Simulating Cooperative Behaviors in a Subsistence Population:

An Agent-Based Modeling Approach

by

Julia Phelps

A Thesis Presented in Partial Fulfillment
of the Requirements for the Degree
Master of Science

Approved April 2023 by the
Graduate Supervisory Committee:

Mark Reiser, Co-Chair
Steven Saul, Co-Chair
Thomas Morgan

ARIZONA STATE UNIVERSITY

May 2023

ABSTRACT

Humans cooperate at levels unseen in other species. Identifying the adaptive mechanisms driving this unusual behavior, as well as how these mechanisms interact to create complex cooperative patterns, remains an open question in anthropology. One impediment to such investigations is that complete, long-term datasets of human cooperative behaviors in small-scale societies are hard to come by; such field research is often hindered both by humans' long lifespans and by the difficulties of collecting data in remote societies. In this study, I attempted to overcome these methodological challenges by simulating individual human cooperative behaviors in a small-scale population. Using an agent-based model tuned to population-level measurements from a real-life marine subsistence population in the southern Philippines, I generated dynamic daily cooperative behaviors in a hypothetical subsistence population over a period of 1500 years and 42 overlapping generations.

Preliminary findings from the model suggest that, while the agent-based model broadly captured a number of characteristic population-level patterns in the subsistence population, it did not fully replicate nuances of the population's observed cooperative behaviors. In particular, statistical models of the simulated data identified reciprocity-based and need-based cooperative behaviors but did not detect kinship-motivated cooperation, despite the fact that kin cooperation traits evolved positively and reciprocity cooperation traits evolved negatively over time in the agent population. It is possible that this discrepancy reflects a complex interaction between kinship and reciprocity in the agent-based model. On the other hand, it may also suggest that these types of statistical models, which are frequently utilized in human cooperation studies in the anthropological literature, do not reliably discriminate between kin-based and reciprocity-based cooperation mechanisms when both exist in a population. Even so, the completeness of the simulated data enabled use of more complex

statistical methodologies which were able to disentangle the relative effects of cooperative mechanisms operating at different decision levels. By addressing remaining pattern-matching issues, future iterations of the agent-based model may prove to be a useful tool for validating empirical research and investigating novel hypotheses about the evolution and maintenance of cooperative behaviors in human populations.

DEDICATION

To my husband, who keeps me well-fed and patiently listens to my late-night mathematical and statistical ramblings.

ACKNOWLEDGMENTS

First and foremost, I would like to thank the people of Linao for their generous participation in interviews, as well as for allowing the Linao research team to observe their daily activities. Without their support, none of this research would be possible. Secondly, I must acknowledge the support of my committee members, Drs. Mark Reiser, Steven Saul, and Thomas Morgan, whose feedback proved instrumental in development of the agent-based model and analysis. Lastly, I must also express my gratitude to Drs. Kim Hill and Rob Boyd for their comments on early iterations of the agent-based model, which helped me improve the initial model specification considerably.

TABLE OF CONTENTS

CHAPTER Page

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Although individuals in many non-human species cooperate within limited contexts [e.g., primates (Mitani, 2006; Silk et al., 2006; Thompson, 2019), feral horses (Cameron et al., 2009), mice (Pillay & Rymer, 2015), and honey bees (Tarpy et al., 2004)], the magnitude and breadth of cooperative behaviors observed in human populations is a defining characteristic of our species. Understanding the evolutionary and cultural mechanisms driving driving human cooperation remains a challenging puzzle for anthropologists, not least because the long average human lifespan (e.g., Gurven & Kaplan, 2007) makes long-term data collection over multiple consecutive generations difficult – if not impossible. While data collected in foraging and other small-scale subsistence societies remain the gold standard for addressing questions related to human evolution (Henrich et al., 2010; Marlowe, 2005), gathering accurate and balanced data samples in these types of populations can also be an intractable problem for a variety of reasons, including the time-costs of gathering sufficiently large samples, irregular availability of participants (especially in nomadic and semi-nomadic populations), and the difficulties inherent in accessing remote study sites. Moreover, increasing market-integration of subsistence populations means that their utility as models of "traditional" human society is diminishing (for a recent example, see Wiessner & Huang, 2022).

Agent-based modeling may offer a path forward for addressing these challenges. Simulating the population structures and behaviors observed in real-world societies can yield large, complete datasets that enable both validation of existing empirical research and exploration of new theories. Additionally, the multi-generational nature of many agent-based data samples opens up the possibility of directly investigating the evolutionary processes underlying currently-observed behaviors. Nonetheless, as I will demonstrate in this analysis, simulating an existing

population – especially a population of humans – can be a daunting task. Accurately capturing the essential characteristics of a population requires both detailed knowledge of the population being modeled and an iterative process of calibration and analysis to identify inconsistencies between the agent-based model and the real-world population.

Agent-Based Models: History and Use in Anthropology

"Agent-based" modeling is a method of simulating a system or population by focusing on the characteristics and mechanisms of the individual components – or "agents" – of that system or population (DeAngelis & Grimm, 2014). That is, by specifying the relevant, "bottom-level" attributes of a system, such as the life-history attributes of an individual agent and the manner in which it interacts with its environment and/or other agents, one can generate complex, dynamic population-level behavior (Grimm et al., 2005). Unlike traditional deterministic mathematical modeling techniques, where simulations are constrained by the tractability of their underlying equations, agent-based models do not necessarily require simplifying assumptions (although such assumptions may be useful for implementation and interpretability) and are instead primarily constrained by computing power alone. As such, agent-based models can easily incorporate additional environmental and spatial characteristics that would likely be challenging (if not impossible) to specify in a mathematical modeling setting.

Agent-based models (ABMs) are also frequently referred to as individual-based models (IBMs) in the literature. Originally, the two terms were used to identify distinct approaches to simulation: Those classified as "agent-based" models focused on evolutionary processes and agent "decisions," while those classified as "individual-based" models mainly focused on modeling individual variation (Rails-

back & Grimm, 2012). However, the two terms have been used interchangeably in recent years, in response to both rapid growth of the field and convergence of modeling strategies. Nonetheless, for the sake of clarity, I will use "agent-based" terminology exclusively going forward.

ABM techniques were first developed in the field of ecology (DeAngelis & Grimm, 2014), where the earliest recognized agent-based model was a "gap-phase replacement" model describing the successions of forest canopy trees (Botkin et al., 1972). Forest ecologists embraced these methods, and many subsequent models emerged in the next years and decades to simulate the dynamics of single- and multi-species forests (e.g., Köhler & Huth, 1998; Rademacher et al., 2004; Shugart Jr & West, 1977; Wissel, 1992). Use of ABM techniques in the field of animal ecology also began to take root; some archetypal examples include fish-schooling models (Huth & Wissel, 1992) and bird-flocking models (Ballerini et al., 2008; Reynolds, 1987). More recently, ABMs have been used extensively to investigate many different aspects of animal populations. Common examples include modeling the growth, distribution, and competition between fisheries (e.g., Clark & Rose, 1997; D. DeAngelis et al., 1993; Saul et al., 2012), simulating ungulate foraging behaviors (Turner et al., 1993), investigating the population structure and spatial dynamics of field voles (Topping et al., 2012), and modeling the spread of rabies between fox populations (Jeltsch et al., 1997) – though this is by no means even close to an exhaustive list of the animal-based applications of ABMs. The impact of foraging behaviors on individual environments themselves have also been explored, such as in Weiss and colleagues' (2014) simulation of the dynamics of species diversity and competition between plants in response to grazing in a simulated grassland environment. Advanced comparative and statistical techniques have been developed for calibration and validation of ABMs across a wide range of applications [for a review, see

Banks & Hooten (2021)]. However, pattern-oriented modeling remains the most commonly-implemented method for parameterizing and verifying agent-based models (Grimm et al., 2005; described more thoroughly in Chapter 2).

Within the field of anthropology, agent-based models have more recently been adopted to study a wide range of theoretical and empirical questions. For example, ABMs have been constructed to investigate foraging behaviors in the context of optimal foraging theory (e.g., Nonaka & Holme, 2007) and multi-level selection of fitness-based cooperative traits (Pepper & Smuts, 2000). Developed ABMs have additionally been used to characterize and investigate human migration patterns, such as in Filho et al.'s (2011) model of the impact of social networks on flows and counter-flows of human migration. Theories surrounding the evolution of human cooperative behaviors have also been examined using ABMs: For example, van Veelen et al. (2012) demonstrated that high levels of cooperation could evolve through the interaction between direct reciprocity and population-structuring assortment mechanisms within an ABM environment. García and colleagues (2014) further showed that individual recognition of shared "tags" denoting cooperative intent allowed for cycles of indiscriminate cooperation and defection in agent populations characterized by low assortment (e.g., low relatedness), but instead promoted cooperation only between those agents who shared the same tag in high-assortment populations (García et al., 2014).

The evolution of human culture has also been explored through an agent-based lens by, for example, modeling the influence of social learning mechanisms on the transmission and cumulative cultural evolution of task-based information (Miu & Morgan, 2020) and investigating the role of cognition in promoting creative recombination of culturally-transmitted ideas (Gabora & Saberi, 2011). In the sub-field of primatology, Acerbi et al. (2022) used an agent-based model to ar-

gue that a "distributional approach to culture" (whereby behaviors that vary between populations are deemed to be cultural via the method of exclusion if no genetic or environmental factors explaining their variation can be identified) cannot be used to identify cultural inheritance in chimpanzees, since similar patterns can arise through purely non-copying socially-mediated processes. Cumulative cultural evolution within the the archaeological record has also been investigated by, for example, using an agent-based approach to examine possible pathways for transmission of skills required in the production of stone projectile points (Garvey, 2018). Patterns demonstrated in ABMs have additionally been used to justify specific anthropological theories such as the Grandmother Hypothesis (Kim et al., 2012), although whether this particular theory is supported by real-world evidence remains an open question (e.g., see Hill & Hurtado 1991; Kaplan et al., 2000).

It is worth noting that agent-based models are most frequently used within anthropology to explore and test specific theories and hypotheses using hypothetical agent populations. Less often have ABMs been used to simulate data from real-world human populations for further analysis, likely because of the complexity involved in accurately developing such models. However, this may be a fruitful new method of researching the behaviors of foraging and other subsistence societies, particularly as the frequency of these types of populations dwindles in response to market-integration and inculturation.

Motivating Study

The current study was initially motivated by statistical challenges that I faced while analyzing data for a previous study of cooperation in a small-scale marine subsistence society (Phelps et al., 2022). In this section, I briefly describe the research population, challenges with collecting data, and statistical issues faced dur-

ing the prior study. In the next section, I will discuss how the current project addresses some – though not all – of these issues.

Study Population and Empirical Data

Data used to parameterize and validate the agent-based model presented in this study were collected between June 2015 and November 2018 and come from a prior study of the cooperative patterns in Linao village (Phelps et al., 2022). Located at the mouth of the Sarangani Bay in Southern Mindanao, Philippines, Linao is a small community of formerly-nomadic marine foragers whose coastal village is situated along a cobbled-together pier built atop mangrove and reef flats. Residents are of primarily Sama ethnic ancestry, an ethnic group/culture which was traditionally characterized by sea-faring nomadism and exploitation of reef and open-ocean marine resources. While the Sama population living in Linao has become more sedentary in recent years, residents are still semi-nomadic and frequently relocate outside of the village for weeks or even months at a time. As such, only a subset of the entire population is in residence within the village at any given time. The Linao people still engage in marine resource exploitation as their main method of resource production, although its form has changed slightly with the increasing sedentism of the community. Economic activity of residents is comprised mainly of coastal and open-ocean fishing, "gleaning" (collecting) intertidal marine resources, and collecting other wild resources from the local environment, along with small, infrequent amounts of opportunistic wage labor in the surrounding area. Linao is also characterized by a high degree of daily cooperation between its residents. Cooperation events in the form of transfers of food, resources, and other material goods between different households (the primary economic unit) are observed many times each day within the community. Importantly, while most individual transfers of food, re-

sources, and other goods are small in terms of monetary value, the sheer frequency with which these transfers occur means that the average quantity of goods received by a household on a given day from other village households amounts to almost two-thirds of the average daily total household income from all resource production activities (Phelps et al., 2022). Clearly, inter-household transfers of goods are a critical part of the economic strategy in Linao.

A total of 36 households are included in the full Linao study population and are used in demographic calculations and descriptive analyses of population structure within this study. Specifically, Linao data utilized in this analysis include daily records of resource transfers between individuals and households, tabulations of daily resource production from all sources, information on village composition, and demographic and genealogical records of all residents. Data was collected via a mix of 24-hour recall interviews, monthly census interviews of all residents, direct observation, and other targeted interviews. Additionally, statistical models developed in the previous Phelps et al. (2022) analysis, which were generated on a 32-household subset for which data on transfers and resource production was more complete, are utilized as a point of comparison for similar models built on the agent-level data simulated in this study. Importantly, while research assistants did their best to balance interview samples from different households during the initial study, some households were directly interviewed more times than others [range: 3-28 interviews per household in the model sample (Phelps et al., 2022)]. This sampling imbalance was mainly due to participant availability/willingness and the intermittent residence of many households in Linao throughout the study period.

Statistical Challenges in Motivating Study

In an initial exploratory analysis of cooperation in Linao, Phelps et al. (2022) modeled daily and long-term mean patterns of cooperation in Linao by using daily inflows and outflows of all material goods (food, money, clothing, etc.) from one household to another as a proxy for inter-household cooperation. One of the major findings from this initial analysis was that the motivation for resource transfers between study households appeared to be dominated by the effects of direct reciprocity (Trivers, 1971), whereby a Household X was incentivized to give to another Household Y that had previously given to them during the same day or in the past, and/or might reasonably be expected to give to them again at some point in the future. Kin selection (Hamilton, 1946) between households was a secondary motivator of transfers during the study period, both in terms of provisioning needy relatives in one-off interactions and in establishing reciprocal relationships with other related households. Needs-based sharing (D. Smith et al., 2019), based on relative differences in per capita income, was also an important predictor of increased transfers of goods between "needy" and "less needy" households (Phelps et al., 2022), although it often functioned within the context of kin selection. Although indirect reciprocity (whereby a focal individual or household helps someone who may not directly reciprocate with them, but who might, in turn, help a third party who then helps the focal individual/household; Alexander, 1987; Nowak & Sigmund, 2005) also likely played a role in promoting cooperation in Linao, it was not explicitly tested in the Phelps et al. (2022) study.

Importantly, due to the nature of the data collected in Linao (Phelps et al., 2022), statistical analyses in the prior study were unable to provide resolution to two major statistical challenges. One, daily measures of the transfer of goods be-

tween community households were zero-inflated, likely as a result of both biased cooperative assortment of community households and broad daily variation in resource acquisition. Two, ethnographic and statistical insights from the community suggest that the quantity and direction of resource transfers on a given day in the study period were likely influenced by earlier transfers in the time leading up to the current day. However, data explicitly describing temporal correlations of resource transfers was not collected during the data collection. Moreover, both the semi-nomadic nature of Linao residents and limitations of the sampling strategy prevented accurate reconstruction of such evidence from the collected data. Below, I will describe initial attempts to address these two statistical challenges (not discussed in Phelps et al., 2022), the failures of which motivate the current study.

Zero Inflation of Collected Data. Zero inflation within the Linao data set can be detected by tabulating the proportions of inflows, outflows, or inflows and outflows between household dyads that equaled zero, which equaled approximately 0.77, 0.80, and 0.69, respectively, within the daily dyadic material good transfers dataset. In particular, while most households had several recorded inflow and outflow events per day, there were almost 500 unique household dyads possible in the household sample. Not all of these households directly interacted with each other, of course – most only interacted with a handful of preferential partners. However, on a given day, all dyads in which both households were currently in residence within the village and at least one household was directly interviewed were recorded to avoid biasing data. As should be plainly obvious, this resulted in a tremendous amount of zero transfer observations within the sample. While the results of the multiple linear regression models considered in the Phelps et al. (2022) paper made good sense from both an ethnographic and a theoretical point of view,

it is also very likely that inclusion of so many zeros in these models resulted in somewhat inaccurate estimates or failure to detect important effects.

In an attempt to address this challenge, I experimented with various alternative modeling strategies with little success (Phelps & Hill, 2021). Single-model strategies for handling zero-inflated data typically enforce constraints on outcome variables that are not always desirable when analyzing continuous data like the Linao cooperative transfer data. For example, commonly-utilized mixture models such as zero-inflated poisson regression ("ZIP" model; Lambert, 1992) and zero-inflated negative binomial regression ("ZINB" model; Greene, 1994) only characterize count data, and as such cannot be used to model continuous variables. While the properties of these two models may be useful for understanding patterns governing different observed *counts* of observed cooperative events, there is no extension to these models that allows for modeling the relative *quantities* of transfers. This is a major limitation, given that variation within the quantities of dyadic transfers is often what researchers are trying to understand.

Two-part hurdle models such as the Cragg hurdle model (Cragg, 1971), the Heckman selection model (Heckman, 1979), and the Blundell double hurdle model (Blundell & Meghir, 1987) all offer promising alternative solutions to this issue. In different ways, each model in this class attempts to account for an overabundance of zeros (or some other boundary value) by separately modeling the processes generating zero and non-zero observations. Surprisingly, these types of models are infrequently implemented within the anthropological literature, despite their attractive properties with respect to modeling human behaviors [but see Kasper & Mulder (2015)]. The Cragg hurdle model (Cragg, 1971) is particularly appealing in the context of the Linao dyadic transfer data, since it first estimates the probability of an observation being zero/non-zero using a probit model, and then separately mod-

els the continuous aspect of a subset of the data which includes only the non-zero values. In the situation where data is non-negative (as is the case with the Linao resource transfer data), Cragg (1971) recommends using a truncated Normal regression model to characterize non-zero observations. Unfortunately, however, attempts to utilize a Cragg hurdle model to overcome zero inflation in the Linao dataset had various problems, including failing to converge. Unequal sample sizes within dyads, uneven sampling of the population, and a high overall level of daily variation in dyadic transfers likely all contributed to this issue.

Temporal Correlations in Collected Data. The second major challenge, i.e., temporal correlations between resource transfers, is equally difficult to overcome in studies of cooperation within small-scale and semi-nomadic societies. The notion that reciprocal cooperative behaviors between humans are temporally-structured makes good sense, and it is additionally broadly supported by prior empirical and theoretical literature from the fields of anthropology (e.g., Allen-Arave et al., 2008; Boyd & Mathew, 2021; Gomes et al., 2009; Gurven et al., 2001; Jaeggi & Gurven, 2013; Wedekind & Braithwaite, 2002) and econometrics (e.g., Davidson et al., 1978; Hendry & Richard, 1983; Paraskevas et al., 2022). Panel models and other time-series statistical modeling methodologies have often been employed to account for the temporal correlations between non-independent observations, and cross-lagged panel models (Kenny, 1975) are particularly attractive in the context of daily resource transfers between individual and household dyads. However, the difficulties of collecting appropriate data in subsistence populations (in terms of time-costs, adequate observation controls, availability/willingness of participants, etc.) means that substantial aggregation of data is often required in order to utilize cross-lagged panel models. In turn, this can have the effect of averaging out important short-term variation in behaviors, limiting the applicability of such mod-

els for detecting subtle behavioral effects. This issue is further compounded when study populations are nomadic or semi-nomadic, since observations of individuals or households may be sparse or unevenly distributed throughout a sample period. Since cross-lagged panel models typically cannot account for missing observations within time periods, using these models with sparsely- or unevenly-sampled data may require severe data reduction and/or aggregation across long time intervals.

The aforementioned issues highlight the exact problem that I faced when attempting to reanalyze the Linao dyadic transfers data with a cross-lagged panel model. While the resultant model broadly supported conclusions from prior analyses (i.e., Phelps et al., 2022), the imbalances in household sampling resulting from both frequent nomadic excursions and participant availability/willingness necessitated aggregation of dyadic transfer data over long time spans (half-year to full year). Additionally, because some households left the community for months at a time and hence were not measured within all time spans, the final model only included a small subset of the total number of households in the Linao study population. As such, the resultant cross-lagged panel model was deemed insufficient for investigating the overall patterns of cooperation that characterize Linao village.

Aims of This Study

In response to the statistical challenges outlined in the previous section, this current study proposes to use an alternate method to investigate cooperation in small-scale societies: Simulation of daily dyadic cooperation data from a hypothetical subsistence population. An agent-based population dynamics modeling approach is adopted for this purpose, and agents are (loosely) modeled to match human life history specifications. To that end, I utilize empirical data collected in Linao during the previous Phelps et al. (2022) study to inform agent structure and

parameterization of the agent-based model (ABM). Furthermore, observed patterns of population structure and statistical models of cooperative interactions in Linao are used to evaluate emergent patterns in the agent-based model, and comparisons of identified mechanisms driving cooperation are drawn between the ABM and the Linao dataset. Importantly, the overarching goal of this agent-based modeling exercise is not to exactly replicate individual behaviors observed in Linao, per se, but rather to capture overall population-level processes that drive the patterns of cooperation observed in the village. This strategy, in turn, will enable future exploration of additional proposed mechanisms (e.g., culturally-defined cooperation norms) that may serve to further stabilize – or destabilize – the unique cooperative dynamics observed in Linao.

The remainder of this manuscript will be organized as follows: First, I will give a qualitative summary of the overarching structure and processes of the agent-based model, following this description with a detailed discussion of model initialization, distinct steps within each iteration of the model algorithm, and the exact mathematical and probabilistic mechanisms used to model agent life history and cooperation dynamics. In the results section, I will briefly discuss early (failed) attempts at the model and then describe the decisions undertaken during parameterization of the final version of the model. An overview of patterns and trends observed in the final ABM will then be discussed, and I will follow this by comparing key characteristic patterns observed in Linao with those generated by the ABM. Lastly, I will utilize data generated by the ABM to develop a Cragg hurdle model, which will explore potential mechanisms driving cooperation between agents in the ABM.

Chapter 2

METHODS

Qualitative Summary of Model

To facilitate interpretation of the quantitative model description laid out in the *Quantitative description of agent-based model algorithm* section below, I will first give a very brief qualitative summary of a typical "day" in the model. Steps in the model algorithm are broadly similar from day to day, so I will focus on a description of what occurs during the first day and then draw attention to differences on subsequent days. Figure 1 at the end of this section contains a flowchart illustrating the overall daily process. Importantly, only major steps in the algorithm are defined here – smaller intermediary steps are described in more detail in the *Quantitative description of agent-based model algorithm* section.

At the start of the model, a set number of lineages are initialized, each containing a set number of individual agents. Agents within each of the lineages are related to one another (although some only very distantly), but unrelated to everyone in other lineages. Population size is fixed at initial size, and all agents are initialized at a similarly young adult age. Right off the bat, a small number of agents in the model may be selected to die in the first step of the day as a result of their age-specific "baseline" mortality rate. (It is certainly grim to start the model off in this way, but it makes for easier coding.) However, given the age of the initialized population, death at this stage is very rare.

Within each of the initialized lineages, all agents are young adults who are at an age where they can select a spouse and start having children. Agents are prohibited from marrying anyone who is their first cousin or closer, so they will typically select a spouse from another lineage. After any agents who die are re-

moved from the population, the next step is for all eligible agents (i.e., those who are single and at marriageable age or older) to select a spouse. Not everyone will be paired off, since some agents may be too related to each other to marry, but most initialized agents will marry at this step. Once two agents marry, they become a reproductive pair and can begin producing offspring immediately. However, since the model has a fixed number of agents at any time step, a reproductive pair can only give birth to a new offspring agent if another agent has died. Specifically, if any agents have died earlier that day, a replacement agent will either be "born" into the population and assigned as a child to one of the reproductive pairs, or they will be "recruited" into the population as an unrelated adult. In practice, however, most replacement agents are newborns of existing agents in the population.

After agents have reproduced (or unrelated migrant agents have joined the population), the real work of the day can begin. Each agent requires a specific amount of resources per day ("need"), which is scaled by age and developmental status. To satisfy this need, all agents above a certain age (i.e., above the age at which they are fully dependent on their parents) produce a random amount of resources each day, which is also scaled based on age. Depending on how many resources an agent produces, they may either have a surplus of resources, in which case they are able satisfy their own need, or a deficit of resources, in which case they do not satisfy their own need. (Note that the probability of an agent producing exactly the amount of resources they need is next to zero.) Deficits of resources are a problem for the agents, since those who do not have enough resources at the end of the current day will suffer higher mortality rates at the start of the next day. An *adult* agent who does not produce enough resources on their own (that is, they are "needy") can potentially make up the difference by soliciting a resource donation from another adult agent who has a surplus. This candidate donor agent may

– or may not – be willing to share with the candidate recipient (the needy agent), based on a few different factors. One, if the donor agent and the recipient agent are sufficiently related, the donor may be motivated by kin selection (Hamilton, 1964) to share with the recipient. Two, if the recipient agent has shared with the donor agent in the past, then the donor agent may be motivated by direct reciprocity (Trivers, 1971) to share with the recipient. Three, there is a small additional probability that the donor agent will share with anyone in the population, regardless of relationship or past cooperative actions. However, the candidate donor agent may also be motivated to keep their extra resources, as they can instead use this surplus to improve their health and decrease their mortality risk to a level below their age-specific baseline rate.

On the other hand, if a *juvenile* agent has a shortfall of resources (either because they didn't produce enough or because they are too young to produce any resources at all), then responsibility for supplementing these resources falls on the juvenile's parents. In particular, if both parents are alive, then each is responsible for an equal portion of the juvenile's resource deficit. If only one parent is alive, then the live parent is responsible for the entirety of their offspring's deficit. On a given day, a parent may produce enough resources to simply cover their juvenile's shortfall on their own. But if not, then the parent will pool their and their offspring's resources and redistribute according to age-specific need. The parent agent is then responsible for securing a cooperative transfer of resources from another agent to cover both their and their offsprings' remaining need.

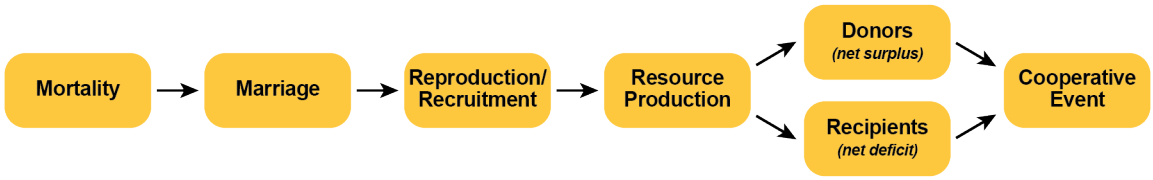
Cooperative interactions between agents are the last step of the day. Each agent can participant in a cooperative interaction action only once per day, either in the role of a recipient or in the role of donor (depending on their net need). Not all agents will engage in a cooperative interaction during a specific day if there is

an unequal amount of candidate donors and candidate recipients. Additionally, only adult agents – or independent agents, in the case of juveniles who have lost both their parents – can solicit cooperative assistance from other adult/independent agents. If a needy agent fails to secure a donation from another agent, then they are stuck with a resource deficit for that day.

After all cooperative interactions have occurred, the day is over and the next one begins. Agents die, marry and reproduce as they are able to. Occasionally, unrelated adult agents join the population. Resources are produced by those who are old enough, and adult/independent agents engage in cooperative interactions with each other. Resource shortfalls at the end of each day increase the risk of mortality at the start of the next day, and resource surpluses decrease these risks. And so it goes, until the end of the model run. The major difference between the first day and all subsequent days is that marriages will be less frequent after the first day, since most agents will either already be married or will be juveniles who are too young to marry. However, as agents die and new reproductive-age agents enter the population either by growing up or by being recruited from elsewhere, there will still be opportunities for single adult agents to marry/remarry.

Figure 1

Flowchart Illustrating Sequence of Major Events That Occur Each Model Day



Notation

Before I go into the details of the agent-based model, I will make note of a few general points on notation. Time steps are denoted with the index t. Individual agents are denoted by the indices i, j, d, and ℓ. The notation a without a subscript is used to make general references to age (in days/time steps) in the model, but the notation $a_{i,t}$ will always be used in reference to the age of an individual agent at a specified time t. Following current mathematical convention in Western/European countries, the set of all natural numbers (i.e., $\{1, 2, 3, ...\}$) is indicated by the blackboard bold symbol $\mathbb{N}$. Indicator functions (i.e., functions that assume a value of 1 if some condition is met, and 0 otherwise) are used periodically throughout the main paper and Appendices A and B, and these are denoted with the blackboard bold symbol $\mathbb{I}$. Wherever possible, I will introduce all other mathematical notation in the context of its first usage. However, it should be noted that this model is complex, involving many different parameters and underlying calculations. For reference, I have included a summary of key indices and parameters in Appendix A, as well as a description of the variables characterizing agent structure in Appendix B.

Quantitative Description of Agent-Based Model Algorithm

In the two subsections that follow, I will first describe how the population is initialized at time $t = 1$, followed by a description of the iterative algorithm that repeats at every time-step from $t = 1$ to $t = T$ (where T is the maximum possible time-steps of the model). Individual iterations of the model may be thought of as "days" in model time.

Population Initialization at Time $t = 1$

The model environment is initialized at $t = 1$. K lineages are each comprised of N_k adults in generation 1 ($g_i = 1$ for each agent i). Note that N_k may be specified as either a single value, indicating that the size of all lineages are equal, or as a vector of values, in which case the k^{th} lineage has N_k members. Population size is fixed throughout the duration of the model at $\sum_{k \in K} N_k$. The life-history of agents within the population is bounded according to several user-specified parameters, which are fixed for the population. First, A controls the average maximum lifespan within the population. Notably, however, this is *not* a maximum possible age in the population: While an agent's baseline mortality is set to 1 at this and older ages, death at this age is not strictly enforced in the model. It is still possible that an agent may decrease their mortality rate sufficiently via additional resources to overcome mortality at this age. However, in practice, the probability of this occurrence is quite small. Second, $\bar{\delta}_m$ controls the age of maturity in the population, i.e., when an agent is considered an independent adult with full resource production capability. Third, a juvenile agent is considered to be fully-dependent on their parents until they reach age $\bar{\delta}_d$. Up until this age, a dependent agent cannot produce any resources by themselves and must rely on their parents for all resource provisioning. At age $\bar{\delta}_d$, an agent can begin to produce an age-scaled amount of resources, which reaches full capacity at age $\bar{\delta}_m$. (This will be described in more detail in the next section.)

Each initialized agent is assigned a randomly-selected "age at first possible reproduction" (α_i), which governs the age at which an agent can first take a spouse and begin reproducing offspring. For convenience, the age $a_{i,t}$ of each agent at initialization is set to their age at first possible reproduction, i.e. $a_{i,t=1} = \alpha_i$.

To generate this age variable for each agent, a random value is sampled from the raised cosine distribution (Chattamvelli & Shanmugam, 2021; Warsza & Korczynski, 2010) with mean parameter $\bar{\alpha}$ and spread parameter s_α. Since this distribution is uncommon in many areas of science and I use it frequently throughout the model, I will briefly digress on its origins, properties, and benefits within a simulation environment.

A standardized version of the raised cosine distribution was originally proposed by Raab and Green (1961), but was limited only to support within the range [-π, π]. More recently, the distribution was developed further by Warsza and Korczynski (2010) to account for non-radian support and incorporate non-standard mean and spread parameters. Since then, multiple parameterizations of the raised cosine distribution have been developed for various engineering applications (a helpful reference can be found in Chapter 7 of Chattamvelli & Shanmugam, 2021), of which I make use of the "Raised Cosine Distribution - Type II" – hereafter, simply "raised cosine distribution" or "RCD".

The raised cosine distribution is a symmetric, bell-shaped distribution that approximates the Normal distribution (Warsza & Korczynski, 2010) but has the additional, computationally-beneficial quality of being bounded at either tail. Specifically, a raised cosine distribution with mean $\bar{\alpha}$ and spread s_α has support only within the range $[\bar{\alpha} - s_\alpha, \bar{\alpha} + s_\alpha]$. In practice, use of this distribution for randomly-sampling α_i values means that no truncation is required in order to keep ages at first possible reproduction within a realistic range. That is, it is possible that α_i values sampled from a Normal distribution could be unrealistically small or large – or even negative, which is not consistent with the range of possible age values. In a computational setting where a Normal distribution was being used to generate α_i values, it would be necessary to address this possibility by artificially increasing

or decreasing the offending values so that they were contained within the realistic range. In turn, this could invalidate distributional assumptions, particularly if such values were frequently observed (which would cause bulge points in the tails). By instead using a raised cosine distribution to generate α_i samples, this problem is side-stepped entirely. This is both statistically and computationally advantageous, particularly since the specified boundaries of the raised cosine distribution allow for easy control over the distribution of possible first reproduction ages.

Importantly, the raised cosine distribution is a continuous distribution, while ages in the model are represented in discrete time-steps. As such, a description of the method used to generate discrete values from the raised cosine distribution is explained below in *Implementation and statistical analysis of agent-based model*. Additionally, it should be noted that the same α_i distribution was used for both females and males in the population as a simplifying assumption. In reality, however, age at first reproduction is typically a few years older for men than women in small-scale societies (e.g., Walker et al., 2006). Moreover, individual samples of α_i may or may not be larger than $\bar{\delta}_m$ (the age of adulthood), depending on the model arguments specified by the user. This is not unrealistic, given that many species including humans reach reproductive age before reaching maturity (e.g., Leigh, 2001).

The sex γ_i of each initialized agent is assigned randomly, following a Bernoulli distribution with probability $p_{\gamma,M}$ of being male. Each agent i also has a vector of cooperation traits, denoted as $\vec{w}_i = \{\omega_{i,B}, \omega_{i,R}, \omega_{i,K}\}$, which control the agent's "willingness to cooperate" in three different contexts: $\omega_{i,B}$ controls an agent's "baseline" willingness to cooperate with anyone, $\omega_{i,R}$ controls their willingness to cooperate with past reciprocity partners, and $\omega_{i,K}$ controls their willingness to cooperate with kin. These "traits" will be described in greater detail below, but for now, it

is important only to note that samples of these three traits are drawn from raised cosine distributions with means and spreads set by the user.

A measure of inbreeding is tracked throughout the model progression, using the coefficient of inbreeding developed by Sewall Wright (1922). For simplicity, however, all initialized agents at time $t = 1$ are assigned a coefficient of inbreeding b_i equal to zero. Nonetheless, members within each of the K distinct lineages are interrelated by definition. To randomly assign non-zero relatedness $r_{i,j}$ within a lineage, a Poisson distribution and the following algorithm are used:

1. A rate parameter r_{exp} for the *exponent* of the coefficient of relatedness is constructed as $r_{\text{exp}} = [\log_{1/2}(\bar{r}) - 1]/2$, where $\bar{r}$ is a parameter set by the user to quantify mean relatedness of the *entire* lineage (not just the initialized agents in generation 1 of the lineage). The base-1/2 logarithm in this calculation extracts the exponent from $\bar{r}$. Then, to account for zeros in the support of the Poisson distribution, 1 is subtracted from the exponent. (This is re-added after sampling.) Lastly, the division by 2 accounts for the requirement that all agents are in the same generation – hence, when no inbreeding is present in the population, all coefficients of relatedness must be of form $(1/2)^{2q+1}$, where $q \in \mathbb{N}$. This is also multiplied back in after sampling.

2. For lineage $k \in \{1, .., K\}$, relatedness calculations are "primed" by selecting a first pair i, j of agents in the lineage. A integer value n is sampled from the Pois(r_{exp}) distribution, and the coefficient of relatedness between i and j is calculated as $r_{i,j} = (1/2)^{2n+1}$.

3. Relatedness for remaining pairs within the lineage is assigned one at a time, using the following rule: If a pair (i, j) is related at 0.5 (i.e. they are full siblings, since they are in the same generation), then the two agents share

the same relations to all other agents in the lineage. As result, all newly-assigned siblings inherit the existing relations of their new sibling. Whenever the relatedness between a new pair of agents breaks this rule, for example because unequal relatedness to a third agent has already been assigned to each of the i, j agents in the new pair, the relatedness of the new i, j pair is resampled until it is less than 0.5.

4. The agent ids of all relatives of agent i at time $t = 1$ (i.e., the entirety of agent i's lineage) are assigned to agent i's C_i parameter, which is primarily for convenience and contains the set of all known relatives of i.

Relatedness between agents in different lineages is initialized at 0, implying that lineages are completely unrelated at time $t = 1$. However, note that intermarriage between members of different lineages is expected as the model progresses in time, so this distinction breaks down quickly.

All other variables used to characterize individual agents (Appendix B) vary at a daily level, and will be described in the following section.

Model Algorithm for $t \geq 1$ (Post-Initialization)

At the start of each $t \geq 1$ iteration (after initialization, if $t = 1$), the following steps happens in sequence.

Step 1. Mortality. At the start of each iteration t, an agent may die if a sample from the Bern($\mu_{i,t}$) distribution is equal to 1. The mortality rate $\mu_{i,t}$ for each agent is a sum of their "baseline" mortality (which comes from a user-specified mortality table) and need-based changes to mortality $\Delta\mu_{i,t-1}$ at time $t - 1$. (When $t = 1$, $\Delta\mu_{i,t-1} = 0$ is assumed.) However, death is not guaranteed even when Bern($\mu_{i,t}$) $= 1$. Instead, the number of agents who die at each iteration can only

be as large as $\bar{\mu}_{\max}$. If $\text{Bern}(\mu_{i,t})$ evaluates to 1 for more than $\bar{\mu}_{\max}$ agents, then the $\bar{\mu}_{\max}$ agents with the highest mortality rates are selected to die. To keep the overall population at a fixed size, agents who die will be replaced later on in the iteration, either by a newborn agent "born" to adult parents or by an adult agent who is "recruited" from outside of the population. If i happens to be the spouse of another agent j, then agent j becomes single (i.e., the set $S_{j,t}$ containing the index of j's spouse becomes an empty set). Similarly, if a dependent juvenile agent dies, then their id variable is removed from their live parents' sets of dependents ($D_{i,t}$).

It is worth noting that the enforcement of an upper limit on mortality at each time step is an unrealistic constraint which does not directly account for seasonal or prolonged famine, environmental catastrophes, epidemics, or other external, population-level drivers of increased mortality. However, one would expect population size to vary over time in response to variation in such "extrinsic mortality." Under such a scenario, replacement births would not likely match the number of deaths, especially in cases where the population-level event causing increased mortality impacted fertility, e.g. via nutritional deficits. However, since population size is already held fixed in the current model, the simplifying assumption of bounded mortality assists with maintaining population structure. In particular, if population size is fixed but mortality numbers are not, individual fertility rates can become unrealistically high very quickly. For example, early versions of the model without fixed mortality were occasionally subject to "death spirals," where too many adults in the population died too quickly by random chance. As result, the remaining population was comprised of too many juvenile dependents assigned to too few parents, and adult production was not high enough to overcome this sudden influx of non-producing "infants." This, in turn, caused most juvenile agents to die at one iteration, only to be reborn at the next. Adult agents also died more frequently,

since individual adult production was often not high enough to cope with the resource needs of an influx of additional dependents per adult. That is, because of a constraint imposed by the model whereby adult agents must pool resources with their offspring when they have a combined deficit, adult agents with many newborn (non-producing) offspring could not keep up with the resource demands. Hence, the probability that the remaining adults would die increased, leading to additional non-producing dependents entering the model. Functionally, this cyclical pattern was equivalent to population collapse, and computationally, it caused the model to crash. Imposing a constraint on maximum deaths per day avoids this issue. (From a theoretical point of view, one might reasonably justify this mortality constraint by considering that individuals unable to secure sufficient resources within their local community might seek out cooperative interactions with acquaintances from outside of the local community.)

Step 2. Marriage. Next, if agent i's age $a_{i,t}$ is greater than or equal to their age at first possible reproduction α_i and they do NOT have a living spouse (i.e. $S_{i,t} = \{\varnothing\}$), then they may be paired at random with another single adult j of the opposite sex γ_j, who is also at least α_j years of age and who has relatedness to agent i of $r_{i,j} < 0.125$. Assignment of spouses is performed by randomly selecting one agent at a time from the subpopulation of single, reproduction-age agents, then identifying the set of all other singletons who fit i's requirements for a spouse. Provided that at least one agent exists who fits i's spouse requirements, the index of that agent is added to $S_{i,t}$ and both agents are removed from the pool of marriageable agents. On the other hand, if no agent exists that fits i's requirements for a spouse, then i remains single and is simply removed from the pool of marriageable agents at time t. Then, a new agent is randomly selected from the remaining pool of marriageable agents, and they are assigned a spouse (or not assigned a spouse)

in the same manner as i. The process repeats until there are no agents left in the marriageable agent pool. All agents with an assigned spouse are considered "reproductive pairs," and can being producing offspring immediately.

Step 3. Independence. An agent who is considered "independent" produces all of their own resources each day and is not dependent on either parent for additional resources. Independent agents are also able to directly seek out cooperative partners when in need of additional resources. Once independent, an agent remains in this state for the rest of their life. Agents automatically become independent once they reach age $a_{i,t} = \bar{\delta}_m$, where $\bar{\delta}_m$ is age at maturity (adulthood) within the population. Additionally, if a juvenile agent is at age $a_{i,t} < \bar{\delta}_m$ but they take a spouse (that is, they are at least age α_i and there is an appropriate spouse available during Step 2), the agent also becomes independent. This second case is a simplifying assumption that avoids coding issues associated with nested dependency structures between grandparents, dependent parents, and offspring.

In the case that both parents of a dependent juvenile agent die in the mortality step (Step 1), the juvenile agent also becomes independent, irrespective of age. On the other hand, if only one parent of a juvenile agent dies and the other remains alive, the juvenile agent continues to be dependent on the remaining live parent (but not the live parent's subsequent spouse). In the former scenario, juvenile agents become independent – rather than simply dying along with their parents – because becoming independent allows juvenile agents a chance to survive to adulthood. That is, independent juvenile agents can directly seek out resource donations from candidate cooperative partners to cover their own daily resource shortfalls. This ability to solicit extra resources is particularly beneficial for older juvenile agents (i.e., $a_{i,t} \geq \bar{\delta}_d$) who are already producing some of their own resources each day and may only need a little extra help on occasion. However, inde-

pendence is less likely to benefit juvenile agents who are too young to produce any resources on their own, since failure to receive a donation covering their entire daily need will substantially increase these agents' likelihood of mortality. Hence, in most cases, very young independent agents die soon after the death of both parents.

Step 4. Reproduction and Recruitment. For each agent that dies in Step 1, a replacement agent is initialized so that the population remains at a fixed size. This is done in one of two ways: Either a new juvenile agent is "born" to a randomly-selected reproductive pair in the population with Bernoulli probability p_{birth}, or an unrelated adult agent is "recruited" into the population from somewhere else with Bernoulli probability $1 - p_{\text{birth}}$. I will describe each of these two methods in turn.

First, suppose that n deaths are recorded in time t. Then n replacement agents must be initialized. Suppose that n_b of these replacement agents will be born, where $n_b \leq n$. The following algorithm is used to assign these births to reproductive pairs in the population: If the number of reproductive pairs is greater than or equal to the number of births n_b, then n_b of these couples are randomly sampled and assigned one new dependent each. If, on the other hand, n_b is greater than the number of reproductive pairs, then new dependent agents are distributed equally across all couples until the remainder of unassigned births is less than the number of couples. At this point, the remaining new dependents are assigned randomly to the reproductive pairs, as before. For example, if n_b is greater than the number of couples, but less than two times the number of couples, then each couple is assigned at least one dependent and some couples randomly get two dependents – "twins." (However, note that this second scenario may happen infrequently or not at all, depending on how $\bar{\mu}_{\max}$ is set by the user.)

New dependents are initialized ("born") at age $a_{i,t} = 0$. The lineage k_i of the individual is assigned as the lineage of their mother and remains fixed throughout life. (At the moment, lineage is simply a convenience variable, but it may be used more intentionally in future versions of the model.) Newborn agents' generation g_i is taken to be 1 plus the minimum of their parents' generations. Newborns are randomly assigned a sex $\gamma_i \sim \text{Bern}(p_{\gamma,M})$, where 0 is female, 1 is male, and the probability of being male is $p_{\gamma,M}$. Age at first possible reproduction, α_i, is drawn randomly from the raised cosine distribution with mean $\bar{\alpha}$ and spread s_α that is specified by the user (i.e., the same one used during model initialization). Each of the new agent's "willingness to cooperate" traits, $\omega_{i,B}$, $\omega_{i,R}$, and $\omega_{i,K}$, are individually inherited from one parent or the other with 0.5 probability, but they are inherited with a very small amount of introduced noise. This noise is also introduced by means of a raised cosine distribution: A newborn's trait is drawn from a raised cosine distribution in which the mean is the value of the inherited trait and the spread is fixed at 0.005. (The value of the distribution's spread is fixed and was arbitrarily selected as a small value in order to constrain the amount of trait "mutation" that can occur from parent to offspring.)

A measure of inbreeding (b_i) is calculated for each newborn agent, using the coefficient of inbreeding measurement proposed by Sewall Wright (1922). Specifically, I utilize Wright's first formulation of the coefficient, which only requires knowledge of the inbreeding coefficients of an agents' parents and the coefficient of relatedness ($r_{M,F}$) between the two parents. Afterwards, the coefficient of relatedness $r_{i,j}$ between dependent i and every other agent j in the current population is calculated. This calculation is performed via an iterative modification of the method proposed by Sewall Wright (1922), which takes advantage of previously-calculated relatedness pathways in the agent population. This iterative approach allows for

the removal of all deceased agents more than a set number of generations back from the relatedness matrix of the model, which drastically improves computational efficiency while still allowing all hereditary information to be included in relatedness calculations (since removed agents' pathways are captured by agents in more recent generations).

Before I move onto describing adult agent recruitment, I will briefly digress again: It is important to note that the coefficient of relatedness between two agents, as formulated by Wright, only includes genetic relatedness by *descent* (Wright, 1922). That is, it does not account for genetic similarity – and the resultant indirect fitness benefits – that might arise through evolutionary and random processes other than common descent from a shared ancestor. For simplifying purposes, however, I only explicitly included genetic relatedness by descent in the measure $r_{i,j}$ and chose to ignore any other potential sources of genetic similarity, both because of the complexity of accounting for such similarity and because of the lack of theoretical development surrounding the impacts of non-descent genetic relatedness on cooperation. Nonetheless, such processes are implicit in the cooperation traits themselves: Two wholly unrelated agents (by Wright's measure) may still have highly similar cooperation traits, even though they did not arrive at this similarity via a shared common ancestor. As a result of the way that I modeled cooperation, such similarity will not factor into the probability that a cooperation event occurs between two agents. Even so, it is important to acknowledge that genetic similarity between agents (with respect to cooperation traits) may occur in the ABM through purely non-descent processes.

To return to the matter at hand, now suppose that n_r out of the n replacement agents for deaths occurring in Step 1 are determined to be new recruited agents, rather than newborns. Recruited agents are initialized in a process simi-

lar to that used during population initialization. Specifically, they are assigned a randomly-generated age $a_{i,t} = \alpha_i$ (where α_i is their age at first possible reproduction), and they are initialized as independent. Sex is assigned randomly according to the Bernoulli probability $p_{\gamma,M}$, and the recruited agents have no spouse. A recruited agent's inbreeding coefficient b_i is set to 0, and importantly, their relatedness to all existing agents in the population (including any other agents that are initialized concurrently) is set to zero. Their lineage k_i is set to a unique number not already included in the model. To avoid inconsistencies in generation number between older live agents in the population and the newly-recruited agent, their generation g_i is set to the maximum generation in the currently-alive adult population. Lastly, each of the recruited agent's cooperative traits are sampled from a raised cosine distribution with mean equal to the mean value of the trait in the current population and spread equal to the user-specified spread (s_α) that is used for initializing agents at the start of the model. This method yields recruited agent traits that are not fully constrained by what is currently observed in the population, but are also not wildly different than what is observed in the current population. Recruited adult agents' traits are selected in this manner to avoid artifically diminishing the variance of the ABM's current cooperative trait distributions while, at the same time, limiting the influence of newly recruited agents' traits on the existing trait distributions.

Step 5. Gross Daily Need. The next measure that is calculated is an agent's "gross" level of need at the particular time t. Given an agent i's age $a_{i,t}$, their age-specific daily gross resource need is the amount of resources that they require on a specific day *before* discounting by any individual resource production or provisioning from others. Using the base level of daily need at birth $\bar{\eta}_\varnothing$, the adult level of daily need $\bar{\eta}_m$ (which is constant for all adults in the population), and the

age at adulthood $\bar{\delta}_m$, an individual i's age-specific gross need at time t is calculated as

$$\bar{\eta}_{a_{i,t}} = \bar{\eta}(a_{i,t} \mid \bar{\eta}_{\varnothing}, \bar{\eta}_m, \bar{\delta}_m) = \bar{\eta}_{\varnothing} + \frac{\bar{\eta}_m - \bar{\eta}_{\varnothing}}{\bar{\delta}_m} y, \tag{1}$$

for $y = \min\{a_{i,t}, \bar{\delta}_m\}$. The value of $\bar{\eta}_{a_{i,t}}$ represents the age-specific amount of resources that an agent i requires at time t in order to have a mortality rate at time $t+1$ that is no greater than their age-specific baseline value.

Step 6. Gross Daily Resource Production. On a given day t, each agent in the community who is at least age $\bar{\delta}_d$ (i.e. they are no longer fully dependent on their parents) will also produce an age-scaled, random amount of resources. The amount of resources that an agent produces may be less than or greater than their own daily age-specific need ($\bar{\eta}_{a_{i,t}}$; see Step 5 above), indicating a daily deficit or a daily surplus in resources, respectively. If an agent i is an independent agent and generates a surplus of resources, they will direct some or all of this surplus to supplementing their and and their dependents' needs. Production $\rho_{i,t}$ for agent i on day t is generated by scaling random draws from the Beta$(\alpha_\rho, \beta_\rho)$ distribution, where $\alpha_\rho > 1$ and $\beta_\rho > 1$. (If the two shape parameters are less than or equal to 1, the mode of the Beta$(\alpha_\rho, \beta_\rho)$ does not have a closed form. $\alpha_\rho = 2$ and $\beta_\rho = 2.5$ are good options for values.) To determine the age-specific scaling factor, an age-specific distributional *mode* $\widetilde{\rho}_{a_{i,t}}$ for agent i must first be calculated:

$$\widetilde{\rho}_{a_{i,t}} = \widetilde{\rho}(a_{i,t} \mid \widetilde{\rho}_m, \bar{\delta}_d, \bar{\delta}_m) = \begin{cases} 0, & \text{if } a_{i,t} < \bar{\delta}_d \\ \dfrac{\widetilde{\rho}_m}{\bar{\delta}_m - \bar{\delta}_d}(a_{i,t} - \bar{\delta}_d), & \text{if } \bar{\delta}_d \leq a_{i,t} \leq \bar{\delta}_m \\ \widetilde{\rho}_m, & \text{if } a_{i,t} > \bar{\delta}_m. \end{cases} \tag{2}$$

In Equation 2 above, $\widetilde{\rho}_m$ is the population mode of adult-aged resource production. If an agent has not yet reached the population age of maturity ($\bar{\delta}_m$) but

is no longer fully dependent on their parents ($a_{i,t} \geq \bar{\delta}_d$), then their mode is linearly scaled by age. This implies that agents between the ages of $\bar{\delta}_d$ and $\bar{\delta}_m$ randomly produce a daily quantity of resources which is sampled from an age-specific Beta distribution with mode $\widetilde{\rho}_{a_{i,t}}$ (described below), where production reaches an adult level with mode $\widetilde{\rho}_{\{a_{i,t} \geq \bar{\delta}_m\}} = \widetilde{\rho}_m$ when an agent reaches the age of maturity/adulthood, $\bar{\delta}_m$. On the other hand, if an agent is fully dependent on their parents (i.e. $a_{i,t} < \bar{\delta}_d$), then their resource production mode is zero and they are unable to produce any resources.

To further understand the Beta distribution governing daily resource production, first observe that $\mu_\rho = \alpha_\rho/(\alpha_\rho + \beta_\rho)$ is the mean of the Beta$(\alpha_\rho, \beta_\rho)$ distribution, while $m_\rho = (\alpha_\rho - 1)/(\alpha_\rho + \beta_\rho - 2)$ is its mode (for $\alpha_\rho, \beta_\rho > 1$). Then, for a scaling factor $s_{\rho,a_{i,t}} = \widetilde{\rho}_{a_{i,t}}/m_\rho$ and a random draw $y \sim \text{Beta}(\alpha_\rho, \beta_\rho)$, the daily random production of agent i at time t can be calculated as:

$$\rho_{i,t} = \rho(\widetilde{\rho}_{a_{i,t}} \mid t, \alpha_\rho, \beta_\rho) = s_{\rho,a_{i,t}} \cdot y = \frac{\widetilde{\rho}_{a_{i,t}}}{m_\rho} y. \tag{3}$$

That is, the random component of each agent's daily resource production is first sampled from the Beta$(\alpha_\rho, \beta_\rho)$ distribution. Then, since this random draw is bounded between 0 and 1, it is scaled up or down to correspond with the agent's age-specific mode of production. (Incidentally, the age-specific mean scaled value of production is then given by $\bar{\rho}_{a_{i,t}} = s_{\rho,a_{i,t}} \cdot \mu_\rho$. For $\alpha_\rho = 2$ and $\beta_\rho = 2.5$, for example, $\bar{\rho}_{a_{i,t}}$ will be approximately $\bar{\rho}_{a_{i,t}}/\widetilde{\rho}_{a_{i,t}} \approx 1.11$ times $\widetilde{\rho}_{a_{i,t}}$.)

One last thing to mention is that the above process implies that long-term economic inequality will not develop within the ABM. That is, because resource production is randomly-generated and surpluses do not carry over from one day to the next, no agent in the population can become resource-wealthy or resource-poor.

Step 7. Initial Resource Calculations. After calculating age-specific daily need and production for all agents, the ABM next calculates each agent's daily "net need." This is the amount of an agent's age-specific daily gross need that is *not* covered by their individual resource production for that day, and it may be positive (indicating a net deficit in resources) or negative (indicating a net surplus in resources). If an agent's net need is positive, meaning that they have not produced enough resources to cover their daily need, then they will need to receive an amount of resources equivalent to their net need from another agent in order to satisfy their gross daily need on that day. However, only independent/adult agents in the population can engage in cooperative events with other agents. As result, parent agents are responsible for at least half of their dependents' net needs. To illustrate, suppose that agent i has dependents $d \in D_{i,t}$ at time t. In the case where the other parent of agent i's dependents is alive at time t, agent i's net need is calculated as

$$\eta_{i,t} = \bar{\eta}_{a_{i,t}} - \rho_{i,t} + \frac{1}{2} \sum_{d \in D_{i,t}} (\bar{\eta}_{a_{d,t}} - \rho_{d,t}); \tag{4}$$

that is, agent i's net need is their gross need minus their own resource production (individual net need), plus half of their dependents' net needs. However, if the other parent of a dependent is not alive, than agent i is responsible for the entirety of that dependent agent's net need. Thus, defining $p_{d_i,t}$ as the proportion of dependent d's net need that adult agent i is responsible for at time t, Equation 4 can be generalized as

$$\eta_{i,t} = \bar{\eta}_{a_{i,t}} - \rho_{i,t} + \sum_{d \in D_{i,t}} p_{d_i,t} \cdot (\bar{\eta}_{a_{d,t}} - \rho_{d,t}). \tag{5}$$

Equation 5 accounts for cases where the other parents of some or all of agent i's dependents are deceased, leaving agent i responsible for all of these dependents' net needs. (Also, note that if an agent has no dependents, the summation term above becomes zero and the agent is simply responsible for their own net need.)

It bears repeating that dependents do not independently engage in cooperative transfers with other agents in the population. Since dependents don't produce any resources until age $\bar{\delta}_d$, and only produce an age-scaled amount after that, they must receive all supplemental provisioning solely from their parents. Then, when a juvenile agent has two live parents, each parent is responsible for half of their dependents' daily net need. However, parents do *not* automatically pool resources with each other. It is possible that the parents of a dependent agent will be paired together during the cooperative step (outlined below) – in which case a transfer of resources from one parent to the other could, in fact, occur – but, since this requires that the two parent agents are randomly paired to cooperate *and* are willing to cooperate, the probability of this occurrance is typically quite low. While somewhat unrealistic, given that most human small-scale subsistence populations have a sexual division of labor whereby one parent is responsible for more than half of their spouse's and offsprings' material provisioning, this decision simplifies the calculations below considerably. It may be partially justified through the existence of domain- and sex-specific resource provisioning of offspring that differs between parents, e.g., where one parent provides hunted goods and the other provides foraged goods (e.g., Gurven & Hill, 2009; Hill & Hurtado, 2009; Kaplan et al., 2000).

Step 8. Cooperative Transfers. Once the net need has been calculated for each adult/independent agent, adult/independent agents with a net deficit may solicit resource donations from other adult/independent agents with a net surplus. Adult and independent agents, as the only cooperatively-engaged members of the

population, are classified as "recipients" if their net need $\eta_{i,t}$ is greater than 0 (indicating a deficit) or as "donors" if their net need $\eta_{i,t} \leq 0$ (indicating a surplus). Note that the probability of exactly $\eta_{i,t} = 0$ is infinitesimally small, so how an agent is classified in this situation is not of particular concern.

The cooperative step that occurs during each iteration of the model proceeds as thus: First, general calculations are performed for both donors and recipients to determine the changes to net need and mortality that would occur if a cooperative transfer of resources was not successful. Then, donors and recipients are randomly-paired, and the success of a potential cooperative event between each donor/recipient pair is evaluated. A general outline of the initial mortality calculations for recipients and donors is described below in Steps 8.1 and 8.2, followed by a description of how cooperative transfer probabilities are calculated and evaluated in Step 8.3.

Step 8.1. Initial Calculations for Recipients. If an adult agent j's net need is positive, then this indicates that they and their dependents have a remaining resource deficit after accounting for their daily production. That is, adult agent j and their offspring have jointly produced less resources than adult agent j requires to satisfy their own daily need and the proportion of their dependents's daily need that they are responsible for. Agent j may supplement their dependents' net need with their own production via resource pooling, and/or they may receive a resource transfer from another individual in the population to cover some or all of their daily shortfall.

If adult agent j is a recipient and receives *no* resource assistance from another adult agent, then they will pool resources produced by themselves and any dependents, which will then be distributed back to j and their dependents as weighted shares of their total resources. Resource distribution weights can be defined as

$$\dot{w}_{j,t} = \frac{\bar{\eta}_{a_{j,t}}}{\bar{\eta}_{a_{j,t}} + \sum_{d_j \in D_{j,t}} p_{d_j,t} \cdot \bar{\eta}_{a_{d_j,t}}} \quad \text{for adults, and} \tag{6}$$

$$\dot{w}_{d_j,t} = \frac{p_{d_j,t} \cdot \bar{\eta}_{a_{d_j,t}}}{\bar{\eta}_{a_{j,t}} + \sum_{d_j in D_{j,t}} p_{d_j,t} \cdot \bar{\eta}_{a_{d_j,t}}} \quad \text{for dependents,} \tag{7}$$

where $p_{d_j,t}$ is the proportion of each dependent's need that agent j is responsible for. Then, given these weights (and assuming no additional resources are transferred to j), each of j and their dependents $d_j \in D_{j,t}$ would receive a share of the pooled resources equal to

$$\dot{\rho}_{x,t} = \Big(\rho_{j,t} + \sum_{d_j \in D_{j,t}} p_{d_j,t} \cdot \rho_{d_j,t}\Big)\dot{w}_{x,t} \quad \text{for } x \in \{j, D_{j,t}\}. \tag{8}$$

It follows that if agent j receives no resource transfers from other adult agents in the population, then the change in *proportional* need for j and their offspring d_j after redistributing the pooled resources would be

$$\Delta\eta_{j,t} = \frac{\bar{\eta}_{a_{j,t}} - \dot{\rho}_{j,t}}{\bar{\eta}_{a_{j,t}}} \quad \text{for agent } j, \text{ and} \tag{9}$$

$$\Delta\eta_{d_j,t} = \frac{p_{d_j,t} \cdot \bar{\eta}_{a_{d_j,t}} - \dot{\rho}_{d_j,t}}{\bar{\eta}_{a_{d_j,t}}} \quad \text{for each of their dependents.} \tag{10}$$

Now, for recipients, $\Delta\eta$ in Equations 9 and 10 above corresponds to the proportional resource shortfall that j and their offspring experience at time t. If one were to assume that mortality and need had a 1:1 linear relationship in this population – that is, if the proportional shortfall in an agent's daily resource production resulted in an equivalent proportional increase in that agent's daily mortality rate at the start of the next time step – then $\Delta\eta$ would simply be an agent's increase in

mortality at time $t + 1$ over their baseline mortality rate. However, initial model tests demonstrated that this assumption is problematic for two reasons:

1. Assuming a 1:1 mortality cost as result of a daily resource shortfall is unrealistically severe, leading to premature adult population crashes. It may be noted that, at least in human populations, a 50% resource shortfall on a single day will *not* produce a 50% increase in mortality during that day. Instead, humans can survive for a number of days with inadequate or even zero resources. Hence, it is necessary to scale the impact of daily resource shortfalls on daily mortality rates such that they are more in line with observed outcomes, or else cooperative behaviors within the population will not have a sufficient chance to evolve before the population crashes.

2. The assumption of a linear relationship between resources and mortality (no matter the slope) does not account for the diminishing benefits of extra resources as an individual ages. Evidence from life history theory suggests that, once an individual reaches an advanced age, their body will repair itself less effectively and their mortality rate will increase quickly (Williams et al., 2006). Importantly, changes to mortality during this period of somatic decline are not driven by resource shortfalls, but rather by the body breaking down. Thus, while higher levels of mortality during this time period can certainly be affected by the availability of resources, it is not possible to fully overcome age-related mortality via increased resource access. As result, the relative benefit of holding onto extra resources (as opposed to sharing them with younger, related individuals) should diminish as an individual ages. There is a substantial body of ethnographic evidence suggesting that this is typical behavior in humans, at least for females (Hawkes &

Coxworth, 2013; Sear & Mace, 2008). Applying the above argument to this model, it should be expected that older donor agents will be more willing to share their extra resources with younger relatives, since keeping these resources should do the older agent little good. However, a direct implication of assuming a linear relationship between resources and mortality is that surplus resources can fully overcome mortality, even during old age. Simply put, an older agent under this scenario would be highly motivated to channel their surplus resources into overcoming their high mortality rate. As a result, an older agent would *never* share with younger relatives, because the benefit (in terms of reduction to mortality) that the elder agent would gain from keeping the resources would always outweigh the benefit that a younger relative would gain from receiving them. To overcome this issue, the relationship between resources and mortality must be non-linear and incorporate diminishing returns as age (and mortality) increase.

Addressing the above concerns in this ABM, resource shortfalls (as well as resource surpluses; see next section) are applied to an agent's mortality rate on the logit scale. The relative effect of resource shortfalls on mortality are further scaled by k_μ to allow for finer control over the impact of resources on mortality. The result of this logit calculation is then transformed back to the probability scale, after which the actual change in an agent's mortality rate is calculated as the difference between the updated rate and the agent's original, "baseline" mortality rate. Since an agent's mortality is evaluated at the start of each time iteration (i.e. well before the resource/need calculations and the cooperative step), an increase in mortality as result of resource shortfalls at time t impacts an agent's chance of dying at the start of time $t + 1$.

To express the change in mortality mathematically, recall that an agent's age-specific "baseline" mortality at time $t+1$ is $\bar{\mu}_{a_{j,t+1}}$. Then, assuming momentarily that no cooperative help is received from other adult agents at time t, the change to mortality at time $t+1$ for agent j and their offspring $d_j \in D_{j,t}$ is

$$\Delta\mu_{x,t} = \frac{1}{1+\exp\left\{-\left[\text{logit}(\bar{\mu}_{a_{x,t+1}}) + k_\mu \Delta\eta_{x,t}\right]\right\}} - \bar{\mu}_{a_{x,t+1}} \quad \text{for } x \in \{j, D_{j,t}\}. \tag{11}$$

If, after the cooperative step below, agent j has indeed received no resources from another agent to cover this shortfall, then the mortality rate of agent j is simply the sum of their mortality rate change at time t and their baseline mortality rate at time $t+1$, i.e.

$$\mu_{j,t+1} = \bar{\mu}_{a_{j,t+1}} + \Delta\mu_{j,t} = \frac{1}{1+\exp\left\{-\left[\text{logit}(\bar{\mu}_{a_{j,t+1}}) + k_\mu \Delta\eta_{j,t}\right]\right\}}. \tag{12}$$

The mortality rate of agent j's offspring is similarly calculated, but also factors in the change to mortality brought about by the other parent when applicable.

On the other hand, if agent j *does* receive some quantity of resources $\rho_{C,j,t}$ during a cooperative interaction with another adult agent, then the change in mortality for agent j and their offspring will be modified accordingly. This, in turn, will affect their final mortality rates at the start of time $t+1$. This is discussed further in Steps 8.3 and 9 below.

Step 8.2. Initial Calculations for Donors. Suppose that agent i is some other adult agent in the population. If $\eta_{i,t} < 0$, it means that an adult agent i and their dependents have generated enough resources to cover their basic daily needs. However, recall that there is a "baseline" level of age-specific mortality, $\bar{\mu}_{a_{i,t}}$, which still exists even after an agent meets its daily needs. If agent i has a daily

surplus at time t (indicated as $\eta_{i,t} < 0$), they may choose to invest some or all of these surplus resources in discounting their and their dependents' base mortality rate ($\bar{\mu}_{a_{i,t+t}}$) at time $t+1$. Alternatively, agent i may choose to give some/all of this away to some other recipient agent who is experiencing a net shortfall in resources at time t.

To calculate the potential amount that agent i may invest in discounting their and their dependents' mortality at time $t+1$, I will temporarily assume that mortality and resources are linearly related. (This assumption will be modified below, in line with the previous section's argument for a non-linear relationship between mortality and resources.) Under this temporary assumption, first observe that survival at any given age (denoted $s = 1 - \mu$) and need at that age are linearly dependent on one another. That is, a change from $s_0 = 0$ to $s_1 = 1 - \bar{\mu}_{a_{x,t+1}}$ (where $\bar{\mu}_{a_{x,t+1}}$ is the agent's baseline mortality rate at time $t+1$) happens when some individual x's available resources at time t increase from 0 to $\bar{\eta}_{a_{x,t}}$ (their age-specific need). Therefore, for any agent x in the model, the rate of change of survival with respect to resources is

$$\frac{ds_{x,t}}{d\eta_{x,t}} = \frac{1 - \bar{\mu}_{a_{x,t+1}}}{\bar{\eta}_{a_{x,t}}}. \tag{13}$$

Inverting this relationship, it follows that the rate of change of *resources* with respect to *survival* is

$$\frac{d\eta_{x,t}}{ds_{x,t}} = \frac{\bar{\eta}_{a_{x,t}}}{1 - \bar{\mu}_{a_{x,t+1}}}. \tag{14}$$

With this in mind, the amount of resources required to increase survival by some specific amount s is

$$\tilde{\rho}_{x,t} = \frac{d\eta_{x,t}}{ds_{x,t}} s = \frac{\bar{\eta}_{a_{x,t}}}{1 - \bar{\mu}_{a_{x,t+1}}} s. \tag{15}$$

Specifically, in order to increase an agent's survival at time $t+1$ from $1 - \bar{\mu}_{a_{x,t+1}}$ to 1 under this linear assumption (i.e., decrease mortality from $\bar{\mu}_{a_{x,t+1}}$ to 0), the target change in survival is $s = \bar{\mu}_{a_{x,t+1}}$. To accomplish this change, the extra resources needed per individual x at time t are

$$\tilde{\rho}_{x,t} = \frac{\bar{\eta}_{a_{x,t}}}{1 - \bar{\mu}_{a_{x,t+1}}} \bar{\mu}_{a_{x,t+1}}. \tag{16}$$

Thus, *under the temporary assumption of a linear relationship between resources and mortality,* the extra resources that a candidate donor i needs at time $t+1$ in order to ensure perfect survival of themselves – as well as half or all of the increase in their offsprings' survival to 1 (depending on whether other parent is alive) – is:

$$\dot{\rho}_{i,t} = \tilde{\rho}_{i,t} + \sum_{d_i \in D_{i,t}} p_{d_i,t} \cdot \tilde{\rho}_{d_i,t}. \tag{17}$$

where $p_{d_i,t}$ is the amount of a dependent d's need that agent i is responsible for (discussed in Step 8.1 above).

Equations 16 and 17 can be used to generate weights used in determining how agent i should apportion the surplus resources available to themselves and their dependents. Specifically,

$$\dot{w}_{i,t} = \frac{\tilde{\rho}_{i,t}}{\dot{\rho}_{i,t}} \quad \text{for adult } i, \text{ and} \tag{18}$$

$$\dot{w}_{d_i,t} = \frac{p_{d_i,t} \cdot \tilde{\rho}_{d_i,t}}{\dot{\rho}_{i,t}} \quad \text{for dependents } d_i \in D_{i,t}. \tag{19}$$

If i does *not* share resources with another agent at time t, then they will redistribute their surplus resources (i.e., the absolute value of their net need, $\eta_{i,t}$, which is negative and includes their dependents' net need) back to themselves and their dependents according to the weights defined above in Equations 18 and 19. That is, agent i will receive $(\text{-}\eta_{i,t})\dot{w}_{i,t}$ resources, and their dependents will receive $(\text{-}\eta_{i,t})\dot{w}_{d_i,t}$ resources. This, in turn, implies that agent i's change in proportional need at time $t+1$ (assuming no sharing) is

$$\Delta\eta_{i,t} = \eta_{i,t}\dot{w}_{i,t}\frac{ds_{i,t}}{d\eta_{i,t}} = \eta_{i,t}\dot{w}_{i,t}\frac{1-\bar{\mu}_{a_{i,t+1}}}{\bar{\eta}_{a_{i,t}}}. \tag{20}$$

The change in i's dependents' resources at time $t+1$, as brought about by i's surplus resources (and not including the other parent's contribution when applicable), is then

$$\Delta\eta_{d_i,t} = \eta_{i,t}\dot{w}_{d_i,t}\frac{ds_{d_i,t}}{d\eta_{d_i,t}} = \eta_{i,t}\dot{w}_{d_i,t}\frac{1-\bar{\mu}_{a_{d_i,t+1}}}{p_{d_i,t}\cdot\bar{\eta}_{a_{d_i,t}}} \quad \text{for } d_i \in D_{i,t}. \tag{21}$$

By substituting the definitions of $\dot{w}$ and $\widetilde{\rho}$ (Equations 18/19 and Equation 16, respectively) into the two definitions of $\Delta\eta$ (Equations 20 and 21), it can be shown that they both reduce to the following:

$$\Delta\eta_{x,t} = \frac{\eta_{i,t}}{\dot{\rho}_{i,t}}\cdot\bar{\mu}_{a_{x,t+1}} \quad \text{for } x \in \{i, D_{i,t}\}. \tag{22}$$

That is, since $\widetilde{\rho}_{i,t}$ (the total resources required to reduce to zero both agent i's mortality and the portion of their dependents' mortality that they are responsible for) is already scaled by the baseline mortality rate of agent i and their offspring, it is sufficient to weight i's redistribution of the actual surplus resources

available to them (-$\eta_{i,t}$) by $\bar{\mu}_{a_{x,t+1}}/\widetilde{\rho}_{i,t}$. Then, the agents' change in proportional need is simply the negative of the redistributed resources. Importantly, observe that the above changes in proportional need will both be negative, since $\eta_{i,t} < 0$ for donors (corresponding to a surplus in resources). It is also important to note that $\Delta\eta_{x,t}$ is not, strictly speaking, a true change in *proportion* – it can be smaller than -1. However, in keeping with the language of the previous section, I will continue to refer to it as thus.

As with the net need of recipients, $\Delta\eta$ should not be used as a direct mortality change here due to the issues resulting from assuming a linear relationship between resources and mortality. Instead, the change in mortality resulting from the net surplus of i and their dependents, *when they do not share resources with any recipient agent,* can be calculated as

$$\Delta\mu_{x,t} = \frac{1}{1+\exp\Big\{-\big[\text{logit}(\bar{\mu}_{a_{x,t+1}}) + k_\mu \Delta\eta_{x,t}\big]\Big\}} - \bar{\mu}_{a_{x,t+1}} \quad \text{for } x \in \{i, D_{i,t}\}. \tag{23}$$

Thus, if agent i does not cooperatively share at time t, their mortality rate at time $t+1$ is

$$\mu_{i,t+1} = \bar{\mu}_{a_{i,t+1}} + \Delta\mu_{i,t} = \frac{1}{1+\exp\Big\{-\big[\text{logit}(\bar{\mu}_{a_{i,t+1}}) + k_\mu \Delta\eta_{i,t}\big]\Big\}}. \tag{24}$$

Again, the mortality rate of agent i's offspring is similar, but may additionally include the change in mortality brought about by the other parent's contribution (if the other parent is alive).

Alternatively, since i is a donor, they may be randomly paired with a recipient adult j who has a need deficit. If i does indeed share some positive quantity of resources with j, then the change in mortality for themselves and their offspring is

updated accordingly to reflect i's decrease in available surplus resources. See Steps 8.3 and 9 below for more details.

Step 8.3. Cooperative Interaction. Assume that adult agent i is a donor, and adult agent j is a recipient. Suppose that recipient agent j has a net need deficit of $\eta_{j,t}$ at time t, and recall that donor agent i has an inherited cooperation trait vector $\vec{\omega}_i = \{\omega_{i,B}, \omega_{i,R}, \omega_{i,K}\}$, where $\omega_{i,B}$ is i's baseline willingness to cooperate with anyone in the population, $\omega_{i,R}$ is their willingness to cooperate with past reciprocity partners, and $\omega_{i,K}$ is their willingness to cooperate with kin. These three traits will be incorporated into a log-odds (logit) probability characterizing donor i's willingness to transfer resources to recipient j. (More on this in just a moment.)

First, suppose that agent j is related to agent i in some way, and suppose that i is evaluating whether to give some $\rho_{C,j,t} > 0$ amount of resources to j at time t. The benefits and costs of this potential cooperative event can be quantified in terms of how the mortality of the involved agents and their dependents at time $t+1$ would change if this transfer of $\rho_{C,j,t}$ resources actually occurred. If recipient agent j does indeed receive $\rho_{C,j,t}$ resources from agent i, then the updated mortality rate at time $t+1$ for agent j and their dependents $D_{j,t}$ can be calculated on the logit scale as

$$\Delta\mu_{C,x,t} = \frac{1}{1 + \exp\left\{-\left[\text{logit}(\bar{\mu}_{a_{x,t+1}}) + k_\mu(\Delta\eta_{x,t} - \rho_{C,j,t}\dot{w}_{x,t}/\bar{\eta}_{a_{x,t}})\right]\right\}}, \tag{25}$$

where $x \in \{j, D_{j,t}\}$. Specifically, the additional resources $\rho_{C,j,t}$ are incorporated into the new mortality rates of agent j and their dependents as $\rho_{C,j,t}\dot{w}_{x,t}/\bar{\eta}_{a_{x,t}}$, which represents the change (i.e., decrease) in j and their dependents' proportional need that would result from the additional resources, weighted by the proportion of each

agent's daily need that agent j is responsible for. (See Step 8.1 for a discussion of recipient weights and change in proportional need.) Hence, under this scenario, the potential benefit $b_{j,t} \geq 0$ that agent j would gain if they were to receive $\rho_{C,j,t}$ resources from agent i would be the sum of their and their dependents' decrease in mortality at time $t+1$ as a result of these additional resources. That is, agent j's potential benefit would be the increased survival for themselves and their dependents when an additional amount of their daily net need is met at time t. Using Equations 12 and 25, this may be formalized as

$$\begin{aligned} b_{j,i} &= \sum_{x \in \{j, D_{j,t}\}} (\mu_{x,t+1} - \Delta\mu_{C,x,t}) \\ &= \sum_{x \in \{j, D_{j,t}\}} \left[(\bar{\mu}_{a_{x,t+1}} + \Delta\mu_{x,t}) - \Delta\mu_{C,x,t} \right]. \end{aligned} \tag{26}$$

Recall that $\mu_{x,t+1} = \bar{\mu}_{a_{x,t+1}} + \Delta\mu_{x,t}$ for $x \in \{j, D_{j,t}\}$ represents the mortality rates of agent j and their dependents at time $t+1$ if no additional resources are transferred to j and they remain at their original net deficit, as calculated in Equation 12. If the cooperative event between agents i and j fails and nothing is transferred to agent j ($\rho_{C,j,t} = 0$), then $\Delta\mu_{C,x,t}$ in Equation 25 simply reduces to the definition of $(\bar{\mu}_{a_{x,t+1}} + \Delta\mu_{x,t})$ in Equation 12. In this case, $b_{j,i} = 0$, indicating that agent j receives no benefit from their cooperative interaction with agent i.

On the other hand, if j does actually receive $\rho_{C,j,t}$ resources from i, this decreases j's mortality towards their baseline age-specific mortality. However, since $\rho_{C,j,t}$ can be no larger than the minimum of recipient j's net need and donor i's available surplus, j's mortality as a recipient will not decrease below their baseline mortality. As for j's dependents, if the other parent of a dependent d_j is still alive, then the maximum decrease in d_j's mortality rate will be halfway between the dependent's original mortality (with a net need and no cooperative assistance) and

their baseline, age-specific mortality. However, if d_j's other parent is deceased, then d_j's mortality can decrease down to their baseline mortality as result of the $\rho_{C,j,t}$ resources transferred to their live parent, agent j.

Now, in order for agent i to provide j with some amount of resources at time t, they must pay a cost in terms of reduced survivorship for themselves and their own dependent offspring. This cost will depend on how much surplus resources i has available (i.e., $\eta_{i,t}$, translated into a change in mortality for themselves and their offspring), the size of the decrease in this surplus ($\rho_{C,j,t}$) as result of transferring these resources to j, and the final mortality rates for themselves and their dependent offspring at time $t+1$ if they choose not to cooperate. Put simply, the cost that i pays for giving $\rho_{C,j,t}$ resources to j is the sum of the increases in their and their dependents' mortality rates at time $t+1$ as result of the loss of the $\rho_{C,j,t}$ surplus resources, relative to the mortality rates they would have if they kept these resources. If agent i does decide to give $\rho_{C,j,t}$ to agent j, then the mortality rates at time $t+1$ of agent i and their dependents $D_{i,t}$ will be updated as follows:

$$\Delta\mu_{C,y,t} = \frac{1}{1+\exp\left\{-\left[\text{logit}(\bar{\mu}_{a_{y,t+1}}) + k_\mu(\Delta\eta_{y,t} + \rho_{C,j,t}\bar{\mu}_{a_{y,t+1}}/\dot{\rho}_{i,t})\right]\right\}}, \tag{27}$$

where $y \in \{i, D_{i,t}\}$. In Equation 27 above, the loss of these $\rho_{C,j,t}$ resources is incorporated as $\rho_{C,j,t}\bar{\mu}_{a_{y,t+1}}/\dot{\rho}_{i,t}$, which equals the weighted *increase* in proportional need that i and their dependents experience from not having access to these additional resources. Since this increases the mortality rate of i and their dependents, this increase in proportional net need is added to $\Delta\eta_{x,t}$ rather than subtracted from it (as compared to Equation 25). Then, the mortality cost that agent i would incur if they decided to help agent j at time t would equal the sum of their and their

dependents' resulting increases in mortality at time $t + 1$. This may be expressed using Equations 24 and 27 as

$$\begin{aligned} c_{i,t} &= \sum_{y \in \{i, D_{i,t}\}} (\Delta\mu_{C,y,t} - \mu_{y,t+1}) \\ &= \sum_{y \in \{i, D_{i,t}\}} \left[\Delta\mu_{C,y,t} - (\bar{\mu}_{a_{y,t+1}} + \Delta\mu_{y,t})\right]. \end{aligned} \tag{28}$$

In the equation above, $\mu_{y,t+1} = \bar{\mu}_{a_{y,t+1}} + \Delta\mu_{y,t}$ for $y \in \{i, D_{i,t}\}$ represents the mortality rate of i and their dependents if i does not cooperate with j (Equation 24). Note that this will be smaller than $\Delta\mu_{C,y,t}$ (Equation 27); hence, $\mu_{y,t+1}$ is subtracted from $\Delta\mu_{C,y,t}$ so that $c_{i,t}$ will always be a positive quantity. It is worth noting, however, that this cost may be extremely close to zero if i's remaining surplus far outweighs the amount that they can transfer to j. Since the change in mortality is calculated on the logit scale, small decreases in otherwise huge resource surpluses have little effect on overall mortality rates.

Now, the above benefits and costs factor into i's willingness to help a *related* agent j. The cost $c_{i,t}$ and benefit $b_{j,t}$ resulting from i donating to j enhances cooperation in the model via the mechanism of kin selection, as first proposed by Hamilton (1964). Specifically, Hamilton theorized that natural selection should favor behavior in which a donor pays a cost c to provide a benefit b to a recipient, provided that the relatedness between the two individuals (r) – relative to the average relatedness within the population as a whole ($\bar{r}$) – satisfies the inequality $(r - \bar{r})b > c$. This prediction has consistently held up across a wide range of species, from primates to social insects (Allen-Arave et al., 2008; Mitani, 2006; Queller, 2000; Seyfarth & Cheney, 2012), and is a foundational component of life history theory.

Within the current model, kin selection is weighted by an agent's inherited "willingness to cooperate with kin" trait, $\omega_{i,K}$. That is, the probability of coopera-

tion via kin selection is evaluated by $\omega_{i,K}[(r_{i,j} - \bar{r})b_{j,t} - c_{i,t}]$, with $\omega_{i,K}$ moderating the strength of an individual agent's tendency toward cooperating with their kin. In this way, it is possible for kin selection to evolve in the population under certain circumstances, provided that agents with lower $\omega_{i,K}$ traits die out and are replaced with agents with larger $\omega_{i,K}$ traits. Furthermore, since this term is incorporated additively into the *logit* probability of cooperation, it will interact with other forms of cooperation in the model to produce the final probability of cooperation (see below).

As an illustrative example of how kin selection operates in this model, suppose that average relatedness in the population at time t is $\bar{r} = 0.0625$. In this time step, i is selected as a donor and j is selected as a recipient for a potential cooperative event. Agent i's inherited "willingness to cooperate with kin" trait is $\omega_{i,K} = 2$, suggesting that they have a relatively strong preference for cooperating with kin. For simplicity, assume that neither agent has a positive coefficient of inbreeding; then, as siblings, the coefficient of relatedness between the hypothetical i and j is $r_{i,j} = 0.5$. Since $r_{i,t} > \bar{r}$, it follows that kin selection will favor cooperation between i and j as long as the benefit to j is sufficiently larger than the cost to i. With this in mind, suppose that the benefit that j would gain from receiving some quantity $\rho_{C,j,t} > 0$ from i is $b_{j,t} = 2$, while the cost incurred by i for giving this amount to j is $c_{i,t} = 0.5$. Then, $\omega_{i,K}[(r_{i,j} - \bar{r})b_{j,t} - c_{i,t}]$ will evaluate to 0.75. Given that this effect is applied on the logit scale, this will positively impact the final probability of cooperation between i and j. On the other hand, if j's potential benefit is smaller – say, $b_{j,t} = 0.5$ – then $\omega_{i,K}[(r_{i,j} - \bar{r})b_{j,t} - c_{i,t}] = -0.5625$. This will negatively impact the final probability of cooperation between i and j.

In a third scenario, agent i may instead be randomly selected as a candidate donor to some unrelated individual, or to an individual who's relatedness to i is

less than the mean relatedness of the population at time t. In this case, $r_{i,t} - \bar{r}$ would be a negative value, which would heavily discount the benefit to j relative to the cost of i. This would result in $(r_{i,j} - \bar{r})b_{j,t} < c_{i,t}$, yielding a negative value of $(r_{i,j} - \bar{r})b_{j,t} - c_{i,t}$. Hence, $\omega_{i,K}[(r_{i,j} - \bar{r})b_{j,t} - c_{i,t}] < 0$, decreasing the final probability of cooperation. Importantly, it is not possible to overcome $r_{i,t} - \bar{r} < 0$ in this model. Since $c_{i,t}$ is calculated as the difference between i and their offspring's mortality rates when i keeps all surplus resources and their mortality rates when i gives away a portion of their resources – and because this change is evaluated on the logit scale – $c_{i,t}$ will always be positive.

Two additional mechanisms the model can increase the likelihood of cooperation, even when relatedness between a cooperative pair is very low. First, the model incorporates a measure of reciprocity [also called "reciprocal altruism" in the literature; Trivers (1971)], whereby agents may preferentially help other agents who can be reasonably expected to help them in return sometime in the future. Reciprocity is also frequently measured in terms of past interactions, since memory of these past interactions may inform discrimination between those likely to help and those likely to defect in the future. Within the context of this model, a donor agent i may be motivated to help a recipient agent j if i's $\omega_{i,R}$ trait (their inherited willingness to cooperate with reciprocity partners) is large *and* if j has typically helped them in the past. That is, their willingness to cooperate with past reciprocity partners ($\omega_{i,R}$) is scaled by a historical record $H_{i,j,t}$ of past interactions between i and j *where i was the recipient and j was the donor.* This historical record is iteratively calculated as

$$H_{i,j,t} = I_{i,j,t-1} + k_H H_{i,j,t-1}, \tag{29}$$

where $k_H \in [0, 1]$ is a fixed decay parameter that controls how far back i effectively "remembers" their past reciprocity history with j (with greater k_H corresponding to a longer "memory"). $I_{i,j,t-1}$, the other component of $H_{i,j,t}$, is a simple pseudo-indicator function that describes the interaction between i and j at time $t-1$ as follows:

$$I_{i,j,t-1} = \begin{cases} 0, & \text{if } i,j \text{ do not interact at time } t-1, \text{ or if } i \text{ is the donor;} \\ \text{-1}, & \text{if } i,j \text{ interact at time } t-1 \text{ with } i \text{ as recipient, but } j \\ & \text{gives nothing to } i; \\ 1, & \text{if } i,j \text{ interact at time } t-1 \text{ with } i \text{ as recipient, and } j \\ & \text{gives some positive amount to } i. \end{cases} \tag{30}$$

Reciprocity is again incorporated into the logit probability of cooperation as an additive effect, i.e., $\omega_{i,R} H_{i,j,t}$. In the case where $H_{i,j,t} = 0$, this means that i and j have never interacted before with i as the recipient. Hence, $\omega_{i,R} H_{i,j,t}$ will evaluate to 0, and i's willingness to cooperate with past reciprocity partners ($\omega_{i,R}$) will not impact the probability of the current cooperative event in either direction. If, however, $H_{i,j,t}$ is positive, this indicates that j has given to i more frequently and/or more recently than they have chosen not to give to i. Provided that $\omega_{i,R}$ is positive, values of $H_{i,j,t}$ greater than 1 will increase i's willingness to cooperate with j above $\omega_{i,R}$, while values between 0 and 1 will result in a positive effect between 0 and $\omega_{i,R}$. Conversely, if $H_{i,j,t}$ is negative, this indicates that j has more frequently and/or more recently chosen *not* to give to i in past time steps. In evolutionary game theory terms, i would consider j to be a "defector" in this case. Given a positive value of i's $\omega_{i,R}$ trait, a negative $H_{i,j,t}$ will decrease i's willingness to cooperate with j in time t. Updating of $H_{i,j,t}$ to $H_{i,j,t+1}$ and $H_{j,i,t}$ to $H_{j,i,t+1}$ occurs during the last step of each iteration t, after the cooperative event (or lack thereof) between i and j occurs. (See Step 9 of this section for more details.)

Lastly, in addition to kin selection and reciprocity traits, each agent in the population has an inherited "baseline willingness to cooperate with anyone" ($\omega_{i,B}$) trait. Suppose that agent i is randomly paired with any agent j in the population. Even if j's relatedness to i and their interaction history with i are not sufficient for kin selection or reciprocity to positively influence the outcome of the cooperative event, agent i may still be willing to help such an agent j if they have a sufficiently large $\omega_{i,B}$ trait. This trait is treated as an intercept in the logit probability of agent i's decision to cooperate, and it applies to anyone in the population – not just strangers and unrelated agents. Hence, in the (rare) hypothetical case where there is no interaction history ($\omega_{i,R}H_{i,j,t} = 0$) and kin selection is ineffectual ($\omega_{i,K}[(r_{i,j} - \bar{r})b_{j,t} - c_{i,t}] = 0$), the probability that i gives to j is simply $\text{logit}^{-1}(\omega_{i,B})$ $= 1/(1 + \exp\{\text{-}\omega_{i,B}\})$.

With all of the above in mind, the probability $p_{C,i,j,t}$ that i will give some amount of resources to j at time t in terms of its logit (log-odds) probability can be expressed as:

$$\begin{aligned} \text{logit}(p_{C,i,j,t}) &= \ln\left(\frac{p_{C,i,j,t}}{1 - p_{C,i,j,t}}\right) \\ &= \omega_{i,B} + \omega_{i,R}H_{i,j,t} + \omega_{i,K}\left[(r_{i,j} - \bar{r}_t)b_{j,t} - c_{i,t})\right], \end{aligned} \tag{31}$$

where the $\omega_{i,\star}$ values are agent i's willingness to cooperate traits, $r_{i,j}$ is the coefficient of relatedness (by decent) for i and j, and $\bar{r}_t$ is the average relatedness in the population at time t. Using a log-odds specification of $p_{C,i,j,t}$ allows for agents to combine different sources of cooperation probability, while maintaining $p_{C,i,j,t} \in [0, 1]$.

To operationalize the above for each candidate cooperative i, j pair, a recipient j and donor i will be randomly selected from the pools of donors and recip-

ients. Each candidate recipient/donor agent is paired at most once with another donor/recipient agent, and some recipient/donors may not be paired with another agent if there are insufficient donors/recipient agents remaining by the time they are randomly selected. While this scenario may be somewhat unrealistic, as individuals in the real world would likely seek out multiple other cooperative partners if an initial candidate donor failed to provision them (or provision them adequately), the assumption of single donor-recipient pairings was adopted in this model for simplification purposes. Nonetheless, allowing agents to seek out additional cooperative partners at each time step is likely important for fully characterizing human cooperative behaviors, and this addition will be pursued in future updates to this model.

Once a donor i is randomly paired with a recipient j, their probability of cooperating (as well as the calculations that this probability is dependent on) will be computed. The outcome of probability $p_{C,i,j,t}$ is evaluated with $\text{Bern}[p_C(i,j,t)]$. Then, depending on the success ($\text{Bern}[p_C(i,j,t)] = 1$) or failure ($\text{Bern}[p_C(i,j,t)] = 0$) of the cooperative event, agent i may share some positive quantity $\rho_{C,j,t}$ of resources with j. Specificially, this quantity will equal the minimum of j's net need and the absolute value of i's net need, and thus agent i will never give away more than they need to maintain their baseline mortality at $t+1$ (as well as the proportion of their dependents' baseline mortality that they are responsible for).

Importantly, if i does indeed share with j, i and j's changes to mortality (as well as those of their offspring, if applicable) must be updated to reflect changes in resource access. Else, if i does not share with j, the mortality changes will remain as previously calculated. Lastly, the two history functions of i and j at time $t+1$ (i.e. $H_{i,j,t+1}$ and $H_{j,i,t+1}$) will be updated in accordance with the outcome of the cooperative event. (More details on this updating process can be found in the next subsection, Step 9.) The two agents i and j will then be removed from the

pools of candidate donors and recipients, and a new donor/recipient pair will be selected and evaluated for cooperation. The cycle will repeat until all remaining donors and/or recipients have been exhausted.

Step 9. Final Mortality and History Calculations. If no cooperative event occurs between a donor/recipient pair i and j at time t, then the change in mortality for i, j, and their dependents remains as initially calculated in Steps 8.1 and 8.2 above (see Equations 11 and 23). However, if a cooperative event *does* occur between i and j at time t, then the calculations for change in proportional need and change in mortality must be updated to reflect the increase/loss in resources as result of the event.

To describe the updated calculations for change in proportional need and mortality as result of a cooperative event, suppose that a donor agent i gives a recipient agent j some positive quantity of resources. Denote $\rho_{C,j,t} > 0$ as the resources that j receives from i, and denote $\rho_{C,i,t} = \text{-}\rho_{C,j,t} < 0$ as i's corresponding loss of available resources. The sole difference between the calculations without a cooperative transfer and those after a cooperative transfer is the change in available resources, which affects the calculation of $\Delta\eta$ for each agent involved. Hence, recalculation of $\Delta\eta$ will be described first.

Since recipient j pools the resources that they and their dependents produce and then redistributes these resources among themselves and their dependents according to need, the pool of resources that j has available to redistribute is increased by $\rho_{C,j,t}$. Hence, the updated quantity of resources that they and their offspring receive is

$$\dot{\rho}_{x,t,\text{upd}} = \Big(\rho_{C,j,t} + \rho_{j,t} + \sum_{d_j \in D_{j,t}} p_{d_j,t} \cdot \rho_{d_j,t}\Big)\dot{w}_{x,t} \quad \text{for } x \in \{j, D_{j,t}\}. \tag{32}$$

The change in proportional need ($\Delta\eta$) of recipient j and their dependents can then be recalculated with these updated redistributed resources, using exactly the same calculation as in Step 8.1:

$$\Delta\eta_{j,t} = \frac{\bar{\eta}_{a_{j,t}} - \dot{\rho}_{j,t,\text{upd}}}{\bar{\eta}_{a_{j,t}}} \quad \text{for agent } j, \text{ and} \tag{33}$$

$$\Delta\eta_{d_j,t} = \frac{p_{d_j,t} \cdot \bar{\eta}_{a_{d_j,t}} - \dot{\rho}_{d_j,t,\text{upd}}}{\bar{\eta}_{a_{d_j,t}}} \quad \text{for each dependent } d_j \in D_{j,t}. \tag{34}$$

On the other hand, a decrease in resources following a cooperative transfer affects donor i's net surplus (i.e. their net need, $\eta_{i,t}$) and the way that this is redistributed to themselves and their dependents. Recall that, as a surplus, $\eta_{i,t} < 0$. Then, for a $\rho_{C,i,t} < 0$ loss in available resources as result of giving to recipient j, donor i's net need/surplus can be updated as

$$\eta_{i,t,\text{upd}} = \eta_{i,t} - \rho_{C,i,t}, \tag{35}$$

where $\eta_{i,t}$ on the right-hand side of the equation is i's original, pre-cooperative transfer amount of net need. Using this updated value of net need, the change in proportional need of agent i and their dependents is simply

$$\Delta\eta_{x,t,\text{upd}} = \frac{\eta_{i,t,\text{upd}}}{\underset{\sim}{\rho}_{i,t}} \cdot \bar{\mu}_{a_{x,t+1}} \quad \text{for } x \in \{i, D_{i,t}\}. \tag{36}$$

Then, the final change in mortality for recipient j, donor i, and their respective offspring can be recalculated exactly as before in Equations 11 and 23. That is, for $x \in \{i, D_{i,t}, j, D_{j,t}\}$, change in mortality is

$$\Delta\mu_{x,t,\text{upd}} = \frac{1}{1 + \exp\{\text{-}[\text{logit}(\bar{\mu}_{a_{x,t+1}}) + k_\mu \Delta\eta_{x,t,\text{upd}}]\}} - \bar{\mu}_{a_{x,t+1}}. \tag{37}$$

Final mortality for any independent/adult agent at time $t+1$, regardless of whether they engaged in a cooperative event at time t, is calculated exactly the same. First, given some adult agent i with a final change in mortality $\Delta\mu_{i,t}$ at time t, their mortality rate at the start of time $t+1$ is

$$\mu_{i,t+1} = \bar{\mu}_{a_{i,t}} + \Delta\mu_{i,t}. \tag{38}$$

Similarly, for any dependent d in the population with parents i and j, the dependent's mortality rate at time $t+1$ is

$$\mu_{d,t+1} = \bar{\mu}_{a_{d,t}} + \Delta\mu_{d_i,t} + \Delta\mu_{d_j,t}, \tag{39}$$

where $\Delta\mu_{d_i,t}$ is the change in mortality brought about by parent i and $\Delta\mu_{d_j,t}$ is the change in mortality brought about by parent j. (If one of i or j is deceased, then the corresponding $\Delta\mu = 0$. Recall that at least one parent of d must be alive in order for d to be considered a dependent.) Note that, since change in mortality $\Delta\mu_{i,t}$ is always calculated on the logit scale, there is no need to forcibly bound $\mu_{i,t+1}$ between 0 and 1 *for independent/adult agents* – it will always fall within this range. However, because dependents with two live parents have two sources of change to their mortality rates, it is possible that a dependent's final calculated mortality rate for time $t+1$ will be outside of the $[0,1]$ range. However, note that this can only occur in the case where both parents are donors. Nonetheless, the ABM truncates as necessary in this case so that probability remains within the 0 to 1 range. (This is primarily for computational purposes, and does not affect the overall outcome of the calculations.)

Once the cooperative transaction (or lack thereof) is over, the collection of partner history functions $H_{i,j,t}$ must also be updated for all recipient i, donor j

pairs in the population (regardless of whether they interact or not). To update a $H_{i,j,t}$ to time $t+1$, it is first necessary to assign a value of 1, 0, or -1 to the interaction $I_{i,j,t}$ at time t in the manner of Equation 30. Then, the corresponding history function is updated for time $t+1$ as:

$$H_{i,j,t+1} = I_{i,j,t} + k_H H_{i,j,t}, \tag{40}$$

where k_H is as defined in Step 8.3. Importantly, for any given pair i and j, both $H_{i,j,t+1}$ and $H_{j,i,t+1}$ must be updated; while i and j can only ever be one of donor or recipient at a specific time t, both directions of the interaction must be recorded.

Lastly, change in mortality must be calculated for all agents who do not engage in a cooperative event at time t. These calculations performed according to whether an agent's net need is positive or negative, using exactly the same calculations as performed for donors/recipients who do not engage in cooperative transfers (Steps 8.1 and 8.2). That is, if an adult agent ℓ's net need is positive (indicating a shortfall), then their and their dependents' change in mortality is calculated as if they were a recipient agent, using the method outlined in Step 8.1. Alternatively, if an adult agent ℓ's net need is negative (indicating a surplus), then their and their dependents' change in mortality is calculated according to Step 8.2. All interaction functions $I_{\ell,j,t}$ and $I_{j,\ell,t}$ involving agent ℓ are assigned a value of 0 (indicating no interaction), and the corresponding history functions are updated accordingly. The interaction and history functions of all dependents are updated in the same way, since as dependents they do not directly interact with other agents yet.

Step 10. Advance to Next Day (Iteration). After Steps 1 through 9 above are complete, t is incremented and the ABM starts back at Step 1. For the coded model, additional optimization tasks such as periodically writing outdated data to file also occur at this time.

Model Selection, Parameterization, and Validation

Initial, iterative development of the agent-based model was performed using a pattern-oriented modeling approach (Grimm et al., 2005). For clarity, I will briefly describe this modeling strategy. Early forms of the pattern-oriented methodology were informally developed within the ecological modeling literature, and it was later formalized and further developed by Grimm and colleagues in a series of papers and books (Grimm et al., 2005; Grimm & Railsback, 2006, 2012; Railsback & Grimm, 2012; Topping et al., 2012). The method utilizes a "bottom-up" approach, in which characteristic patterns within multiple levels of a real-life system are used to inform development, parameterization, and discrimination between agent-based models used to model the real-life system (Grimm et al., 2005; Railsback & Grimm, 2012).

I initially began model development by hypothesizing which population-level and individual-level patterns were critical to characterizing the observed systems within Linao. Identified population-level patterns that informed model design and development included population size, age structure, interrelatedness, mortality rates, and overall cooperative patterns. Individual-level patterns used for this purpose included age-timing of critical lifestage events, i.e., length of period of full dependency on parent, length of juvenile period, timing of resource production skill acquisition, age at maturity, age at first reproduction, and age at death. Additionally, mechanisms that were associated with cooperative transfers during the prior analysis of Linao dyadic transfer data (Phelps et al., 2022) were used as building-blocks for constructing cooperative functionality in the ABM. Wherever possible, calibration of parameter values for the agent-based model was directly informed by relevant observational measurements made within Linao village. In cases where this

was not possible, I instead based parameter selection on published estimates from the human behavioral ecology literature. In a few specific cases, estimates were also scaled to reflect unit choices made during design of the model. (For example, while energetic need and resource production are typically defined in the literature with respect to kCals required/produced, I chose to define these variables in dimensionless units ranging from 0 to around 40. However, this was just a matter of convenience/interpretability, and the model could just as easily be reformulated to take kCal estimates instead.)

Discrimination between early versions of the ABM was performed iteratively by examining the aforementioned patterns in the ABM and discarding or modifying those models that did not replicate observed patterns at least reasonably well. Verification of the final model (as well as identification of remaining pattern mismatches) was performed via direct descriptive and statistical comparisons with the empirical data from Linao. A brief discussion of early model attempts (which were falsified via pattern-matching) is provided in the first section of the results (Chapter 3), and exact parameterization of the final ABM is discussed in the second section of the results. Verification of the final model against Linao data is described in the third and fourth sections of the results, *Overview of Final Agent-Based Model* and *Comparison between final agent-based model and field data from Linao.*

Implementation and Statistical Analysis of Agent-Based Model

All agent-based models, statistical analysis, and visualization were performed in the `R` statistical software, version 4.2.1. In addition to base `R` functionality, utilized packages critical to this analysis include `data.table` [v.1.14.2], `dplyr` [v.1.0.10], `fda` [v.6.0.5], `ggplot2` [v.3.3.6], `ggpubr` [v.0.4.0], and `truncreg` [v.0.2-5]. (In the preceding list, package version numbers are indicated in brackets.)

The ABM itself was coded directly in `R`, using a functional approach. Nested functions were written for various agent- and environment-level processes, and the final model algorithm was constructed inside an overarching function. This function took user-supplied starting parameters (see Appendix A), and ran autonomously once initialized. The model wrote out data every 25 iterations to either *.csv or *.Rds files. I performed checks of the data every 2500 iterations, in which I calculated various descriptive statistics in order to monitor population structure, relatedness, and cooperation. The model function returned its current state to the global environment at the end of a pre-specified number of iterations (T). Models could then be resumed from current state data using a secondary process, which continued the same ABM algorithm. Implementation was structured in this way to facilitate data generation via batched runs, avoiding the need for time-intensive "all in one go" runs of the model. Additionally, this strategy enabled continued data generation from the existing final model at any time, which will allow for further investigation of the current model's evolving cooperation dynamics in future.

Importantly, while most probabilistic and statistical calculations were implemented via provided functions within base `R` or the packages mentioned above, the raised cosine distribution has not yet been implemented in any `R` packages that I could find. To utilize this distribution, I instead constructed functions to calculate measurements of its cumulative density function and to sample from it. In particular, to facilitate easier sampling, I created functions which binned probability estimates over very tiny ranges which were characterized by their midpoints, and then used the base `R` function `sample()` to generate random samples of the midpoints. In the specific case of sampling from daily age distributions (which are discrete by default, given that they are represented with respect to iterations of the model), I calculated a binned range of probabilities of form $[a - 0.5, a + 0.5)$ for each age

a. Samples were then selected from the range of possible ages, using the vector of probabilities assigned to each age.

Functional data objects were created and managed using the `fda` package. Descriptive data aggregations were performed using base `R`, `dplyr`, and `data.table` functionality, while exploratory and statistical visualizations were created primarily with the `ggplot2` and `ggpubr` packages. Where appropriate, simple statistical tests were performed to evaluate descriptive statistics. In particular, to avoid assuming a distribution, non-parametric Mann-Whitney-Wilcoxon tests (also known as Wilcoxon rank-sum tests) were used to compare age distributions between the model and the empirical data collected in Linao. Significance of these tests was assessed at $\alpha = 0.05$. On the other hand, comparison between the Linao mortality function and a mortality function estimated from the simulated data was made via simple visual inspection, since more than one functional estimate per group is required in order to perform a conventional functional permutation test (Kokoszka & Reimherr, 2017).

Four different types of regression models were used in this analysis: Simple linear regression, multiple linear regression, probit generalized multiple linear regression link, and truncated Normal generalized multiple linear regression. The probit and truncated Normal models were specifically used as components of Cragg two-part hurdle models (Cragg, 1971). Ordinary simple and multiple linear regression models were fit using the base `R` function `lm()`, and the probit model was fit using `glm()`. Truncated Normal models were fit using the `truncreg()` function from the package of the same name (Croissant & Zeileis, 2018; for implementation of truncated Normal models in `R`, the help file for the `truncreg` function and the *Truncated Regression | R Data Analysis Examples* (n.d.) reference was also useful). All explanatory variables were standardized to permit cross-comparison

of effect sizes, although the outcome variables were left unstandardized. The type of test used to assess coefficient significance in models varied depending on model form. For simple and multiple linear regression models (hereafter MLR models), conventional t-tests were used. In probit GMLR models, a Wald Normal approximation test was used, yielding a z-statistic. Finally, an approximated t-test was also used to assess significance in the truncated Normal GMLR models. To avoid issues with multiple testing, the significance of model coefficients was assessed via p-values controlled by an overall family-wise error rate of $\alpha = 0.05$ with the Bonferroni correction (Dunn, 1961; VanderWeele & Mathur, 2019). That is, for g regressors, the individual significance level used to evaluate each coefficient would be $\alpha_g = \alpha/g = 0.05/g$. While the Bonferroni correction is often criticized as overly-conservative, I believe that the limited number of explanatory variables that I considered – combined with the larger-than-typical sample size that the ABM yielded – provides sufficient justification for using the Bonferroni correction here. In particular, given that each of the models in the main text include between 34153 and 58880 observations, p-values in the neighborhood of α (that is, not significant at α_g) might have been relatively easy to achieve via sheer sample size alone, necessitating a family-wise error correction.

Overall model fits for the MLR models were assessed via standard Wherry/McNemar adjusted R^2 statistics (hereafter, adjusted R^2), which are implemented by default in `R` for ordinary least squares models. For the probit GMLR models, a McKelvey & Zavoina pseudo-R^2 was calculated to give a measure of model fit (McKelvey & Zavoina, 1975; Veall & Zimmermann, 1994), although, importantly, this measurement is only approximately comparable to regular OLS R^2 estimates. It appears that no formalized assessment of model fit has been developed for truncated Normal regression models yet, so an estimate of fit based on correlations be-

tween outcome and predicted values was developed following the process recommended in *Truncated Regression | R Data Analysis Examples* (n.d.).

Code files for the agent-based model and all statistical analyses are available from the author (myself) upon request.

Ethics

Permission to use previously-collected empirical data from Linao in this study was granted by members of the Linao research team. Research design, interview protocols, and all observational and interview-based data collection methods used to conduct the Linao study were approved by the Arizona State University Institutional Research Board, IRB# STUDY00001593. Linao community leaders, in consultation with all village residents, granted the original study's research team permission to collect data in Linao. All interview participants provided informed consent before interviews took place, and interview participants were compensated with interview fees (varying at a pre-agreed-upon hourly rate) at the conclusion of each interview. Additionally, the Linao research team provided frequent donations to the community, e.g., large food donations for annual religious celebrations, materials for community building projects, payment of most medical expenses during the study period, and payment of other incidental expenses as needed.

Chapter 3

RESULTS

Development of Final Agent-Based Model via Iterative Pattern-Matching

Multiple early versions of the agent-based model were developed with varying degrees of success. In the earliest attempts, the ABM typically ran well for a short period of time, but then drifted into regions of the modeling environment which proved terminal (i.e., the model crashed for one reason or another). Later attempts were somewhat more successful (they didn't crash), but didn't fully replicate important characteristic patterns. I will not describe every modeling attempt made here, as many of these early versions overlap with respect to the problems that occurred. Instead, I will classify early attempts into a few sequential stages, and I will describe the broad strokes of model development and falsification within each of these stages.

1. Stage 1 was characterized by (supposed) population stability for the first year or so, followed by a sudden "death spiral" and subsequent crash of the ABM. In this first stage, the level of daily resource production that could be achieved by agents was too low, relative to daily need, resulting in daily shortfalls for much of the population. More importantly, the maximum number of deaths per iteration was not constrained in any way, and changes in mortality were linearly related to resource shortfalls. As result, both juvenile and adult agents died far too frequently, only to be replaced by fully-dependent juvenile offspring. Ultimately, the population consisted of only one pair of reproductive adults with a huge amount of offspring for whom they could not generate enough resources. As soon as one of the two adults

died (which occurred within an iteration or two), the model crashed. This ABM clearly did not replicate observed patterns of human resource production, reproduction, and mortality, so it was quickly discarded.

2. In Stage 2 of model development, constraints on mortality were implemented in the model. Specifically, change in mortality was reformulated as a logit-change (instead of a linear change), which dampened the daily impact of resource shortfalls on mortality. This change also served to limit the impact of additional resources when baseline mortality was already very low (in which case agents are already protected from mortality risks) or very high (e.g., at end of life, where extra resources should not be able to overcome species-level senescence patterns, and hence agents should not be overly motivated to hang onto their surpluses). Instead, extra resources were most beneficial to agents with moderate levels of mortality risk. As an extra layer of protection against death spirals, I also constrained the maximum number of agents who could die per day so that the model could not randomly drift into terminal spaces. Lastly, a better ratio of resource production to need was introduced to ensure that a fair number of agents could produce resource surpluses each day. Despite these improvements, the model performed quite well for about 300 years and then also abruptly crashed. Inspection of the underlying data revealed that, while the population age-structure as a whole was reasonable, adult agents in the population had become too highly related to one another over time. Once this occurred, the number of reproductive pairs in the population shrank quickly, since agents were restricted from selecting spouses who were first cousins or closer (i.e., $r_{i,j} \geq 0.125$). Remaining reproductive pairs had increasingly high juvenile dependency rates,

resulting in larger resource shortfalls and subsequent deaths. Shortly before crashing, the model only included one remaining reproductive pair with over 35 juvenile offspring; it crashed as soon as one of the two parents died, since newborn agents could no longer be assigned to a reproductive pair. Given that the relatedness pattern observed in the last years of this stage of the model was not at all representative of relatedness within Linao or other small-scale societies, it was also quickly falsified.

3. To combat issues with over-relatedness in the agent population, I introduced a mechanism in Stage 3 that occasionally replaced deceased agents with "recruited" adult agents who were totally unrelated to all members of the current population, rather than with additional newborn agents. This did an excellent job of controlling adult relatedness within the population, resulting in adult agents marrying at normal rates again. Initially, however, the cooperative traits of recruited agents were randomly-generated using the mean and spread observed in the current agent population. This caused rapid artificial decreases in the variance within cooperative traits, which inhibited evolutionary processes and resulted in premature stabilization of traits at various levels. I also noticed that mortality rates were still a bit too high (due to misspecification of parameters), which had the effect of amplifying the speed with which trait distributions shrank toward their means. Hence, while this model was a vast improvement over previous versions, it was also falsified for failing to capture the expected level of variance within cooperative traits.

In the final model (Stage 4), I improved the specification of mortality parameters and generated cooperative traits for recruited agents using the fixed spread

of cooperative traits ($\vec{s}_\omega$) that was initially provided to the model. These two changes stabilized mortality to within reasonable levels (improving the overall age distribution of the population in the process) and allowed cooperative traits to evolve more realistically. Hence, this ABM was provisionally accepted. However, as I will describe below, this final ABM did not entirely meet the standards imposed by my choice of characteristic patterns.

Parameter Specification of Final Agent-Based Model

Table 1 lists all scalar and vector parameter values selected for the final agent-based model. A total of $t_y = 25$ iterations per year were simulated. The model was run for $T = 37{,}500$ iterations (1,500 years, in model time), split into five consecutive batches of between 2,500 and 10,000 iterations each (100-400 years). Six separate lineages (K) were initialized, each of which contained $N_k = 25$ interrelated agents, for a total of 150 live agents per time step.

Ages at first possible reproduction were randomly-sampled from a raised cosine distribution with mean $\bar{\alpha}_y = 18$ years of age and spread $s_{\alpha,y} = 3$ years of age (Table 1), meaning that individual agents could begin to marry and reproduce as early as 15 years of age and as late as 21 years of age, depending on their assignment of α_i. This is broadly consistent with female age at first reproduction in many forager and subsistence populations (Walker et al., 2006), including Linao village (Phelps et al., 2022), though males in these populations typically had their first child in their early 20s. Agents in the starting population were initialized at their age of first possible reproduction (α_i). I also assumed a lack of sex bias in the population, i.e., that any agent in the population (born or recruited) had a $p_{\gamma,M} = 0.5$ of being male. This assumption is not totally consistent with data from Linao, as a male-favored sex bias within the population as a whole has been observed in re-

cent years (Phelps & Hill, 2021). However, this bias may be due, at least in part, to in-migration of adult males and out-migration of adult females as a result of marriages. In any case, since the observed sex bias in Linao remains an open question and migration out of the agent-based population was not modeled, I chose to assume a sex-balanced population in the final agent-based model.

Table 1

Parameter Specification of the Final ABM

Parameter	Specified Value	Description
T	37,500	Total number of iterations in the model
t_y	25	Number of days (iterations) per year
K	6	Number of distinct lineages initialized
N_k	25	Number of agents per initialized lineage
$\bar{r}$	$(1/2)^3$	Mean $r_{i,j}$ between lineage members at initialization
A_y	90	Age at which baseline mortality rate equals 1
$\bar{\delta}_{m,y}$	18	Age at maturity, i.e. when agent is an adult
$\bar{\delta}_{d,y}$	4	Age up to which agent is fully-dependent on parents
$\bar{\alpha}_y$	18	Mean age at first possible reproduction
$s_{\alpha,y}$	3	Spread of age at first possible reproduction
p_{birth}	0.95	Probability of birth (vs. recruited adult)
$p_{\gamma,M}$	0.5	Probability that a new agent is male
$\bar{\mu}_{\max}$	1	Maximum number of agents that can die per iteration
$\bar{\eta}_m$	10	Mean daily gross need of adults
$\bar{\eta}_\varnothing$	2	Mean daily gross need of agents at birth
α_ρ	2.5	First shape parameter of resource production Beta distr.
β_ρ	3.5	Second shape parameter of resource production Beta distr.
$\tilde{\rho}_m$	15	Mode of adult-aged resource production
$\vec{\omega}$	{-2.2, 2, 2}	Vector of mean cooperative traits ($\{\bar{\omega}_B, \bar{\omega}_R, \bar{\omega}_K\}$)
$\vec{s}_\omega$	{0.4, 1, 1}	Vector of cooperative trait spreads ($\{s_{\omega,B}, s_{\omega,R}, s_{\omega,K}\}$)
k_μ	0.7	Weight of resources on mortality rate
k_H	0.8	Decay of influence of past cooperative interactions

Note. All time-based parameters (e.g., age) are given in years to facilitate interpretation of the values.

The mean relatedness parameter at model initialization was $\bar{r} = 0.125$, corresponding to relatedness at the level of first cousins. In the initialized population of the final agent-based model, however, the mean relatedness of agent-pairs within lineages ranged between approximately 0.135 and 0.196. This is due to both random sampling of $r_{i,j}$ values and to the constraints imposed as result of building a consistent relatedness matrix in which full siblings were all equally-related to other relatives, and was expected. (See *Quantitative description of agent-based model algorithm* for an explanation of the relatedness assignment algorithm.) Once agents started pairing off and reproducing, the probability that a replacement agent was born (versus recruited as an adult) was $p_{\text{birth}} = 0.95$, implying that unrelated adult agents were recruited into the population to replace deceased agents approximately 5% of the time. Agents who were born into the population stopped being fully-dependent on their parents at age $\bar{\delta}_{d,y} = 4$, at which point they started producing an age-scaled amount of resources on a daily basis. This is relatively consistent with Linao data, in which children as young as five years old were observed collecting intertidal foods (Phelps et al., 2022); assuming that this resource production started at age four in the model accounted for the "ramp-up" period in which juvenile agents produced very little. The age at which agents were considered to be adults was $\bar{\delta}_{m,y} = 18$, following convention in most recent anthropological literature.

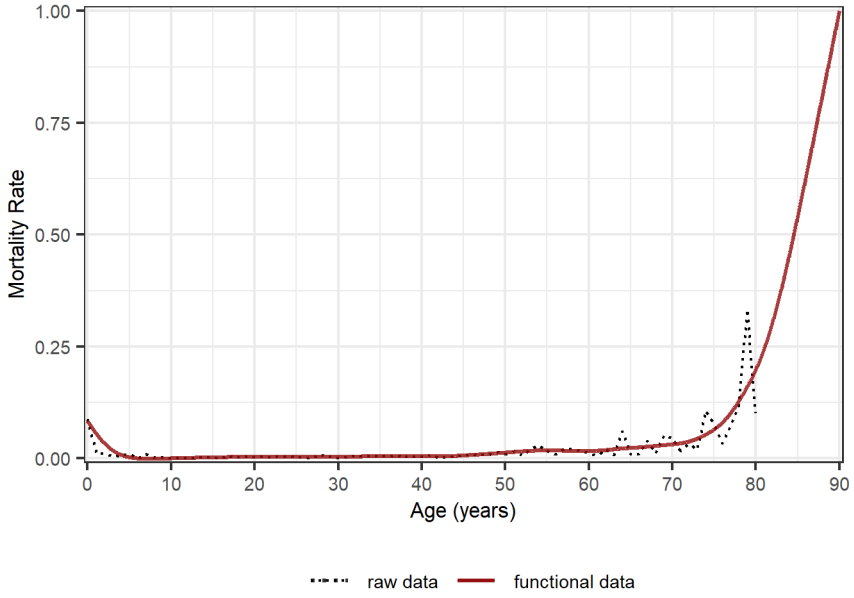
The functional data object of baseline mortality (Mort_{fd}) was generated using mortality estimates from an ongoing demography project in Linao (Phelps & Hill, 2021). Only mortality estimates within the range of ages between 0 and 80 years old were used to build the function, since mortality estimates in higher age ranges were calculated with very small samples. (Few Linao residents in recent history have survived past 80.) However, because a few people in Linao did live into their late 80s and the oldest recorded age was 92, I set the age at which baseline

mortality equals 1 to $A_y = 90$ years old. Mortality rate estimates between 80 and 90 were interpolated. To smooth the final mortality curve, B-spline bases of order 4 were fit to the raw mortality data, using a penalized smoothing approach (Kokoszka & Reimherr, 2017) with smoothing parameter $\lambda = 10^{5.7}$. Figure 2 shows the final result of this functional data process. The final mortality function fits the raw data fairly well, only deviating substantially in regions where raw mortality rate estimates were based on smaller sample sizes or shifted quickly (indicating possible inaccuracies in the raw data). Notable is the low rate of mortality during childhood and most adult ages – this appears to be true in Linao, but is unusual among other studied subsistence populations (e.g., Gurven & Kaplan, 2007). In fact, only 0 to 4.17 deaths per 150 people per year were recorded in Linao over all years in which accurate historical population data exists. It is unknown why this is the case, since Linao residents have only had access to modern medical care in recent decades (Phelps et al., 2022). It could be that some quality of the lifeway in Linao is protective against pathogens, or that cooperation between Linao individuals reduces overall mortality rates to levels much lower than what is observed in other small-scale societies. Given this uncertainty, I decided to set the maximum number of agents allowed to die per day at $\bar{\mu}_{\max} = 1$, meaning that up to 25 agents can die per year in the model. While this might result in higher agent mortality than what is observed in Linao, I felt that relaxing the constraints on mortality rates was appropriate in the context of such uncertainty.

No measure of age-specific caloric intake has yet been developed within Linao, so I estimated daily need parameters based on data from other forager societies. Gross resource need per day of adult agents was set at $\bar{\eta}_m = 10$, while the minimum gross resource need per day for agents at birth (that is, $a_{i,t} = 0$ for a newborn agent i) was set to $\bar{\eta}_\varnothing = 2$. Although scaled to unit measurements, the difference between

Figure 2

Fitted Versus Raw Mortality Curves From Linao

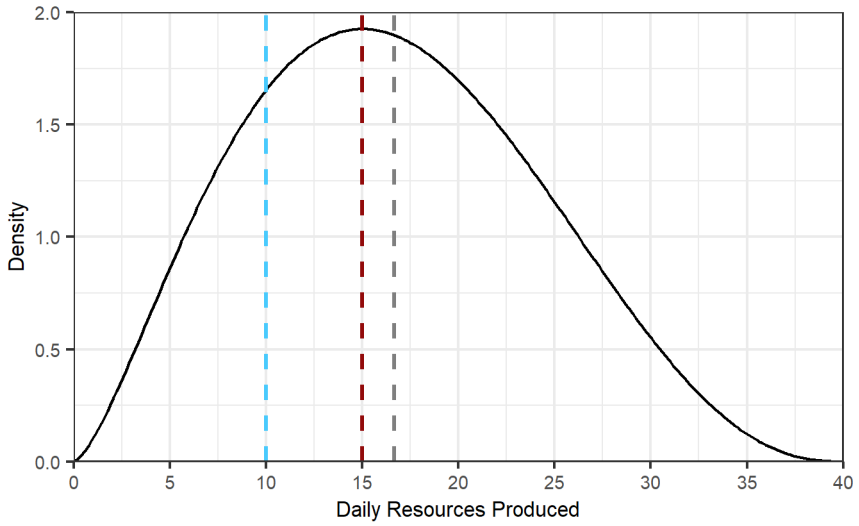


Note. Black dotted line is the raw mortality rate data from Linao. Red line is the fitted mortality function provided to the agent-based model (Mort_{fd}).

$\bar{\eta}_m = 10$ and $\bar{\eta}_\varnothing = 2$ loosely reflects differences in caloric intake between newborns and adults in Ache and Ju/'hoansi populations (Gurven & Walker, 2006). Given that adults in human populations frequently produce substantial daily resource surpluses beyond what is required for individual energy intake requirements – but the magnitude of these surpluses across populations vary based on economic mode, local ecology and other factors (e.g., Kraft et al., 2021) – I set the mode of adult daily resource production at 1.5 times daily adult need, i.e., $\tilde{\rho}_m = 15$. The shape parameters of the Beta distribution underlying random sampling of daily resource production quantities were set at $\alpha_\rho = 2.5$ and $\beta_\rho = 3.5$. The density of this distribution, scaled up to adult resource production values, is shown in Figure 3.

Figure 3

Density of Beta Distribution Controlling Daily Adult Resource Production



Note. Density (y-axis) is generated via a Beta(α_ρ, β_ρ) distribution with $\alpha_\rho = 2.5$ and $\beta_\rho = 3.5$. Daily production quantities (x-axis) are scaled up to the ranges observed in the model. Blue dashed line indicates daily adult gross need ($\bar{\eta}_m = 10$), red dashed line indicates daily adult mode of resource production ($\widetilde{\rho}_m = 15$), and gray dashed line indicates the mean daily adult resource production ($\bar{\rho}_m \approx 16.67$).

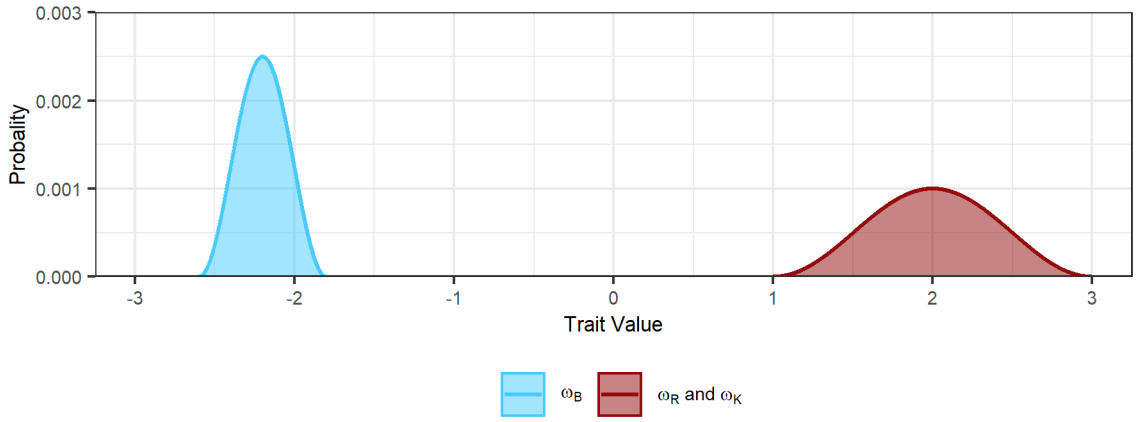
In particular, the selected parameters produce a positively-skewed, left-leaning resource production distribution in which a substantial proportion of the overall probability is centered at or below daily adult gross need (Figure 3, blue dashed line). Adult agents can produce sizable resource surpluses at times, although surpluses in excess of 30 are rare. However, since most adults in the population have multiple dependent offspring whose needs they are (at least partially) responsible for, even daily production values well in excess of 15 (the mode) may not cover the summed net daily needs of an adult and their children. Hence, a production distribution of this shape still necessitates frequent cooperation between adult agents in the population in order to meet daily resource demands. That being said, the relative impact of resource deficits/surpluses on the logit-scaled change to an agent's

mortality rate was set at $k_\mu = 0.7$. This parameter was chosen arbitrarily to reflect the fact that the impact of resource deficits on immediate daily mortality rates is not particularly strong – although physiologically harmful, humans can survive for a number of days without food.

Figure 4 illustrates the densities of the initial raised cosine distributions for each of the three "willingness to cooperate" traits. I assumed that baseline cooperation in the agent population was initially quite low and less variable, with a distributional mean of $\bar{\omega}_B$ = -2.2 and and a distributional spread of $s_{\omega,B} = 0.4$. When holding the other two traits fixed at 0, the mean of the baseline cooperation distribution (at initialization) is equivalent to an approximate 10% chance of sharing resources with anyone. The distribution of baseline cooperative traits was initialized in this way to reflect the fact that there is no evidence to suggest that Linao residents share resources intensively with strangers: Most cooperation appears to be motivated primarily by reciprocity and kinship (Phelps et al., 2022). Since both reciprocity and kinship are important motivators of transfers in Linao, I initialized both trait's distributions with a mean of $\bar{\omega}_R = \bar{\omega}_K = 2$, reflecting a tendency to cooperate with other agents who are reciprocity partners or kin. However, in an effort to not initially bias the model too much in favor of these types of cooperation, I selected larger spreads ($s_{\omega,R} = s_{\omega,K} = 1$) for both traits. When newborn agents inherited traits directly from one of their parents, a small amount of raised-cosine noise not exceeding ± 0.005 was added to the inherited value of the trait. The tight constraints on the amount of introduced noise prevented too much shift in traits from parents to their offspring, preserving meaningful heritability (and hence, evolution) of traits over time in the population. Lastly, I selected a history decay parameter of $k_H = 0.8$ for the model, which suggests that while reciprocity history is important to agents, it is not always remembered or perfectly accounted for.

Figure 4

Initialized Densities of Each of the Three Cooperative Traits, ω_B, ω_R, and ω_K



Note. Baseline willingness to cooperate with anyone (ω_B) has a raised cosine distribution with mean $\bar{\omega}_B = -2.2$ and spread $s_{\omega,B} = 0.4$ (support in range $[-2.6, -1.8]$). Willingness to cooperate with reciprocal partners (ω_R) and kin (ω_K) are identically-distributed according to raised cosine distributions with mean $\bar{\omega}_R = \bar{\omega}_K = 2$ and spread $s_{\omega,R} = s_{\omega,K} = 1$ (support in range $[1,3]$). Reciprocity and kinship are plotted together, since they have identical distributions.

Overview of Final Agent-Based Model

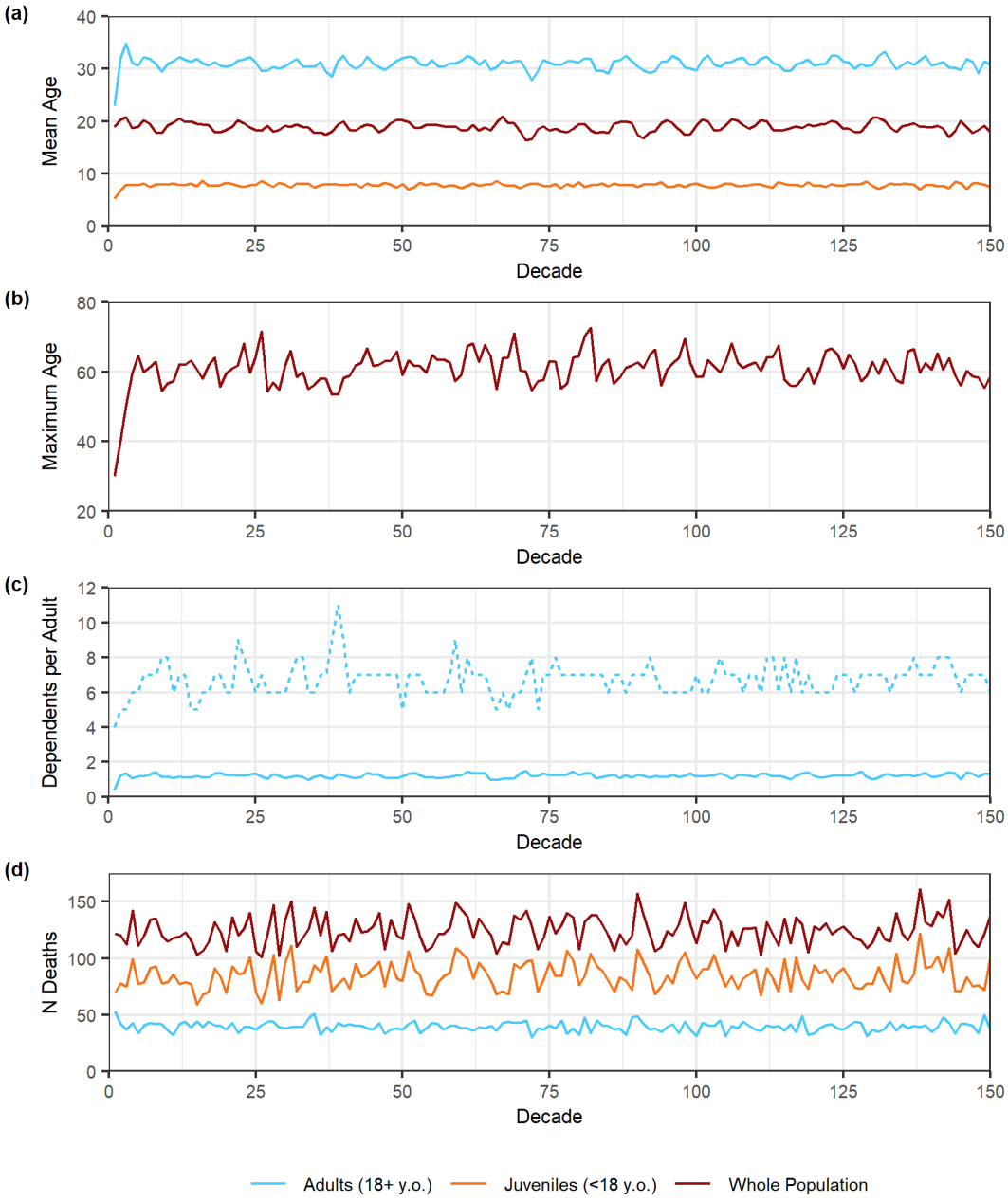
To introduce the results of the final agent-based model, I will begin by discussing population-level temporal trends in the simulated data. As mentioned in the previous section, the model was run for 1,500 years (37,500 time steps), corresponding to 42 overlapping generations of individual agents. As such, inspection of the data at a daily level – or even at a yearly level – is difficult, due to the time-density of the sample. To make sense of the trends occurring over time in the simulated data, I will instead report overall population statistics aggregated by decade (10 model years) in this section. Additionally, I will report age-related trends in years of age, rather than days of age, since the former is more interpretable.

Figure 5 shows population structure measurements of the model at each decade. In this figure, a brief period of population instability was detected for some measures within the first 5-10 decades of the model run (Figure 5a-c). This instability is particularly apparent in the higher variation of adult mean and maximum age in the first few decades of the model (Figure 5a, blue line; Figure 5b), as well as in the lower mean and maximum number of dependents per adult in the first 10 decades (Figure 5c). This was likely a result of the agent population being entirely adults at initialization, since the model needed a "burn-in" period to settle into a population structure/dynamic more typical of small-scale societies. Investigation of yearly-level aggregations of the same metrics (not shown) corroborate this observation: The model took around 100 years in order for overall population structure to stabilize into the patterns of variation seen in later decades. With this in mind, data from the first 10 decades (100 years) will not be reported in further investigations of the model.

Mean age of adult agents (age $a_{i,t} \geq 18$ years old) within a decade typically hovered around 28 to 33 years of age (Figure 5a), and this trend was stable across most decades. Similarly, mean age of juvenile agents (age $a_{i,t} < 18$ years old) ranged between approximate 7 and 8.5 years old. Mean age of *all* agents observed in the population during a decade was generally between 18 and 20 years of age (Figure 5a), although observations as low as 16.3 and as high as 20.8 mean years of age were observed and reflect periods in which the age-structure of the population was more variable via purely random processes. The maximum age observed within each decade ranged between 53 and 73 years of age (Figure 5b; mean across decades $\approx$ 61 years old), suggesting that the high mortality risks imposed by the model at advanced ages made survival of old agents ($\geq$ 55-60 years old) rare.

Figure 5

Trends in Age, Dependency Load, and Mortality by Decade



Note. In all plots, red lines include entire population, blue lines include only adult agents ($\geq$ 18 years of age), and orange lines include only juvenile agents ($<$ 18 years of age). (a) Mean age of group per decade; (b) Maximum age observed in population per decade; (c) Mean dependents (solid line) and maximum dependents (dashed line) of adults per decade; (d) Number of deaths observed per decade.

At any one time within a decade, the average number of juvenile dependents assigned to an adult agent (the parent) hovered between 1 and 1.5 individuals (Figure 5c). Within each decade, juvenile dependency load ranged from 0 offspring to between 5 and 11 offspring. The mean of the maximum across decades was ≈ 6.8 juvenile offspring, which is relatively consistent with completed female interbirth interval estimates of about 3 years in many foraging populations (e.g., Hill & Hurtado, 1996). That is, three years between births in which offspring survive implies that a human female can have around 6 juvenile offspring at any time (not counting additional offspring resulting from twin births, etc.). Additionally, males can have higher dependency loads resulting from multiple consecutive or concurrent female partners, which may help to justify higher juvenile dependency observations within the model. (Importantly, however, constraints on female reproduction were not included in this model, so high dependency load estimates are observed for both males and females in the agent population.)

Mortality rates remained fairly stable for both adults and juveniles over the course of the model, with deaths averaging around 124 individuals per decade (Figure 5d). Notably, however, this estimate is higher than would be expected in a stable population of this size. Higher-than-expected mortality rates were primarily driven by juvenile mortality, and inspection of the raw data suggests that deaths in this age category were frequently young offspring of adult agents who were "recruited" into the population. Since the parent (the adult recruit) was unrelated to anyone in the population by design, negative kin selection effects typically inhibited cooperative transfers from other agents in the population, and the resulting mortality increases from this lack of cooperation most strongly affected juvenile offspring of the unrelated adult. Some of the increased mortality was also due to aging adults in the population, which can be explained by the sharp increase in base-

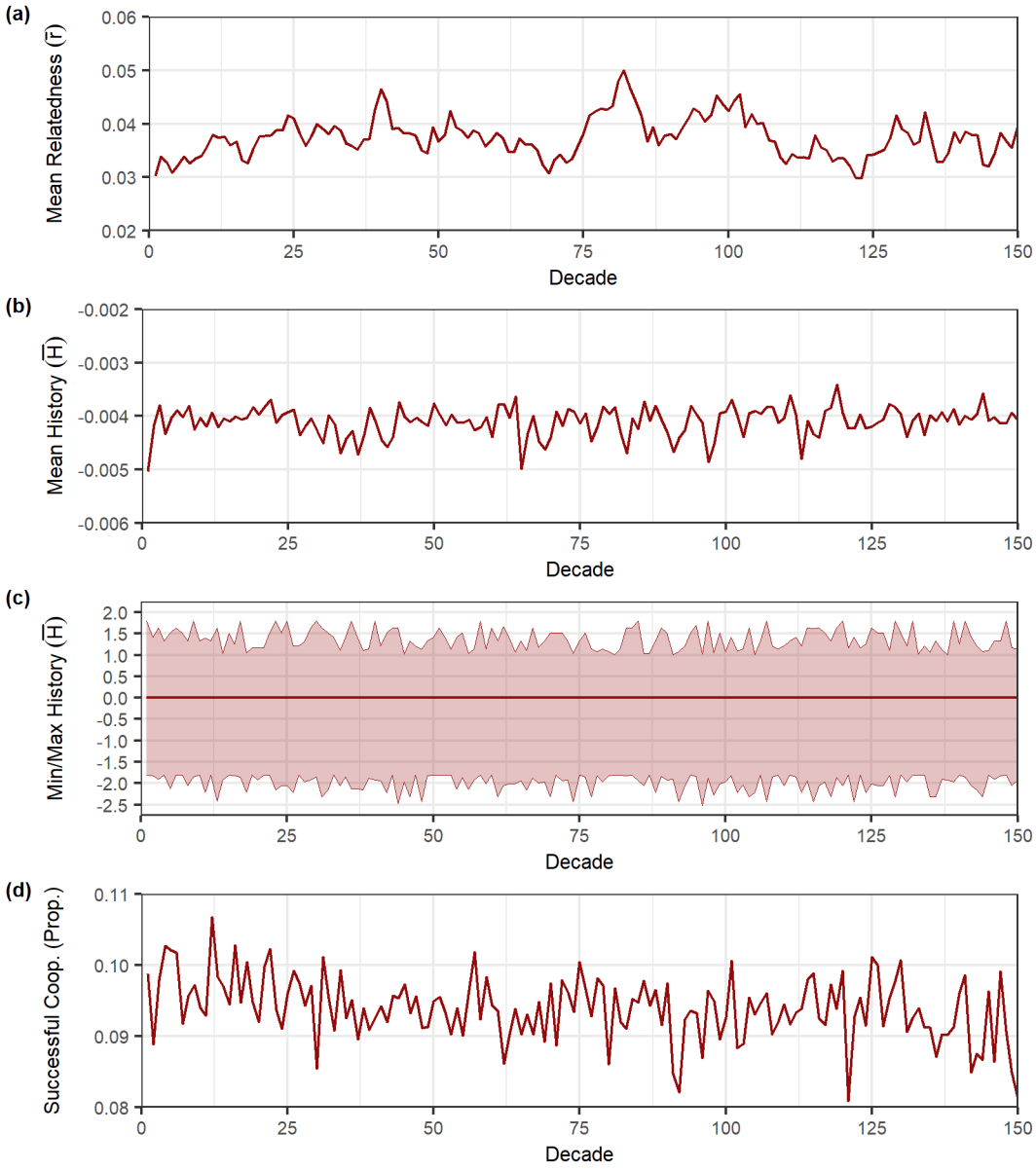
line mortality at advancing ages (Figure 2). The overall mortality pattern observed here may also be driven to an extent by the manner in which mortality was calculated. Specifically, using a logit function to derive changes to mortality means that additional resources had a diminishing effect on very high or very low baseline mortality rates, which could bias the model in favor of ages at which baseline mortality rates were very low (Figure 2: age-specific mortality rates between 5 and 45 years of age vs. other ages).

Additional metrics related to population structure and cooperation are given in Figure 6. Note that an unstable burn-in period of approximately 100 years was again detected, particularly within average relatedness and cooperation history (Figures 6a and 6b), so I will again restrict my discussion to decade 11 and higher.

Mean relatedness between live agent pairs within a decade ranged between approximately $\bar{r} = 0.03$ and $\bar{r} = 0.05$ (Figure 6a), which is relatively consistent with an observed mean pairwise relatedness of approximately $r = 0.035$ in Linao during the study period. This is somewhat lower than what would be expected in a closed population and reflects the periodic recruitment of unrelated adult agents into the ABM population. (The offspring of these adult recruits would have lower-than-average relatedness to the rest of the agent population as well.) Moderate variation in this trend with respect to time is evident (e.g., mean relatedness in decades 50 to 70 vs. decades 75 to 100), suggesting that random shifts in population structure occurred at various points throughout the model run. However, when running a standardized simple linear regression model of mean pairwise relatedness by decade, the correlation between decade number and mean pairwise relatedness was quite low (Appendix C, Table C1: $\beta_{\text{decade}} \approx 0.01650$). Importantly, note that while this coefficient estimate was not significant, I do not report p-values here because they may be meaningless in the context of simulated data.

Figure 6

Trends in Relatedness, Cooperation History, and Proportion of Successful Cooperative Events by Decade



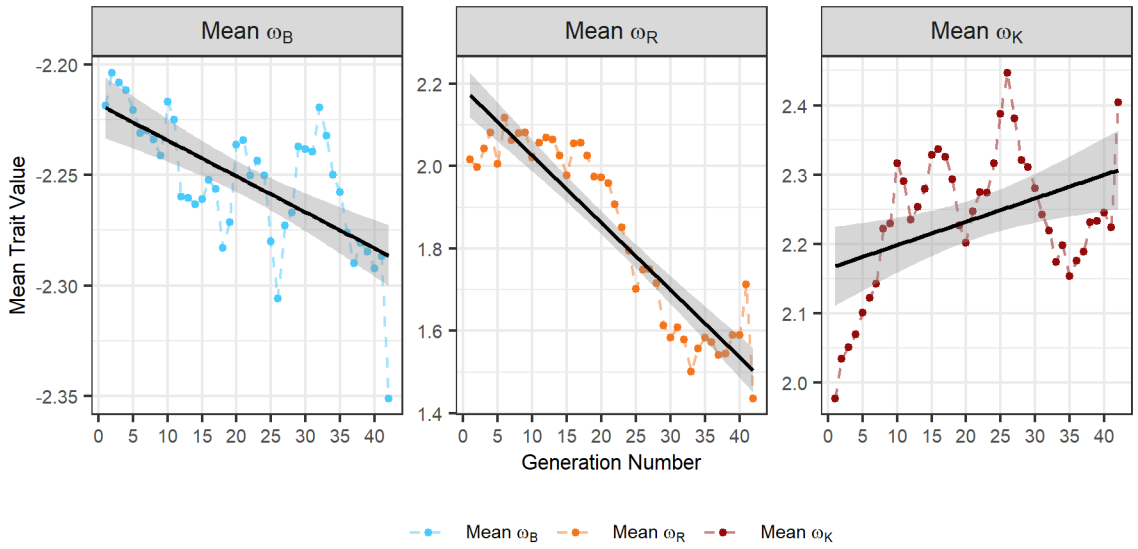
Note. All subplots include entire population recorded within decade. (a) Mean pairwise relatedness of all agents alive during decade; (b) Mean pairwise history of all live agent pairs during decade; (c) Minimum and maximum pairwise history of all live agent pairs during decade (band), with mean denoted as red line for reference; (d) Proportion of all cooperative events in which a transfer of resources successfully occurred.

Cooperative history between agents was similarly stable (Figures 6b and 6c), with a pairwise mean ranging between -0.0050 and -0.0034. As reflected in Figure 6c, these shifts were minimal within the overall range of pairwise cooperative history values observed (range of minimum by decade: [-2.5259, -1.8000]; range of maximum by decade: [1.0038, 1.8000]), and mean pairwise cooperative history and decade were only very weakly correlated (Appendix C, Table C2: $\beta_{\text{decade}} \approx 0.11512$).

On the other hand, the proportion of cooperative events that resulted in a successful transfer of resources trended slightly negatively over time (Figure 6d) and was weakly to moderately correlated with decade (Appendix C, Table C3: $\beta_{\text{decade}} \approx$ -0.32822). To investigate this further, I also examined changes in the distributions of the three cooperative traits (Figure 7). Due to the complexity of separating overlapping generations by decade, I chose instead to calculated the mean of each of the three cooperation traits within each generation of agents observed. Interestingly, both the average of the baseline willingness to cooperate trait (ω_B) and the average of the willingness to cooperate with reciprocity partners trait (ω_R) trended downward over successive generations. The correlation between generation and baseline cooperation was moderately negative (Appendix D, Table D1: $\beta_{\text{gen}} \approx$ -0.66889), but only represented an approximate decrease of a little less than 0.01 in the overall baseline probability of cooperation from generation 1 to generation 42 (Figure 7). The correlation between mean willingness to cooperate with reciprocity partners and generation number was quite a bit stronger (Appendix D, Table D2: $\beta_{\text{gen}} \approx$ -0.91615), and more importantly, it represented a much larger shift in overall population behavior. According to the average behavior of agents' reciprocity trait, selection did not strongly favor cooperation with past reciprocity partners in this model.

Figure 7

Evolution of "Willingness to Cooperate" Traits Over Agent Generations



Note. Dots indicate mean values within each generation. Black lines are simple linear regressions of mean trait value per generation (outcome) by generation number. Gray bands around regression lines are 95% confidence intervals.

In contrast, the mean of the willingness to cooperate with kin trait (ω_K) broadly increased with respect to agent generation, though it was only weakly to moderately correlated with generation number (Appendix D, Table D3: $\beta_{\text{gen}} \approx 0.41315$). It is worth noting that there is a lot of noise in this trait's progression, so estimates should not be interpretted too closely. Overall, however, it appears that the model's selection mechanism favored cooperation between kin, resulting in increases in the magnitude of this trait over time. Of particular interest is the higher variability among mean estimates of the kin cooperation trait, relative to the other two traits. In fact, the valley in mean ω_K at generation 20 and the peak at generation 26 correspond with the dip in mean relatedness at around decade 65 and the peak and subsequent stabilization in mean relatedness from decades 80 to 100,

respectively (Figure 6a). This suggests that changes in overall population composition due to random processes, as evinced by the noted decreases and increases in mean population relatedness, may have had a strong influence on the direction of selection for kin cooperation.

Finally, it is worth mentioning that trends in the final agent-based model's cooperation traits were notably different than corresponding trends found in an early version of the model in which unrelated agents were *not* recruited. (See Stage 2 of model development in *Development of final agent-based model via iterative pattern-matching*.) The early version of the model had much higher population relatedness (mean pairwise relatedness in excess of $\bar{r} = 0.2$), and interestingly, this resulted in a somewhat different cooperation dynamic. Specifically, baseline and reciprocal cooperation appeared to be favored in the early model, while kin cooperation was disfavored (Appendix E). While these results are only suggestive, as the model was not run for very long before it crashed, the underlying implication may be that high levels of interrelatedness create a situation in which it is often less beneficial for a donor to cooperate with kin than non-kin. Unfortunately, the high relatedness of agents within this model was also its downfall, as is described in Stage 2 of model development (*Development of final agent-based model via iterative pattern-matching* section).

Comparison Between Final Agent-Based Model and Linao Field Data

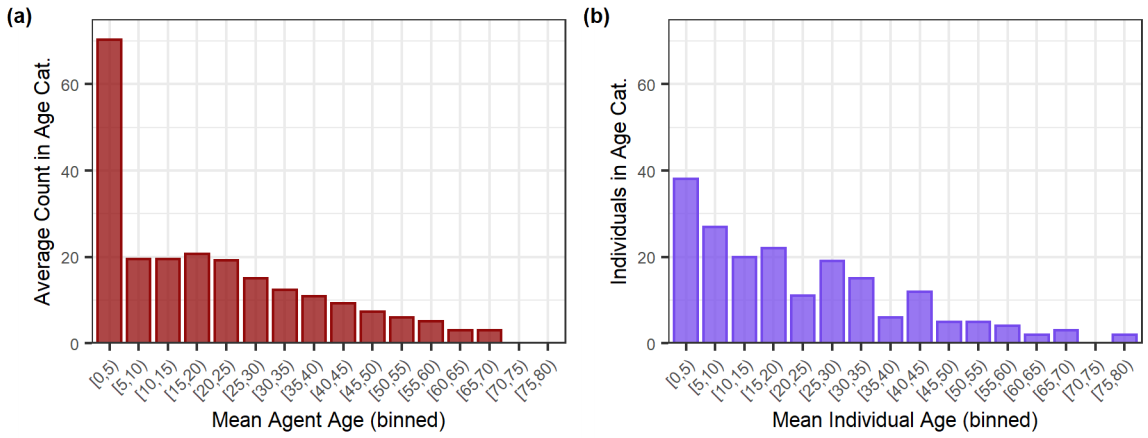
I now turn to a comparison between characteristic patterns observed in the empirical data from Linao village and corresponding patterns in the agent-based model data. To simplify comparison, remaining analyses will be constructed from a 100-year subset of the full 1,500 year dataset – specifically, years 901 to 1,000. This subset was selected because it was a particularly stable period in the final ABM

with respect to mean pairwise agent relatedness (Figure 6a) and average number of dependents per adult agent (Figure 5c).

I will begin with a comparison between the age-structured population distributions of the 100-year subset and the residents in Linao village. Note that I am omitting historical Linao data from this comparison, due to the fact that the historical data predominantly only includes information about ancestors of current Linao residents and may not be representative of the population as a whole at each time period. Since the complete data on current/recent Linao residents covers only about 3.5 years, I split the 100-year agent subset into consecutive 4-year subsamples and estimated an overall population age structure by first tabulating the count of observations at each age within each 4-year subsample, then averaging these age-specific counts across the 4-year subsamples. This method was adopted (instead of simply summing across the entire 100-year sample) to avoid multiplicative over-inflation of estimates at lower ages as result of the larger number of observations within these age ranges. Importantly, the mean age of agents in the 4-year estimation sample coincided closely with the overall mean age of the agent population in the full 100-year subset (mean age $\approx$ 18.50 in the 4-year mean sample vs. mean age $\approx$ 18.46 in the 100-year subset). The result of this estimation process is shown in Figure 8a, and Figure 8b shows the corresponding observed age structure of the Linao sample. Ages are binned into 5-year categories to facilitate observation of the underlying trends. The age-patterns of the two distributions are broadly consistent across most age categories, although the mean age in observed in Linao during the sample period was approximately 21.50, and a Mann-Whitney-Wilcoxon test comparing the age structures of the two distributions approached significance at $\alpha = 0.05$ (estimated location shift: $\hat{\theta} \approx 2$; MWW test statistic: $W = 23438$; $p \approx 0.05272 > \alpha = 0.05$).

Figure 8

Comparison Between Mean Population Age Structure in Agent-Based Model and Population Age Structure in Linao



Note. (a) Average age structure of the ABM population over consecutive 4-year subsets of data between year 901 and year 1000. (b) Age structure of Linao village over 3.5 years of observation. In both plots, age is binned into 5-year categories to facilitate observation of underlying trend.

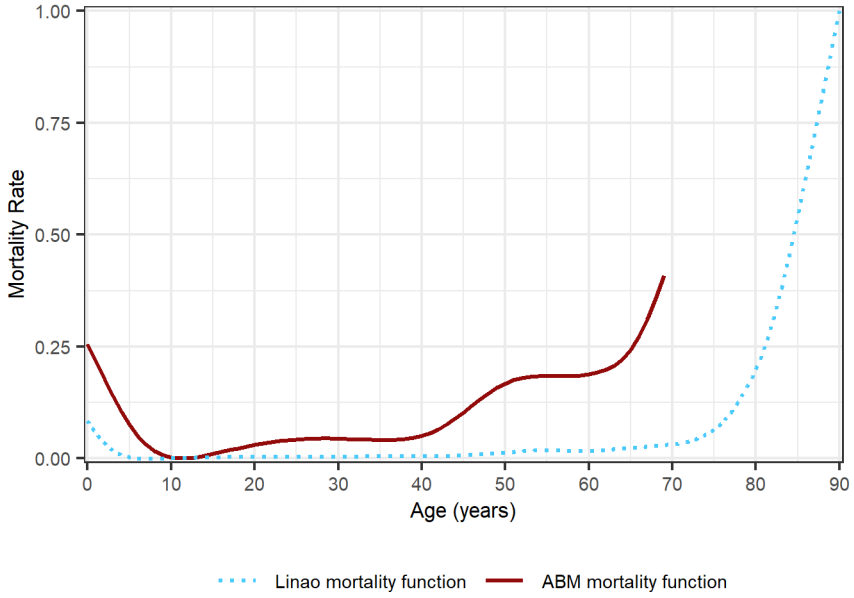
Notably, the major difference between the two age distributions (Figure 8) comes from the fact that the estimated distribution of agent ages in Figure 8a has an excess of young juveniles (mean age less than 5 years old). Specifically, there are almost double the amount of young juveniles in the estimated agent sample as compared to the Linao sample. This is evident in the difference between the sample size of the estimated agent age structure (221 agents) and the sample size of the empirical Linao sample (191 residents). Inspection of the underlying data suggests that the estimated count of young juvenile agents was inflated by increased mortality rates of juvenile agents between 0 and 2 years of age, relative to those observed in the current Linao population. That is, while the Linao mortality curve included a large number of infant deaths, many of these were historical records predating the current sample period. In contrast, recent years in Linao have seen far fewer in-

fant deaths, likely due in part to increased access to modern medical care. However, since the agent-based model was informed by the long-term mortality curve from Linao, juvenile mortality rates were likewise higher. Thus, since most deaths in the model were immediately compensated by a birth (with probability $p_{\text{birth}} = 0.95$), counts of young juvenile agents in the simulated population became inflated beyond what is observed in the current Linao population. These juvenile agents were, in turn, at an increased risk of mortality due to resource shortfalls, compounding the problem and resulting in rates of juvenile mortality even higher than those observed in the historical Linao record.

With this hypothesis in mind, I ran a second version of the Mann-Whitney-Wilcoxon test in which infant deaths at ages below 1 year were excluded (Appendix F). A threshold at age 1 – rather than age 2 – was chosen out of an abundance of caution, and only 17 juvenile agents were excluded from the estimated age sample. Even so, this version of the test was unambiguously non-significant (estimated location shift: $\hat{\theta} \approx 1$; MWW test statistic: $W = 20302$; $p \approx 0.46985 > \alpha = 0.05$), suggesting that it was, in fact, the increased infant mortality rate in the model that drove observed differences in age structure between the agent sample and the Linao sample. Indeed, a comparison between the estimated agent mortality function and the Linao mortality function suggests that infant mortality was almost three times higher in the agent population (Figure 9). Also notable is the increased mortality in the agent population after around age 40: While this explains the slightly steeper fall-off in ages in the agent population (Figure 8a), the similarity of overall population age structure between the agent population and the current Linao population (Figure 8) may also indicate that recent Linao mortality estimates in advanced ages are closer to the estimated agent mortality curve than they are to the historical mortality function (Figure 8b vs. Figure 2).

Figure 9

Comparison of Estimated Agent Mortality Function and Linao Mortality Function



Note. Solid red line is an estimated mortality function for the agent population observed between years 901 and 1000. Blue dotted line is the fitted mortality function from Linao, which was provided to the agent-based model (i.e., Mort_{fd}).

The last characteristic pattern that I investigated was a comparison between the mechanisms driving cooperation in the agent-based model and the mechanisms driving cooperation in Linao. To make this comparison, I used the 100-year subset of agent data (years 901 to 1,000) to build daily and long-term mean dyadic models of cooperative resource inflows that were similar to those examined by Phelps et al. (2022) in the earlier Linao study.

In the model used by Phelps et al. (2022) to characterize daily inflows of resources from one household in a dyad to another, cooperation within unique household "dyads" was modeled with a multiple linear regression (MLR) model. In particular, the MLR model sought to explain daily cooperative resource transfers re-

ceived by a specific *focal* "Household X" from another specific "Household Y." Dyads including 32 of the 36 total households in the study population were examined, although not all household-dyads were considered on each interview day due to the challenges with data collection described in Chapter 1 (*Statistical Challenges in Motivating Study*). The independent variables considered to explain daily dyadic inflows of resources included: 1) The household-level coefficient of relatedness; 2) the difference in ages of the two households; 3) the difference in per capita resource production income between the two households (a proxy for relative need); 4) spatial proximity of the two households' physical houses; 5) outflow of resources from Household X to Household Y on the same day (a measure consistent with daily reciprocity); 6) mean outflow of resources from Household X to Household Y over the sample period (a measure of long-term reciprocity); and 7) the interaction between household-level relatedness and difference in incomes, which was a proxy for need-based kin selection effects. The model also included a control for the mean daily inflow of resources received by Household X over the study period to eliminate the effect of households who simply received more in general from others. [Complete descriptions of the original variable calculations for the Linao dataset can be found in Phelps et al. (2022).]

In order to implement a similar model for daily dyadic resource inflows between agent dyads, I had to make a few minor modifications to the model structure. First of all, agents are not combined into full "households" in the agent-based model, since each member of a reproductive pair independently sought resources from or provided resources to other agents. (Moreover, it was possible in the ABM for an agent with a net deficit resource value to be randomly paired to cooperate with their spouse if that spouse had a net surplus in resources – although the probability of this occurrence was quite low.) However, adult agents with offspring may

be considered "pseudo-households," since the adult agents were responsible for at least half of the net needs of their offspring and engaged in cooperative interactions on their offsprings' behalf. Under this definition, independent agents with no current offspring would be considered singleton "pseudo-households."

Another important modification was motivated by differences in the way that reciprocity was calculated in the Linao dataset versus how it was modeled in the agent-based model. Households dyads observed in the empirical study could – and frequently did – transfer resources in both directions on a given day (Phelps et al., 2022). Typically, these parallel transfers were made across different currencies (e.g., one household giving fish and the other giving rice), satisfying different forms of need for each of the two households in a dyad on a particular day. However, only one type of currency (unspecified "resources") was included in the agent-based model, and agents always had either a deficit or a surplus in this currency. The result was that two-way cooperative transfers between agent dyads could never be observed on the same day. To get around this, I utilized the cooperation history variable (i.e., $H_{j,i,t}$) between a donor j and a recipient i on a given day t as a measure of reciprocity. Recall that $H_{j,i,t}$ is a measure of past cooperation interactions in which i was the donor and j was the recipient. Hence, when modeling resource inflows where j is the donor and i is the recipient, $H_{j,i,t}$ corresponds to the reciprocal effect of past outflows from i to j. Additionally, while $H_{j,i,t}$ captures past interactions between cooperative pairs, it varies on a daily scale as new cooperative interactions (or absence of cooperative interactions) between an (j, i) dyad are incorporated. This implies that $H_{j,i,t}$ captures a measure of both daily and long-term reciprocity between an agent pair. With this in mind, I was able to replace both the daily and long-term mean reciprocity effects in the Linao model with the daily measure of $H_{j,i,t}$ from the agent-based model.

Since the difference in resource production income in the Linao model was intended as a proxy for relative need between households in a dyad, I chose to calculate it explicitly as the difference in net need between recipient agent i and donor agent j for this model (a direct measure of relative need). I also included a measure of the difference in dependency loads between agent j (the focal recipient) and agent j (the donor) to account for the per capita scaling of relative need in the original Linao model (Phelps et al., 2022). Equivalent versions of all other variables from the Linao model were included in the model of daily inflows between agent dyads, apart from the distance between households' physical houses. Since spatial distance between agents was not incorporated into the agent-based model and did not influence cooperative interactions between agents in any way, I was not able to measure this effect and excluded it entirely from the dyadic agent inflow model. (However, in the Linao model, spatial proximity was never an important predictor with respect to its relative effect size anyway. Thus, I believe that its exclusion here is reasonable.) Finally, only independent agents were included in the model, since dependent juvenile agents were not permitted to seek resources from another agent on their own in the model. Similarly, daily observations of independent agents who were *not* paired with another agent for a cooperative transfer on that day were also omitted, since the agent in question was not part of a dyad on that day. Both of these omissions are consistent with the daily dyadic inflow model in the previous study (Phelps et al., 2022).

In keeping with the methods used in the Phelps et al. (2022) study, I constructed the final model of daily dyadic inflows of resources between agents as a MLR model. A total of 58,880 observations of cooperative inflow transfers between agents over the period from year 901 to year 1,000 were included in the model, and all explanatory variables were standardized to permit cross-comparison of effect

sizes. Significance of regressor coefficients was assessed using conventional t-tests, and a Bonferroni-corrected significance level of $\alpha_{g=8} = 0.05/8 = 0.00625$ was used as the threshold for significance. Importantly, p-values may not be meaningful in the context of statistical models of simulated data, since the dataset is complete and significance of weak effects can easily be achieved simply by increasing the sample size. Hence, while I will use p-values to guide identification of significant parameters with relatively larger effect sizes, care must be taken to not interpret p-values too closely. Additionally, since the mean daily inflow received by a focal agent was used as a control variable, I will not directly interpret its results. (It is not wholly clear what this control variable represents in either this model or the previous Linao model. It may be measuring "generosity," overall willingness to cooperate with anyone, or something else.)

Table 2 presents the results of this first model. The effect size control variable (*Mean Daily Inflow to* i) was the largest out of all significant explanatory variables (estimate ≈ 0.101, $t \approx 22.262$, $p < 0.000005 < 0.00625 = \alpha_{g=8}$), suggesting that some unknown agent-level process may be occurring. After discounting the control effect, only the effects of reciprocity ($H_{j,i,t}$) and mean differential need (*Mean Diff Need,* $i - j$) remained as significant predictors of increased daily inflow ($H_{j,i,t}$: estimate ≈ 0.022, $t \approx 4.745$, $p < 0.000005 < 0.00625 = \alpha_{g=8}$; *Mean Diff Need,* $i - j$: estimate ≈ 0.042, $t \approx 8.970$, $p < 0.000005 < 0.00625 = \alpha_{g=8}$). That is, increases in the value of the reciprocity history $H_{j,i,t}$, which indicate increased success of past outflows from the focal agent i to the other agent j, had a positive impact on the amount of resources that j transferred to i during a given event. This is the relationship between resource inflows and past cooperative history that is expected by reciprocity theory (Trivers, 1971). Similarly, positive long-term mean differences between the recipient i's need and the donor j's need (indicating that i is

"needier") were associated with increases in the amount of resources given to i from j each day. This accords with the expectations of need-based altruism theories (D. Smith et al., 2019). Importantly, however, no kin-related effects were detected in either the direct measurement of relatedness between agents ($r_{i,j}$: estimate $\approx$ -0.001, $t \approx$ -0.283, $p \approx 0.77708 > 0.00625 = \alpha_{g=8}$) or in the need-based kin selection term (i.e., the interaction term $r_{i,j}\times$ *Mean Diff Need,* $i-j$: estimate $\approx$ -0.002, $t \approx$ -0.518, $p \approx 0.60415 > 0.00625 = \alpha_{g=8}$. Although one should refrain from interpreting a lack of evidence as support for a null hypothesis, it is still worth noting that *if* this detected pattern is true (and not a Type-II error), then it suggests that more highly inter-related agent dyads did *not* transfer more resources to each other. This is in direct contrast with kin selection theory (Hamilton, 1964). Additionally, if true,

Table 2

MLR Model of Daily Dyadic Resource Inflows From Agent j to Agent i

Response Variable: Daily Dyadic Inflow, j to i ($\rho_{C,i,j,t}$)				
Effect (Standardized)	Estimate	Est. SE	t-Stat	Pr($>\|t\|$)
(Intercept)	0.26803	0.0045	59.203	< 0.00001
Mean Daily Inflow to i	0.10139	0.0046	22.262	< 0.00001
$r_{i,j}$	-0.00128	0.0045	-0.283	0.77708
$H_{j,i,t}$	0.02157	0.0045	4.745	< 0.00001
Mean Diff Need, $i-j$	0.04161	0.0046	8.970	< 0.00001
Diff Dependency Load, $i-j$	0.01127	0.0060	1.880	0.06007
Diff Age, $j-i$	-0.01087	0.0061	-1.786	0.07408
$r_{i,j}\times$ Mean Diff Need, $i-j$	-0.00242	0.0047	-0.518	0.60415
Observations = 58880; DF = 58872				
Residual SE $\approx$ 1.098				
Multiple $R^2 \approx 0.011$; Adjusted $R^2 \approx 0.011$				

Note. Diff = Differential; SE = Standard Error; DF = Degrees of Freedom. All independent effects are standardized, and response variable is unstandardized.

it is not clear why kin cooperation traits still evolved positively over time in the agent-based model (Figure 7).

Overall, the results of this first statistical model (Table 2) are somewhat different than what was detected in the Linao study. In particular, reciprocity on both the daily and long-term scale was the largest significant predictor of daily dyadic resource inflows in Linao village, while kinship, relative need, and the interaction between the two effects (kin selection) played smaller secondary – but nonetheless important – roles in driving cooperation (Phelps et al., 2022). Notably, however, the model of daily resource inflows between agent dyads (Table 2) had an extremely small adjusted R^2 (≈ 0.011). Since some of the day-to-day variation in the ABM may be due to random noise introduced via numerous probabilistic calculations, I followed Phelps et al. (2022) in developing a second, long-term mean version of the dyadic inflow model. The structure of this second model (Table 3) is the same as in the daily model (Table 2), except that all daily measurements were averaged within agent dyads.

The results of the long-term dyadic inflow model are broadly similar to those observed in the daily dyadic inflow model, with one notable exception: Long-term mean cooperation history was no longer significant at $\alpha_{g=8} = 0.00625$, and its estimated effect size was somewhat reduced (*Mean* $H_{j,i,t}$: estimate ≈ 0.013, $t \approx 2.487$, $p \approx 0.01288 > 0.00625 = \alpha_{g=8}$). This directly conflicts with the results of the long-term mean dyadic inflow model calculated with the Linao data, in which reciprocity was the main driver of resource inflows at both the daily and long-term mean scale (Phelps et al., 2022). However, the decrease in efficacy of long-term reciprocity within the agent dyadic inflow model may reflect an issue with averaging the $H_{j,i,t}$ variable over longer time spans. That is, since $H_{j,i,t}$ values were roughly centered around zero, averaging over the long-term might have the effect of pulling estimates

Table 3

MLR Model of Long-Term Mean Dyadic Inflows From Agent j to Agent i

Response Variable: Mean Daily Dyadic Inflow, j to i (Mean $\rho_{C,i,j,t}$)				
Effect (Standardized)	Estimate	Est. SE	t-Stat	Pr($>\|t\|$)
(Intercept)	0.26803	0.0045	59.203	< 0.00001
Mean Daily Inflow to i	0.10139	0.0046	22.262	< 0.00001
$r_{i,j}$	-0.00128	0.0045	-0.283	0.77708
Mean $H_{j,i,t}$	0.02157	0.0045	4.745	< 0.00001
Mean Diff Need, $i - j$	0.04161	0.0046	8.970	< 0.00001
Mean Diff Dependency Load, $i - j$	0.01127	0.0060	1.880	0.06007
Mean Diff Age, $j - i$	-0.01087	0.0061	-1.786	0.07408
$r_{i,j}\times$ Mean Diff Need, $i - j$	-0.00242	0.0047	-0.518	0.60415
Observations = 34153; DF = 34145				
Residual SE ≈ 1.098				
Multiple $R^2 \approx 0.016$; Adjusted $R^2 \approx 0.016$				

Note. Diff = Differential; SE = Standard Error; DF = Degrees of Freedom. All independent effects are standardized, and response variable is unstandardized.

closer to a mean of zero and reducing overall spread of the measurement. Measuring cooperation history in terms of actual transferred resource amounts, rather than in terms of event indicators, may be a future solution to this issue. Nonetheless, the adjusted R^2 of the long-term mean model was still pitiful (Table 3: adjusted $R^2 \approx 0.016$).

Additional Statistical Analysis of Final Agent-Based Model

In the previous section, I statistically modeled the effects driving resource inflows between agent dyads and compared these models to existing models of dyadic transfers between households in Linao village. While the effects of history and differential need that were detected in the two agent inflow models corresponded rea-

sonably well with effects observed in Linao (Phelps et al., 2022), the daily agent inflow model did not detect an effect of kinship (Table 2). This contrasts with the Linao models, where kinship was a small but important predictor of increased daily inflows (Phelps et al., 2022), and with the positive evolution of the "willingness to cooperate" traits (ω_B) that was observed in the ABM. Additionally, both models developed for the agent-based dataset characterized very little of the overall variation in agent cooperative behaviors. The very low adjusted R^2 values estimated for both the daily and long-term mean models of dyadic resource inflows suggest that there are other forces driving the lion's share of the variation observed in agent resource transfer patterns. As mentioned previously, random noise may be at least partially to blame for this lack of model fit, since both resource production and the final evaluation of whether or not a donor gave to a recipient were randomly-sampled from probability distributions. However, massive zero inflation was also detected in this model [moreso even than in the dyadic inflow model developed in Phelps et al. (2022)]. Issues with estimation resulting from this zero inflation may have inhibited accurate detection of effects, which might offer a partial explanation for the high degree of unexplained variance in both models (Tables 2 and 3). In this final section of the analysis, I will briefly describe the zero-inflation observed in the agent data and then develop models to address this issue.

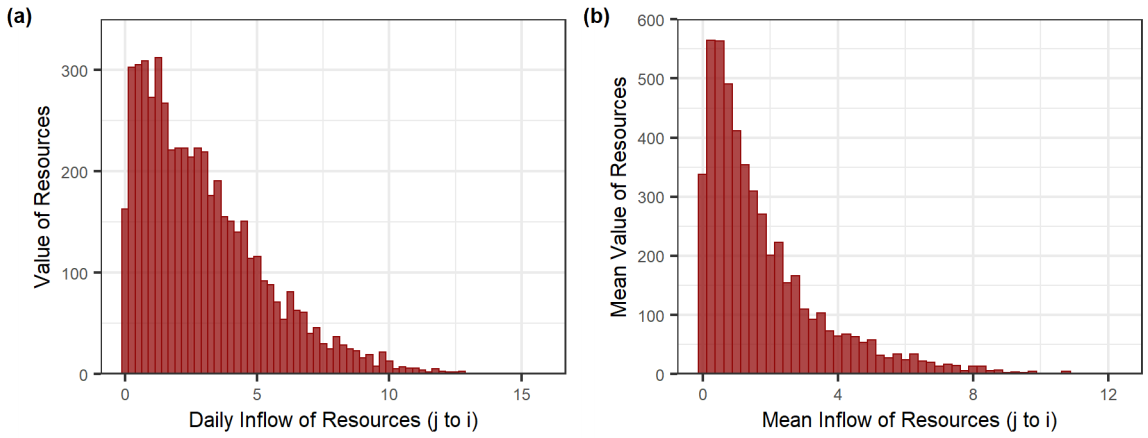
Using the 100-year dataset from the previous section, I first sought to characterize the degree of zero-inflation present in the agent data. A quick inspection of all possible cooperative events (that is, observations in which a needy agent was paired with an agent who had a surplus of resources) revealed that only about 9.1% of all possible cooperative events were "successful," i.e.,resulting in a transfer of resources. This implies that almost 91% of all possible transfers failed and zero resources were transferred. [For comparison, zero transfers were observed in ap-

proximately 69.4% of all potential transfer events recorded in Linao (Phelps et al., 2022).] This is obviously problematic, since such drastic zero inflation almost certainly prevents accurate identification and estimation of model effects.

On a hunch, I next investigated the distribution of the quantity of resources transferred during non-zero inflow events in the agent dataset. As suspected, Figure 10 demonstrates that this distribution (when omitting observations of zeros) is highly non-Normal. Instead, the distribution of dyadic resource inflows more closely approximates a truncated Normal distribution with a boundary at zero. This makes sense, given that most transfers were of small quantities and a recipient agent could not receive a quantity of resources less than zero. (The same holds true in the way that inflows and outflows were defined in the Linao data, since inflows and outflows of resources were measured separately and were frequently small in value.)

Figure 10

Histograms of Daily and Long-Term Mean Resource Outflows, Excluding Zero-Inflow Observations



Note. Binwidths equal 0.25 resource units in both plots. (a) Daily non-zero resource inflows. (b) Long-term mean non-zero resource inflows.

With the zero-inflation and non-Normality of the agent inflow data in mind, I reformulated the daily dyadic inflow model from the last section as a Cragg hurdle model (Cragg, 1971). Although his model was initially developed to study consumer spending patterns with respect to infrequently-purchased goods, Cragg's eponymous two-part model anticipates the issues inherent in measuring dyadic transfers of subsistence resources across a population. Cragg's solution was to first model the probability that a transfer was non-zero versus zero (typically using a probit model); then, the magnitude of non-zero observations which cleared that first "hurdle" (not being zero) were regressed in a second, nested model according to the distributional properties of the non-zero observations. He proposed a truncated Normal regression for cases where the non-zero amounts clustered near the zero threshold, a strategy which I adopted here. Importantly, since the quantity of zero and non-zero observations are modeled separately in this formulation, each of the two hurdle submodels can include the same explanatory variables (or different ones), the effects of which are estimated separately within each submodel.

The anthropological theory surrounding cooperative transfers does not necessarily discriminate between mechanisms driving larger transfer amounts and mechanisms driving "any transfer at all," so I chose to include all of the regressors considered previously in both parts of the hurdle model of dyadic agent inflows. Results of the daily version of this model are in Table 4. Since the two parts of the model were estimated separately, I chose to use a Bonferroni correction of $1/g = 1/8$ to the family-wise error rate for significance assessment in each model.

For the probit part of the model (Table 4a), only reciprocity between a focal agent i and another agent j was associated with increased probability of a non-zero inflow from agent j to agent i, apart from the control variable (Table 4a, $H_{j,i,t}$: estimate ≈ 0.060, $z \approx 6.610$, $p < 0.000005 < 0.00625 = \alpha_{g=8}$). In fact, when ignoring

Table 4

Cragg Hurdle Model of Daily Dyadic Inflows From Agent j to Agent i

(a) Probit GLRM

Response Variable: Probability of Non-Zero Daily Dyadic Inflow ($\mathbb{I}\{\rho_{C,i,j,t} > 0\}$)

Effect (Standardized)	Estimate	Est. SE	z-Stat	$\Pr(>\|z\|)$
(Intercept)	-1.34628	0.0073	-183.483	< 0.00001
Mean Daily Inflow to i	0.10604	0.0070	15.131	< 0.00001
$r_{i,j}$	0.00104	0.0073	0.143	0.88641
$H_{j,i,t}$	0.05979	0.0090	6.610	< 0.00001
Mean Diff Need, $i - j$	-0.00751	0.0074	-1.010	0.31270
Diff Dependency Load, $i - j$	-0.01185	0.0096	-1.234	0.21717
Diff Age, $j - i$	-0.01309	0.0098	-1.339	0.18048
$r_{i,j}\times$ Mean Diff Need, $i - j$	-0.00433	0.0075	-0.578	0.56307
Observations = 58880; DF = 58872; Dispersion = 1				
Null Deviance $\approx$ 35815; Residual Deviance $\approx$ 35538				
McKelvey & Zavoina Pseudo-$R^2 \approx 0.015$				

(b) Truncated Normal GLRM

Response Variable: Non-Zero Daily Dyadic Inflow ($\rho_{C,i,j,t}$, where $\rho_{C,i,j,t} > 0$)

Effect (Standardized)	Estimate	Est. SE	t-Stat	$\Pr(>\|t\|)$
(Intercept)	-0.51469	0.2589	-1.988	0.04681
Mean Daily Inflow to i	0.87918	0.0691	12.721	< 0.00001
$r_{i,j}$	-0.01578	0.0897	-0.176	0.86028
$H_{j,i,t}$	0.02247	0.1102	0.204	0.83840
Mean Diff Need, $i - j$	1.21087	0.0936	12.938	< 0.00001
Diff Dependency Load, $i - j$	0.59282	0.1157	5.125	< 0.00001
Diff Age, $j - i$	-0.16239	0.1152	-1.409	0.15872
$r_{i,j}\times$ Mean Diff Need, $i - j$	0.00846	0.0881	0.096	0.92347
σ	3.74694	0.1014	36.941	< 0.00001
Observations = 5340; DF = 5331				
Correlation Estimate of Explained Variance $\approx$ 0.087				

Note. Diff = Differential; SE = Standard Error; DF = Degrees of Freedom. The σ parameter is equivalent to an OLS estimate of residual standard error. All independent effects are standardized, and response variable is unstandardized.

the control variable, the estimated effect size of $H_{j,i,t}$ was the largest of any explanatory variable in the model, significant or non-significant. This contrasts with the MLR model of daily dyadic agent inflows developed in the last section, in which differential need between an agent pair was a stronger predictor than cooperation history (Table 2). Interestingly, no such effect for $H_{j,i,t}$ was detected in the truncated Normal model (Table 4b, $H_{j,i,t}$: estimate ≈ 0.022, $t \approx 0.204$, $p \approx 0.83840 > 0.00625 = \alpha_{g=8}$). Instead, differences in both need and dependency loads were the primary (significant) drivers of increased daily dyadic inflows within the truncated Normal model (*Mean Diff Need,* $i - j$: estimate ≈ 1.211, $t \approx 12.938$, $p < 0.000005 < 0.00625 = \alpha_{g=8}$; *Diff Dependency Load,* $i - j$: estimate ≈ 0.593, $t \approx 5.125$, $p < 0.000005 < 0.00625 = \alpha_{g=8}$). Differential need had the larger effect size of the two (almost double that of differential dependency load), but in both cases this suggests that recipient agents who had higher resource deficits and numbers of offspring typically received more resources on a given day. Additionally, the effect size estimated for relative need in the truncated Normal model was around three times the size of the corresponding effect size estimated in the MLR model, and no significant effect was detected for differential dependency load in the MLR model (Table 2). Assuming that the estimates in the truncated Normal model are more accurate, this implies that modeling zeros in the data separately from non-zero values facilitates more nuanced detection of effects.

Results were mostly similar in the long-term mean version of the dyadic inflow hurdle model (Table 5). In the probit portion of the model, long-term mean history was the sole significant regressor (*Mean* $H_{j,i,t}$: estimate ≈ 0.040, $z \approx 4.140$, $p \approx 0.00003 < 0.00625 = \alpha_{g=8}$) out of all non-control variables. Notably, this effect was not detected in the long-term mean GLM model (Table 3), at least not at significance level $\alpha_{g=8} = 0.00625$. Interestingly, there was also a small effect for

Table 5

Cragg Hurdle Model of Long-Term Mean Dyadic Inflows From Agent j to Agent i

(a) Probit GLRM

Response Variable: Probability of Non-Zero Mean Dyadic Inflow ($\mathbb{I}\{\rho_{C,i,j,t} > 0\}$)

Effect (Standardized)	Estimate	Est. SE	z-Stat	$\Pr(>\|z\|)$
(Intercept)	-1.05513	0.0084	-125.658	< 0.00001
Mean Daily Inflow to i	0.12109	0.0081	14.862	< 0.00001
$r_{i,j}$	0.00528	0.0083	0.638	0.52365
Mean $H_{j,i,t}$	0.03977	0.0096	4.140	0.00003
Mean Diff Need, $i - j$	0.01297	0.0085	1.522	0.12794
Mean Diff Dependency Load, $i - j$	0.01163	0.0116	1.003	0.31595
Mean Diff Age, $j - i$	0.02617	0.0118	2.226	0.026025
$r_{i,j}\times$ Mean Diff Need, $i - j$	-0.00424	0.0085	-0.500	0.61690
Observations = 34153; DF = 34145; Dispersion = 1				
Null Deviance $\approx$ 28598; Residual Deviance $\approx$ 28354				
McKelvey & Zavoina Pseudo-$R^2 \approx 0.016$				

(b) Truncated Normal GLRM

Response Variable: Non-Zero Mean Dyadic Inflow ($\rho_{C,i,j,t}$, where $\rho_{C,i,j,t} > 0$)

Effect (Standardized)	Estimate	Est. SE	t-Stat	$\Pr(>\|t\|)$
(Intercept)	-17.73181	3.6782	-4.821	< 0.00001
Mean Daily Inflow to i	2.14875	0.3077	6.984	< 0.00001
$r_{i,j}$	-0.19400	0.3264	-0.594	0.55232
Mean $H_{j,i,t}$	0.09121	0.4337	0.210	0.83344
Mean Diff Need, $i - j$	3.64006	0.6363	5.720	< 0.00001
Mean Diff Dependency Load, $i - j$	0.69342	0.4290	1.617	0.10599
Mean Diff Age, $j - i$	-2.28563	0.5530	-4.134	0.00004
$r_{i,j}\times$ Mean Diff Need, $i - j$	0.17142	0.3195	0.537	0.59159
σ	5.95051	0.5266	11.299	< 0.00001
Observations = 5044; DF = 5035				
Correlation Estimate of Explained Variance $\approx$ 0.071				

Note. Diff = Differential; SE = Standard Error; DF = Degrees of Freedom. The σ parameter is equivalent to an OLS estimate of residual standard error. All independent effects are standardized, and response variable is unstandardized.

difference in agent ages (donor j minus recipient i), though it was not significant at the Bonferroni-controlled error rate (Table 5a, *Mean Diff Age,* $j - i$: estimate ≈ 0.026, $z \approx 2.226$, $p \approx 0.02600 > 0.00625 = \alpha_{g=8}$). However, differential age *was* a significant effect in the long-term mean truncated Normal model (Table 5b, *Mean Diff Age,* $j - i$: estimate $\approx$ -2.286, $t \approx$ -4.133, $p \approx 0.00004 < 0.00625 = \alpha_{g=8}$). This suggests that donors who were older relative to recipients typically gave the recipient less over the long-term. This effect may be driven by increased dependency loads in older ages, although why that effect was not detected is not clear. Regardless, differential need was again a strong positive predictor of increased inflows from j to i (Table 5b, *Mean Diff Need,* $i - j$: estimate ≈ 3.640, $t \approx 5.72$, $p < 0.000005 < 0.00625 = \alpha_{g=8}$), implying that recipients who were more needy than donors, on average, received more from those donors over time.

Lastly, while the Cragg models appeared to improve overall coefficient estimation, these hurdle still explained relatively little of the overall variance in the inflow data. In the daily dyadic inflow hurdle model, the estimated McKelvey-Zavoina Pseudo-R^2 for the probit part was only 0.015. The correlation estimation of explained variance in the truncated Normal portion of the daily model was higher at approximately 0.087, but it is important to stress that this is a linear approximation only and should not be interpreted too closely. (The long-term mean version of the hurdle model yielded similar values, at 0.016 and 0.071, respectively.) Nevertheless, it does appear that adoption of the Cragg hurdle model method allowed for disentanglement of effects working at different scales, which may prove useful for understanding the nuances surrounding cooperative behaviors.

Chapter 4

DISCUSSION

Broadly speaking, the agent-based model that I present in this paper is somewhat effective in capturing the general patterns which characterize human population structure and cooperative interactions. However, when compared directly with data from the Linao population, there were some important differences with respect to both age-structure and mortality rate. Specifically, more juvenile agents were present in the ABM population relative to the current Linao population (Figure 8), and the mortality risk for each of these juvenile agents was higher (Figure 9). Likewise, mortality rates for elderly agents were also higher than what is observed in Linao, resulting in a slightly steeper drop-off in population counts at advanced ages. This resulted in distinct differences between the ABM's expected age distribution and the age distribution observed in Linao.

Despite these differences in population structure, statistical models of resource inflows between agent dyads appeared to be reasonably consistent with similar models describing dyadic resource inflows in Linao (Phelps et al., 2022), at least with respect to detecting effects for reciprocity and needs-based assistance. In both the multiple linear regression models (Tables 2 and 3) and Cragg hurdle models (Tables 4 and 5), reciprocity and differential need were demonstrated to be strongly associated with increased inflows of resources from one agent to another, although the relative strengths of these effects varied between model type. In particular, by utilizing a Cragg hurdle model to disentangle effects predicting the magnitude of a non-zero transfer versus whether a transfer occurred at all, I demonstrated that while reciprocal history between agent dyads strongly influenced the probability of an inflow transfer at both the daily and long-term scale (Tables 4a and 5a), it was

not associated with increases in the specific quantity of resources transferred during a successful cooperative event (Tables 4b and 5b). Instead, the quantity of non-zero transfers was best predicted by differential levels of need between the recipient and donor agent in a dyadic transaction (Tables 4b and 5b). This suggests that reciprocity may have been operating primarily as a partner selection mechanism in the agent-based model, where it informed initial decisions of whether to cooperate at all but did not influence how much was given once initial cooperative intent was established. This is a notable observation, with potentially broad implications for understanding how reciprocal cooperation functions. As more complete empirical datasets are collected, it will be interesting to examine whether this pattern also holds true in Linao village and in other subsistence populations.

However, one major difference between the statistical models of dyadic inflow between agents and dyadic inflow between Linao households is that kinship was not detected as a significant predictor of increased inflow in any of the agent models (Tables 2 through 5), despite being a small but significant predictor of Linao inflow transfers (Phelps et al., 2022). This is especially puzzling, given that selection within the agent-based model appeared to be favoring increased kin cooperation, as was evident in progressive increases in the population mean of the ω_K trait over generations (Figure 7 and surrounding text). Conversely, while the "willingness to cooperate with reciprocity partners" trait (ω_R) evolved negatively as the simulation progressed, it showed up as a strong, significant *positive* predictor of dyadic resource inflows between agents in all statistical models of agent inflows.

It is not totally clear why this discrepancy between agents' cooperative traits and statistically-detected behavior occurred, but one possible hypothesis is that kinship could be "priming" reciprocity in the ABM by initially influencing which preferential partners an agent should begin reciprocal interactions with, and is thus

not independently detectable in statistical models. If so, this could be consistent with both the Phelps et al. (2022) analysis and other recent studies of dyadic cooperation in subsistence populations (Allen-Arave et al., 2008; Gurven et al., 2001; Thomas et al., 2018) which suggest that reciprocity and kinship positively correlate in small-scale societies. Under this hypothesis, however, one would not expect the effect of kinship to simply disappear from the model, since it would be important for initiating first interactions between agent dyads. Indeed, an independent effect for kinship typically showed up in the models utilized in the aforementioned studies, even if the relative effect size of kinship was much smaller than that of the utilized reciprocity variable(s). The implication of this is that a hypothesis suggesting that the priming effect of kinship on reciprocity are inhibiting detection of kinship effects may *not* sufficiently explain why statistical models of the agent-based data were unable to detect an effect for kinship. Instead, this discrepancy raises the question of whether multiple linear regression and other related models can reliably detect the true effects of kinship on cooperation, particularly in the presence of reciprocity effects. This is a major concern, given that these sorts of models have been used frequently in past studies of cooperation. An inability to statistically disentangle and detect kinship effects could result in undue emphasis being placed on reciprocity as an important motivator of cooperation, as well as potentially false claims that reciprocity is a more important cooperation mechanism than kinship.

It is, nonetheless, worth mentioning that there remain some important statistical limitations in this analysis that may have impeded accurate detection and estimation of effects, including that of kinship. While I addressed the substantial zero-inflation issue that was detected in the agent-based dataset, I did not address the possibility of temporal correlations using a cross-lagged panel model. It is possible that expanding the current statistical modeling strategy to account for temporal

correlations between dyads might assist with detection of subtle effects. However, such a model might again require considerable data reduction, given that different dyads were sampled different numbers of times due to random cooperative partner assignment within the agent-based model, and zero-inflation in the agent data would still need to be accounted for. Additionally, it is likely that the observations within particular agent dyads are themselves non-independent (this is certainly true within the household dyads in Linao), which might further compound issues with detection and estimation. I did not address this issue by implementing a random or fixed effect for dyad in the current analysis, mainly due to the nature of the current simulated data: Since cooperative agent dyads were randomly-assigned at each iteration of the model, some agents only interacted with each other once, which could cause issues with estimation of dyad-level effects. However, improvements to the cooperative mechanisms in the model (discussed below) may alleviate these statistical challenges, as may adopting a network analysis modeling strategy. In any case, future work using alternative statistical models will be required to validate the current statistical results and to determine whether there is truly cause to be concerned about the efficacy of more commonly-utilized statistical modeling approaches.

As alluded, there are also some limitations in the way that the current agent-based model was implemented that may have affected its ability to capture nuances of daily cooperative behaviors occurring in Linao. Firstly, adult agents in the ABM were not able to directly discriminate between different candidate partners because agent interactions were randomly assigned each day. As such, data from the ABM did not – and cannot – capture any sort of strong assortment over time: Agents randomly interacted with each other, so while they preferentially chose to cooperate with specific partners (e.g., past reciprocity partners) over others when given the

opportunity, they did not increasingly restrict their interactions to only preferred partners. This is reflected in the low percentage of successful cooperative events in the agent-based model (between 8-10% on average; see Figure 6) which resulted from frequent random interactions with less-preferred candidate partners. On the other hand, ethnographic and statistical evidence from Linao suggests that residents interact more often with a specific set of preferred partners than with other, less-preferred individuals in the community (Phelps et al., 2022).

Secondly, individuals in the agent-based model were only able to interact with one candidate cooperative partner per day, and these interactions were unidirectional (i.e., an agent could only be a donor or a recipient during a cooperative event on a given day). This is in direct contrast to cooperative patterns observed in Linao, in which individuals frequently both receive goods from and give goods to multiple different people on a given day (Phelps et al., 2022). By not allowing agents to interact with multiple different cooperative partners on a given day, the ABM may be artificially decreasing observed rates of cooperation in the agent population. Thirdly, parent agents did not jointly pool their resources with each other and their offspring, meaning that each parent could only provision and/or solicit up to half of the daily net needs of their offspring. This is not realistic, given that resources are typically pooled within full households in small-scale societies (which are typically comprised of both parents and all offspring, at minimum). This inconsistency with real-world human populations might have contributed to increased mortality rates, particularly for young juveniles and older adults who are already at an increased mortality risk. Fourth, no measure of affinal relationship between agents was included in the ABM. That is, although a donor might be related to the *spouse* of a candidate recipient, this did not factor into kinship-based decisions on whether or not to help the recipient. This is an important limitation in the con-

text of modeling human populations, given that there exists a plethora of evidence suggesting that humans' recognition of affines and other "fictive kin" can mediate interactions and promote cooperation with these individuals (e.g., Alvard, 2011; Nimmo, 2001; Thomas et al., 2018; Tonkinson, 1991). For example, Linao individuals often justified transfers of goods based on their affinal relationship with the recipient (Phelps et al., 2022).

Moreover, ethnographic evidence suggests that there may be additional adaptive mechanisms and cultural norms driving the complex cooperation patterns observed in Linao (although there is not yet sufficient quantitative data to test these hypotheses). For example, Linao residents have strong opinions about the qualities of households that they believe to be giving and receiving the most (Phelps et al., 2022), suggesting that reputation may be an integral part of cooperative partner choice in Linao. Additionally, there is some limited evidence suggesting that Linao residents preferentially cooperate with distant kin in ways *not* predicted by kin selection theory, implying that additional cultural norms may be prescribing increased cooperative efforts in these scenarios. Thus, in order to more accurately replicate the cooperative patterns that characterize the Linao community, it will be necessary to carefully consider additional mechanisms that may be promoting or hindering cooperation within the village.

Chapter 5

CONCLUSION

In this paper, I described preliminary results from an agent-based model simulating the dynamic cooperative behaviors in a hypothetical subsistence population, which was informed by observations from Linao village in the Philippines. Despite some current limitations, this ABM was moderately successful in simulating realistic patterns of population structure and cooperative tendencies, and I statistically demonstrated that several of these patterns were broadly similar to what is observed in Linao. Even so, there are still many opportunities for improvement in the model. Modifications to the way in which mortality rates are implemented at each stage of an agent's life may assist with stabilizing population age structures and allowing agents to survive to older ages, while allowing agents and their spouses and offspring to function as "agent households" may facilitate more direct comparisons with human populations. Incorporating multiple daily attempts for needy agents to secure extra resources (as well as two-way interactions between agent dyads) may also improve the realism of daily cooperative interactions. Additionally, giving agents more "free will" with respect to partner choice could allow for complex patterns of assortment to emerge from the model, as are generally expected in small-scale human populations (K. M. Smith et al., 2018). However, it is critical to determine statistical modeling strategies that are more effective for the type of dyadic data that the ABM produces, since inappropriate modeling choices may inhibit accurate detection of mechanisms driving cooperation, resulting in false conclusions about the relative importance of specific cooperative mechanisms within the simulated environment.

There are a number of additional cooperative mechanisms that might be included in future iterations of the agent-based model. In particular, explicitly modeling cooperation between affines and between distant kin via cultural norms and/or indirect reciprocity effects could improve the model's ability to replicate the complex cooperation dynamics observed in Linao. Adding a reputational component to donors' evaluations of candidate recipients might also improve realism by providing another pathway for agents to select cooperative partners. It will certainly also be important to avoid over-complicating the model to the point where it becomes cumbersome, but employing calibration methodologies such as Approximate Bayesian Computation (Banks & Hooten, 2021; van der Vaart et al., 2016) to evaluate competing models against real-world data may provide a useful way forward as model complexity increases. Nevertheless, incorporating some or all of these changes should serve to improve structural realism with respect to cooperative patterns, allowing this agent-based model to become a "virtual laboratory" (Railsback & Grimm, 2012) within which I can validate existing empirical research and explore new theories concerning the evolution and maintenance of human cooperation.

REFERENCES

Acerbi, A., Snyder, W. D., & Tennie, C. (2022). The method of exclusion (still) cannot identify specific mechanisms of cultural inheritance. *Scientific Reports, 12*(1), 21680. https://doi.org/10.1038/s41598-022-25646-9

Alexander, R. D. (1987). *The biology of moral systems.* A. de Gruyter.

Allen-Arave, W., Gurven, M., & Hill, K. (2008). Reciprocal altruism, rather than kin selection, maintains nepotistic food transfers on an Ache reservation. *Evolution and Human Behavior, 29*(5), 305–318. https://doi.org/10.1016/j.evolhumbehav.2008.03.002

Alvard, M. (2011). Genetic and Cultural Kinship among the Lamaleran Whale Hunters. *Human Nature, 22*(1–2), 89–107. https://doi.org/10.1007/s12110-011-9104-x

Ballerini, M., Cabibbo, N., Candelier, R., Cavagna, A., Cisbani, E., Giardina, I., Lecomte, V., Orlandi, A., Parisi, G., Procaccini, A., Viale, M., & Zdravkovic, V. (2008). Interaction ruling animal collective behavior depends on topological rather than metric distance: Evidence from a field study. *Proceedings of the National Academy of Sciences, 105*(4), 1232–1237. https://doi.org/10.1073/pnas.0711437105

Banks, D. L., & Hooten, M. B. (2021). Statistical Challenges in Agent-Based Modeling. *The American Statistician, 75*(3), 235–242. https://doi.org/10.1080/00031305.2021.1900914

Blundell, R., & Meghir, C. (1987). Bivariate alternatives to the Tobit model. *Journal of Econometrics, 34*(1–2), 179–200. https://doi.org/10.1016/0304-4076(87)90072-8

Botkin, D. B., Janak, J. F., & Wallis, J. R. (1972). Some Ecological Consequences of a Computer Model of Forest Growth. *The Journal of Ecology, 60*(3), 849. https://doi.org/10.2307/2258570

Boyd, R., & Mathew, S. (2021). Arbitration supports reciprocity when there are frequent perception errors. *Nature Human Behaviour, 5*(5), 596–603. https://doi.org/10.1038/s41562-020-01008-1

Cameron, E. Z., Setsaas, T. H., & Linklater, W. L. (2009). Social bonds between unrelated females increase reproductive success in feral horses. *Proceedings of the National Academy of Sciences, 106*(33), 13850–13853. https://doi.org/10.1073/pnas.0900639106

Chattamvelli, R., & Shanmugam, R. (2021). *Continuous Distributions in Engineering and the Applied Sciences—Part I.* Springer International Publishing. https://doi.org/10.1007/978-3-031-02430-6

Clark, M. E., & Rose, K. A. (1997). Individual-based model of stream-resident rainbow trout and brook char: Model description, corroboration, and effects of sympatry and spawning season duration. *Ecological Modelling, 94*(2–3), 157–175. https://doi.org/10.1016/S0304-3800(96)00010-5

Cragg, J. G. (1971). Some Statistical Models for Limited Dependent Variables with Application to the Demand for Durable Goods. *Econometrica, 39*(5), 829. https://doi.org/10.2307/1909582

Croissant, Y., & Zeileis, A. (2018). *truncreg: Truncated Gaussian Regression Models.* https://CRAN.R-project.org/package=truncreg

Davidson, J. E. H., Hendry, D. F., Srba, F., & Yeo, S. (1978). Econometric Modelling of the Aggregate Time-Series Relationship Between Consumers' Expenditure and Income in the United Kingdom. *The Economic Journal, 88*(352), 661. https://doi.org/10.2307/2231972

DeAngelis, D. L., & Grimm, V. (2014). Individual-based models in ecology after four decades. *F1000Prime Reports, 6.* https://doi.org/10.12703/P6-39

DeAngelis, D., Shuter, B., Ridgway, M., & Scheffer, M. (1993). Modeling growth and survival in an age-0 fish cohort. *Transactions of the American Fisheries Society, 122*(5), 927–941.

Dunn, O. J. (1961). Multiple Comparisons among Means. *Journal of the American Statistical Association, 56*(293), 52–64. https://doi.org/10.1080/01621459.1961.10482090

Filho, H. S. B., de Lima Neto, F. B., & Fusco, W. (2011). Migration and social networks – An explanatory multi-evolutionary agent-based model. *2011 IEEE Symposium on Intelligent Agent (IA),* 1–7. https://doi.org/10.1109/IA.2011.5953616

Gabora, L., & Saberi, M. (2011). An agent-based model of the cognitive mechanisms underlying the origins of creative cultural evolution. *Proceedings of the 8th ACM Conference on Creativity and Cognition,* 299–306. https://doi.org/10.1145/2069618.2069667

García, J., van Veelen, M., & Traulsen, A. (2014). Evil green beards: Tag recognition can also be used to withhold cooperation in structured populations. *Journal of Theoretical Biology, 360,* 181–186. https://doi.org/10.1016/j.jtbi.2014.07.002

Garvey, R. (2018). Current and potential roles of archaeology in the development of cultural evolutionary theory. *Philosophical Transactions of the Royal Society B: Biological Sciences, 373*(1743), 20170057. https://doi.org/10.1098/rstb.2017.0057

Gomes, C. M., Mundry, R., & Boesch, C. (2009). Long-term reciprocation of grooming in wild West African chimpanzees. *Proceedings of the Royal Society B: Biological Sciences, 276*(1657), 699–706. https://doi.org/10.1098/rspb.2008.1324

Greene, W. H. (1994). *Accounting for excess zeros and sample selection in Poisson and negative binomial regression models* (NYU Working Paper No. EC-94-10). New York University. https://ssrn.com/abstract=1293115

Grimm, V., & Railsback, S. F. (2006). Agent-Based Models in Ecology: Patterns and Alternative Theories of Adaptive Behaviour. In F. C. Billari, T. Fent, A. Prskawetz, & J. Scheffran (Eds.), *Agent-Based Computational Modelling* (pp. 139-152). Physica-Verlag. https://doi.org/10.1007/3-7908-1721-X_7

Grimm, V., & Railsback, S. F. (2012). Pattern-oriented modelling: A 'multi-scope' for predictive systems ecology. *Philosophical Transactions of the Royal Society B: Biological Sciences, 367*(1586), 298–310. https://doi.org/10.1098/rstb.2011.0180

Grimm, V., Revilla, E., Berger, U., Jeltsch, F., Mooij, W. M., Railsback, S. F., Thulke, H.-H., Weiner, J., Wiegand, T., & DeAngelis, D. L. (2005). Pattern-Oriented Modeling of Agent-Based Complex Systems: Lessons from Ecology. *Science, 310*(5750), 987–991. https://doi.org/10.1126/science.1116681

Gurven, M., Allen-Arave, W., Hill, K., & Hurtado, A. M. (2001). Reservation food sharing among the Ache of Paraguay. *Human Nature, 12*(4), 273–297. https://doi.org/10.1007/s12110-001-1000-3

Gurven, M., & Hill, K. (2009). Why Do Men Hunt?: A Reevaluation of "Man the Hunter" and the Sexual Division of Labor. *Current Anthropology, 50*(1), 51–74. https://doi.org/10.1086/595620

Gurven, M., & Kaplan, H. (2007). Longevity Among Hunter-Gatherers: A Cross-Cultural Examination. *Population and Development Review, 33*(2), 321–365. https://doi.org/10.1111/j.1728-4457.2007.00171.x

Gurven, M., & Walker, R. (2006). Energetic demand of multiple dependents and the evolution of slow human growth. *Proceedings of the Royal Society B: Biological Sciences, 273*(1588), 835–841. https://doi.org/10.1098/rspb.2005.3380

Hamilton, W. D. (1964). The genetical evolution of social behaviour. I. *Journal of Theoretical Biology, 7*(1), 1–16. https://doi.org/10.1016/0022-5193(64)90038-4

Hawkes, K., & Coxworth, J. E. (2013). Grandmothers and the evolution of human longevity: A review of findings and future directions. *Evolutionary Anthropology: Issues, News, and Reviews, 22*(6), 294–302. https://doi.org/10.1002/evan.21382

Heckman, J. J. (1979). Sample Selection Bias as a Specification Error. *Econometrica, 47*(1), 153. https://doi.org/10.2307/1912352

Hendry, D. F., & Richard, J.-F. (1983). The Econometric Analysis of Economic Time Series. *International Statistical Review / Revue Internationale de Statistique, 51*(2), 111. https://doi.org/10.2307/1402738

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences, 33*(2–3), 61–83. https://doi.org/10.1017/S0140525X0999152X

Hill, K., & Hurtado, A. M. (1991). The evolution of premature reproductive senescence and menopause in human females: An evaluation of the "grandmother hypothesis." *Human Nature, 2*(4), 313–350. https://doi.org/10.1007/BF02692196

Hill, K., & Hurtado, A. M. (1996). *Aché life history: The ecology and demography of a foraging people.* Aldine de Gruyter.

Hill, K., & Hurtado, A. M. (2009). Cooperative breeding in South American hunter-gatherers. *Proceedings of the Royal Society B: Biological Sciences, 276*(1674), 3863–3870. https://doi.org/10.1098/rspb.2009.1061

Huth, A., & Wissel, C. (1992). The simulation of the movement of fish schools. *Journal of Theoretical Biology, 156*(3), 365–385. https://doi.org/10.1016/S0022-5193(05)80681-2

Jaeggi, A. V., & Gurven, M. (2013). Reciprocity explains food sharing in humans and other primates independent of kin selection and tolerated scrounging: A phylogenetic meta-analysis. *Proceedings of the Royal Society B: Biological Sciences, 280*(1768), 20131615. https://doi.org/10.1098/rspb.2013.1615

Jeltsch, F., Müller, M. S., Grimm, V., Wissel, C., & Brandl, R. (1997). Pattern formation triggered by rare events: Lessons from the spread of rabies. *Proceedings of the Royal Society of London. Series B: Biological Sciences, 264*(1381), 495–503. https://doi.org/10.1098/rspb.1997.0071

Kaplan, H., Hill, K., Lancaster, J., & Hurtado, A. M. (2000). A theory of human life history evolution: Diet, intelligence, and longevity. *Evolutionary Anthropology: Issues, News, and Reviews, 9*(4), 156–185. https://doi.org/10.1002/1520-6505(2000)9:4¡156::AID-EVAN5¿3.0.CO;2-7

Kasper, C., & Mulder, M. B. (2015). Who Helps and Why?: Cooperative Networks in Mpimbwe. *Current Anthropology, 56*(5), 701–732. https://doi.org/10.1086/683024

Kenny, D. A. (1975). Cross-lagged panel correlation: A test for spuriousness. *Psychological Bulletin, 82*(6), 887–903. https://doi.org/10.1037/0033-2909.82.6.887

Kim, P. S., Coxworth, J. E., & Hawkes, K. (2012). Increased longevity evolves from grandmothering. *Proceedings of the Royal Society B: Biological Sciences, 279*(1749), 4880–4884. https://doi.org/10.1098/rspb.2012.1751

Köhler, P., & Huth, A. (1998). The effects of tree species grouping in tropical rainforest modelling: Simulations with the individual-based model Formind. *Ecological Modelling, 109*(3), 301–321. https://doi.org/10.1016/S0304-3800(98)00066-0

Kokoszka, P., & Reimherr, M. (2017). *Introduction to functional data analysis.* CRC Press, Taylor & Francis Group.

Kraft, T. S., Venkataraman, V. V., Wallace, I. J., Crittenden, A. N., Holowka, N. B., Stieglitz, J., Harris, J., Raichlen, D. A., Wood, B., Gurven, M., & Pontzer, H. (2021). The energetics of uniquely human subsistence strategies. *Science, 374*(6575), eabf0130. https://doi.org/10.1126/science.abf0130

Lambert, D. (1992). Zero-Inflated Poisson Regression, with an Application to Defects in Manufacturing. *Technometrics, 34*(1), 1. https://doi.org/10.2307/1269547

Leigh, S. R. (2001). Evolution of human growth. *Evolutionary Anthropology: Issues, News, and Reviews, 10*(6), 223–236. https://doi.org/10.1002/evan.20002

Marlowe, F. W. (2005). Hunter-gatherers and human evolution. *Evolutionary Anthropology: Issues, News, and Reviews, 14*(2), 54–67. https://doi.org/10.1002/evan.20046

McKelvey, R. D., & Zavoina, W. (1975). A statistical model for the analysis of ordinal level dependent variables. *The Journal of Mathematical Sociology, 4*(1), 103–120. https://doi.org/10.1080/0022250X.1975.9989847

Mitani, J. C. (2006). Reciprocal exchange in chimpanzees and other primates. In P. M. Kappeler & C. P. van Schaik (Eds.), *Cooperation in Primates and Humans* (pp. 107–119). Springer Berlin Heidelberg. https://doi.org/10.1007/3-540-28277-7_6

Miu, E., & Morgan, T. J. H. (2020). Cultural adaptation is maximised when intelligent individuals rarely think for themselves. *Evolutionary Human Sciences, 2,* e43. https://doi.org/10.1017/ehs.2020.42

Nimmo, H. A. (2001). *Magosaha: An Ethnology of the Tawi-Tawi Sama Dilaut.* Ateneo de Manila University Press.

Nonaka, E., & Holme, P. (2007). Agent-based model approach to optimal foraging in heterogeneous landscapes: Effects of patch clumpiness. *Ecography, 30*(6), 777–788. https://doi.org/10.1111/j.2007.0906-7590.05148.x

Nowak, M. A., & Sigmund, K. (2005). Evolution of indirect reciprocity. *Nature, 437*(7063), 1291–1298. https://doi.org/10.1038/nature04131

Paraskevas, J., Eckerd, S., & Grimm, C. M. (2022). Driving cooperative actions: A multimethod study of the temporal duration of unilateral commitments. *Journal of Supply Chain Management, 58*(3), 3–22. https://doi.org/10.1111/jscm.12273

Pepper, J. W., & Smuts, B. B. (2000). The evolution of cooperation in an ecological context: An agent-based model. *Dynamics in Human and Primate Societies: Agent-Based Modeling of Social and Spatial Processes,* 45–76.

Phelps, J. R. & Hill, K. (2021). [Demographic data from Linao][Unpublished raw data]. Arizona State University.

Phelps, J. R., Pitogo, K. M. E., Emit, A. T., & Hill, K. (2022). Inter-household transfers of material goods among Sama "sea nomads" of the Philippines: Reciprocity, helping, signaling, or something else? *Manuscript Submitted for Publication.*

Pillay, N., & Rymer, T. L. (2015). Alloparenting enhances the emotional, social and cognitive performance of female African striped mice, Rhabdomys pumilio. *Animal Behaviour, 99,* 43–52. https://doi.org/10.1016/j.anbehav.2014.10.003

Queller, D. C. (2000). Relatedness and the fraternal major transitions. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 355*(1403), 1647–1655. https://doi.org/10.1098/rstb.2000.0727

Raab, D. H., & Green, E. H. (1961). A cosine approximation to the normal distribution. *Psychometrika, 26*(4), 447–450. https://doi.org/10.1007/BF02289774

Rademacher, C., Neuert, C., Grundmann, V., Wissel, C., & Grimm, V. (2004). Reconstructing spatiotemporal dynamics of Central European natural beech forests: The rule-based forest model BEFORE. *Forest Ecology and Management, 194*(1–3), 349–368. https://doi.org/10.1016/j.foreco.2004.02.022

Railsback, S. F., & Grimm, V. (2012). *Agent-based and individual-based modeling: A practical introduction.* Princeton University Press.

Reynolds, C. W. (1987). Flocks, herds and schools: A distributed behavioral model. *Proceedings of the 14th Annual Conference on Computer Graphics and Interactive Techniques - SIGGRAPH '87*, 25–34. https://doi.org/10.1145/37401.37406

Saul, S., Die, D., Brooks, E. N., & Burns, K. (2012). An Individual-Based Model of Ontogenetic Migration in Reef Fish Using a Biased Random Walk. *Transactions of the American Fisheries Society, 141*(6), 1439–1452. https://doi.org/10.1080/00028487.2012.697091

Sear, R., & Mace, R. (2008). Who keeps children alive? A review of the effects of kin on child survival. *Evolution and Human Behavior, 29*(1), 1–18. https://doi.org/10.1016/j.evolhumbehav.2007.10.001

Seyfarth, R. M., & Cheney, D. L. (2012). The Evolutionary Origins of Friendship. *Annual Review of Psychology, 63*(1), 153–177. https://doi.org/10.1146/annurev-psych-120710-100337

Shugart Jr, H., & West, D. C. (1977). Development of an Appalachian deciduous forest succession model and its application to assessment of the impact of the chestnut blight. *Journal of Environmental Management.*

Silk, J. B., Alberts, S. C., & Altmann, J. (2006). Social relationships among adult female baboons (Papio cynocephalus) II. Variation in the quality and stability of social bonds. *Behavioral Ecology and Sociobiology, 61*(2), 197–204. https://doi.org/10.1007/s00265-006-0250-9

Smith, D., Dyble, M., Major, K., Page, A. E., Chaudhary, N., Salali, G. D., Thompson, J., Vinicius, L., Migliano, A. B., & Mace, R. (2019). A friend in need is a friend indeed: Need-based sharing, rather than cooperative assortment, predicts experimental resource transfers among Agta hunter-gatherers. *Evolution and Human Behavior, 40*(1), 82–89. https://doi.org/10.1016/j.evolhumbehav.2018.08.004

Smith, K. M., Larroucau, T., Mabulla, I. A., & Apicella, C. L. (2018). Hunter-Gatherers Maintain Assortativity in Cooperation despite High Levels of Residential Change and Mixing. *Current Biology, 28*(19), 3152-3157.e4. https://doi.org/10.1016/j.cub.2018.07.064

Tarpy, D. R., Gilley, D. C., & Seeley, T. D. (2004). Levels of selection in a social insect: A review of conflict and cooperation during honey bee (Apis mellifera) queen replacement. *Behavioral Ecology and Sociobiology, 55*(6), 513–523. https://doi.org/10.1007/s00265-003-0738-5

Thomas, M. G., Ji, T., Wu, J., He, Q., Tao, Y., & Mace, R. (2018). Kinship underlies costly cooperation in Mosuo villages. *Royal Society Open Science, 5*(2), 171535. https://doi.org/10.1098/rsos.171535

Thompson, N. A. (2019). Understanding the links between social ties and fitness over the life cycle in primates. *Behaviour, 156*(9), 859–908. https://doi.org/10.1163/1568539X-00003552

Tonkinson, R. (1991). *The Mardu Aborigines: Living the Dream in Australia's Desert* (2nd ed). Holt, Rinehart and Winston.

Topping, C. J., Dalkvist, T., & Grimm, V. (2012). Post-Hoc Pattern-Oriented Testing and Tuning of an Existing Large Model: Lessons from the Field Vole. *PLoS ONE, 7*(9), e45872. https://doi.org/10.1371/journal.pone.0045872

Trivers, R. L. (1971). The Evolution of Reciprocal Altruism. *The Quarterly Review of Biology, 46*(1), 35–57. https://doi.org/10.1086/406755

Truncated Regression | R Data Analysis Examples. (n.d.). UCLA: Statistical Consulting Group. Retrieved February 28, 2023, from https://stats.oarc.ucla.edu/r/dae/truncated-regression/

Turner, M. G., Yegang Wu, Romme, W. H., & Wallace, L. L. (1993). A landscape simulation model of winter foraging by large ungulates. *Ecological Modelling, 69*(3–4), 163–184. https://doi.org/10.1016/0304-3800(93)90026-O

van der Vaart, E., Johnston, A. S. A., & Sibly, R. M. (2016). Predicting how many animals will be where: How to build, calibrate and evaluate individual-based models. *Ecological Modelling, 326,* 113–123. https://doi.org/10.1016/j.ecolmodel.2015.08.012

van Veelen, M., García, J., Rand, D. G., & Nowak, M. A. (2012). Direct reciprocity in structured populations. *Proceedings of the National Academy of Sciences, 109*(25), 9929–9934. https://doi.org/10.1073/pnas.1206694109

VanderWeele, T. J., & Mathur, M. B. (2019). Some Desirable Properties of the Bonferroni Correction: Is the Bonferroni Correction Really So Bad? *American Journal of Epidemiology, 188*(3), 617–618. https://doi.org/10.1093/aje/kwy250

Veall, M. R., & Zimmermann, K. F. (1994). Evaluating Pseudo-R2's for binary probit models. *Quality & Quantity, 28*(2), 151–164. https://doi.org/10.1007/BF01102759

Walker, R., Gurven, M., Hill, K., Migliano, A., Chagnon, N., De Souza, R., Djurovic, G., Hames, R., Hurtado, A. M., Kaplan, H., Kramer, K., Oliver, W. J., Valeggia, C., & Yamauchi, T. (2006). Growth rates and life histories in twenty-two small-scale societies. *American Journal of Human Biology, 18*(3), 295–311. https://doi.org/10.1002/ajhb.20510

Warsza, Z. L., & Korczynski, M. J. (2010). Shifted up cosine function as unconventional model of probability distribution. *Journal of Automation Mobile Robotics and Intelligent Systems, 4,* 49–55.

Wedekind, C., & Braithwaite, V. A. (2002). The Long-Term Benefits of Human Generosity in Indirect Reciprocity. *Current Biology, 12*(12), 1012–1015. https://doi.org/10.1016/S0960-9822(02)00890-4

Weiss, L., Pfestorf, H., May, F., Körner, K., Boch, S., Fischer, M., Müller, J., Prati, D., Socher, S. A., & Jeltsch, F. (2014). Grazing response patterns indicate isolation of semi-natural European grasslands. *Oikos, 123*(5), 599–612. https://doi.org/10.1111/j.1600-0706.2013.00957.x

Wiessner, P., & Huang, C. H. (2022). A 44-y perspective on the influence of cash on Ju/'hoansi Bushman networks of sharing and gifting. *Proceedings of the National Academy of Sciences, 119*(41), e2213214119. https://doi.org/10.1073/pnas.2213214119

Williams, P., Day, T., Fletcher, Q., & Rowe, L. (2006). The shaping of senescence in the wild. *Trends in Ecology & Evolution, 21*(8), 458–463. https://doi.org/10.1016/j.tree.2006.05.008

Wissel, C. (1992). Modelling the mosaic cycle of a Middle European beech forest. *Ecological Modelling, 63*(1–4), 29–43. https://doi.org/10.1016/0304-3800(92)90060-R

Wright, S. (1922). Coefficients of Inbreeding and Relationship. *The American Naturalist, 56*(645), 330–338. https://doi.org/10.1086/279872

APPENDIX A

LIST AND DESCRIPTION OF PARAMETERS AND INDICES USED IN AGENT-BASED MODEL

The following is a list of the various population- and environment-level parameters and components used in the model, accompanied by brief descriptions of each element. Said list is for the purpose of general reference, and additional details can also be found in the function code file (provided upon request). This list is broken into three sections: A short list of the various indexes used throughout the paper; a list of all of the parameters specified during model initialization; and a list of all the major variables calculated and used internally by the model. Importantly, intermediary variables (such as those used within multi-step calculations) are not included in the list below. Definitions of such variables can be found in the main text (Chapter 3, *Quantitative description of agent-based model algorithm*).

List A1.

General Indexing Notation Used in Model Specification

$t \coloneqq$ current model time-step (in days), where $t \in \{1, ..., T\}$
$i \coloneqq$ unique index for agent/individual, with $i \in \mathbb{N}$
$k \coloneqq$ unique index for each lineage, with $k \in \{1, ..., K\}$
$a \coloneqq$ a general, population-level index of age in days, where $a \in \{1, ..., A\}$

The following are parameters set by the user when initializing the simulation function. Some functional parameters were specified with default arguments, which are noted in the list below.

List A2.

Population- and Environment-Level Parameters Set by User

T $\coloneqq$ Max possible time-steps (days) in simulation, with $t \in \{1, ..., T\}$ indicating the current model time-step. Required that $T \in \mathbb{N}$. The specified value of T should be sufficiently large as to permit the desired minimum number of overlapping generations (i.e., it is recommended that $T \geq A \times$ number of desired generations).

t_y $\coloneqq$ Time-steps per year. Must be set s.t. $t_y \in \mathbb{N}$. It is practical to set t_y such that $T \bmod t_y = 0$, but this is not required.

K $\coloneqq$ Number of lineages (sub-populations) at initialization. Required that $K \in \mathbb{N}$.

N_k $\coloneqq$ Initial population size of the k^{th} lineage, where $k \in \{1, ..., K\}$. Specified either as a single integer value, or as a numeric vector of length K, where $N_k \in \mathbb{N}$ for all $k \in \{1, ..., K\}$. Can be constant (if specifying a single value) or variable across lineages (if specifying a length-K vector).

$\bar{r}$:= Mean coefficient of relatedness between all members *within* a lineage at initialization. (Relatedness between individuals in different lineages will be initialized at 0.) Used as a starting point for generating relatedness between agents in the same lineage; however, note that the actual mean relatedness of initialized lineages may be somewhat different, due to random sampling and the constraints necessarily imposed by the process of building a consistent relatedness network within the lineage. See Methods (Chapter 2) for more details. Required that $\bar{r} \in (0, 1/2]$, and recommended that $\bar{r}$ be specified as $(1/2)^n$ for some $n \in \mathbb{N}$. Default is $(1/2)^3 = 0.125$.

A_y := Maximum age, in years, that is included in the calculation of the baseline mortality function (Mort_{fd}; see below). Required that $A_y \in \mathbb{N}$. At this age and older, the baseline mortality rate of agents is set to 1. (Note that this is *not* an enforced maximum age in the population; see Chapter 2, Methods for more details.)

$\bar{\delta}_{m,y}$:= Age (in years) at which an agent is considered an "adult" and has full resource production capability. Fixed for population. Required that $0 \leq \bar{\delta}_{d,y} \leq \bar{\delta}_{m,y} \leq A_y$, and recommended that $\delta_{m,y}$ is significantly less than A_y (as fits the life history of the population being modeled). To simplify code within the agent-based model function, it is also required that $\bar{\delta}_{m,y} \in \mathbb{N}$.

$\bar{\delta}_{d,y}$:= Age (in years) *up to* which individual is *fully* dependent on their parents, with no resource production capability. Individuals with age greater than or equal to $\delta_{d,y}$ will have at least minimal age-specific resource production capability. Required that $0 \leq \bar{\delta}_{d,y} \leq \bar{\delta}_{m,y} \leq A_y$: If $\bar{\delta}_{d,y} = 0$, juvenile agents will be able to produce an age-scaled amount of resources starting on the day that they are born, while if $\bar{\delta}_{d,y} = \bar{\delta}_{m,y}$, an individual will be fully dependent on their parents (i.e. no individual production) until they are an adult. (See calculation of $\bar{\rho}_a$ in next section below for more details.) Also required that $\bar{\delta}_{d,y} \in \mathbb{N}$ for coding purposes. Fixed for population.

$\bar{\alpha}_y$:= Population-level mean age (in years) of first *possible* reproduction, i.e., age at which an individual is first able to select a reproductive mate with whom they have a non-zero probability of reproducing. Required that $\bar{\alpha}_y \in [\bar{\delta}_{d,y}, A_y]$, and recommended that $\bar{\alpha}_y << A_y$. Can be less than or greater than $\bar{\delta}_{m,y}$, as befitting the life history of the population being modeled. This parameter is the mean/median/mode of a raised cosine distribution of individual age at first possible reproduction (in years), and is used in conjunction with $s_{\alpha,y}$ below to randomly-generate individual values of this threshold age for all agents in the population. Consulting life history data for the population being modeled is suggested, as it will inform the combined selection of $s_{\alpha,y}$ and $\bar{\alpha}_y$. For additional details, see the description of α_i sampling in Chapter 2.

$s_{\alpha,y}$ $\coloneqq$ Spread (in years) around mean of individual age of first *possible* reproduction ($\bar{\alpha}_y$), assuming a raised cosine distribution. Required that $s_{\alpha,y} \in \mathbb{R}^+$ (a positive real number). It is also highly recommended that this parameter be set such that the entire support of the raised cosine distribution of age of first possible reproduction (i.e., $[\bar{\alpha}_y - s_{\alpha,y}, \bar{\alpha}_y + s_{\alpha,y}]$) lies within $[\bar{\delta}_d, A_y)$. While the model will not break if this recommendation is not followed, setting $s_{\alpha,y}$ in this way will avoid aberrant model behaviors such as fully-dependent juvenile agents who are already able to marry/reproduce (but are not able to produce any resources) and adults agents who can never marry/reproduce. Consulting life history data for the population being modeled is suggested, as it will inform the combined selection of $s_{\alpha,y}$ and $\bar{\alpha}_y$. For additional details, see the description of α_i sampling in Methods (Chapter 2).

p_{birth} $\coloneqq$ Bernoulli probability that a replacement agent is "born" versus "recruited". Born agents are assigned as juvenile dependents of a live reproductive pair of agents (the "parents"), while recruited agents are newly-initialized adult agents who are unrelated to all current agents in the population. Required that $p_{\text{birth}} \in [0, 1]$, and recommended that p_{birth} be set as a small probability, e.g., $p_{\text{birth}} \leq 0.1$.

$p_{\gamma,M}$ $\coloneqq$ Bernoulli probability of being male at birth, with $1 - p_{\gamma,M} \coloneqq$ probability of being female at birth. Each agent's sex at birth will be determined by a sample from the Bern($p_{\gamma,M}$) distribution, where an outcome of 0=female and an outcome of 1=male. Setting this value to something other than 0.5 will artificially change the expected sex ratio in the population. If sex ratio effects are not of interest, the default value is $p_{\gamma,M} = 0.5$, which will yield a population with an approximately balanced sex ratio. Required that $p_{\gamma,M} \in (0, 1)$, but note that it is possible that the population will crash for $p_{\gamma,M}$ set near boundaries (due to lack of reproductive adults of one sex or the other). Fixed for population.

Mort_{fd} $\coloneqq$ An `fd` object containing the "functional data object" fit of an age-specific "baseline" mortality table, as returned by the provided convenience function `mort_table_fun()`. Note that this function generates an `fd` object via the `fda` package in `R`, and that the age-specific mortality rate for each possible age is calculated on the daily scale ($[0, A]$), not the yearly scale. (This is all performed within the function `mort_table_fun()`, so the user need not make the conversion themselves.) Within the model function itself, "baseline" age-specific mortality rates are extracted from Mort_{fd} for each agent at each time step, using the `eval.fd()` function from the `fda` package. Fixed for population.

$\bar{\mu}_{\max}$:= Max number of agents who can die at each timestep. Required that $\bar{\mu}_{\max} \in \mathbb{N}$ such that $0 < \bar{\mu}_{\max} < \sum_k N_k$. Recommended that this be set to a value that is consistent with the mortality data used to generate Mort_{fd}. Fixed for population.

$\bar{\eta}_m$:= Mean daily *gross* need of adults in population, before discounting by any production or provisioning from others. Required that $\bar{\eta}_m \geq 0$. Note that, for simplicity, it is assumed that need does not decline in later adult ages. (This may be modified in future versions of the ABM.) Fixed for population.

$\bar{\eta}_\varnothing$:= Mean daily *gross* need of population agents at age 0, i.e. minimum daily resources needed by individuals at birth, before discounting by any production or provisioning from others. Required that $0 \leq \bar{\eta}_\varnothing \leq \eta_m$, and recommended that $0 < \bar{\eta}_\varnothing < \eta_m$. Fixed for population.

α_ρ, β_ρ := Parameters given to control the shape of a $\text{Beta}(\alpha_\rho, \beta_\rho)$ distribution used for generating random daily production of resources. Required that $\alpha_\rho, \beta_\rho > 1$; default values are $\alpha_\rho = 2$ and $\beta_\rho = 2.5$, although these must be fine-tuned alongside the $\bar{\rho}_m$ parameter below to produce the desired population resource abundance/scarcity. See Methods (specifically *Step 5. Gross Daily Need*) for more details on this calculation and how it is adjusted to reflect age-specific production rates. Fixed for population.

$\widetilde{\rho}_m$:= The *mode* of daily adult resource production in population, which is used to scale the $\text{Beta}(\alpha_\rho, \beta_\rho)$ distribution used for generating random daily production of resources by an age-specific amount. Required that $\bar{\rho}_m > \bar{\eta}_m$. Recommended that $\bar{\rho}_m$ be larger than $\bar{\eta}_m$ in order to prevent premature population crashes, particularly if the initial "willingness to cooperate" parameters (see $\bar{\omega}$ below) are set to lower values. See Methods (specifically *Step 5. Gross Daily Need*) for more details on this calculation and how it is adjusted to reflect age-specific production rates. Fixed for population.

$\vec{\omega}$:= A vector $\{\bar{\omega}_B, \bar{\omega}_R, \bar{\omega}_K\}$ of population means at initialization for the three heritable cooperative traits: $\bar{\omega}_B$ is the baseline willingness to cooperate with anyone in the population, $\bar{\omega}_R$ is willingness to cooperate with reciprocity partners, and $\bar{\omega}_K$ is willingness to cooperate with kin. Values must be specified in the order above. Additionally, all three traits are on the log-odds scale – so, while no firm requirement is made about their values, it is recommended that $\vec{\omega}$ be specified such that $-3 \leq \bar{\omega}_* - s_{\omega_*} < \bar{\omega}_* + s_{\omega_*} \leq 3$ for $* = B, R, K$ and spread vector $\vec{s}_\omega$ (described below). Values in $\vec{\omega}$ are used in conjunction with their corresponding values in $\vec{s}_\omega$ to sample initialized agent traits from raised cosine distributions. (After initialization, traits are inherited with a small amount of noise from one parent.)

$\vec{s}_\omega$:= A vector $\{s_{\omega_B}, s_{\omega_R}, s_{\omega_K}\}$ of the distribution spreads around the population means of the three heritable traits in $\bar{\omega}$. Must be specified in the given order. See $\bar{\omega}$ above for more details.

k_μ := A weight parameter which controls the impact of resources on mortality rate (on the logit scale). Required that $k_\mu \geq 0$, and recommended that $k_\mu > 0$. Default is $k_\mu = 1$. See *Step 8. Cooperative Transfers* subsection of the Methods section for more details.

k_H := A decay parameter controlling the influence of past cooperative interactions (relative to a current cooperative interaction) in the calculation of an agent-pair's "cooperative history" value ($H_{i,j,t}$). This parameter can be thought of as how well a prospective donor agent i remembers a candidate recipient j's past cooperative actions towards i. (That is, how well i remembers past decisions by j to either help or deny help to i.) Smaller values of k reduce the importance of past actions on the current $H_{i,j,t}$ calculation, while larger values increase the importance of past actions. For more information, see the *Step 8. Cooperative Transfers* subsection of the Methods section. Required that $k \in [0, 1]$, but note that $k = 0$ implies no memory (i.e., no reciprocity), while $k = 1$ implies perfect memory. Fixed for population.

The following are internally-computed by the model function. Most are parameters that are simply converted to the age-in-days scale; however, the last three are age-specific calculations derived from the user-specified variables.

List A3.

Population- and Environment-Level Internal Parameters/Variables Derived From User-Specified Parameters

A := Maximum age, in days, that is included in the calculation of the baseline mortality function (Mort_{fd}). Computed internally as $A = A_y t_y$, with $a \in \{1, ..., A\}$ used as a general, population-level index of age in days.

$\bar{\delta}_m$:= Age, in days, at which individual is considered an adult with full resource production capability. Computed internally as $\bar{\delta}_m = \bar{\delta}_{m,y} t_y$.

$\bar{\delta}_d$:= Age, in days, up to which individual is considered fully dependent on their parents, with no resource production capability. Computed internally as $\bar{\delta}_d = \bar{\delta}_{d,y} t_y$.

$\bar{\alpha}$:= Population-level mean age, in days, of first *possible* reproduction. Computed internally as $\bar{\alpha} = \bar{\alpha}_y t_y$.

s_α $\coloneqq$ Spread around population-level mean age (in days) of first *possible* reproduction. Computed internally as $s_\alpha = s_{\alpha,y} t_y$.

$\bar{\mu}_a$ $\coloneqq$ Age-specific "baseline" mortality rate, i.e. mortality rate at age a (in days), when agent's needs are fully met with no surplus. This value is fixed for all members of the population at each age a. When specified in reference to a particular agent i with age $a_{i,t}$ at time t, this is denoted as $\bar{\mu}_{a_{i,t}}$. The value of $\bar{\mu}_a$ for each possible age a is extracted from the Mort_{fd} functional data object provided by the user.

$\widetilde{\rho}_a$ $\coloneqq$ Age-specific mode of daily production, i.e. the *mode* of the distribution of production per day per individual of age a. In reference to a particular agent i with age $a_{i,t}$ at time t, it is denoted as $\widetilde{\rho}_{a_{i,t}}$. Computed as follows:

$$\widetilde{\rho}_{a_{i,t}} = \widetilde{\rho}(a_{i,t} \mid \bar{\rho}_m, \bar{\delta}_d, \bar{\delta}_m) = \begin{cases} 0, & \text{if } a_{i,t} < \bar{\delta}_d \\ \dfrac{\bar{\rho}_m}{\bar{\delta}_m - \bar{\delta}_d}(a_{i,t} - \bar{\delta}_d), & \text{if } \bar{\delta}_d \leq a_{i,t} \leq \bar{\delta}_m \\ \bar{\rho}_m, & \text{if } a_{i,t} > \bar{\delta}_m. \end{cases}$$

See *Step 6. Gross Daily Resource Production* in the Methods section for more details on this calculation.

$\bar{\eta}_a$ $\coloneqq$ Age-specific level of "gross need," i.e. population average need per day per individual as a function of age a, calculated before discounting by any individual production or provisioning from others. When referencing the particular value of an agent i with age $a_{i,t}$ at time t, it is denoted as $\bar{\eta}_{a_{i,t}}$. Computed as follows:

$$\bar{\eta}_{a_{i,t}} = \bar{\eta}(a_{i,t} \mid \bar{\eta}_\varnothing, \bar{\eta}_m, \bar{\delta}_m) = \bar{\eta}_\varnothing + \frac{\bar{\eta}_m - \bar{\eta}_\varnothing}{\bar{\delta}_m} y, \quad \text{for } y = \min\{a_{i,t}, \bar{\delta}_m\}.$$

APPENDIX B

LIST OF VARIABLES DEFINING STRUCTURE OF EACH AGENT

At time $t \in 1:T$, agent $I_{i,t}$ with unique identifier i and lineage $k \in 1:K$ is characterized by the following set of variables. Convenience variables (i.e., those used in the code but not important to the overall model) are denoted in red and are not described in the main text.

$$
I_{i,t} = \begin{cases}
k_i & := \text{lineage (k) of individual, inherited from mother} \\
g_i & := \text{generation number (minimum of parents' generation numbers plus 1)} \\
\gamma_i & := \text{sex of individual (female = 0; male = 1)} \\
\alpha_i & := \text{age at first possible reproduction} \\
\vec{\omega}_i & := \text{vector of willingness to cooperate with others, inherited from parents} \\
b_i & := \text{coefficient of imbreeding of agent } i \\
r_{i,j} & := \text{coefficient of relatedness to each agent } j, \text{ where } j \neq i \\
M_i & := \text{index of } i\text{'s mother} \\
F_i & := \text{index of } i\text{'s father} \\
R_i & := \text{set of indexes of all known relatives of } i \\
a_{i,t} & := \text{age at time } t \\
\mathbb{I}_{\mu,i,t} & := \text{indicator of whether agent } i \text{ is alive(=0) or dead(=1) after mortality step at time } t \\
\mathbb{I}_{\mu,MF,i,t} & := \text{indicator of whether at least one of agent } i\text{'s parents is alive(=0), or if both parents are dead(=1)} \\
\mathbb{I}_{\text{indep},i,t} & := \text{indicator of whether agent } i \text{ is a juvenile dependent(=0) of another agent or independent(=1) at time } t \\
S_{i,t} & := \text{index of } i\text{'s current spouse, if they exist; else, empty set} \\
D_{i,t} & := \text{set of indexes of dependent juvenile offspring at time } t, \text{ if they exist; else, empty set} \\
\rho_{i,t} & := \text{individual resource production at time } t \\
\eta_{i,t} & := \text{net need at time } t \text{ (including dependent needs, if any)} \\
C_{i,t} & := \text{index of agent } i\text{'s cooperative partner at time } t, \text{ if one has been randomly assigned} \\
\rho_{C,i,t} & := \text{amount transferred between } i \text{ and } C_{i,t} \text{ at time } t \\
\Delta\mu_{i,t} & := \text{change in mortality at time } t+1 \text{ due to unmet need at time } t \\
H_{i,j,t} & := j\text{'s history of cooperative assistance given to } i \text{ at time } t, \text{ as recalled by agent } i
\end{cases}
$$

APPENDIX C

SIMPLE LINEAR REGRESSION MODELS OF AVERAGE RELATEDNESS, COOPERATIVE HISTORY, AND PROPORTION OF SUCCESSFUL COOPERATIVE EVENTS BY DECADE

The following are ordinary simple linear regression models that examine possible time trends in population mean relatedness, mean history, and proportion of cooperative interactions. Data is aggregated at the decade level and shown in Figure 6, main text. Regressors and outcomes are standardized so that Pearson correlations between the two can be inferred directly from parameter estimates. Note that p-values are given only for reference, and should not be closely interpreted due to the nature of the simulated data.

Table C1

Simple Linear Regression of Mean Pairwise Relatedness by Decade

Response Variable: Mean $r_{i,j}$ Within Decade				
Effect	Estimate	Est. SE	t-Stat	Pr($>\|t\|$)
(Intercept)	< 0.00001	0.0819	< 0.001	> 0.99999
Decade	0.01650	0.0822	0.201	0.84116
Observations = 150; DF = 148				
Residual SE ≈ 1.003				
Multiple $R^2 < 0.001$; Adjusted $R^2 \approx$ -0.006				

Note. SE = Standard Error; DF = Degrees of Freedom.

Table C2

Simple Linear Regression of Mean Pairwise Cooperative History by Decade

Response Variable: Mean $H_{i,j}$ Within Decade				
Effect	Estimate	Est. SE	t-Stat	Pr($>\|t\|$)
(Intercept)	$>$ -0.00001	0.0814	< 0.001	> 0.99999
Decade	0.11512	0.0817	1.410	0.16069
Observations = 150; DF = 148				
Residual SE ≈ 0.998				
Multiple $R^2 \approx 0.013$; Adjusted $R^2 \approx 0.007$				

Note. SE = Standard Error; DF = Degrees of Freedom.

Table C3

Simple Linear Regression of Proportion of Cooperative Events That Were Successful (Resources Transferred) by Decade

Response Variable: Proportion Successful Events Within Decade				
Effect	Estimate	Est. SE	t-Stat	$\Pr(>\|t\|)$
(Intercept)	< 0.00001	0.0774	< 0.001	> 0.99999
Decade	-0.32822	0.0776	-4.227	0.00004
Observations = 150; DF = 148				
Residual SE ≈ 0.948				
Multiple $R^2 \approx 0.108$; Adjusted $R^2 \approx 0.102$				

Note. SE = Standard Error; DF = Degrees of Freedom.

APPENDIX D

SIMPLE LINEAR REGRESSION MODELS OF WILLINGNESS TO COOPERATE BY AGENT GENERATION

The following are ordinary simple linear regression models that examine possible generational trends in the three cooperation traits. Mean estimates of the three traits are aggregated within generations and shown in Figure 7, main text. Regressors and outcomes are standardized so that Pearson correlations between the two can be inferred directly from parameter estimates. Note that p-values are given only for reference, and should not be closely interpreted due to the nature of the simulated data.

Table D1

Simple Linear Regression of Mean Baseline Willingness to Cooperate ($\bar{\omega}_B$) by Agent Generation

Response Variable: $\bar{\omega}_B$ Within Generation				
Effect	Estimate	Est. SE	t-Stat	Pr($>\|t\|$)
(Intercept)	< 0.00001	0.1161	< 0.001	> 0.99999
Generation Number	-0.66889	0.1175	-5.691	< 0.00001
Observations = 42; DF = 40				
Residual SE ≈ 0.753				
Multiple $R^2 \approx 0.447$; Adjusted $R^2 \approx 0.434$				

Note. SE = Standard Error; DF = Degrees of Freedom.

Table D2

Simple Linear Regression of Mean Willingness to Cooperate With Reciprocity Partners ($\bar{\omega}_R$) by Agent Generation

Response Variable: $\bar{\omega}_R$ Within Generation				
Effect	Estimate	Est. SE	t-Stat	Pr($>\|t\|$)
(Intercept)	$>$ -0.00001	0.0626	< 0.001	> 0.99999
Generation Number	-0.91615	0.0634	-14.456	< 0.00001
Observations = 42; DF = 40				
Residual SE ≈ 0.406				
Multiple $R^2 \approx 0.839$; Adjusted $R^2 \approx 0.835$				

Note. SE = Standard Error; DF = Degrees of Freedom.

Table D3

Simple Linear Regression of Mean Willingness to Cooperate With Kin ($\bar{\omega}_K$) by Agent Generation

Response Variable: $\bar{\omega}_K$ Within Generation				
Effect	Estimate	Est. SE	t-Stat	$\Pr(>\|t\|)$
(Intercept)	$>$ -0.00001	0.1423	< 0.001	> 0.99999
Generation Number	0.41315	0.1440	2.869	0.00654
Observations = 42; DF = 40				
Residual SE ≈ 0.922				
Multiple $R^2 \approx 0.171$; Adjusted $R^2 \approx 0.150$				

Note. SE = Standard Error; DF = Degrees of Freedom.

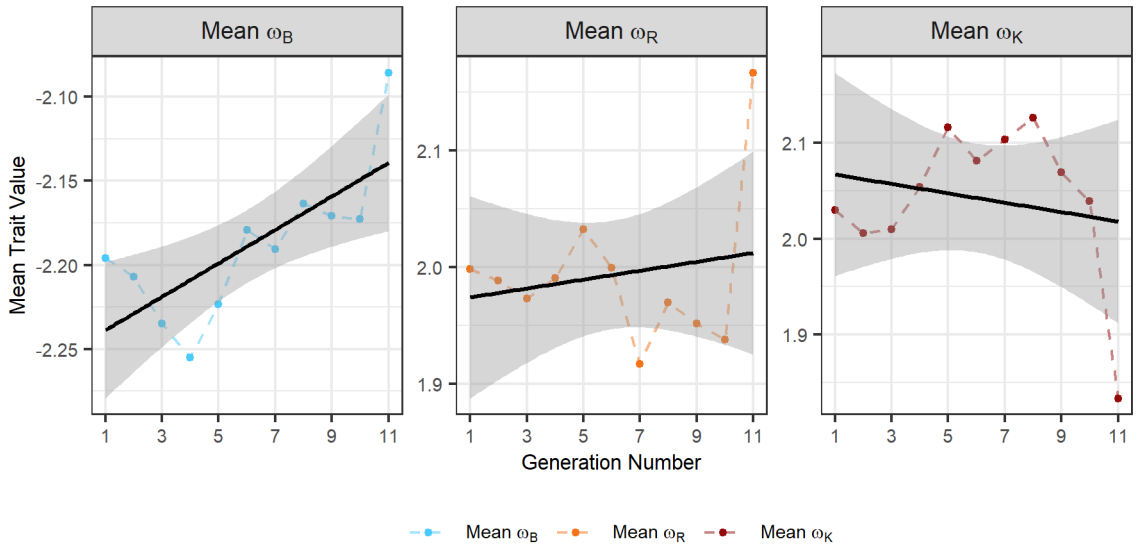
APPENDIX E

COOPERATION TRAITS IN EARLY VERSION OF ABM WHERE ADULT AGENTS WERE NOT RECRUITED

Cooperation traits by generation for an early version of the model, in which unrelated adult agents *were not* recruited into the model. This model had a higher-overall pairwise relatedness between agents ($\bar{r} > 0.2$).

Figure E1

Evolution of "Willingness to Cooperate" Traits Over Agent Generations in Early Version of ABM



Note. Dots indicate mean values within each generation. Black lines are simple linear regressions of mean trait value per generation (outcome) by generation number. Gray bands around regression lines are 95% confidence intervals.

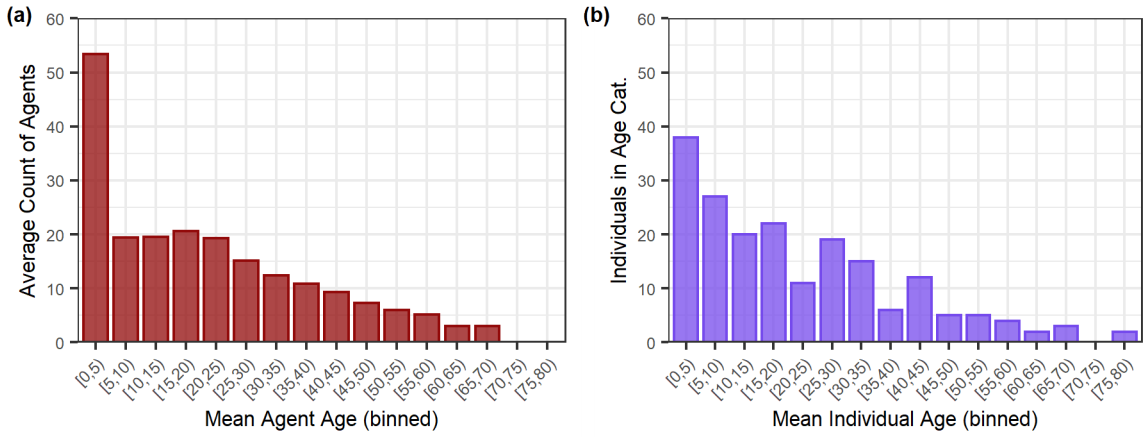
APPENDIX F

REPLICATION OF POPULATION AGE DISTRIBUTION COMPARISON, OMITTING INFANT AGENT DEATHS

The following is a replication of Figure 8 and its associated Mann-Whitney-Wilcoxon test in the main text. In the following figure and test, agents who die at less than one year of age are omitted from the average agent age distribution to determine if these infant agent deaths are driving the statistically-significant differences in population age distribution detected in the main text. A total of 17 juvenile agents are omitted from the average agent age distribution, which raises the mean age of estimated agent population to approximately 20.4 years old. See the *Comparison between final agent-based model and field data from Linao* section of Chapter 3 (Results) for additional details.

Figure F1

Comparison Between Mean Population Age Structure in Agent-Based Model and Population Age Structure in Linao, Excluding Agent Deaths at Ages Below One Year



Note. (a) Average age structure of the ABM population over consecutive 4-year subsets of data between year 901 and year 1000, when excluding agents who die at age $a_{i,t} < 1$ year old. (b) Age structure of Linao village over 3.5 years of observation. In both plots, age is binned into 5-year categories to facilitate observation of underlying trend.

Table F1

Mann-Whitney-Wilcoxon Test Comparing Linao Age Distribution With Distribution of Estimated Agent Population Ages When Omitting Infant Agent Deaths

H_0: true location shift θ is equal to zero			
H_a: true location shift θ is not equal to zero			
Estimated Location Shift ($\hat{\theta}$)	95% Confidence Interval	W-Stat	Pr($>\|W\|$)
0.99995	(-1.99995, 3.99998)	20302	0.46985

Note. Test performed using a Normal approximation.