethinking Poor Numbers

The imperative to collect high-quality, clean data (terms whose precise meanings are elaborated in chapter 1) is at the core of survey research and underlies demographers' dreams of data production on a well-oiled assembly line. The harmonizing efforts of survey projects aim to combat the problems of data quality that have long plagued similar endeavors in colonial and post-colonial African contexts. Talk about data from and within Africa since the colonial period has trafficked in metaphors of scarcity, lack, and poor quality (Hill 1990). In the classic volume *The Demography of Tropical Africa*, Lorimer (1968, 3) calls for a shift from cruder sources of demographic information (such as tax registration) to more systematic efforts such as surveys or censuses, and van de Walle (1968, 13, 59) observes that the inability of Africans to know their exact ages or to identify dates without being accustomed to calendars leads to poor data quality.

Many reflections on data in Africa implicitly place responsibility for poor data quality on the figure of the African enumerator, not unlike Martin's

(1949) comments on the 1948 census. This trend dates from the colonial period. Lord Hailey, reflecting on the immense need for population statistics in the pages of his *African Survey* (published in 1938), suggested, “There is still much to be learnt of the technique of sampling in African conditions, and it must, moreover, be recognized that whatever the advance made in technique, there will remain the problem of securing enumerators who can elicit the information required” (Hailey 1957, 139). Phyllis Deane (1953, 10), in her analysis of data collected in the late 1930s on economic transactions in Northern Rhodesia (present-day Zambia) and Nyasaland, suggests that “deficiencies in data” were attributable to the lack of trained African research assistants. In the present, meanwhile, the costliness in time and money of the intensive prefeldwork training sessions for fieldworkers illustrates researchers’ enduring perception that fieldworkers are likely to mess up their data.

Researchers have likewise long associated Africa itself with bad population data. Ocho and Ayiemba (1995, 44) suggest that prior to the 1970s, “the African continent was a desert in terms of availability of accurate and reliable demographic data.” In an annual review article, Zuberi et al. (2003) note that understandings of Africa’s demography up to the present day are based on the unsystematic analysis of data from different sources and periods. The head of the Statistics and Survey Unit at the APHRC suggests that it aims to fill a “data gap” in Africa, where the paucity of “accurate, reliable, and timely data” has constrained effective monitoring of development programs and interventions on the continent (Beguy 2016). Responding to this discourse of data scarcity and problems, economic historian Morten Jerven’s (2013, 32) important book-length analysis of the poor quality of statistics pertaining to economic development in Africa is an effort to “gauge the size of errors and evaluate the direction of bias in [statistical] evidence,” which are often obscured by data users’ blind faith in the experts who produce or interpret numbers.

Amid a sea of poor numbers, however, it should be noted that the data collected by the survey projects described here yield—by demographic criteria—better numbers than, for example, censuses or DHS surveys, because they provide localized surveillance in a smaller area over time. Unlike the census, which aims to provide a full, comprehensive count of a nation’s population for the government, or DHS surveys, which yield nationally representative data, the surveys discussed here collect responses from individual agents in a sample—a portion of the total population drawn from the same enumeration units employed by the census and DHS. Data from the surveys in this book complement census data by administering comprehensive and directed questions to a random sample.

Despite the generally critical or antagonistic relationship of anthropology to demography, the book does not endeavor to prove that demographic data fail to represent rural Malawian realities or to expose their uncertainty, but rather takes a more sympathetic tack toward quantifying projects (Colvin 2015). Rather than dismissing numbers as simply false, socially constructed, or inaccurate, the book aims to critically examine the criteria and metrics that help numbers attain their legitimacy and authority by presenting a fine-grained account of data's life course and handling by many diverse actors. Others have sufficiently critiqued the quality of Africa's poor numbers, showing how and why quantifying projects often get things wrong or miss out on what is really going on. Building on this work, I analyze in depth the social lives and cultural work that numerical data do, even before they appear as statistics. Numerical data's provisional and uncertain status, I show, is often well known to those who make it. Following Lampland (2010, 2): "Provisional and false numbers can only function if there is some sort of agreement about their status as temporary or fragile symbols" and "false numbers appear when the primary task is to learn how to deploy numbers, making the relative accuracy of the numerical sign less important than the attempt to master the logic of formal procedures." As Erikson (2012, 373) points out, even if numbers are "hollow" they enable other forms of value to be produced.

This book does not aim to determine how accurate estimates of HIV prevalence or other statistical phenomena are, but carefully considers how demographers tell (themselves and others) convincing stories about AIDS and other social, economic, and health issues in Malawi through numbers (Setel 2000, 10). These stories reveal some things clearly and hide others, not unlike the stories anthropologists tell about their field sites (Wendland 2016, 60). We should remember that, even as numbers and surveillance are at the heart of colonial and present-day governance projects of racialized bodies, they nonetheless can do important work in turning uneventful suffering into aggregate suffering and making it visible (Povinelli 2011, 14; Livingston 2012; Stevenson 2014, 186). Furthermore, being counted in an impoverished context such as Malawi might entail forms of incorporation, recognition, and support that would be otherwise unavailable (Ferguson 2015, 85).

In this sense, this book critically examines the criteria and metrics that underscore data's production and consumption. These standardizing criteria, rather than being stable, are invented, embodied, and negotiated in the everyday practices of research worlds. Like other recent work by scholars engaged in critical global health studies, this book considers how large-scale outsider-led projects in Africa are situated in and rely on local regimes of economic,

cultural, and social capital. However, the emphasis is on showing how a particular set of epistemic criteria creates the human and social scaffolding for its implementation and to what ends. Importantly, it challenges the abstract universality of data unanchored from its site and relations of production by showing how Malawi and Malawians shaped it.

Scholars have shown how numbers, categories, and statistics are taken up, critiqued, or negotiated by those they claim to represent; this book contends that understanding how those who make numerical data handle and engage with it can shed new light on the politics, stakes, and unintended consequences of quantification in sub-Saharan Africa. While the book is an account of enumeration practices in academic-demographic research, my analysis of how these practices operate in the field should resonate with those involved in implementing operations research and monitoring evaluation projects—which often face more time constraints and are less well funded than the projects in this book—as well. This book reflects the potential of anthropology’s commitment to “slow research” amid the value placed on speed, efficiency, standards, and comparability in global health, development, NGO worlds, and population science (Adams, Burke, and Whitmarsh 2014), but also prompts anthropologists to reflect on how our own data activities likewise cook data, with important implications for the claims it is possible to make (my own attempt to do this appears in the conclusion). A granular analysis of research worlds in a particular place at a particular time, the book suggests, encourages us to more critically engage with the kinds of evidence we too often take for granted, whether inside or outside our discipline or training.

Assembling Data: A Road Map

In chapter 1, I introduce the work that must be done before survey research projects enter the field where data will be collected. The chapter interprets survey design, the first step in assembling data, as an exercise that attempts to amalgamate the idealized categories of insider (local) and outsider (foreign) expertise. In analyzing debates between Malawian and foreign collaborators around cultural and linguistic translation and fine-tuning of survey concepts and questions, around plans for where surveys should be administered, and around what should be the objectives of research, the chapter draws attention to the different material and academic investments of foreign and Malawian researchers in data collection, which are often obscured by partnership rhetoric. Chapter 1 illustrates how demographers’ dreams of an assembly line for data take shape in the office, before data collection begins in the field.

The middle three chapters of the book center their attention on data collectors, whose practices in the field determine the quality of data. Chapter 2 introduces the Malawian secondary school or college graduates employed as data collectors by research projects, unskilled middlemen who have been overlooked in accounts of science in Africa, despite their central roles in producing and handling data. Since the earliest surveys were enacted in sub-Saharan Africa, these individuals have been portrayed as menial laborers, as interchangeable cogs in the machinery, and as liabilities to the collection of good data. Chapter 2 challenges such assumptions by showing how data collectors, through serial research project jobs, acquire particular forms of expertise that ensure projects run smoothly. I describe fieldworkers' interests in maintaining ownership over the local knowledge foreign researchers expect them to possess. I also examine prefieldwork training sessions as an important site where fieldworkers are initiated into new professional identities and where social and spatiotemporal boundaries that undergird data collection are performed. Throughout, the chapter takes interest in how fieldworkers come to live from project to project, enabling them some measure of access to social, cultural, and economic capital, and producing new forms of value and expertise.

Chapter 3 examines the transactions that undergird the administration of household-level surveys. Centering the encounters between fieldworkers and their rural Malawian research subjects, it explicitly considers the value of data for different actors in research worlds. In line with international human subjects research ethics that privilege informed consent and prohibit provision of inducements that might endanger it, research participants were given bars of soap as a gift in exchange for information they provide to research teams. I interpret this standardized gift as a central site where people negotiate political, ethical, and moral questions that arise in research worlds. This standard research gift facilitates the recognition that bits of information are tangible items with a negotiable value and highlights the role of small-scale transactions in stabilizing—and potentially unraveling—data as they move through their life course. Chapter 4 argues that producing high-quality data necessitates standardization of habits, scripts, and social interactions across thousands of research encounters in the field. I employ ethnographic analysis to show how demographers' epistemic investment in clean data that is accurate, reliable, and timely not only guides the movements and agendas of survey research teams in the field but also produces categories, identities, and practices that reinforce and challenge these standardizing values.

Chapter 5, the final empirical chapter, is an ethnographic study of downstream sites where data in their clean and finished forms are performed to

and consumed by audiences. It is concerned with how the kind of data represented as raw (survey responses, HIV tests), discussed in previous chapters, is validated as evidence in the policy-research arena. Drawing on participant observation at a number of Malawi-based, regional, and international AIDS research conferences where quantitative health data were presented, at policy-making sessions and meetings, and on interviews with survey researchers working in multiple African contexts, I show how knowledge is made and evaluated in contingent social performances that employ scripts, props, leading actors, special effects, and supporting actors. I interpret these sites as contingent end points in data's life story, and show how even data in their final finished form as evidence are further cooked in their re-presentations and in social relations. The chapter also critically analyzes the discourse of the policy-research gap—conceived of as a chasm of blocked communication or knowledge sharing between researchers and policy makers. I show how this gap is better analyzed as a confluence of multiple interests that determine the kinds of evidence that gain authority in the policy-research nexus, and the efficacy of its translation between the two spheres.

The book's conclusion is a meditation on the meanings, intentions, and assumptions embedded in the anthropological project to critique global health and other research institutions in Africa. I present vignettes from my field notes that did not make it into the empirical chapters of the book to turn the lens on the anthropologist among the demographers. I take up long-standing concerns of anthropologists—complicity, the field, and the compulsion to make our work useful—from the perspective of a contemporary ethnographer of global health. The conclusion takes full circle the main interest of the book by showing how data—whether demographic or ethnographic, quantitative or qualitative—reflect and cohere the social worlds they claim to represent.

This page intentionally left blank