



California Center for Population Research
University of California - Los Angeles

**The Mekong Island Population
Laboratory (MIPopLab), A
Demographic Surveillance System
in Rural Cambodia (2000-06)**

Patrick Heuveline, Brooks Ambrose, Sothy Eng, Gail Yen,
Bunnak Poch, and Savet Hong

PWP-CCPR-2015-012

January 6, 2016

***California Center for Population Research
On-Line Working Paper Series***

**The Mekong Island Population Laboratory (MIPopLab),
A Demographic Surveillance System in Rural Cambodia
(2000-06)**

**Patrick Heuveline,ⁱ Brooks Ambrose, Sothy Eng, Gail Yen,
Bunnak Poch, and Savet Hong**

**California Center for Population Research
University of California, Los Angeles (UCLA)**

Abstract

From 2000 to 2006, the Mekong Island Population Laboratory (MIPopLab) collected basic demographic information on a population of roughly 10,000 individuals in Central Cambodia. This paper accompanies the release of these 13 waves of demographic-surveillance data. The data files are being released with English translations of all questionnaires (administered in Khmer, the official language of Cambodia) and codebooks that also include simple frequency tables. In this paper, we provide additional background information on the data collection operations. We also present the different types of data files, together with some basic tabulations derived from these, such as the population by gender and age-groups from the benchmark censuses and the number of births, deaths, and migrants by gender and age-groups from the subsequent demographic updates. Whenever possible, we compare these distributions with those from nationally representative data and provide references to further analyses based at least in part on MIPopLab data.

Project Background and Rationale

The Mekong Island Population Laboratory (MIPopLab) conducted 13 rounds of data collection between 2000 and 2006, at which point MIPopLab was integrated into the Mekong Integrated Population-Registration Areas of Cambodia (MIPRAoC). At this writing, MIPRAoC has conducted 4 rounds of data collection, with ongoing planning for a 5th round to be fielded in early 2016.

MIPopLab emerged from two need-assessment missions to Cambodia undertaken by the first author in 1997 and 1999. The first one followed a visit to a demographic surveillance system (DSS) in the Tonga population of Southwestern Zambia (Clark et al. 1995), a project member of the International Network of field sites with continuous Demographic Evaluation of Population and Their Health (INDEPTH). Between these two missions, the commitment of international agencies and the weakening of the military opposition carried out by the Khmer Rouge (KR) in all but a few local strongholds finally made possible the conduct of the 1998 General Population Census (National Institute of Statistics 1999, GPC 1998 thereafter), providing the first near-national-scale demographic data for the country since 1962. Demographic levels remained uncertain, however. An administrative structure for recording basic demographic events had been set up by the first post-Khmer-Rouge government, the People's Republic of Kampuchea (PRK, 1979-89), but the records were not, or no longer, consistently collected and consolidated.

The concrete planning and design of a DSS that would address the uncertainty surrounding demographic trends in Cambodia began in collaboration with Lany Trinh and a team of Cambodian researchers at the Royal University of Phnom Penh (RUPP).¹ From 1995 to 2000, Lany Trinh directed the RUPP training program in basic and applied demography supported by the United Population Fund (UNFPA). Involving the program alumni in field work was of mutual benefit, strengthening their academic training with a practical research experience and using the experience results to seek future extramural funding, without which most of the alumni would likely be unable to engage in research activities related to their

demographic training. To improve the prospects for extramural funding, the DSS in MIPopLab was envisioned as the project's foundation on which other, topical, so-called "rider" surveys could be built to address more basic-science issues.

While conducting research in post-Khmer-Rouge Cambodia presented many logistical challenges, the unusual features of its recent history and current demography also provided unique research opportunities. The initial rider survey was designed to assess the role of the mortality crisis in subsequent fertility changes in Cambodia. The study of historical mortality crises has provided a different perspective (e.g. Lindstrom and Berhanu 1999; Lee 1997; Palloni, Hill and Aguirre 1996; Galloway 1988; Watkins and Menken 1985; Eversley 1957) on the larger issue of the contribution of mortality declines to the late 20th -century fertility transitions (Hirschman 1994; Mason 1997; Cleland 2001). But historical time series cannot render the perceptions and decision framework of past populations. The recent experience of Cambodia presented an opportunity to converse with the survivors of one of the most intense mortality crises in modern times. MIPopLab was thus set up to provide both retrospective, quantitative and qualitative data on reproductive behavior and prospective follow-up data on population dynamics. This data release only concerns the quantitative data, which is described in further details below, but individual interviews and focus-group discussions with males and females at different stages of their reproductive lives were also conducted. Additional information on these data can be provided upon request.

Lany Trinh first identified and suggested to use as the catchment area for the future DSS an island situated on the Mekong River, hence the project's name, the Mekong Island Population Laboratory. The island constitutes a single administrative unit referred to as a "Commune." Administratively, the Cambodian territory is divided into Provinces and the Municipality of Phnom Penh, the capital city (Heuveline forthcoming). Provinces are divided into rural Districts and Cities, with the rural Districts then divided into Communes. There are 1,600 Communes in Cambodia. With roughly 10,000 people at the time of the GPC 1998, this

particular Commune's population was larger than average, representing just about one thousandth of the national population size. The Commune was then part of a rural District located in the Kandal Province, whose territory forms a ring around the Municipality of Phnom Penh.

Several features of this particular Commune were attractive for setting up the project. First, a certain proximity to the capital city, Phnom Penh, was a necessity at the time as the volatile political situation still prevented travel to some parts of the country and because human resources were highly concentrated in the capital city. Given road conditions at the time, households at the southern tip of the island could be reached from the city in about an hour by ferry and motorcycle, while those at the northern tip of the island required an additional 45 minutes by motorcycle. The reasonably short distance to Phnom Penh allowed for the collaboration with the staff of the demography program at the RUPP, the only such program in the country. For the island residents, however, the physical proximity of Phnom Penh was balanced by the difficulty of transportation and associated costs that reduced the frequency of back and forth population movements. For the owner of a motorcycle, the roundtrip costed about \$1; otherwise it added up to \$3 with the "moto-taxi" fares. These were not trivial costs for Cambodia, where the Gross National Product per capita was under one dollar a day at the time (Prescott and Pradhan 1997). The island residents were thus close enough to travel back and forth to Phnom Penh on any given day, but unlikely to do so on a daily (commuting) basis. Second, the insularity of the Commune provided the advantage that the boundaries of the site and the resident population could be defined unambiguously and durably, even in the event of administrative redistricting. Indeed, to account for the continuing expansion of the capital city's urban agglomeration, the Municipality of Phnom Penh has since gained several Districts previously classified as rural, including the one in which the Commune is located, but the administrative boundaries of this particular Commune have so far remained intact. Finally, even though we did not have the pretense for MIPopLab to yield nationally representative data, the

GPC 1998 exhibited strong urban-rural gradients in nearly all socio-demographic indicators, suggesting that the characteristics of a rural community relatively close to the capital city might be fairly close to national averages. The expectation that socio-demographic indicators in MIPopLab are not too far off national averages is further assessed below through comparisons with data from the GPC 1998 and the first Cambodian Demographic and Health Survey conducted in 2000 (National Institute of Statistics 2001, CDHS 2000 thereafter).

Even though Communes are themselves divided into Villages, the Commune is the lowest administrative level with any administrative authority. Under the PRK, even Villages were further divided into Solidarity Groups. These Groups were introduced to organize production and monitor population movements as the government was still fighting the remnants of the Khmer-Rouge army (Frings 1994; Ledgerwood and Vijghen 2002). The prerogatives of these Groups and of their leaders gradually faded out after the land privatization and demilitarization. While the Commune chiefs used to be full-time political appointees and have been elected officials since 2002, Village chiefs and Group leaders were selected mostly for their literacy in order to perform administrative duties. For Group leaders, these duties nominally included reporting vital events to their Village chiefs. Receiving only minor compensation for their administrative work, however, Village and Group officials had to engage in other professional activities and administrative tasks were performed sporadically at best.

At the time of our first visit, the vital registers in the Commune had not been updated for three years. However, many adult rural residents were still aware of the existence of Villages and Groups and generally knew who their Village chief and Group leader were. An average-size Group consisted of about a dozen adjacent households. In rural areas, the common practice of establishing a new household next to kin implied that individual members of the households in a given Group were often related to some degree. In the absence of reliable maps to canvas the territory of the commune, our data collection strategy relies in part on this dormant administrative structure below the Commune level. First, the Commune's five Villages were

mapped into Groups with the assistance of the different Village chiefs. Second, on days of data collection, our field workers were paired with one of the Group leaders, and interviewed each household head in his or her group at the respondent's home.

MIPRAoC Data Collection Timeline

DSS registration started in 2000 with a baseline census of a single Village. At the time of registration, each household head provided for each household member their name (later replaced by a unique identifier), gender, birth date, relationship to head, and parental information (is the mother/father alive, and if so, where does s/he lives, else when did s/he die). The amount of data collected during the baseline census was thus kept to a minimum in order not to compromise the objective of setting up a DSS for the long-term. For the scientific aims of the initial rider survey, however, we also recorded complete marriage and birth histories from all eligible women, i.e., women between the ages of 15 and 74. To increase the quality of these retrospective reports, we used a version of the calendar method of data collection (Goldman, Moreno and Westoff 1989), with references to the salient dates of the recent Cambodian history (e.g., Khmer Rouge regime, Vietnamese withdrawal, U.N. presence). To increase the quality of age reporting, field workers were provided with a calendar indicating the Christian year number, the Buddhist year number, and the Chinese zodiacal signs, the latter being typically better remembered than the year of birth. This was intended to avoid estimating the year of birth from the age reported by the respondent. Age is commonly reported not as the number of years between birth and the last birthday, but by starting at one at the time of birth and adding one year on each Cambodian New Year, which is celebrated in mid-April. Subtracting the age reported by the respondent from the year of the survey may result in a bias, most visible in the under-reporting of children under age one in all the Cambodian censuses to date.

DSS registration was extended to a 2nd Village in 2001, and with the award of a *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD) grant to

the first author in March 2002, registration was completed in the remaining three Villages of the Commune in July 2002. After their respective benchmark censuses, the list of households in each Village and each household's composition was then updated biannually until 2006. The updates consisted of birth, death and migration registration, and for women aged 15 to 49, pregnancy and marital status updates. The total number of rounds varies by Village, depending on the date of the benchmark census. In the first Village, the benchmark census took place in 2000 and the first update in April 2001. There are thus a total of 13 rounds for this Village (the census plus two updates a year from 2001 to 2006). For the last three Villages having their census in mid-2001 and their first update at the end of 2001, there are only 10 rounds (the census and nine updates).

The raw data for each of these rounds is being released although some early rounds before the NICHD award were administered to pilot the forms. The first update in the first Village in April 2001, in particular, is only a partial update and only a fraction of the Village households could be visited on that date. Specific item questions and responses on the registration form for each of the events have continued to evolve over time in an effort to incorporate feedback from the field and to streamline the DSS. The raw data files and their documentation (questionnaire et codebook) are provided by date, since the same questionnaire was used across Villages at a given date even though it might correspond to, say, the 2nd update after the census in the last three Villages (round 3) but to the 5th update in the first Village (round 6). While the chronological presentation is more appropriate for a precise description of the files, to avoid repetitions, the summary description below glosses over minor round-to-round variation and present each type of file instead.

Datafiles

Census files (round 1)

Census files contain information about the household interviews, the residents in each household, and the marriage and birth histories of eligible women in each household. The household interview and resident data are each provided in a separate file. The marriage and birth data are presented in two or three additional file (see below).

Interview

The *Interview* file lists each of the households visited, with the visit date, the individual identification code of the respondent providing the household information, and a few summary variables that can also be recovered from the *Residents* file. We expect that the only analytical use of this file is to be merged with other files in order to append the date of the interview (for instance, to calculate the age of the respondent at the time of the interview from his or her birth date).

Residents

The *Residents* file lists each of the Village residents at the time of the benchmark census, by household and with his or her relationship to the head. Also provided are the respondent's birth date (month and year), gender, and parental information (place of residence or, if deceased, date of death).

A total of 2,146 residents were so registered in the 1st Village in 2000, with another 1,938 residents being registered a year later in the 2nd Village and 7,031 residents being registered six months later in the remaining three Villages of the selected Commune. As shown in Table 1, the *Residents* files provide a breakdown of these 11,115 residents by age and gender which can be compared to the GPC 1998 data for the country and the two "closest" areas, the Province of Kandal to which it was allocated at the time and Municipality of Phnom Penh to which it has been reallocated since. Demont (2011) has conducted a thorough investigation of the quality of age reporting in MIPopLab. Based on the low value of the modified Whipple index (Shryock and

Siegel 1973), the author reports that there does not seem to be a strong preference for any particular age. Overall, she concludes that the data quality is relatively high compared to what is typical of South Asia and similar to what has been documented elsewhere in East and Southeast Asia.

The GPC 1998 data displays three main features that we should be able to observe in MIPopLab data as well. First, the population pyramid exhibits the demographic scars of the 1975-79 Khmer-Rouge period and the following baby-boom (Huguet et al. 2000). The high mortality of the late 1970s particularly affected adult males (Banister and Johnson 1993; Heuveline 1998 & 2015; Neupert and Prum 2005; DeWalque 2005; Lognard 2012). For the country, the male to female ratio is 71 males per 100 females among individuals age 40 and over, that is, among those who were 17 years and older at the outset of the Khmer Rouge regime. The ratio is nearly the same in the Province of Kandal and, likely due to migration, higher in the Municipality of Phnom Penh (76 males per 100 females). For the same age group, the ratio across MIPopLab Villages averages to 69 males per 100 females, thus showing the same gender imbalance in these generations.

The 1975-79 Khmer-Rouge regime also affected fertility levels which reached a low in 1976-78 (the Khmer Rouge took over in April 1975 so the regime change had little impact on the number of births in 1975) before rebounding markedly in the 1980s (Heuveline and Poch 2007). The birth cohorts of the late 1970s are thus smaller than in adjacent years, which is clearly visible in the deficit of people aged 20 to 24 in the GPC 1998 age distribution. The phenomenon is slightly less visible in the 5-year age distributions of the 2000, 2001 and 2002 residents as these birth cohorts are now partially included in two adjacent age groups, 20 to 24 and 25 to 29 years. However, year of birth data shows the expected patterns. For years 1976, 1977 and 1978, the number of births is 18% lower than for 1973-75 and 113% lower than for 1979-81.

Finally, the baby-boom of the 1980s was followed by a fertility decline that has continued to the present. This is visible in a smaller number of individuals under age 5 than between ages 5

and 9 in the GPC 1998 data for the country and for the Province of Kandal. In the Municipality of Phnom Penh, the numbers between ages 5 and 9 are also lower than the numbers between 10 and 14. Among the MIPopLab residents, we find that the larger five-year age-group is 10 to 14 year-olds. This is consistent with the fact that they were registered a few years after the GPC 1998 and a few more years after the onset of the fertility decline in the Commune. In spite of the fertility decline that contributed to smaller birth cohorts in the previous 10 years, the age structure remains very young, with 47.5% of the population being under age 20 at the time of registration in MIPopLab compared to 54.6% nationwide in the GPC 1998. As was the case for fertility levels, the age structure of MIPopLab residents at the time of registration (2000-2002) is closer to that of the Phnom Penh population in 1998 (47.6% under age 20) than that of the country, suggesting that fertility decline in this Commune might have started about three years later than in the Municipality of Phnom Penh.

Residents' records indicate individuals' relationship to the household head. Demont and Heuveline (2008) used these data in part to study household structure in Cambodia, combined with nationally-representative data from the GPC 1998 and CDHS 2000. The prevalence of various types of household composition suggests the nuclear family remains the norm, but with a high tolerance and pragmatism for alternative living arrangements. Again, the authors show that the prevalence of nuclear families observed in MIPopLab is intermediate between the levels observed in the GPC 1998 and CDHS 2000 for urban and rural areas (in Cambodia, untypically, nuclear families are more prevalent in rural than in urban areas). However, the advantage of using MIPopLab data is to allow for a dynamic study of living arrangements over the live course using data from 2001-20006 updates, as described below.

For women over the age of 15, residents' records also their marital status. Heuveline et al. (2012) combined the GPC 1998 cross-sectional data on age, relationship to the head and marital status to study the transitions out of the parental home or out of the single state by age.

MIPopLab update data could also be used to study the dynamics of the transition to adulthood in Cambodia and its change over time.

Residents' records also provide information on parental survival. These data have not been comprehensively analyzed, but unpublished estimates show that among residents whose father was alive in 1975, 27% lost their father under the KR. Similar proportion for residents' mothers alive in 1975 is 17% for mothers. These proportions are slightly lower than national-level estimates for adult males and females (Heuveline 1998, Fig. 4), which conforms to the expectation given the variations in mortality between rural and urban residents under the KR (Kiernan 1996).

Finally, residents' record provide information on the location of surviving parents. Demont (2011) used these data to confirm that nuclear rather than multi-generational families tend to be the norm in Cambodia. The authors finds that 68% of married couples had formed at independent household at the time of the interview, and when do not, a greater tendency to live with or near relatives on the bride's than on the groom's side (20% v. 12%). However, cross-sectional data may mask normative, but short stages in the life course. Among younger married individuals, 40% still live with the bride's parents. The demographic updates allow for dynamic analyses of living arrangements (see below).

Marriage and Birth History Data

As part of the first rider survey, marriage and birth data were also collected at the time of the benchmark censuses in each of the Villages. Marriage histories were collected in each Village, providing for all ever-married women up to the age of 75 the date of marriage, the husband's birth date, and if applicable, the date the marriage ended and how it ended (death, divorce or separation). These data are provided in a separate *Marriage* file in each Village.

Heuveline and Poch (2006) used primarily the CDHS 2000 to study marriage stability by marriage cohorts, with a particular focus on the KR-era marriage cohorts (1975-79). Marriage under the KR has been described as "forced marriage," sometimes conducted without the

consent of the bride or groom, although the term is problematic in a setting where arranged marriages were the norm at the time (LeVine 2010). The CDHS 2000 data has several limitations for studying possible changes in marriage stability. First, the CDHS 2000 only samples women up to the age of 50, which raises well-known issues of age-censoring when studying more distant past events (Rindfuss et al. 1982). The issue was a real concern here since the main focus was on marriages that took place 21 to 25 years before the survey. The second issue is that the CDHS 2000 questionnaire does not include full marriage histories, but only records the age at first marriage and data on the current husband. With full marriage histories collected for all women up to the age of 75, MIPopLab allowed for simulations of the effects that the age-censoring and missing marriage information could have on the results obtained from the CDHS 2000 data. The study main result is that, surprisingly perhaps, KR-era marriage cohorts do not exhibit a higher probability of divorce or separation than temporally-adjacent marriage cohorts. Instead, the probability appears to have been slowly increasing among recent cohorts whose marriages are also less likely to have been “arranged” by relatives.

Birth data differ from the first Village to the rest of the Commune. In the first Village, only Children Ever Born/Children Surviving information was collected to allow for the estimation of child mortality as suggested by Brass (1975). These data are presented in a separate *CEB* file. Birth data was augmented in the other four Villages to provide the birth dates of all children ever born and to indicate the current location of all those still alive. For these Villages, the data are provided in two separate files. The *Births* file contains Children Ever Born/Children Surviving information, whereas the *Children* file list the birth dates of all the children ever born, and if surviving, their current residence.

Heuveline and Poch (2007) used the birth histories from the last four Villages to study fertility during the KR and the post-KR baby-boom. Again, the age censoring issue made the CDHS 2000 data inappropriate for studying fertility this far back. For the more recent period, MIPopLab data yield fertility rates that are lower than the CDHS-2000 national average, with

values intermediate between the CDHS-2000 estimates for the Municipality of Phnom Penh and those for the Province of Kandal. Despite the differences in levels, the trend in MIPopLab data is extremely consistent with the CDHS-2000 trend for the country, with MIPopLab values remaining at 77% to 78% of the CDHS-2000 ones from 1980 to 2000 (Heuveline and Poch 2007, p.414). Moreover, the difference originates almost entirely in different marriage rates, and MIPopLab marital fertility rates remain very close to the national averages between 1980 and 2000 (Heuveline and Poch 2007, p.419). Going further back, the MIPopLab data suggest a one-third decline in total fertility during the KR, followed by a substantial rebound. Combining MIPopLab marriage and birth histories allows the authors to further identify a two-year marriage “bubble” and a decade-long surge in marital fertility.

Update files (rounds 2 to 13)

Interview

As for the Census files, the *Interview* file lists each of the households visited during an Update round, with the visit date, the individual identification code of the respondent providing the household information, and a count of the event forms filled during the visit (e.g., birth, death, migration). We expect that the only analytical use of this file is to be merged with other files in order to append the date of the interview (for instance, to calculate the age of the respondent at the time of the interview from his or her birth date). *Interview* files were only generated up to the mid-2003 rounds of data collection. For the subsequent rounds, the following “central” dates are suggested: December 31st, 2003; June 30th, 2004; December 16th, 2004; June 30th, 2005; December 16th, 2005; June 30th, 2006; and December 16th, 2006.

Birth

The *Birth* file is a record of any live birth that has occurred between the last and the current visits to a woman residing in the household at the time of the current visit. The record simply indicates the birth date, but also contains household information that can be linked to the *Interview* file to recover the visit date, and paternal and maternal information that can be linked

to the *Residents* file to recover the parents' age at birth, for instance. The *Birth* file also provides the baby's relationship to the head of the household at the time of birth. For all rounds except the December-2005 and June-2006 rounds, there is one *Birth* file per village per round. For these two rounds, there is a single *Birth* file per round for all villages combined.

A total of 1,039 live births have been recorded between 2001 and 2006, with a sex ratio of 104 male to 100 female births. This total corresponds to a 2001-06 Crude Birth Rate (CBR) of 19.2 births per 1,000 person-years, which is 75% lower than the national average of 25.6 for the three-year period preceding the 2005 Cambodian Demographic and Health Survey (National Institute of Statistics 2006, CDHS 2005 thereafter). This difference is consistent with the above-mentioned birth-history estimates which showed MIPopLab total fertility remained at 77% to 78% of national averages, while marital fertility remained very close to national averages.

Table 2 shows the distribution of births by age of the mother, compared with the urban, rural and national distributions for the three-year period preceding the CDHS 2005. The distributions are fairly close with a possible exception of the proportions under age 25, with only 8% of birth to mothers under age 20 and 37% between ages 20 and 25 compared to 10% and 32% nationally. The MIPopLab distribution suggests a later age at first birth than in the rest of the country, contributing to the lower fertility levels observed in MIPopLab.

Note that for the purpose of estimating fertility rates, the criteria for making a birth record assume that women that move into the household between visits have been observed for the entire period between visits. Ideally, we would observe in-migrant women giving birth only from the time they move into the household and observe out-migrant women giving birth up to the time they move out of the household. In practice, however, we do not typically observe women who moved out of the household during visits and could not reliably record whether they have given birth before moving out of the household. To avoid a bias in our birth count and exposure measure, we thus use the entire period between visits for in-migrants and assume that

out-migrants were lost to follow-up at the time of the previous visit. In any event, a move and a birth within the same between-visit period (half a year) are relatively rare occurrences.

A more important concern for avoiding a bias in fertility estimation is to make sure that live births that do not survive to the time of the visit are still recorded as both a birth and a death since last visit. To reduce the risk that such births are not recorded, we also followed-up the pregnancy status of all eligible women at each visit, and asked the outcome of any previously recorded pregnancy at our next visit. In case the pregnancy resulted in a live birth but the child did not survive to the time of the survey, field workers were instructed to record both the live birth and the child's death. Comparing MIPopLab 2001-2006 data with data based on annual visits with the same population in subsequent years, Demont (2011) concludes that bi-annual visits reduced the bias due to unrecorded live births that do not survive to the time of the next visit, but that an underestimation of infant mortality (and thus fertility) is not entirely absent even in MIPopLab data.

Pregnancy

The *Pregnancy_Ong* file is a record of any on-going pregnancy at the time of the current visit. It simply identifies the woman reporting a pregnancy and its duration. At the next visit, this record is used as a reminder to update the pregnancy status of women reporting an ongoing pregnancy in the previous round. Though rare, it is possible that the pregnancy is still ongoing (observed, for example, in the 3rd and 9th month), in which case a 2nd on-going pregnancy record is made. More often during the 2nd visit, a record is made of the pregnancy outcome. These records are provided in the *Pregnancy_End* file, listing the woman's identifying information with how it ended (e.g., live birth, still birth, spontaneous or induced abortion) and when (month and year). Pregnancy outcomes are recorded for all pregnancies that have ended since the last visit regardless of whether an on-going pregnancy had previously been recorded. Whether it had been recorded during the previous visit is also indicated on the pregnancy outcome record with the pregnancy duration at the time.

For all rounds except the December-2005 and June-2006 rounds, there is one *Pregnancy_Ong* file and one *Pregnancy_End* file per village per round. For the December-2005 round, there is a single *Pregnancy_Ong* file and a single *Pregnancy_End* file for all villages combined. For the June-2006 round, there is only a single *Pregnancy_Ong* file for all villages combined. For the 3rd round in one village (December 2001), there is also a full reproductive update file, *Women_update*, which indicates if a reproductive-age woman has experienced any change in pregnancy or marital status. In subsequent rounds, these women can be assumed to have experienced no change unless recorded in one of the *Pregnancy* or *Marital status* files.

Death

The *Death* file is a record of any death that has occurred between the last and the current visits to any resident of the household at the time of the last visit. The record simply indicates the date of death, household information that can be linked to the *Interview* file to recover the visit date, and individual information that can be linked to the *Residents* file to recover the individual's age at death, for instance. For all rounds except the December-2005 and June-2006 rounds, there is one *Death* file per village per round. For these two rounds, there is a single *Death* file per round for all villages combined.

A total of 370 deaths have been recorded between 2001 and 2006, with a sex ratio of 109 male to 100 female deaths. The total corresponds to a 2001-06 Crude Death Rate (CDR) of 6.9 deaths per 1,000 person-years, which is low but consistent with the very young age distribution showed in Table 1 (median age just under 20 years for males and about 22-23 years for females). Table 3 shows the distribution of deaths by age and sex for all deaths. In the absence of a comparable nationally-representative age-and-sex distribution of deaths, the distribution of adult deaths (ages 15 to 49) can be compared to the distribution for the 6 years before the CDHS 2005. The two distributions are not directly comparable, however, as the CDHS 2005 adult death data are from sibling histories.

Based on MIPopLab data (2001-06) and the first round of MIPRAoC data collection in 2008 for the same population, Demont (2011) observes that infant and child mortality seem relatively low compared to overall mortality, which indicates a best fit to the age pattern of the *East* pattern among the Coale and Demeny (1983) model schedules of mortality. The author also provides age-specific mortality rates by gender and derives life tables with a life expectancy at birth of 60.7 for males and 65.1 for females (2011, pp.418-9), levels that are similar to the national average obtained for 2008. Unfortunately, there is no other source of direct estimates of mortality rates to which the MIPopLab estimates could be compared, and national estimates have been based on indirect techniques applied to on children ever born/children surviving reports and on sibling histories.

For the purpose of estimating mortality rates, note that exposure needs to be measured slightly differently than for fertility rates. The criteria for making a death record is based on an individual's presence at the round prior to the current one. It is of course possible that a new resident moves into the household between rounds and dies before the current visit. However, our protocol is such that these individuals are unlikely to be recorded as both a migration and a death. For this reason, individuals registered as migrants between the last two visits are considered not to contribute to the measure of exposure to the risk of death. That measure only counts the exposure of individuals registered as residents at the prior visit, with their exposure censored at the time of out-migration or death if either has occurred between the last two visits.

The 2001-06 CBR and CDR together yield a Crude Rate of Natural Increase of 12.3 persons per thousand person-years. This is only slightly lower than the intercensal rate of population change of 1.54 percent per year (15.4 per thousand person-years) for the whole country between 1998 and 2008 measured in the General Population Census 2008 (National Institutes of Statistics 2010, GPC 2008 thereafter). Most of the population change, however, originates from migration to and from the MIPopLab catchment area. In particular, a high level

of migration was expected due to the proximity to the capital city and the fact that most of the country's rural areas experience substantial outmigration (NCPD 2009).

Migration

The *Migration_In* file is a record registering an individual as a resident of the household, while the *Migration_Out* file is a record registering an individual as no-longer a resident of the household. The residency status can sometimes be difficult to assess due to the high frequency of temporary and circular migration in this population. In rural Cambodia, there is a long tradition of temporary migration during periods of lesser agricultural needs (Delvert 1961), which, in order not to conflict with these needs, coincide with the timing of our visits. We defined a resident as someone who stays overnight more than 50% of the time, say, 4 nights a week or more, or for longer duration of back and forth migration, 3 months or more in the last 6 months (typical duration between visits). These *de facto* criteria partly conflicted with the residents' perception of who was or wasn't a resident of the household. A child attending school away from home and coming back on weekends might still have been considered as a household resident, whereas an employee who had been living in the household for a year might not. In a nutshell, household heads tended to report on whom they considered to be members of the household and to discount temporary absence. We believe that, over the years, we were able to have the records better match our residency definitions. In some instances, this involved reporting migration that had occurred several years earlier. To report those who pre-dated the establishment of MIPopLab at the end of 2000, we used the convention of dating them as happening in the year 2000 (the actual date being irrelevant as long as those are excluded from our measures of exposure). A total of 304 such migration records (236 in, 68 out) are thus to be treated as corrections to the residency records at the time of the baseline census rather than actual moves since the time of the census. For all rounds except the December-2005 and June-2006 rounds, there is one *Migration_In* file and one *Migration_Out* file per village per round.

For the December-2005 round, there is a single *Migration_In* file and a single *Migration_Out* file for all villages combined. There is no *Migration_In* file for the June-2006 round.

A total of 4,552 individual moves have been recorded between 2001 and 2006, with slightly more move out of the area (2,393) than in (2,140). The excess of departures over arrivals between 2001 and 2006 is thus relatively small (253). However, this excess amounts to about one third of natural increase during the period (669). The overall population growth during the period is thus positive but modest, less than .8 percent per year.

The age distributions of individual moves to and from the catchment area are shown in Table 4 and generally conform to the age schedule of migration elsewhere. Most moves correspond to young males (20-to-29 year-olds) and females (15-to-24 year-olds), and their children under the age of 5. Compared to the distribution of individuals having moved between the GPC 1998 and GPC 2008, the young-adult peak is much more visible in our distributions, with proportions in the modal age group (20-to-24 year-olds) reaching 30% or more. This is due in part to the fact that the GPC 2008 is based on the age at the time of the census rather than at the time of the move so an individual moving at age, say, 23, could be anywhere between 23 and 33 at the time of the census. If we simulate observing the migrant's age at the end of a ten-year period rather than at the time of the move, our distribution would peak in the 25-to-29 age-group rather than the 20-to-24 age group and only reach 23% at the mode (results not shown). The mode is still higher than in the GPC 2008 distribution which only reaches 14%. As ours are distributions of moves rather than of individuals, this suggests, plausibly, that young adults are also the most likely to make multiple moves in any given period.

Another difference is that the sex ratio of moves in MIPopLab is about 115 male moves per 100 female moves, whereas the national figures show roughly equal numbers of male and female migrant. This difference might originate in the economic activities of the MIPopLab residents. While males are primarily engaged in agriculture, many women are engaged in crafts, silk weaving in particular. Over the years, we observed a shift in the relative profitability of the

two economic sectors, with some men abandoning agricultural work for silk weaving on the island or for salaried work elsewhere. In the early 2000s at least, silk weaving seemed to have provided young women an alternative to seeking employment in one of the many garment factories that have sprung around the capital city (Chea and Sok 2001; Lim 2008). A second factor is that marriage remains one of the primary reasons for migration and, while a married couple is expected to form an independent household eventually, it is fairly common for newlyweds to first cohabit with either the groom's parents or, more frequently, the bride's parents (Heuveline forthcoming). The GPC 2008 data show that marriage was the reason for migration for 18.6% of male migrants versus 10.6% of female migrants.

As for the volume of migration, the ratio of moves to residents approached 8 percent in MIPopLab between 2001 and 2006. Because the same individual may have contributed multiple moves between 2001 and 2006 (an internal move, in particular, should be recorded both an out-migration and an in-migration), this ratio cannot be directly compared with the GPC 2008 proportion of individuals having changed residence since the GPC 1998 (26.5%). The MIPopLab ratio would suggest 80 individual moves per 100 residents over a ten-year period, three times more than individual "movers" in the GPC 2008, or an average of one internal move (recorded both at origin and at destination) and one move outside of the catchment area per mover. Tentative though they are, these comparisons suggest a level of mobility in MIPopLab comparable to or possibly a little higher than the national average. As mentioned earlier, the proximity to the capital city led us to expect extensive, labor-related migration when the area was selected. The nearly equal level of in-migration is more surprising, and indicates that most migration is temporary rather than permanent.

Marital Status

To avoid double counting, changes in marital status (marriage, marriages ending in divorce or death, physical separation of still married partners, and reunion ending such a period of separation) are only recorded for women. In 2001 and 2002, these records are provided in a

single *Marital_Change* file per village and per round, identifying the identity of the two partners, the date of the change and the type of change. From 2003 on, records are provided in separate *Marriage*, *Divorce* (actually including all marriages ending in either divorce or husband's death), *Separation*, and *Reunion* files. For the December-2005 round, there is a single file for all villages combined for each of these four types of marital-status change. For the June-2006 round, there is only a *Marriage* file for all villages combined. From 2001 to 2006, a total of 593 changes in marital status have been recorded.

Demont (2011) provides detailed analyses of marriage data showing that marriage remains nearly universal for men and women and that a first marriage occurs around age 25 on average for men and two to three years earlier for women. The author also shows that 12.6% of men and 15.0% of women marry with the Commune, and similar proportions marry outside the Commune but within the District. Less than half the men and less than half the women marry outside the Province, which is slightly surprising given the extent of circular migration to the capital city (which was at the time located in a different Province). The analyses also confirm the above-mentioned tendency to cohabit with the bride's parents rather than with groom's when the newly-weds do not transition to an independent household right after marriage.

Headship Changes

Changes in the identity of the head of the household were recorded in order to update the household members' relationship to the head. However, separate records were only created for these changes in the two rounds conducted in 2002. A total of 63 changes recorded during these two rounds are shown in separate *Headship_Change* files per round and per village.

Demont (2011) combines information on relationship to the head at the time of the census in the *Resident* files with information on subsequent changes resulting from birth, death, migration and headship changes to study the dynamics of living arrangements and life course trajectories from 2000 to 2006.

Further Analyses

Further analyses can be carried on MIPopLab data, which this release will hopefully facilitate. Again, MIPopLab data are not necessarily representative of the country or any particular region of the country. As illustrated by some of the above-mentioned analyses, when nationally-representative, comparable yet cross-sectional data are available, it becomes possible to take advantage of the longitudinal nature of MIPopLab data and to compare MIPopLab data and national levels at one particular point in time. Analyses so far suggest that in spite of some notable differences in levels (lower rates of fertility and mortality, higher rates of in- and out-migration), demographic trends in MIPopLab trends seem to parallel those in the rest of the country.

References

- Banister, Judith and E. Paige Johnson. 1993. After the nightmare: The population of Cambodia, in Benedict Kiernan (ed), *Genocide and Democracy in Cambodia: The Khmer Rouge, the United Nations and the International Community*. New Haven, CT: Yale University Southeast Asia Studies, pp. 65–140.
- Brass, W. 1975. *Methods for Estimating Fertility and Mortality from Limited and Defective Data, Based on Seminars Held 16-24 September 1971 at the Centro Latinoamericano De Demographia (CELADE), San Jose, Costa Rica*. Chapel Hill, NC: International Program of Laboratories for Population Statistics.
- Chea, H., and H. Sok. 2001. The Cambodia Garment Industry. *Cambodia Development Review* 5 (3): 1–8.
- Clark, S., E. Colson, J. Lee and T. Scudder. 1995. “Ten Thousand Tonga: A Longitudinal Anthropological Study from Southern Zambia,” *Population Studies* 49(1): 91-109.
- Cleland, John. 2001. The effects of improved survival on fertility: A reassessment, *Population and Development Review* 27(Supplement: Global Fertility Transition): 60–92.
- Coale, Ansley and Paul Demeny, with Barbara Vaughan. 1983. *Regional Model Life Tables and Stable Populations*. New York, NY: Academic Press.
- Delvert, Jean. 1961. *Le paysan cambodgien*. Paris, France: Mouton et Compagnie
- Demont, Floriane. 2011. *Dynamiques Démographiques et Parcours de Vie au Cambodge, 1998-2008 : Des Trajectoires Individuelles et Familiales à l'Épreuve du Développement*. Genève, Switzerland : Université de Genève, Doctoral Thesis.
- Demont, Floriane and Patrick Heuveline. 2008. “Diversity and Change in Cambodian Households (1998-2006),” *Journal of Population Research* 25(3): 287-313.
- De Walque, Damien. 2005. Selective mortality during the Khmer Rouge period in Cambodia, *Population and Development Review* 31(2): 351–368.

- Eversley, David E. C. 1957. "A Survey of Population in an Area of Worcestershire From 1660-1850 on the Basis of Parish Records," *Population Studies* 10(3):253-79.
- Frings, Viviane. 1994. "Cambodia After Decollectivization (1989-1992)," *Journal of Contemporary Asia* 24 (1): 49–66.
- Galloway, Patrick R. 1988. "Basic Patterns in Annual Variations in Fertility, Nuptiality, Mortality, and Prices in Pre-Industrial Europe," *Population Studies* 42(2):275-302.
- Goldman, Noreen, Lorenzo Moreno, and Charles F. Westoff. 1989. *Peru Experimental Study: An Evaluation of Fertility and Child Health Information*. Columbia, Maryland: Demographic and Health Surveys, Institute for Resource Systems/Macro Systems.
- Heuveline, Patrick. 1998. 'Between one and three million': Towards the demographic reconstruction of a decade of Cambodian history, *Population Studies* 52(1): 49–65.
- _____. 2015. "The Boundaries of Genocide: Quantifying the Uncertainty of Death Toll During the Pol-Pot Regime (1975-1979)," *Population Studies* 69(2):201-218.
- _____. Forthcoming. "Cambodian Families and Households," in S. Springer and K. Brickell (Eds.) *Handbook of Contemporary Cambodia*. Abington, United Kingdom: Routledge
- Heuveline, Patrick, Floriane Demont and Bunnak Poch. 2012. "The Transition to Adulthood in Cambodia," Population Working Paper PWP-CCPR-2012-016, California Center for Population Research, University of California, Los Angeles.
- Heuveline, Patrick and Bunnak Poch. 2006. "Do Marriages Forget Their Past? Marital Stability in post-Khmer-Rouge Cambodia," *Demography* 43(1): 99-125.
- _____. 2007. "The Phoenix Population: Demographic Crisis and Rebound in Cambodia," *Demography* 44(2): 405-26.
- Hirschman, Charles. 1994. "Why Fertility Changes?" *Annual Review of Sociology* 20:203-233.

- Huguet, Jerrold W, Apichat Chamrathirong, Nott Rama Rao, and San Sy Than. 2000. Results of the 1998 population census in Cambodia, *Asia-Pacific Population Journal* 15(3): 3–22.
- Kiernan, Benedict. 1996. *The Pol Pot Regime*. New Haven, CT: Yale University Press.
- Ledgerwood, Judy, and John Vijghen. 2002. "Decision-Making in Rural Khmer Villages," pp.109-150 in Judy Ledgerwood (Ed.) *Cambodia Emerges from the Past*, DeKalb, IL: Center for Southeast Asian Studies, Northern Illinois University.
- Lee, Ronald D. 1997. "Population Dynamics: Equilibrium, Disequilibrium, and Consequences of Fluctuations," pp.1063-1115 in Mark R. Rosenzweig and Oded Stark (Eds.) *Handbook of Population and Family Economics, vol. 1*, Amsterdam: Elsevier Science B.V.
- LeVine, Peg. 2010. *Love and Dread in Cambodia: Weddings, Births, and Ritual Harm Under the Khmer Rouge*. Singapore: National University of Singapore.
- Lim, Sovannara. 2008. "Youth Migration in Cambodia," *Cambodia Development Review* 12 (4): 1–5.
- Lindstrom, David P. and Betemariam Berhanu. 1999. "The Impact of War, Famine, and Economic Decline on Marital Fertility in Ethiopia," *Demography* 36(2):247-61.
- Lognard, Marie-Odile. 2012. "Conséquences démographiques de la période khmer rouge," [Demographic consequences of the Khmer Rouge period] paper presented at the Chaire Quételet, Impact démographique des conflits. Last accessed on 12/12/2013 at <http://www.uclouvain.be/cps/ucl/doc/demo/documents/Lognard.pdf>
- Mason, Karen O. 1997. "Explaining Fertility Transitions," *Demography* 34(4):443-54.
- National Committee for Population and Development (NCPD) .2009. *Migration and Development in Cambodia*. Phnom Penh, Cambodia: The Office of the Council of Ministers and United Nations Population Fund.

- National Institute of Statistics. 1999. *General Population Census of Cambodia 1998: Final Census Results*. Phnom Penh, Cambodia: National Institute of Statistics, Ministry of Planning.
- _____. 2001. *Cambodia Demographic and Health Survey 2000*. Phnom Penh, Cambodia: National Institute of Statistics and Directorate General for Health, and Calverton, MD: ORC Macro.
- _____. *Cambodia Demographic and Health Survey 2005*. Phnom Penh: National Institute of Statistics and Directorate General for Health, and Calverton, MD: ORC Macro.
- National Institute of Statistics. 2010. *General Population Census of Cambodia 2008: Final Census Results*. Phnom Penh: Ministry of Planning.
- Neupert, Ricardo F. and Virak Prum. 2005. "Cambodia: Reconstructing the demographic stab of the past and forecasting the demographic scar of the future," *European Journal of Population* 21(2/3): 217–243.
- Palloni, Alberto, Kenneth Hill, and Guido P. Aguirre. 1996. "Economic Swings and Demographic Changes in the History of Latin America," *Population Studies* 50(1):105-32.
- Prescott, Nicholas M. and Menno Pradhan. 1997. *A Poverty Profile of Cambodia*. Washington, D.C.: The International Bank for Reconstruction and Development.
- Rindfuss, R., J. Palmore and L. Bumpass. 1982. "Selectivity and the Analysis of Birth Intervals from Survey Data," *Asian and Pacific Census Forum* 8(3):5-16.
- Shryock, Henry S. and Jacob S. Siegel. 1973. *The Methods and Materials of Demography*. Washington, DC: U.S. Department of Commerce.
- Watkins, Susan C. and Jane Menken. 1985. "Famines in Historical Perspective," *Population and Development Review* 11(4):647-75.

Table 2: Distribution of births by age of mother, MIPopLab (2001-06) and CDHS 2005 (2002-05)

	MIPopLab (2001-06)		CDHS 2005 (2002-05)
	All women	15-44 only	National
Under 15	0.4		
15-9	7.9	8.0	10.2
20-4	36.8	37.0	32.0
25-9	22.8	22.9	22.2
30-4	17.3	17.4	17.8
35-9	9.9	9.9	12.2
40-4	4.3	4.3	5.2
45-9	0.5	0.5	0.5
Over 50	0.1		
All	100.0	100.0	100.0
Total	1039	1034	4995

Table 3: Distribution of deaths by age and gender, MIPopLab (2001-06) and CDS 2005 (6 years prior survey)

	All Ages		Ages 15-49 only			
	MIPopLab (2001-06)		MIPopLab (2001-06)		CDHS 2005 (6 years prior survey)	
	Males	Females	Males	Females	Males	Females
Under 5	13.1	3.5				
5-9	2.9	1.4				
10-4	1.3	2.1				
15-9	0.7	1.1	2.2	3.8	7.2	6.6
20-4	2.0	3.2	6.5	11.3	10.7	10.5
25-9	0.7	2.5	2.2	8.8	14.0	13.6
30-4	5.9	4.6	19.4	16.3	23.4	21.1
35-9	9.2	6.0	30.1	21.3	17.6	17.8
40-4	5.9	3.5	19.4	12.5	16.4	17.2
45-9	6.2	7.4	20.4	26.3	10.8	13.2
50-4	6.5	7.1				
55-9	2.3	4.3				
60-4	2.6	6.7				
65-9	4.9	6.7				
70-4	9.2	11.7				
75-9	10.5	8.5				
80-4	10.1	8.5				
85-9	2.3	6.4				
90 and over	3.9	4.6				
Total	100.0	100.0	100.0	100.0	100.0	100.0

Table 4: Age distribution of migrants, MIPopLab (2001-06), in-migrants and out-migrants, and GPC 2008 (1998-2008), all migrants

	All ages				Ages 10 and over only					
	MIPopLab (2001-06)				MIPopLab (2001-06)				GPC 2008 (1998-2008)	
	In-Migration		Out-Migration		In-Migration		Out-Migration		All migrants	
	Males	Females	Males	Females	Males	Females	Males	Females	Males	Females
Under 5	11.4	11.3	8.4	7.8						
5-9	5.6	6.8	6.4	8.1						
10-4	5.2	5.9	6.0	7.6	6.2	7.2	7.0	9.0	5.9	5.7
15-9	9.7	20.3	14.5	19.9	11.7	24.8	17.1	23.7	9.4	10.2
20-4	28.8	25.5	27.4	24.7	34.8	31.1	32.1	29.4	12.8	13.4
25-9	16.4	9.7	14.9	12.2	19.8	11.8	17.5	14.6	14.3	12.7
30-4	8.2	6.4	8.8	6.0	9.8	7.8	10.3	7.1	9.1	7.8
35-9	5.2	3.3	5.1	3.3	6.2	4.0	6.0	3.9	11.6	10.0
40-4	2.3	2.5	1.8	2.0	2.7	3.0	2.2	2.4	10.0	9.0
45-9	2.4	1.9	2.1	2.0	2.9	2.3	2.4	2.4	8.6	8.5
50-4	2.0	2.1	1.2	1.5	2.4	2.6	1.4	1.8	5.8	7.2
55-9	0.6	0.4	1.1	1.1	0.7	0.5	1.3	1.3	4.5	5.3
60-4	0.3	0.9	0.9	1.3	0.4	1.1	1.0	1.5	3.0	3.5
65-9	0.3	0.9	0.6	0.6	0.4	1.1	0.7	0.7	2.2	2.6
70-4	0.6	0.8	0.5	0.7	0.7	1.0	0.6	0.8	1.4	1.9
75-9	0.7	0.9	0.2	0.8	1.2	1.7	0.3	1.4	1.5	2.3
80-4	0.3	0.3	0.1	0.3						
85-9	0.0	0.2	0.0	0.1						
90 and over	0.0	0.1	0.0	0.0						
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Sex ratio	116		114		117		115		102	

Note

¹ The Mekong Island Population Laboratory (MIPopLab) was established in collaboration with a team of Cambodian researchers, most of them affiliated at the time with the Royal University of Phnom Penh (RUPP), including Bun Serey, Deup Channarith, Ken Veasna, Long Bunnath, Nguon Srieng, Prum Virak, Puch Sina, and Tuot Sovannary. Foremost, Lany Trinh-Bo substantially contributed to the project design and served as the founding project manager. Later MIPopLab collaborators include Chea Chantha, Chhim Chankev, Lay Huoy, Long Yav, Nuon Horn, Saing Hay, and Tong Soprach. The project was developed with the support of several seed grants to the first author: (1997) from the Population Studies Center, University of Pennsylvania, the Mellon Fund for Training and Research in the Demography of the Less Developed Countries; (1999-2000) from the Population Research Center, NORC and the University of Chicago, infrastructure grant P30HD18288 from the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD); and (2000-2001) from the University of Chicago, J. David Greenstone Memorial Fund and Office of the Provost. From 2002 to 2005, MIPopLab was supported by NICHD research grant R03HD41537 to the first author. MIPopLab was eventually integrated into the Mekong Integrated Population-Registration Areas of Cambodia (MIPRAoC) project, which was developed with the support of a seed grant to Alan Kolata and the first author (2005) from the Population Research Center, NORC and the University of Chicago, NICHD infrastructure grant P30HD18288. Benjamin Clark, Jeffrey Eaton, David Kirk, Floriane Demont, Savet Hong, Eng Sothy, Gail Yen and Brooks Ambrose have contributed to the coding, cleaning and release of the MIPopLab datafiles. Preparation of this paper made use of the facilities and resources at the California Center for Population Research, UCLA, which is supported by infrastructure grant R24HD041022 from NICHD. The first author expresses his deep gratitude to all the above institutions and individuals, and the many more who could not all be named here but nonetheless provided ideas and encouragements over the years. The authors remain solely responsible for the content of this paper.