**Appendix 1.**

Overview, Design Concepts, and Details (ODD) of the Chitwan ABM.

## OVERVIEW

Following the Overview, Design concepts, and Details (ODD) model description framework (Grimm et al. 2006, 2010), this section describes the purpose, entities, state, variables, and scales, and the processes of the Chitwan ABM. The model code is available online at http://www.openabm.org/model/3640.

### Purpose

The Chitwan ABM was constructed to explore the role of feedbacks between land-use and demographic change in determining decadal scale dynamics of land-use and land-cover change in the Chitwan Valley. The model is intended to be used as a laboratory for testing the relative effect of different drivers of demographic and land use and cover change, and for the exploration of alternative development scenarios.

We represent the sample of the Chitwan Valley Family Study (CVFS) using a nested hierarchical structure of individual people nested within households, nested within their broader neighborhood (~ 500 m radius spatial scale). Consistent with this analysis and with past findings from the literature, we assume that individual-level decision-making processes regarding fertility, marriage, and migration are shaped by neighborhood characteristics, and that neighborhoods that are close together (spatially) are more strongly linked than those that are more distant (for example, people are likely to build new households near their old neighborhoods). Furthermore, and consistent with empirical observations, we assume that the dominant land-use change in the settled area of the Valley is land conversion out of agriculture and into the built environment, and that new household formation is the major driving force behind this change.

## Entities, state variables, and scales

The Chitwan ABM represents the population of the Chitwan Valley using a multi-level structure mirroring that used in the CVFS (individuals, households, and neighborhoods, see Table 2) as well as the observed social context (electrification, existence of non-family organizations or NFOs), and land-use patterns. The lowest-level agent is a “person” agent. The person agent represents a single individual person from the CVFS survey. The model is initialized with 8,245 person agents. Each person agent has a number of attributes assigned based on the survey data (see Table A1.1 for a complete list). Each person agent is in turn assigned to one of 1,522 households, by matching the household ID number of each individual in the CVFS survey to the appropriate household number in the survey. Each household has a set of attributes assigned from empirical data (Table A1.2). Household agents are assigned to one of 151 neighborhoods, by linking household and neighborhood ID numbers from the CVFS survey. Neighborhoods, in turn, have a set of attributes (Table A1.3), including land use within each neighborhood.

Agricultural land is the predominant land use category within the cleared area of the study site. As of 1996, when the first CVFS mapping was conducted agricultural land occupied greater than 80% of the total non-forested land area. The CVFS collected land use data at three points in time (1996, 2000, and 2007) at the neighborhood-level. Each neighborhood was mapped by hand, with compasses and tape measures. The original survey mapped 18 different classes – for simplicity, and consistency with prior analyses (such as Yabiku 2006; Axinn and Ghimire 2007; Ghimire and Hoelter 2007), we collapse these 18 classes into 5 classes: agricultural vegetation, non-agricultural vegetation, private buildings, public buildings, and other (canals, ponds, rivers, silted land, and undeveloped land).

Land use data was collected at the neighborhood-level through the CVFS – individual land parcels are not connected to individual households (as no unique IDs exist in the mapping data to match a household plot with its owner). The ABM tracks land use at the neighborhood-level, as the total land area within each class. Summary statistics on LULC within the CVFS sample neighborhoods are given in Table 1. Land use within the model itself is therefore stored as point data – land areas within each category, associated with a particular neighborhood.

The model runs with a monthly time step, beginning in January 1997 (the first month of the CVFS household registry data collection), and ending in December, 2050. A monthly time step was chosen in order to match the available data from the CVFS, while minimizing computational complexity (a smaller time step would significantly increase the computational requirements for running the model).

**Process overview and scheduling**

As seen in Figure A1.1, the Chitwan ABM runs with a monthly time step, and a series of submodels run in succession at each time step. Each submodel is briefly discussed here; see the Submodels subsection of the Details section for full details on each submodel. Each time step begins with the fertility submodel, which handles women's first children after marriage, and then handles subsequent births to women who have already had their first child. The mortality submodel runs next. The marriage submodel follows, and then the migration submodel which allows in and out-migration at the individual and household levels. The last submodels to run are the divorce submodel and the education submodel. Landscape change due to new household formation is modeled in both the marriage and migration submodels (see these subsections for details).

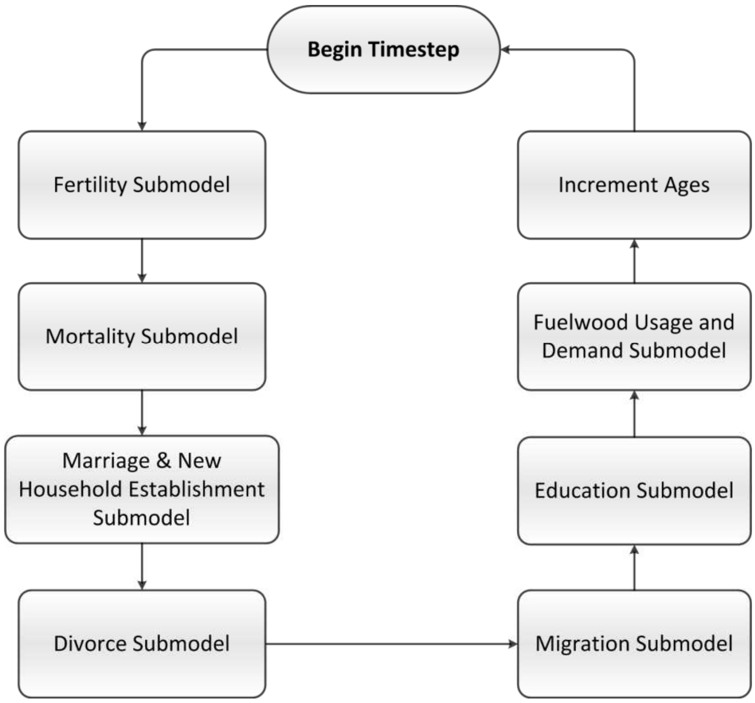


**Figure A1.1: Process scheduling in the Chitwan ABM.**

## DESIGN CONCEPTS

In this section, we outline the key design concepts behind the Chitwan ABM, under the eleven major subheadings of the ODD protocol.

### Basic principles

The Chitwan ABM draws on theory from the literature on the drivers of land-use and land-cover change, and from the sociological literature on the key determinants of demographic change. Existing work has shown the influence of land-use change (such as the prevalence of agriculture in an individual's neighborhood) on human behaviors including fertility, migration, and marriage. Drawing on this body of evidence, the Chitwan ABM uses separate submodels to model demographic and land-use change, allowing for reciprocal causation, whereby changes in land use can influence human behaviors, which, in turn, determine land use patterns.

**Emergence**

Population dynamics and land-use change emerge based on the lower-level interactions and decisions of person and household agents. These decisions are governed by the process models build into the Chitwan ABM (see Submodels subsection of Details section).

**Adaptation**

Adaptation is not modeled directly in the present version of the Chitwan ABM. Though agents responses vary with environmental and demographic conditions, the models representing agent responses are fixed (the coefficients in the marriage model, for example, are fixed over the course of a model run).

**Objectives**

Agents do not seek to maximize “success” or the probability of a particular objective, but are instead modeled as autonomous agents encountering a number of decision-making points (marriage, education, migration, etc.). At each decision point, agents make decisions in accordance with the observed data, with a set of stochastic techniques (empirically derived probability distributions, regression models, heuristic models, etc.) used to model the decision making process.

**Learning**

Learning is not represented in the Chitwan ABM – the rules agents use to make decisions to not change over time. Agent’s responses evolve over time based on changing demographic and environmental conditions; however, the decision rules and stochastic models are fixed.

**Prediction**

Agents in the Chitwan ABM do not use prediction models in their decision making.

**Sensing**

All person agents are assumed to know their own demographic and socioeconomic characteristics. Each person agent is also aware of the characteristics of their household and neighborhood, including land use in their local area. This information informs person agent's decisions. Global characteristics (total population, total land use) cannot be observed by the agents in the model.

**Interaction**

Agents interact with each other (marriage, divorce, and fertility) and with their environment (through new household establishment). The primary interaction between agents is marriage. In the marriage submodel, two eligible person agents may marry and may establish a new household with probability dependent on the state variables associated with each person, and on the coefficients of the marriage timing model and the household fission rate.

**Stochasticity**

Stochasticity plays a large role in the Chitwan ABM. Every submodel discussed in the Submodels subsection of the Details section includes a stochastic component. See this subsection for additional details on stochasticity in the Chitwan ABM.

**Collectives**

Two types of collectives exist in the Chitwan ABM: neighborhood agents and household agents. Neighborhood agents are composed of household agents, and household agents are composed of person agents. Both neighborhood and household agents also possess their own unique, higher-level state variables beyond those derived from their member households (they could be considered 'imposed' collectives as they are forced to exist by the definitions of the modelers).

**Observation**

Our primary result of interest for this paper is how feedbacks between natural and human systems affect spatial and temporal patterns of land use in the Chitwan Valley, and spatial and temporal patterns in human actions (marriage and first-births). As these phenomena are related in a complex manner, with feedbacks between human decisions and environmental change, the spatial patterns resulting from changes in first birth timing and marriage timing cannot be easily predicted. In order to observe the impact of feedbacks between first birth timing and land use and feedbacks between marriage timing and land use, we observe a set of model outcomes: first birth times, marriage ages, land use, and total population. Other outcomes, including household size, fuelwood usage, migration patterns, birth rates, mortality rates, and spatial patterns of demographic change can also be monitored, but are not considered for this paper.

**DETAILS**

Following the ODD description framework, in this section, we describe the details of model initialization, input data used to model exogenous processes (not applicable to the Chitwan ABM), and provide a detailed description of the submodels included in the ABM.

**Initialization**

The Chitwan ABM is initialized using household registry and survey data from the CVFS survey. The number of agents of each type follows the distribution of the CVFS data (see Table 2). There are four different types of agents (persons, households, neighborhoods, and regions), with the population of each type of agent determined by the number of respondents in the CVFS survey data. The respondents from the CVFS survey are represented in the model using a one-to-one mapping, in which each agent in the model represents a single person, household, or neighborhood from the CVFS sample. There is only one region agent, of which all neighborhoods are a member (regions agents are included in the model code for usage in later experiments).

**Input data**

Other than the initialization data mentioned above, no exogenous input data is used in the Chitwan ABM to "model processes that change over time" (Grimm et al. 2010, 2765).

**Submodels**

Marriage, fertility, mortality, migration, education, and land-use change are modeled based on relationships derived from existing empirical works, and from our analysis of time series data from the CVFS. The model is designed with flexibility in mind, to allow testing of different hypotheses regarding feedbacks between demographic change and land-use. A modular design allows switching between alternative sub-models and parameterizations so that users may pick the most appropriate process representations for a particular research question.

*Marriage*

The marriage sub-model uses regression results to parameterize the marriage process based on a number of covariates. Prior studies in Chitwan has shown that individuals within primarily agricultural neighborhoods marry sooner than those in primarily urban neighborhoods, even when controlling for other factors (Yabiku 2006a). To represent this in the model, the marriage sub-model uses regression coefficients derived from a discrete-time-hazard analysis predicting probability of marriage within a given month depending on a number of covariates (Table 3). The covariates included in the marriage timing model are chosen based on the work of Yabiku (2006a).

The regression model used in the Chitwan ABM is based on the work of Yabiku (2006a). The model is based on the discrete-time event history model published by Yabiku (Table 2, model 5, on page 456 of Yabiku 2006). The regression model was rerun for this paper based on the same modeling strategy and controls Yabiku used, in order to obtain the intercept for the regression equation, and the values of the monthly dummy variables included in the model to represent the baseline hazard (these parameters were estimated, but not published in Yabiku, 2006a). As different statistical software was used, the regression coefficients used here vary slightly from those in the original paper, though the significance and effect sizes for the coefficients generally agree.

In Table 3 the regression coefficients are converted to odds ratios (exponentiated regression coefficients) so that they can be more easily interpreted. A positive odds ratio on a covariate therefore means that an increase in the covariate will lead to a decrease in marriage age (by increasing the monthly probability of marriage). In each time step, the marriage sub-model calculates the monthly probability of marriage for each eligible individual based on the values of these covariates during that time step.

Following the work of Yabiku (2006a), the marriage timing regression model includes the log of agricultural vegetation to be consistent with theoretical expectations of the feedback of land-use change on marriage timing. The log scale represents, for example, the expectation that the transition from 10% to 20% agriculture (a 10% change) would be expected to have more of an effect on individuals than the transition from, for example 80% to 90% agriculture (again a 10% change). While the magnitude of both these changes is identical, the relative impact of a doubling of agricultural land use (10% - 20%) would represent a greater qualitative change in a neighborhood than going from 80%-90% agriculture.

To model the marriage process, at the beginning of each time step, a list is made of all unmarried individuals (both male and female). The marriage process in the Chitwan begins with the model first running through this list and checking if each person (male or female) has reached the minimum marriage age (15), and is below the maximum marriage age (35, both are empirically determined). For those meeting these requirements, the probability of marriage within that time step is calculated for each individual using whichever of the two marriage sub-models has been chosen for the model run. Once this probability is calculated, a random number is drawn on the interval of [0, 1], and, if the number is less than or equal to the probability of marriage for that individual, the person is added to a list of individuals to be married in that time step. To model in-migration through marriage, individuals may marry an in-migrant with a probability of .1, as calculated from the 48 months of the CVFS household registry data.

Next, males and females within the list are paired up, subject to the requirement that all individuals marry within their own ethnic group. Marriage outside of one's ethnic group is exceedingly rare in the Chitwan Valley. The empirical data shows that spouses tend to be fairly close in age, with the man tending to be slightly older than the woman. To account for the expected age differential between spouses, we assign a probability of marriage for each potential spouse based on the age difference between the individuals (see Figure A1.2 for details).

*Fertility*

There are three parts to the fertility model in the Chitwan ABM. First is the determination of a woman's desired family size. Second, a submodel is used to model the first birth after marriage. Third, a separate process models all subsequent births, up until a woman's desired family size is reached (or until she is no longer married, or dies). These models are developed using empirical data from the CVFS household registry, which contains monthly records of live births for all individuals in the survey.

Desired family size is determined based on the desired family size reported in the CVFS survey data. For new agents born into the model, a desired family size is assigned to each woman at marriage based on the histogram of the empirically reported data (see Figure A1.3). While future work may take into account covariates associated with desired family size, this is not done at present.

First birth timing is modeled in a similar manner to marriage timing, by using the results of a regression model to predict the monthly probability of birth for each eligible woman for each time step. Following the results of Ghimire and Hoelter (2007) and Ghimire and Axinn

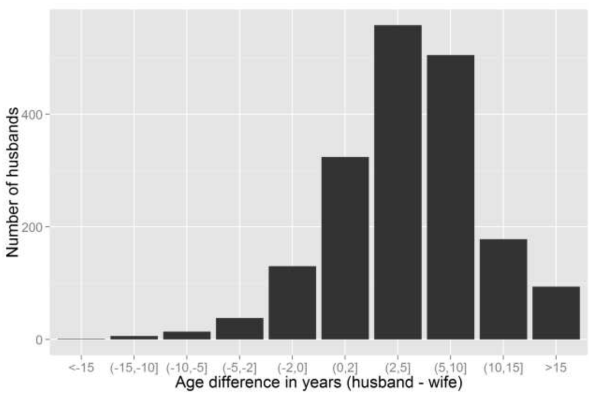


**Figure A1.2: Empirically derived histogram of spouse age differences (in years) used to weight the probability of a given person agent marrying a particular potential spouse.**

(2010), we use discrete time event-history analysis to develop a regression model relating a series of covariates, including agricultural land use at the neighborhood level, to the monthly odds of a newly married woman having her first birth (Table 4). A positive odds ratio on a covariate therefore means that an increase in the covariate will lead to a decrease in first birth timing (by increasing the monthly probability of a first birth). The regression model used in the Chitwan ABM was rerun for this paper based on the modeling strategy and controls used in panel A, model 3 in Table 2 in Ghimire and Hoelter (2007, 309), in order to obtain the intercept for the regression equation (which was not published in Ghimire and Hoelter 2007). As different statistical software was used, the regression coefficients used here vary slightly from those in the original paper, though the significance and effect sizes for the coefficients generally agree.

At the beginning of each time step, a list is constructed of all married women who have not yet had their first birth, and who have a desired family size greater than zero. For each woman, a probability of first birth in that month is then calculated based on the values of all of

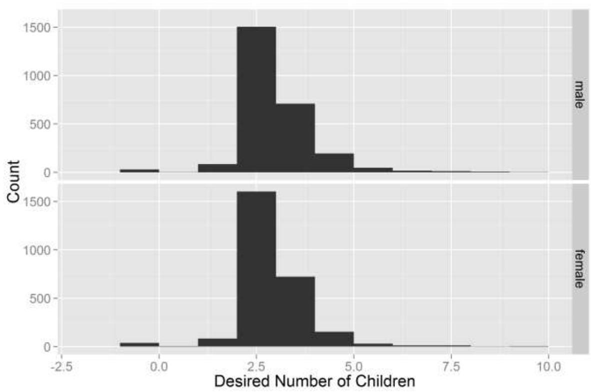


**Figure A1.3: Desired number of children, plotted for male and female respondents, based on data from the CVFS.**

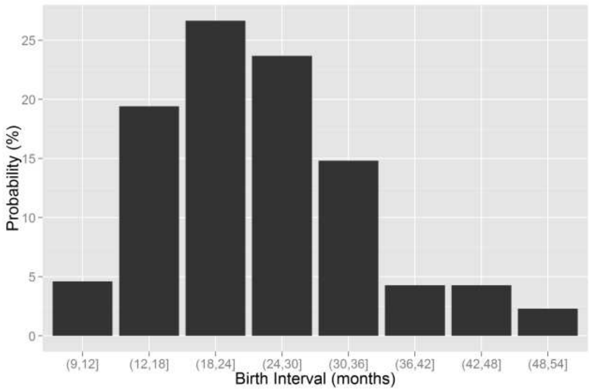


**Figure A1.4: Time between births, plotted from survey data from the CVFS.**

the model covariates for that month. A random number is drawn, and if it is less than the woman's probability of first birth, she will give birth in that month.

Following the first birth, additional births occur with birth intervals that are chosen from the empirically observed histogram of birth intervals (Figure A1.4). The maximum total number of births per woman is limited by the woman's desired family size. Once a woman is either too old to give birth (defined as age 45) or has reached her desired number of children, she is excluded from the fertility submodel.

*Mortality*

We model mortality as a simple random process, with age and sex dependent mortality rates. We calculate mortality rates using the first 48 months of the CVFS household registry data, which records all deaths that occur within the CVFS sample. Probability of mortality is calculated independently for males and females over identical bins (age in years, inclusive on the high end) of: 0-3, 3-6, 6-12, 12-20, 20-40, 40-60, 60-80, 80-90, 90 and above (Figure A1.5). When an individual dies, they are removed from the model, and their spouse (if any) may marry again, as they are returned to the pool of individuals eligible for marriage on the next time step.

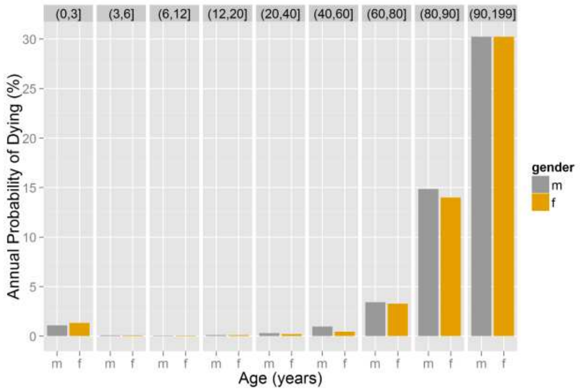


**Figure A1.5: Mortality model from the Chitwan ABM. The annual probability of dying is empirically derived based on monthly data from the CVFS, and calculated independently for males and females in a number of different age groups.**

*Migration*

We model individual and household-level out-migration in the Chitwan ABM. At the individual-level, a logistic regression model is used to calculate the probability of out-migration for any individual above the minimum age of migration (default of 15 years). The regression model (based on Massey, Axinn and Ghimire, 2010) takes into account a series of person-level, household-level, and community-level covariates to model the probability of an individual out-migration from the Chitwan Valley (Table A1.4) in a given month. We define individual migration following the "long distance migration" definition used by Massey, Axinn and Ghimire (2010) – as an individual "departure from the Chitwan Valley lasting at least a month".

Once a probability of migration is calculated for an individual, we draw a random number, and if that number is less than the calculated probability for that individual, then that individual out-migrates starting in that time step. The vast majority of migrants from Chitwan return within a relatively short period (as seen in Figure A1.6). To model this process, we include

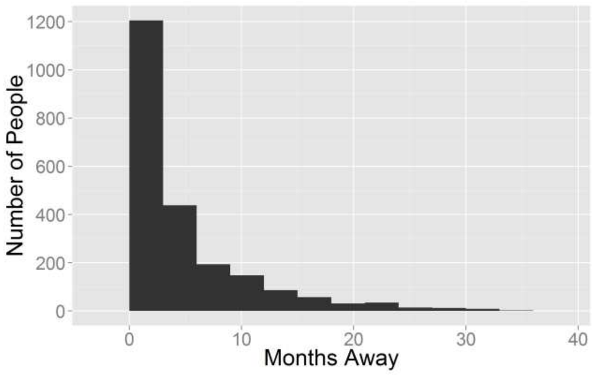


**Figure A1.6: Empirically derived histogram of individual out-migration times (time from out-migration until return to Chitwan) in the Chitwan ABM.**

a second component in the migration model – a model of migration length. Once an individual becomes a migrant, an outmigration length is calculated. First, a random number is drawn to see if an out-migration is permanent. If the random number is less than the *permanent out-migration probability*, then that agent will be removed from the model, and the household will be tagged as having a permanent out-migrant member. If the person does not become a permanent out-migrant based on this calculation, a second random number is drawn, and is used to draw a random out-migration length from the empirically observed distribution of migration times (Figure A1.6). The out migrant agent is then added to a special group of person agents that, while not present in Chitwan, are still part of the model (their age increments every time step and they are still subject to the mortality submodel). If the person survives until a number of time steps equal to their outmigration length has passed, the person is then returned to their original household. Each household tracks a list of temporary out-migrants who are not currently present in the household, but who will be returning after an amount of time equal to their outmigration duration has passed. This list is used to ensure that a household is retained in the model even if it becomes entirely empty for a time due to its members making temporary migrations.

Household level in- and out-migration is also modeled, whereby we allow an entire household to in- or out-migrate with probability specified by a fixed empirically derived parameter. Future work will explore variability in the demographic characteristics and environmental determinants of in- and out-migrating households, but for the present model we do not differentiate household-level in- or out-migration based on demographic or environmental state variables.

*Education*

Due to the importance of education as a predictor of demographic behavior, we model education endogenously in the Chitwan ABM. The education model is a two part process. First, using the results of an ordinal logistic regression (Table A1.5), a "final schooling level" (in years) is calculated for each person agent at age six. The schooling regression model takes into account gender, ethnicity, and access to non-family services (including schooling access) in each person's neighborhood. The final schooling level is the highest level of schooling that person may achieve in their lifetime (provided they live long enough to complete schooling, and do not permanently migrate outside of Chitwan (thus leaving the model). Also at age 6, the "schooling status" variable associated with each person is set to "in school", for each person whose calculated final schooling level is greater than zero. Every month, the model checks each person's schooling status. If the schooling status is "in school", and a person has not yet reached their final schooling level, a month is added to their schooling level, to simulate attending school during that month. Once a person reaches their final schooling level, their school status variable is set to "out of school", and the model will no longer increment their schooling level.

*Land-use and land-cover change*

When a new household is constructed in Chitwan, the land area it occupies is determined from a probability distribution of household land areas taken from the CVFS mapping data (Figure A1.7). To simulate land use conversion, an amount of land equal in area to the area of the new household is deducted from either the agricultural land use category (preferentially) or the non-agricultural vegetation land use category (if insufficient agricultural land exists) and is added to the private buildings category of the appropriate neighborhood. The household is preferentially built in the husband's parent's neighborhood. If there is not enough land available in that neighborhood for this to happen, the household will be located in the closest neighborhood to the husband's parents household that does have available land.

**CHANS characteristic features**

Coupled Human-Natural Systems (CHANS) have been noted to have characteristic features that can impact management policies: reciprocal effects and feedback loops, nonlinearity and thresholds, surprises, legacy effects and time lags, resilience, and heterogeneity (An et al. in press; Liu, Dietz, Carpenter, Alberti, et al. 2007). The primary feedbacks we focus on for this paper are between land use and the marriage timing and first birth timing submodels. As noted in

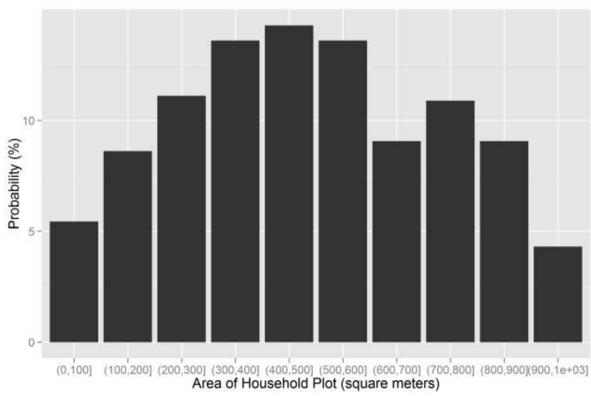


**Figure A1.7: Empirically derived histogram of new household plot sizes (area in sq. m).**

the paper, feedbacks between land use and marriage timing and land use and first birth timing can lead to small, though statistically significant, differences in model outcomes in the Chitwan ABM.

Time lags are also are common in CHANS, and are observed in the Chitwan ABM. A prominent example is the lagged effect of changing first birth timing on agricultural land conversion. As seen in Figure 3, 2020 land use does not differ significantly between the calculated, half, and double effect scenarios in scenario set C (where the size of the first birth timing feedback is varied). The effects of changing first birth timing are only visible in the 2050 model results, as it takes some time for the effects of changed first birth timing to appear on the landscape due to the lag between additional births leading to an increase in birth leading to a subsequent increase in new household establishment as the "extra" children born due to shortened birth timing reach marriage age.

In Chitwan land-use change is resilient to moderate changes in fertility or marriage rates that affect the population balance. The primary determining factor behind land-use change that is represented in the model is new household establishment. This process is largely a legacy effect of past demographics – due to past high fertility rates, there was a large youth bulge in Chitwan in 1996. As this population ages and begins to establish new households, the effect is land transition out of agriculture. While large sudden shocks (substantial in/out-migration following a crisis for example) could affect land-use change temporarily, it does not affect this general dynamic. Heterogeneity of model outcomes is also a feature if the Chitwan ABM. Ethnic variation in marriage rates and fertility preferences, and spatial variation in non-family service access all influence the patterns we observe in the outcomes of the Chitwan model.

**Verification and validation**

Validation of the results of agent-based models is often difficult (Brown et al. 2005). To verify our results, we follow the process described in An et al. 2005. First, we conduct a progressive debugging and testing of the model, to catch errors in the model code. The model includes simplified versions of the key submodels (a first birth timing model that is purely a histogram for example, for example). The results of model runs with these simplified submodels are compared to runs with the more complex submodels. Large differences between these model runs indicate errors may exist in the model code. Second, the model is verified using extreme value tests and extreme combination tests, to test that the model responds as expected to extreme values of model parameters. Third, we conduct a sensitivity analysis, testing the sensitivity of model outcomes to small changes in key model parameters. Lastly, while the availability of a time series of social data from the CVFS facilitates model parameterization, it also allows validation of model outcomes as we can compare simulated outcomes to observed data. A subset of the monthly CVFS data (data from the 1996-2002 surveys) was used to calibrate the model, while leaving data from later surveys (2002-2008) for validation (evaluation of the extent to which the model results reflect the observed data).

## REFERENCES

An, L., M. A. Linderman, J. Qi, A. Shortridge, and J. Liu. 2005. Exploring complexity in a human environment system: An agent-based spatial model for multidisciplinary and multiscale integration. *Annals of the Association of American Geographers* 95 (1):54–79.

An, L., A. Zvoleff, J. Liu, and W. G. Axinn. in press. Agent based modeling in coupled human and natural systems (CHANS): Lessons from a comparative analysis. *Annals of the Association of American Geographers*.

Axinn, W. G., and D. J. Ghimire. 2007. *Social organization, population, and land use*. University of Michigan: Population Studies Center.

Axinn, W. G., A. Thornton, J. S. Barber, S. A. Murphy, D. J. Ghimire, T. Fricke, S. Matthews, D. R. Dangol, L. D. Pearce, A. Biddlecom, S. Shrehtha, and D. Massey. 2011. *Chitwan Valley Family Study*. University of Michigan, Population Studies Center and Survey Research Center.

Brown, D. G., S. Page, R. Riolo, M. Zellner, and W. Rand. 2005. Path dependence and the validation of agent-based spatial models of land use. *International Journal of Geographical Information Science* 19 (2):153–174.

Ghimire, D. J., and W. G. Axinn. 2010. Community context, land use, and first birth. *Rural Sociology* 75 (3):478–513.

Ghimire, D. J., and L. F. Hoelter. 2007. Land use and first birth timing in an agricultural setting. *Population & Environment* 28:289–320.

Grimm, V., U. Berger, F. Bastiansen, S. Eliassen, V. Ginot, J. Giske, J. Goss-Custard, T. Grand, S. K. Heinz, G. Huse, A. Huth, J. U. Jepsen, C. Jørgensen, W. M. Mooij, B. Müller, G. Pe'er, C. Piou, S. F. Railsback, A. M. Robbins, M. M. Robbins, E. Rossmanith, N. Rüger, E. Strand, S. Souissi, R. A. Stillman, R. Vabø, U. Visser, and D. L. DeAngelis. 2006. A standard protocol for describing individual-based and agent-based models. *Ecological Modelling* 198 (1–2):115–126.

Grimm, V., U. Berger, D. L. DeAngelis, J. G. Polhill, J. Giske, and S. F. Railsback. 2010. The ODD protocol: A review and first update. *Ecological Modelling* 221 (23):2760–2768.

Liu, J., T. Dietz, S. R. Carpenter, M. Alberti, C. Folke, E. F. Moran, A. N. Pell, P. Deadman, T. Kratz, J. Lubchenco, E. Ostrom, Z. Ouyang, W. Provencher, C. L. Redman, S. H. Schneider, and W. W. Taylor. 2007. Complexity of coupled human and natural systems. *Science* 317 (5844):1513–1516.

Massey, D., W. G. Axinn, and D. J. Ghimire. 2010. Environmental change and out-migration: Evidence from Nepal. *Population and Environment*.

Yabiku, S. T. 2006. Land use and marriage timing in Nepal. *Population & Environment* 27 (5):445–461.

# TABLES

| Attribute | Possible Values | Initial Value |
|---|---|---|
| Person ID | 001001001-151999999 | Unique value assigned at birth |
| Age (months) | 0 – age at death | 0 |
| Gender | Male, female | Random, either male or female |
| Ethnicity | High-caste Hindu, Low-caste Hindu, Newar, Hill Tibeto-Burmese, Terai Tibeto-Burmese | Father's ethnicity |
| Mother ID | 001001001-151999999 | Derived from mother |
| Father ID | 001001001-151999999 | Derived from father |
| ID of household at birth | 001001-151999 | Derived from mother |
| ID of neighborhood at birth | 001-151 | Derived from mother |
| Date of last migration | 1/1995 – 12/2049 | None† |
| Date of last birth (women) | 1/1995 – 12/2049 | None† |
| Spouse | None, reference to spouse 'person' instance | None† |
| Marriage time | | |
| Child list | List of children 'person' instances | None† |
| Ever work for pay | True, False | Set to 'yes' with probability .205 for females and .450 for males‡ |
| ***Education:*** | | |
| Schooling status | Undetermined, in school, out of school | Undetermined |
| Final schooling level | 0 – 15 | None† |
| Years of schooling | 0 – 15 | 0 |
| ***Mother's characteristics:*** | | |
| Years of schooling | 0 – 15 | Derived from mother |
| Ever work for pay | Yes, No | Derived from mother |
| Number of children | 0 – 10 | Derived from mother |
| ***Father's characteristics:*** | | |
| Years of schooling | 0 – 15 | Derived from father |
| Ever work for pay | Yes, No | Derived from father |
| **Attributes assigned only to female person agents:** | | |
| Desired number of children | 0 – 10 | Taken from CVFS data |
| Time of last birth | 1/1997 – 12/2049 | Reset after each birth |

**Table A1.1: Attributes of the 'person' agent type.**

Note that the initial values listed here and in Table A1.2 are the values given to agents that are endogenously generated during a model run (for people, through birth or in-migration). The initial values of attributes for person agents present in 1997 at the beginning of the model are taken directly from the survey data. †Parameters with initial values of 'None' are set equal to the 'None' value in the model code, and conditional statements are used in the model to handle cases where attributes are set to 'None' (for example, for the date of last migration, a value of 'None' would mean that no prior migration has occurred)., and for the education model, a final schooling level of 'None' would indicate the education submodel has not yet been run for that agent. ‡From CVFS data

| Attribute | Possible Values | Initial Value |
|---|---|---|
| Household ID | 001001-151999 | Unique value assigned at household establishment |
| Own household plot | Yes, No | Set to 'yes' with probability .93[†] |
| Own land | Yes, No | Set to 'yes' with probability .83[†] |
| Rent out land | Yes, No | Set to 'yes' with probability .61[‡] |
| Use any non-wood fuel | Yes, No | Set to 'yes' with probability .11[‡] |
| Time of last migration | 1/1995 – 1/2050 | None |

**Table A1.2: Attributes of the 'household' agent type.**

[†]From CVFS data [‡]Axinn and Ghimire (2007)

| Attribute | Possible Values | Initial Value |
|---|---|---|
| Neighborhood ID | 001-151 | Taken from CVFS data |
| x, y coordinates | UTM45N coordinates of neighborhood | Taken from CVFS data |
| Elevation (m) | 143 – 207 | Taken from 90 m SRTM DEM |
| Electricity available | Yes, No | Taken from CVFS data |
| Land area: agriculture (sq. m) | 0 – 280,915 | Taken from CVFS data |
| Land area: non-agriculture (sq. m) | 0 – 36,130 | Taken from CVFS data |
| Land area: private buildings (sq. m) | 253 – 20,386 | Taken from CVFS data |
| Land area: public buildings (sq. m) | 100 – 80,448 | Taken from CVFS data |
| Land area: other (sq. m) | 0 – 14,484 | Taken from CVFS data |
| Distance to urban center (km) | 0 – 28 | Calculated from coordinates |
| Distance to nearest school (minutes on foot) | 0 – 30 | Taken from CVFS data |
| Distance to nearest health post (minutes on foot) | 0 – 90 | Taken from CVFS data |
| Distance to nearest bus stop (minutes on foot) | 0 – 75 | Taken from CVFS data |
| Distance to nearest market (minutes on foot) | 0 – 120 | Taken from CVFS data |
| Distance to nearest employer (minutes on foot) | 0 – 480 | Taken from CVFS data |

**Table A1.3: Attributes of the 'neighborhood' agent type.**

| | Odds Ratio | p | Sig. |
|---|---|---|---|
| Enrolled in school | 0.82 | 0.062 | . |
| Years of schooling | 1.075 | <.001 | *** |
| Female | 0.624 | <.001 | *** |
| **Physical capital** | | | |
| Market access | 1.026 | 0.516 | |
| Farmland | 0.93 | 0.528 | |
| **Age** | | | |
| 15-25 | 2.797 | <.001 | *** |
| 25-35 | 1.544 | 0.021 | * |
| 35-45 | 0.94 | 0.751 | |
| 45-55 | 1.05 | 0.807 | |
| **Ethnicity**† | | | |
| Low-Caste Hindu | 1.109 | 0.477 | |
| Hill Tibeto-Burmese | 1.222 | 0.083 | . |
| Newar | 0.854 | 0.339 | |
| Terai Tibeto-Burmese | 0.617 | <.001 | *** |
| **Duration** | | | |
| Month‡ | -0.565 | <.001 | *** |
| Month squared‡ | 0.116 | 0.05 | . |
| **Intercept** | 0.007 | <.001 | *** |

**Table A1.4: Model of local-distant migration (following the work of Massey et al. 2010).**

Odds-ratios are from a discrete time event history analysis predicting monthly hazard of making a move outside of the Chitwan Valley. Two sided P-values. Significance is coded as: · for P < 0.1, * for P < .05, ** for P < 0.01, *** for P < 0.001. †The odds-ratios for ethnicity are relative to Upper-caste Hindus. ‡For month and month-squared coefficients are listed rather than odds ratios as for these two covariates odds ratios are not directly interpretable (e.g. month cannot be held constant without varying month-squared).

| | Odds Ratio | p | Sig. |
|---|---|---|---|
| **Neighborhood level** | | | |
| Avg. years non-family services (15 min. ft.) | 1.1 | < 0.001 | *** |
| **Individual level** | | | |
| Female | 0.1 | < 0.001 | *** |
| **Ethnicity** | | | |
| Low-Caste Hindu | 0.1 | < 0.001 | *** |
| Hill Tibeto-Burmese | 0.2 | < 0.001 | *** |
| Newar | 1.0 | 1 | |
| Terai Tibeto-Burmese | 0.1 | < 0.001 | *** |
| **Intercepts** | | | |
| Years schooling greater than 0, less than 4 | 8.0 | < 0.001 | *** |
| Years schooling greater than 4, less than 8 | 3.9 | < 0.001 | *** |
| Years schooling greater than 8, less than 11 | 1.00 | 0.900 | |
| Years schooling greater than 11 | 0.2 | < 0.001 | *** |
| Avg. years non-family services (15 min. ft.) | 1.1 | < 0.001 | *** |

**Table A1.5: Education model used in the Chitwan ABM.**

The model is derived from analysis of a sample of CVFS data including all individuals aged 25-30 in 1996. Odds-ratios are from an ordinal logistic regression predicting education level as 0, 0-4, 4-8, 4-11 or >11 years. P-values are two-sided. Significance is coded as: · for $P < 0.1$, * for $P < .05$, ** for $P < 0.01$, *** for $P < 0.001$. n=715, pseudo $R^2$=.435. †The odds-ratios for ethnicity are relative to Upper-caste Hindus.