Supplementary A

GLOBE cases

1. Magliocca, N.R. “Bouami”. GLOBE Georeferenced Case Study. Published June 26, 2015; doi:10.7933/K13N21B6.
2. Magliocca, N.R. “Muongmuay”. GLOBE Georeferenced Case Study. Published June 26, 2015; doi:10.7933/K1C24TCJ.
3. Magliocca, N.R. “Phadeng”. GLOBE Georeferenced Case Study. Published June 26, 2015; doi:10.7933/K17D2S2J.
4. Magliocca, N.R. “Sao Pedro”. GLOBE Georeferenced Case Study. Published June 26, 2015; doi:10.7933/K1GT5K3Z.
5. Magliocca, N.R. “Mashete”. GLOBE Georeferenced Case Study. Published June 26, 2015; doi:10.7933/K1V40S4H.
6. Magliocca, N.R. “Ulumi”. GLOBE Georeferenced Case Study. Published June 26, 2015; doi:10.7933/K1ZW1HVH.

GLOBE Analyses

1. Magliocca, N.R. “Similarity Analysis of Bouami, Laos test case”. GLOBE Similarity Analysis. Published June 26, 2015; doi:10.7933/K1FT8HZ5.
2. Magliocca, N.R. Similarity Analysis of Moungmuay, Laos test case. *GLOBE Similarity Analysis*. Published May 29, 2015; doi:10.7933/K1W37T8B.
3. Magliocca, N.R. Similarity Analysis of Phadeng, Laos test case. *GLOBE Similarity Analysis*. Published May 29, 2015; doi:10.7933/K1RB72J0.
4. Magliocca, N.R. Similarity Analysis of Sao Pedro, Brazil test case. *GLOBE Similarity Analysis*. Published May 29, 2015; doi:10.7933/K1MK69T8.
5. Magliocca, N.R. Similarity Analysis of Mashete, Tanzania test case. *GLOBE Similarity Analysis*. Published June 26, 2015; doi:10.7933/K1B56GNG.
6. Magliocca, N.R. Similarity Analysis of Ulumi, Tanzania test case. *GLOBE Similarity Analysis*. Published June 26, 2015; doi:10.7933/K16D5QXR.

Supplementary B

B1 ODD Protocol Model Description

B1.1. Purpose

This is a generalized agent-based model (ABM) of land-use and livelihood decision-making developed for the purpose of providing mechanistic explanations of heterogeneous household responses to changing local and exogenous economic, environmental, and demographic conditions. The overall objective of this modeling framework is to support the application of a broadly applicable ABM architecture as a standardized experimental and observational tool for cross-site comparison, referred to as the agent-based synthesis system (ABSS) approach. The generalized ABM can be parameterized to a given location using fine-grained (e.g., case study) as well as global datasets. By using the same modeling framework for each site, similarities across and local contingencies within land-use and livelihood decision-making in response to changing global market forces can be compared systematically and with a high degree of statistical rigor. Differences in modeled responses to experimental manipulations across the test cases will provide insights into the relative importance of particular factors and processes in driving land use and livelihood outcomes in each context.

B1.2. Entities, state variables, and scales

B1.2.1. Agents

Each agent represents a household, and the number of households represented in the model depends on values taken from a global population density dataset.

Table S1. Combined labor and input costs.

Attribute	Brief description
Number of households	The number of household agents is the sum of all population density values across the model landscape divided by the total area of the model landscape and household size.
Age structure	The population is divided evenly between children and adults. Age structure is held constant and determines a settlement agent's labor supply and food and income demands.
Household size	All households are composed of two adults and two children, and aggregated to the settlement level according to population density.
Stocks	Initial food and income stocks are allocated based on minimum subsistence requirements. These are dynamically updated based on agents' land-use and livelihood decisions.
Subsistence requirements	Minimum subsistence requirements (860 kg yr^{-1} $person^{-1}$) consist of grain for household food consumption and livestock feed [1]. Minimum monetary income requirements equal annual farm input costs plus the cost of a year's worth of food should crops fail. An agent's minimum subsistence and income requirements equal those of an individual household multiplied by the number of households in the population. Children require half of the subsistence needs of an adult [2].
Labor supply	Total available labor (96 person-weeks) is calculated by multiplying a year's worth of labor net of required "home" time (15, e.g. leisure, home maintenance, home textiles, *etc.*) by household size [3].
Risk preferences	A parameter ranging from 0 to 1, heterogeneous across agents, that weighs the potential pay-off of an activity against the certainty equivalent pay-off from a risk-neutral [4,5]. Agents are assigned heterogeneous risk preferences drawn randomly from a normal distribution with mean of 0.5.
Land-use preferences	Agents are assigned uniform preferences across land-uses. However, future model versions can differentiate land-use preferences according to agricultural suitability, cultural preferences, or the relative contribution of each land use to the agent's income, for example.
Initial household subsistence stock	Initial food stocks are assumed to cover a year's subsistence requirements (2580 kg, grain equivalents).
Initial household money stock	Combined farm input costs and the cost of a year's subsistence needs at the long-term average crop price with a market influence of 0.5 (1426 US$).
Subjective aspiration levels	The wage rate of the livelihood activity forgone, which must be met or exceeded by the chosen livelihood activity (*i.e.*, opportunity cost).

B1.2.2. Spatial Units

Stylized landscapes of 100 by 100 irregular grids of cells are generated, with each cell representing one hectare. Each cell has a number of biophysical attributes.

Table S2. Accounting of structural elements, endogenous processes, and agent characteristics currently represented in the model, as well as those that can be introduced in future work.

Attribute	Brief description
Topography	Percent slope is derived from a Digital Elevation Model (DEM). Slope is a proxy for soil suitability for agriculture [6].
Precipitation constraints	Precipitation constraints are varied uniformly across the landscape as a proxy for number of growing days [7].
Agricultural suitability	Slope and precipitation constraints impose zero to 100 percent reductions in agricultural yield according to agricultural suitability classes [6,7].

B1.2.3. Land Uses

Land-uses are modeled as functional groups, rather than specific crops, to maintain generality across land-use systems. Land-use/cover categories include five productive uses, intensive and extensive cultivation, pasture, multi-cropping, and cash cropping, which vary in their potential productivity, degradation/regeneration rates, and labor and input costs. Agricultural yields are taken from Monfreda *et al.* [8] and averaged across the specific crops and cultivation intensities as specified by the case study. Non-productive uses include forest/fallow and non-use areas (e.g., water bodies). Only the most intensive cultivation system reported in each test case (e.g., intensive upland crops or short fallow shifting cultivation) is used for model evaluation and compared to global land cover data. This is due to uncertainties in the percent cropland category in global data related to limitations of remotely sensing and reliably classifying extensive cultivation (e.g., long fallow shifting cultivation or non-timber forest products). The model landscape is initialized with forest in areas classified as unsuitable for agriculture due to slope, the lowest labor input agricultural use (*i.e.*, extensive cultivation) in the highest quality cells, and fallow for remaining cells.

Table S3. Parameter descriptions for settlement agents that apply equally to all households within a village.

Attribute	Brief description
Potential yield	Crop or livestock yields per hectare (Table S1).
Degradation rate	Yields decline through continual use at varying rates depending on the type of land use (intensive agriculture, 0% annually; extensive agriculture, 25% annually; pasture, 18% annually, [9–12]).
Regeneration rate	Yields recover during fallow periods at different rates depending on the type of land use (intensive agriculture, 0% annually; extensive agriculture, 4% annually; pasture, 50% annually, [9–12]).
Labor costs	Labor costs varying according to the intensity of land-use, and are expressed in person-weeks per hectare (Table S2).

B1.2.4. Environment

Table S4. Model inputs and observed variables[1] used to test, understand, and analysis model output. Pearson's correlations (R) between variables are whited-out if less than 0.5, and color-coded based on their level of emergence[2].

Attribute	Brief description
Population density	Sampled from global population density dataset and expressed as people/km^2 (see Table S2).
Market influence/access index	The global/regional market setting of the focal landscape. Market influence determines relative farm-gate crop prices, farm input costs, non-farm wage rates, transportation costs to market, and non-farm employment transaction costs (see Table S2, 3, and Section B3.4).

A set of cost functions are hypothesized that link global market influence index values to local farm-gate and food prices, farm input costs, and non-farm wages and transaction costs. Global commodity prices and U.S. minimum wage represent agricultural commodity prices and non-farm wages realized by a farmer in locations with a market influence index at or near 1. Local product and factor prices and costs in locations with market influence less than 1 vary according to the cost functions below.

Farm-gate prices ($P_{j,t}$, Table S3) for agricultural products produced by land-use j at time t are a function of mean agricultural commodity price (P_0) [13,14], market influence index value (*MII*), and the crop price factor (β_{pcrop}).

$$P_{j,t} = P_0 * MII^{\beta_{pcrop}} \tag{S1}$$

The baseline non-farm wage rate (W_o) is determined by the relationship between a benchmark non-farm wage (*NFW*; U.S. minimum wage at a MII of 1) and the global MII for the location subject to the non-farm wage factor (β_{nfwage}).

$$W_o = NFW * MII^{\beta_{nfwage}} \tag{S2}$$

Farm input costs (C_{farm}) change proportionally with the baseline non-farm wage (W_0) and market access (*MA*) at a rate determined by the farm cost factor (β_{fcost}).

$$C_{farm} = \beta_{fcost} \frac{W_o}{1+MA} \tag{S3}$$

If intensive cultivation is performed for sale on the market, labor time costs (*LC*) are converted to a monetary value to represent both the costs of non-labor inputs (e.g. fertilizer) and the opportunity cost of forgone non-farm wage labor.

Transaction costs associated with locating, securing, and maintaining non-farm wage employment (C_{nfarm}) change proportionally with the baseline non-farm wage (W_0) and market access (*MA*).

$$C_{nfarm} = \beta_{nfcost}(1 - MA)W_o \tag{S4}$$

The effective non-farm labor wage rate is equal to the gross income net of transaction costs (Equation (S4)) per unit of labor time required to convert to or maintain in land use j from i ($LC_{i\to j}$) which produces the given agents' most profitable agricultural commodity.

$$W_{nfarm} = \left(W_o - C_{nfarm}\right)LC_{i\to j} \tag{S5}$$

B1.2.5. Spatial and Temporal Scales

One model time step represents one year, and the model is run over a twenty-year period (with the first ten as model spin-up). Total landscape area depends on the study site, and each grid cell represents 1 ha, for a total area of 100 km^2.

B1.3. Process Overview and Scheduling

The model uses a discrete event-sequencing framework (Figure S1) in which each agent makes land-use decisions over their entire cultivation area sequentially, but the states of landscape cells across agents' holdings are updated synchronously. The main processes in operation each time step are biophysical regeneration/degradation, agents' labor allocation, land-use selection and harvest, and yield and price expectation formation for next period.

Pseudo-Code

```
FOR 1 to the number of cells
        Calculate Yield based on potential yield subject to the time in use of the current land-
                use multiplied by the Regeneration/Degradation Rate.
END FOR
FOR 1 to number of agents
        FOR 1 to number of land uses
                Calculate the Marginal Expected Yield and Marginal Expected Return based
                        on Expected Yield and Expected Price net of input costs per land-use
                        specific labor costs.
        END FOR
        Calculate Farm Wage, Food Price, Shadow Price of Production, Aspiration
                Level, and Non-Farm Wage based on Market Influence Index,
                Market Access Index, and Risk Preferences.
        Based on the level Money and Food Stocks relative to subsistence and income
                requirements, Total Labor is allocated to between Home Labor, Farm
                Labor, and Non-Farm Labor.
        Based on subsistence and income requirements, allocate Farm Labor between
                Subsistence Labor and Market Labor.
        FOR 1 to number of land uses
                Using Expected Yield and Marginal Expected Return for each cell in
                        the agent's holdings, calculate per cell Expected Production.
                Calculate the Expected Utility of each land use in each cell based on
                        Expected Production. Expected Utility is based on a 'satisficing'
                        framework for subsistence production and a profit maximizing
                        framework for market production.
                Select land use in each cell with highest Expected Utility in subsistence
                        production first until subsistence requirements are met or all
                        Subsistence Labor is allocated, whichever comes first. Next,
                        select land use in each cell with the highest Expected Utility in
                        market production until income requirements are met or all
                        Market Labor is allocated, whichever comes first.
                Actual yields and prices realized.
                Money and Food Stocks updated based on agricultural production and
                        livelihood decisions.
                Expected Yields and Expected Prices formed for next time step using
                        prediction models.
        END FOR
        Landscape cells updated with new land uses.
END FOR
```

Figure S1. Process overview and scheduling presented as pseudo-code.

B1.4. Design concepts

B1.4.1. Basic Principles

A central theory explaining the dynamics of agricultural land use by smallholders is "induced intensification" [15], which relates changes in farming systems to the behavioral responses of smallholders to dynamic demographic, economic, and technological factors. Early descriptions of agricultural intensification by Boserup [16] and Chayanov [17] described a process through which smallholders were forced to increase the labor-intensity of cultivation through techno-managerial innovations to meet increasing production demands from rising population density. A wide range of disciplines expanded on these insights to consider the roles of environmental suitability [18] and commercial agricultural activities [19–21] in driving agricultural intensification, which became more broadly labeled as "induced intensification" theory [22].

This model attempts to enrich induced intensification theory by explicitly linking global market influence to land-use and livelihood decisions. The role of local economic conditions, especially non-farm wage opportunities in relation to land-use choices, has yet to be systematically linked to land-use intensity across locations globally. Applicable theory, however, has developed from the accumulation of case-study knowledge. De Janvry and colleagues [19] offered a generic explanation for variations in market participation across sites relating to local farm-gate prices, internal costs of production, and food prices. The relationship between internal costs of production and farm-gate prices, which are dependent on local market influence (*i.e.*, both physical access to markets and purchasing power), determine the value of agricultural products (*i.e.*, shadow price) for a given household. The shadow price of agricultural products, relative to the costs of purchasing food on the market, structure the consumption and production decisions of households, and consequently their degree of market participation. Additionally, access to non-farm wage opportunities influences the intensity of land-use, as non-land-based income sources can supplement or fulfill food and income requirements. Combined, these theoretical strands provide a potential framework for representing the land-use consequences of household decision-making in response to local economic conditions.

The model is designed to reproduce the observed patterns of land-use in response to demographic, economic, and agro-ecological conditions. Agent-level behavioral rules are based on the theoretical frameworks of Boserup [16] and Chayanov [17], which provide a smallholder household rationale for cultivation choices in response to population pressure and labor and land constraints. However, Boserup and Chayanov stop at describing intensification of subsistence agriculture, and more recent literature describes the importance of further transitions within rural agriculture to market-based production. The "livelihoods" perspective within the field of development economics provides a means for extending existing intensification theories by considering the role of market opportunities in agricultural production choices. The model's design incorporates livelihood diversification concepts [23–25] to explicitly represent non-farm wage opportunities and factors influencing agricultural production for the market. Integration of these household-level theoretical frameworks informs agents' behavioral rules for livelihood diversification, labor allocation, agricultural production mode, and land-use choices, and through the interaction of many agents with their environment, attempts to reproduce the system-level agricultural dynamics described by induced intensification theory from the bottom-up.

B1.4.2. Emergence

This model is designed to explore the decision-making processes of agents in response to varying demographic, economic, and environmental conditions and the land-use patterns that result. In addition, the livelihood choices of agents are analyzed with respect to the diversity of labor allocation between on- and off-farm and subsistence- *versus* market-oriented livelihood activities. Labor allocation arises from the decisions of individual agents based on their expectations of pay-offs from each livelihood activity, individual risk tolerances, and larger-scale demographic, economic, and environmental conditions. Although larger-scale factors influencing livelihood decisions are specified exogenously and held constant throughout a given simulation, agents learn to predict and adapt to dynamic local conditions. Livelihood choices are subject to some path-dependence and individual agents' learning abilities. Therefore, agents' final labor allocations and system-level land-use outcomes cannot be predicted from the model's initial conditions.

B1.4.3. Adaptation

Agents make livelihood and land-use decisions each period based on the success of past decisions and their expectations for pay-offs in the current period. Agents select the best livelihood activities and land-uses according their expected utilities, and can adapt to declining or improving yields from land-based activities resulting from past cultivation choices. Agents can also adjust their subjective wealth aspirations as they learn possible returns from various livelihood activities, some of which can produce economics of scale. Extensions of the current model could include additional sources of adaption. For example, agents could adapt their preferences for particular land-uses based on the proportion of revenue each produces.

B1.4.4. Objectives

Agents attempt to maximize expected utility in their livelihood and land-use decisions. Agents allocate labor to on- and non-farm activities proportionally to the ratio of expected wage rates. Land-uses choices are made cell-by-cell based on the highest expected utility among possible land-uses. Subsistence-oriented land-uses take precedent over market-oriented land-uses. Expected utility for subsistence land-uses is calculated as the marginal return per unit labor, and the best land-use is selected using a satisficing framework [26,27]. Expected utility for market-oriented land-uses is calculated as the marginal return per unit labor net of production costs, and the best land-use is selected using a profit maximization framework.

B1.4.5. Learning

Agents have a set of prediction models for forming expectations of future yields and crop prices that they update each period as new information becomes available (see Section B1.4.6 below for description of the prediction models). The performance (*i.e.*, error) of each model is tracked every period, and the agent acts on the prediction of the currently most successful model (*i.e.*, the "active" model). In the next period, actual yields and prices are realized and model performances are updated. Agents are therefore able to learn which models best predict yield and price trends, and can adaptively switch to following the predictions of a previously "dormant" model if it out-performs the current "active" model when conditions change.

B1.4.6. Prediction

Agents form expectations of agricultural yields and prices by detecting trends in past observations, which are extrapolated one period into the future to form expectations. Agents use a set of "backward-looking" expectation models that have been adapted from their original use in financial agent-based markets [28,29] to consider non-monetary and spatially explicit information. Each agent is randomly given a set of twenty prediction models that vary in the prediction method and time span over which past observations are considered. Each prediction model may use one of six different prediction methods that map past and present crop prices (P) and yields (given by substituting Y for P) into the next period using various extrapolation methods:

(1) *Mean model*: predicts that *P(t+1)* will be the mean price of the last *x* periods.

$$P(t+1) = \frac{\sum_{i=t-x:t} P(t_i)}{x} \tag{S6}$$

(2) *Cycle model*: predicts that *P(t+1)* will be the same as *x* periods ago (cycle predictor).

$$P(t+1) = P(t-x) \tag{S7}$$

(3) *Projection model*: predicts that *P(t+1)* will be the least-squares, non-linear trend over the last *x* periods.

$$P(t+1) = aP(t_s)^2 + bP(t_s) + c \tag{S8}$$

where t_s is the time span of *t-x* to *t*, and *a*, *b*, and *c* are coefficients of fit.

(4) *Mirror model*: predicts that *P(t+1)* will be a given fraction ξ of the difference in this period's price, *P(t)*, from price *t-x* periods ago, *P(t-x)*, from the mirror image around half of *P(t)*.

$$P(t+1) = 0.5P(t) + [0.5P(t) - (1-\xi)(P(t) - P(t-x))] \tag{S9}$$

(5) *Re-scale model*: predicts that *P(t+1)* will be a given factor ζ of this *x* period's price bounded by {0,2}.

$$(t+1) = \zeta P(t-x) \tag{S10}$$

(6) *Regional model*: predicts that *P(t+1)* is influenced by regional price information coming from neighboring agents.

The performance (*i.e.*, error) of each model is tracked every period, and the agent acts on the prediction of the currently most successful model (*i.e.*, the "active" model). In the next period, actual yields and prices are realized and model performances are updated. Agents are therefore able to learn which models best predict yield and price trends, and can adaptively switch to following the predictions of a previously "dormant" model if it out-performs the current "active" model when conditions change.

B1.4.7. Sensing

Agents are assumed to know the suitability and potential yields of all possible land-uses on all cells within their cultivation area. Actual yields and prices are known only after agents engage in a particular land-use or livelihood activity. Agents keep a record of past yields and prices for all of their cultivated cells and livelihood activities, which is used to update their prediction models.

B1.4.8. Interaction

Agents interact directly with the landscape through the selection of a cultivation method and corresponding land use. Interaction between agents occurs during the land allocation process during model initialization, because the use of land by one agent excludes its use by another. During model execution, however, agents do not interact. Future extensions of the model could include spatial interactions through land tenure rules and/or land markets, as well as the exchange of information and cultural norms through social networks.

B1.4.9. Stochasticity

Prediction methods and time horizons and land holdings are randomly assigned among each agents' set of prediction models. No other sources of stochasticity currently exist. Extensions of the current model can explore the effects of stochasticity in crop prices and/or yields on agents' livelihood strategies and land-use choices.

B1.4.10. Collectives

Agents are themselves an aggregate representation of a number of individual households, which reasonably approximates the household context.

B1.4.11. Observation

The data used for testing, understanding, and analyzing system-level model outcomes include modeled land-use patterns and dynamics, household labor allocation among livelihood options, and household production and consumption levels. This data is compared to target patters used for the evaluation phase of the model (Section 3.4. and Table 4), and to measure effect sizes of experimental treatments in the form of response ratios.

B1.5. Initialization

If spatially explicit site maps are provided in their source publications, a minimum of ten control point links are used to trace the site boundaries and align with the map with the regional WGS 1984 UTM projection. If spatially explicit site maps are not provided (e.g., just latitude and longitude coordinates), the GLOBE Land Unit in which the georeferenced point is contained is used as the model landscape. All site geometries were then rasterized to a resolution of 100 m and represented in the model as irregular grids.

The number of household agents is then determined by the sum of all population density values over the model landscape—taken from a global population density dataset—divided by the total area of the model landscape (km^2) and the average household size. Land is then allocated to each household agent based on a simple random seeding and area-growing algorithm. Agents are located in the landscape in random order with the probability of any location being chosen based on population density and suitability of the land for agriculture (*i.e.*, high population density locations on the most suitable land has the highest probability).

Assignment of specific cells to agents occurs in two phases. The size of land holdings per agent is determined by a random draw from a random exponential distribution with a mean specified through a genetic algorithm used with the pattern-oriented modeling (POM) approach (see Section 3.4 and

Table 3). The mean size of land holdings takes on values at or above the minimum of land required to meet subsistence needs (in hectares). In the first phase of land allocation, agents are chosen in random order and assigned all adjacent cells with the highest available suitability for agriculture (*i.e.*, "best land"). The assignment of cells to each agent in this phase stops when either the assigned size of land holdings is met or all available, contiguous "best land" cells are taken by other agents. In the second phase, agents are again selected in random order and select from the remaining cells a cell at a time until all cells that allow cultivation are assigned. This simple algorithm provides a generic land allocation scheme in which agents manage land units relatively close to their dwellings, agents can choose to expand to or abandon marginal or excess land holdings, and land holdings are subject to land suitability constraints and competition from other agents.

The landscape is initialized with the least labor-intensive land use on highest quality land, forest for all other usable land, and built or unusable in locations corresponding to those uses in the real landscape. Each agent is randomly assigned a set of twenty prediction models that vary in the prediction method and time span over which past observations are considered (see Section B1.4.6). Risk preferences are assigned randomly from a normal distribution with mean of 0.5 ranging from one to zero. Initial food and income stocks are set to minimum subsistence levels for all agents.

B1.6. Input Data

Agents responded to constraints imposed and opportunities afforded by population density, environment, and market forces, which were represented by model relationships based on generalized empirical data for agricultural productivity, labor and transaction costs, agro-ecological dynamics, and a global index of market influence (Table 1). Input data used to parameterize agricultural productivities and biophysical processes or degradation and regeneration are described in Section B1.2.3. Labor costs for specific land-uses are adapted from case studies of land change and presented in Table S1. Local farm-gate and food prices, farm input costs, and non-farm wages and transaction costs in relation to the global market influence index are specified according to the procedure described in Magliocca and Ellis [30].

B1.7. Submodels

The main submodels include biophysical processes, yield and price expectation formation, expected utility calculation, labor allocation, and land-use selection. Yields from each land use are calculated for every landscape cell dependent on the time in the current land use and land-use-specific regeneration/degradation rates. Expectation models are described in Section B1.4.6.

Agents derive utility from subsistence and monetary income. In this generalized context, income is defined as cash and food contributions to the welfare of the village derived from the set of livelihood activities in which village members are engaged. Utility from subsistence production follows a "satisficing" framework and is derived as the marginal return from labor. In contrast, utility from market production follows a profit-maximizing framework and is calculated as marginal production net input costs.

For subsistence production, expected marginal utility is given by:

$$EU(a,j) = \beta_{a,j} \frac{EY_{j,t}}{LC_{i\to j}} \tag{S11}$$

where the expected marginal utility from subsistence production of agent a for land-use j is the product of the a's preference, β, for land-use j and the marginal return of expected yield, EY, at time t subject to labor costs, LC, of converting from land-use i to j. For market production, expected marginal utility is given by:

$$EU(a,j) = \beta_{a,j}\left(EP_{j,t}EY_{j,t} - w_f LC_{i\rightarrow j}\right) \tag{S12}$$

where the expected marginal utility from market production is additionally a function of the expected price, EP, for production from land-use i, and the farm labor wage rate w_f.

Labor allocation (Figure S2), expected utility calculation, and land-use selection are described as part of the model algorithm below.

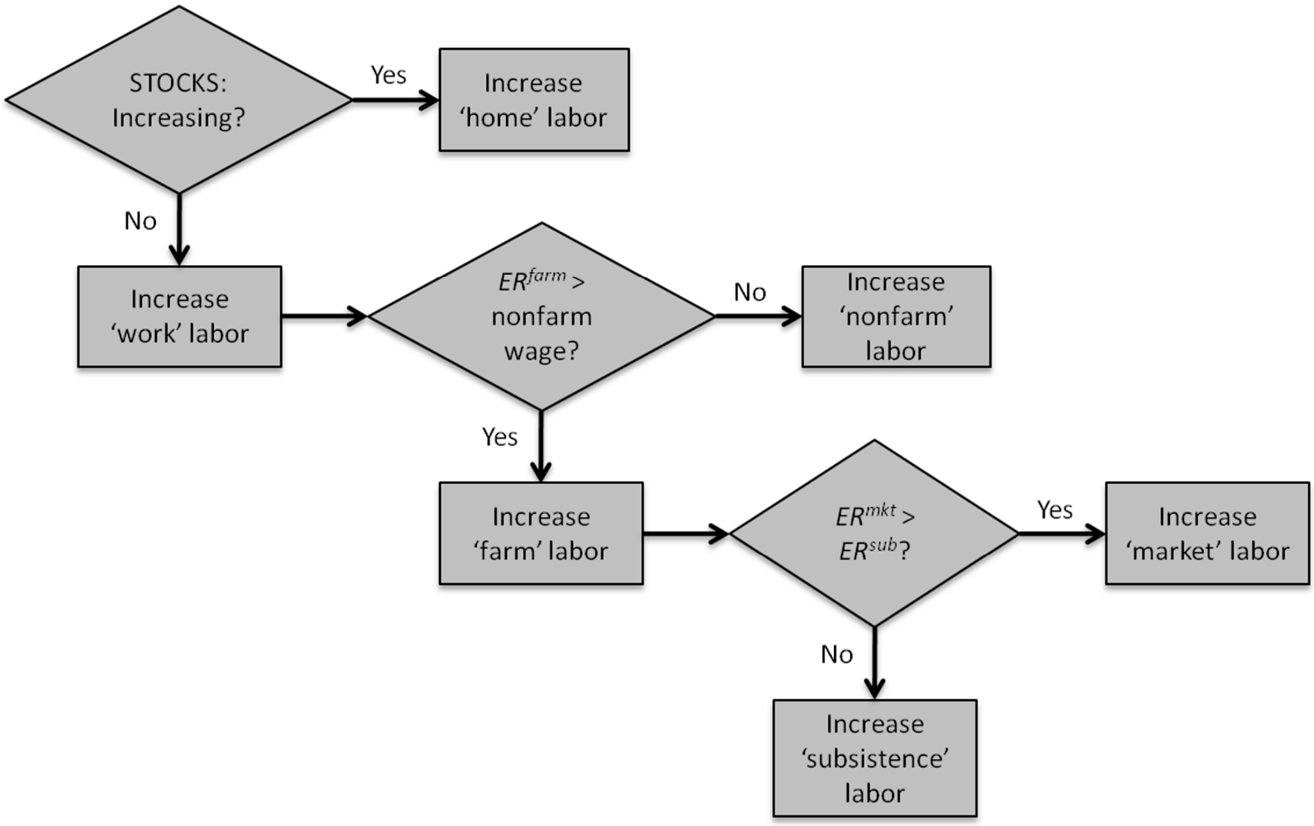

Figure S2. Heuristic decision tree of agents' labor allocation process.

The following algorithm describes the sequence of events for one simulated time period (year). The model is programmed in MATLAB. A decision tree of the labor allocation process is provided in Figure S2.

1) Agents determine the minimum amount of labor needed, L^0_{sub}, to meet minimum subsistence needs, $\delta_{\min}$, with the long-term average yield, Y_j^*, of the most productive land-use, j, on their land:

$$L^0_{sub} = \frac{\delta_{\min}}{Y_j^*} \tag{S13}$$

2) Each agent calculates their risk-neutral expected returns, $ER_{rn|a}$, of conservative activities (farm work and subsistence production) based on the discounted average observed yield and agricultural commodity prices ($P_{i,t}$)of their most productive land use, farm wage (W_{farm}), labor costs for maintaining land-use j.

$$ER_{rn|a}^{farm} = 0.5\left(\frac{P_{j,t}Y_j^*}{W_{farm}LC_{i\to j}}\right) \tag{S14}$$

$$ER_{rn|a}^{sub} = 0.5\left(P_{j,t}Y_j^* - W_{farm}LC_{i\to j}\right) \tag{S15}$$

3) Each agent calculates their risk-averse expected returns, $ER_{ra|a}$, of conservative activities (non-farm work and market production) based on the expected yield ($EY_{j,t}$) and agricultural commodity prices ($EP_{i,t}$) of their most productive land-use, farm wage (W_{farm}), non-farm wage (W_{nfarm}), labor costs for maintaining land-use j, which are discounted by idiosyncratic risk preferences (α_a).

$$ER_{ra|a}^{nfarm} = \alpha_a W_{nfarm} \tag{S16}$$

$$ER_{ra|a}^{mkt} = \alpha_a\left(EP_{j,t}EY_{j,t} - W_{farm}LC_{i\to j}\right) \tag{S17}$$

4) Based on the change in food (S^{food}) and money (S^{mon}) stocks, allocate proportion λ_h of total labor (L_{TOT}) to "home activities" (L_h).

$$\lambda_h = 1 + (1 - MI)\frac{\left(S_{a,t}^{food} - S_{a,t-1}^{food}\right)}{\delta_{\min}} + MI\frac{\left(S_{a,t}^{mon} - S_{a,t-1}^{mon}\right)}{\delta_{inc}} \tag{S18}$$

$$L_h = \lambda_h L_{TOT} \tag{S19}$$

5) Based on risk-discounted expected returns from farm production (Equation (S14)) and non-farm labor (Equation (S16)), allocate labor to farm (L_f) *vs.* non-farm (L_{nf}).

$$\lambda_f = \frac{ER_{rn|a}^{farm}}{ER_{ra|a}^{nfarm}} \tag{S20}$$

$$L_f = \lambda_f L_{TOT} \tag{S21}$$

$$L_{nf} = L_{TOT} - L_h - L_f \tag{S22}$$

6) Based on risk-discounted expected returns from subsistence (Equation (S15)) and market (Equation (S17)) production, allocate labor to subsistence (L_{sub}) *vs.* market (L_{mkt}) farm production.

$$\lambda_{mkt} = \frac{ER_{ra|a}^{mkt}}{ER_{rn|a}^{sub}} \tag{S23}$$

$$L_{mkt} = \lambda_{mkt} L_f \tag{S24}$$

$$L_{sub} = L_f - L_{mkt} \tag{S25}$$

7) For all possible land uses in each of the cell in an agents' landholdings, calculate expected marginal return on labor from subsistence production, and expected net marginal return from market production, and weight by land-use preferences to obtain expected marginal utility (Equations (S11) and (S12), respectively).
8) Agents first allocate subsistence labor (Equation (S25)) to cells that maximize marginal expected utility from subsistence production until subsistence labor or land constraints are met. Market labor (Equation (S24) is allocated to remaining cells that maximize expected marginal utility from market production until market labor or land constraints are met.

References

1. De Vries, F.W.T.P.; Rabbinge, R.; Groot, J.J.R. Potential and attainable food production and food security in different regions. *Philos. Trans. R Soc. Lond B* **1997**, *352*, 917–928.
2. Evans, T.; Manire, A.; de Castro, F.; Brondizio, E.; McCracken, S. A dynamic model of household decision-making and parcel level landcover change in the eastern Amazon. *Ecol. Model.* **2001**, *143*, 95–113.
3. Macmillan, W.; Huang, H. An agent-based simulation model of a primitive agricultural society. *Geoforum* **2008**, *39*, 643–658.
4. Ellis, F. *Peasant Economics: Farm Households and Agrarian Development*; Cambridge University Press: Cambridge, UK, 1993.
5. Netting, R. *Smallholders, Householders: Farm Families and the Ecology of Intensive, Sustainable Agriculture*; Standford University Press: Standford, CA, USA, 1993.
6. Global Agro-Ecological Zones (GAEZ). Terrain Constraints Available online: http://www.iiasa.ac.at/Research/LUC/GAEZ/index.htm (accessed on 6 June 2011).
7. Global Agro-Ecological Zones (GAEZ). Agro-climatic Constraints. Available online: http://www.iiasa.ac.at/Research/LUC/GAEZ/index.htm (accessed on 6 June 2011).
8. Monfreda, C.; Ramankutty. N.; Foley, J. Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary productivity in the year 2000. *Glob. Biogeochem. Cycl.* **2008**, *22*, doi:10.1029/2007GB002947.
9. Siebert, S.; Doll, P. Quantifying the blue and green virtual water contents in global crop production as well as potential production losses without irrigation. *J. Hydrol.* **2010**, *384*, 198–217.
10. Tian, G.; Kang, B.T.; Kolawole, G.O.; Idinoba, P.; Salako, F.K. Long-term effects of fallow systems and lengths on crop production and soil fertility maintenance in West Africa. *Nutr. Cycl. Agroecosyst.* **2005**, *71*, 139–150.
11. Tiessen, H.; Salcedo, I.H. Sampaio, E.V.S.B. Nutrient and soil organic matter dynamics under shifting cultivation in semi-arid northeastern Brazil. *Agric. Ecosyst. Environ.* **1992**, *38*, 139–151.
12. Tyson, K.C.; Roberts, D.H.; Clement, C.R.; Garwood, E.A. Comparison of crop yields and soil conditions during 30 years under annual tillage or grazed pasture. *J. Agric. Sci.* **1990**, *115*, 29–40.
13. Food and Agricutlure Organization of the United Nations (FAO). Trade and Markets: Commodities markets monitouring and outlook. Available online: http://www.fao.org/economic/est/est-commodities (accessed 1 June 2015).
14. Indexmundi. Commodity Prices: Price charts, data, and news. Available online: http://www.indexmundi.com/commodities/ (accessed 1 June 2015).
15. Turner, B.L. II; Hyden, G.; Kates, R.W. *Population Growth and Agricultural Change in Africa*; University of Florida Press: Gainesville, FL, USA, 1993.
16. Boserup, E. *The Conditions of Agricultural Growth: The Economics of Agrarian Change under Population Pressure*; Aldine: Chicago, IL, USA, 1965.
17. Chayanov, A.V. *The Theory of Peasant Economy*; Thorner, D., Kerblay, B., Smith, R.E.F., Eds.; R.D. Irwin: Homewood, IL, USA, 1966.
18. Turner, B.L., II; Hanham, R.; Portararo, A. Population pressure and agricultural intensity. *Ann. Assoc. Am. Geogr* **1977**, *67*, 384–396.

19. Boserup, E. *Population and Technological Change*; University of Chicago Press: Chicago, IL, USA,1981.
20. Dorsey, B. Agricultural intensification, diversification, and commercial production among smallholder coffee growers in Central Kenya. *Econ. Geogr.* **1999**, *75*, 178–195.
21. Goldman, A. Agricultural innovation in three areas of Kenya: Neo-Boserupian theories and regional characterization. *Econ. Geogr.* **1993**, *69*, 44–71.
22. Turner, B.L., II; Ali, A. Induced intensification: Agricultural change in Bangladesh with implications for Malthus and Boserup. *Proc. Nat. Acad. Sci. USA* **1996**, *93*, 14984–14991.
23. De Janvry, A.; Fafchamps, M.; Sadoulet, E. Peasant household behavior with missing markets: Some paradoxes explained. *Econ. J.* **1999**, *101*, 1400–1417.
24. Barrett, C.; Reardon T.; Webb P. Nonfarm income diversification and household livelihood strategies in rural Africa: Concepts, dynamics, and policy implications. *Food Policy* **2001**, *26*, 315–331.
25. Reardon, T.; Berdegue, J.; Barrett, C.; Samoulis, K. Household income diversification into rural nonfarm activities. In *Transforming the Rural Nonfarm Economy*; Haggblade, S., Hazell, P., Reardon, T., Eds.; Johns Hopkins University Press: Baltimore, MD, USA, 2006; pp. 115–140.
26. Brown D.; Aspinall R.; Bennett D. Landscape models and explanation in landscape ecology—A space for generative landscape science? *Prof. Geogr.* **2006**, *58*, 369–382.
27. Schreinemachers, P.; Berger, T. Land use decisions in developing countries and their representation in multi-agent systems. *J. Land Use Sci.* **2006**, *1*, 29–44.
28. Arthur, W. Inductive reasoning and bounded rationality. *Am. Econ. Rev.* **1994**, *84*, 406–411.
29. Arthur, W.; Durlauf, S.; Lane, D. *The Economy as an Evolving Complex System II*; Addison-Wesley: Sante Fe, NM, USA, 1997.
30. Magliocca, N.R.; Ellis, E.C. Using pattern-oriented modeling (POM) to cope with uncertainty in multi-scale agent-based models of land system change. *Trans. GIS* **2013**, *17*, 883–900.