Received May 27, 2018, accepted June 12, 2018, date of publication June 27, 2018, date of current version July 25, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2849731

An Agent Based Model Approach for Perusal of Social Dynamics

KARANDEEP SINGH AND CHANG-WON AHN

Department of Computer Software, Korea University of Science and Technology, Daejeon 34113, South Korea
IDX Cognitive Informatics Research Section, SW. Contents Research Laboratory, ETRI, Daejeon 34129, South Korea

Corresponding author: Karandeep Singh (karandeep.singh@etri.re.kr)

This work was supported in part by the Ministry of Science and ICT, South Korea, through the R&D Program of NST and the ETRI R&D Program under Grant Cognitive Informatics Research: 18ZS1230

ABSTRACT Agent-based modeling has recently gained popularity in the field of simulation and modeling. Due to their characteristic properties, agent-based models (ABMs) allow for an improved and flexible way of modeling complex systems. Social dynamics is one such a complex system, which has complex components, such as demography, sociology, economics, psychology, health, and so on. Demography is one of the bigger and more important sub-systems of this complex system. In this paper, we have focused on utilizing the potential of ABM techniques to analyze the underlying processes in social demography. We propose and implement a holistic ABM that can be used for analysis, understanding, and prediction of socio-demographic processes as well as a tool for policy design and evaluation. The proposed model incorporates well-known factors affecting demography and provides ease and flexibility of adding newer factors. In this paper, we considered the use case of Korea and utilized the Korean census data for model development. Many ABMs have been proposed for demography but most of them not only have limited functionalities, but also lack the suitable usage of agent-based modeling itself. The proposed approach is wide-ranging, flexible, and general and can be reused by simply making changes in the initial data set for the analysis of the target society. We exhibit the validation of the proposed model, as well as its usage by executing virtual experiments for socio-demographic analysis and prediction.

INDEX TERMS Agent based modeling, policy evaluation, population dynamics, population prediction, social demography, discrete event system.

I. INTRODUCTION

Demography can be generalized as the study of human populations and associated factors that affect population size and structure. In addition, social demography could incorporate various sub-factors such as the economy, society, culture, aging societies, race, gender etc. These demographic processes are studied vis-à-vis the description, understanding, prediction, and analysis of changes in the size and structure of the populations. The changes and variations brought about by the various factors are spatial and temporal in nature. In this paper, we focus on the demographic processes from the perspective of size and the age structure.

A complex system like demography consists of various intertwined and interlinked sub-systems. In general, the model building of a complex system is approached by aggregating simpler sub-systems. More often than not, the heterogeneous nature of these sub-systems is ignored and they are modeled uniformly, or heterogeneity is present only to a very limited extent. Popular examples are statistical approaches and microsimulation (MiSiM) models [1]. Statistical and MiSiM models are well-developed and widely used in demographic simulations. In statistical models, the relationship between the developed model and the target is generally very clearly understood [2]. The target system is modeled by mathematically derived equations, because of which, these techniques are brittle and cannot adjust to changes in system structure over time. MiSiM models individuals with microdata from real population data and provides the final output by accumulation of individual simulated outputs. MiSiM, as well as statistical approaches, do not attempt to enable interactions between the constituent units in the model. Statistical and MiSiM techniques are difficult to implement for a complex system, especially where non-linearity is involved. These approaches are not suitable for understanding and analysis of various sub-systems, and how they interact to affect the final outcome, but are suitable tools for prediction. Although MiSiM and agent-based models (ABMs) have similar features like heterogeneity and being bottom-up approaches, they show some key differences. In MiSiM, the final output is the addition of individual outputs from its component units, whereas in ABM, the final output can be more or less than the addition of individual outputs, because agents in the ABMs can interact and influence each other's characteristic behavior.

The ability of ABMs to enable interactions between the agents or constituent units, which reflects the essence of any complex system, has made them overcome some of the inherent drawbacks in statistical and MiSiM approaches. Agent-based social simulation, or agent based social modeling is a generalization of the approach that contains all the aspects of social demography. When a complex system is implemented, various assumptions are made and only the required components of the target complex system are implemented. However, if such assumptions are not made diligently and carefully, they flaw the very purpose of the simulation, such as understanding and prediction of the target system.

Recently, the research on the socio-demographic processes has received a special attention from governments of developed countries. These countries face problems of aging societies and declining fertility rates. Countries with very low fertility rates include Singapore (0.82), Korea (1.25), Japan (1.4), Italy (1.43), Greece (1.42), and numerous others [3]. Fig. 1 (source [4]) shows the population pyramids of Korea, India, and the USA. Fig. 2a shows Korean population age structure over time and dependency ratio (DR) of population. Korea, which has faced the problem of low-fertility and ageing societies for some years now, has a top-heavy graph. This shows that the elderly population is larger as than the youth population. On the other hand, India has a larger youth population than the elderly population, and hence a bottom-heavy graph. The population pyramid of the USA is comparatively balanced, with rather comparable youth and elderly population. Fig. 2b shows stark difference in median ages of some developing and developed countries. Low fertility leads to various economic and social problems in the long run, which we will discuss in Section II of the paper. A better understanding of various processes involved in social demography is urgently needed so that issues pertaining to aging societies and declining birth rates can be handled by planning appropriate strategies and policies.

On the basis of the above discussion, we developed an ABM for social demography which includes demographic processes of matrimony, education, fertility, the economy, mortality, and immigration. The model is general in nature as it takes as input a sample of the population, including males, females and child agents. Calibrated with the real data, education, income, and rates of demographic activities like mortality, fertility, and divorce evolve continuously. The

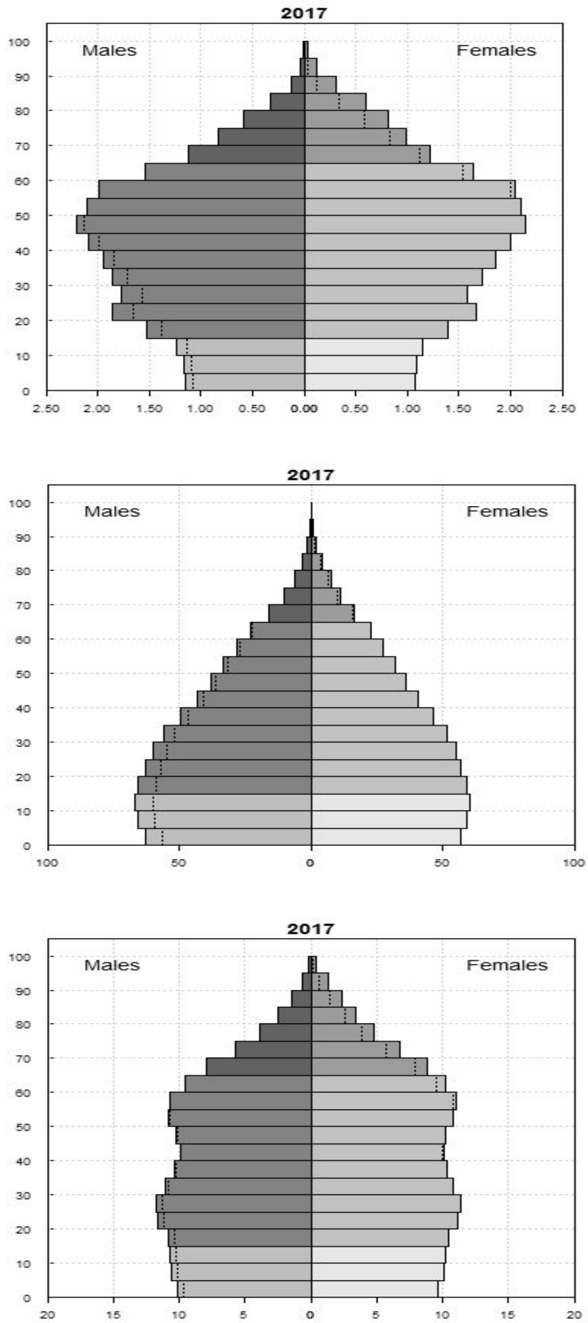

FIGURE 1. Population Pyramids (from top to bottom) of Korea, India and USA. The countries with the problem of low fertility rate and ageing societies are 'top-heavy'. India has high youth population, hence bottom heavy. USA has a comparatively balanced graph. The vertical axis is the age numbers in years, horizontal axis represents population numbers.

updates for every agent characteristic with simulation time give a very dynamic nature to our model. Moreover, the interlinking of characteristics of the agents makes the model simulation emergent and complex in nature. The agents' social interactions with other agents, and with the environment, have an effect on the decision process of the individual agents, such as fertility, income, education. Thus, not only are agents' interactions carried out at the micro (individual) level, but also at the macro (society or environment) level.

The important choices made by the agents for demographic events are deterministic in nature and are determined by the agents' rule set, derived either from the accepted social norms or from the real data. Some of the arguments and features that prove the novelty of our model are as follows: an improved general model of the social demography, flexible nature of the model, improved ABM nature, direct inclusion of population data in the simulation, dynamic nature of all the related demographic events, usage of high-quality census data for initialization and agent rule-set design, and ability to analyze as well as predict (generally a property of MiSiM, due to usage of real data). Although we have utilized Korean data for this particular work, owing to the flexible nature of this model, it can be put to use by modelers and researchers with desirable data. Additionally, newer features can easily be added or extended as per need to include other desirable components like health care, economics, and psychology.

The proposed model aims and overcomes most of the drawbacks in popular related works, discussed in the next section. As stated previously, the proposed model is built to be general purpose in nature. It does not target a specific demographic activity, unlike the related works, and implements all the major demographic activities such as matrimony, fertility, education, economy, mortality, and immigration. The approach of targeting a particular sub-system of a larger complex system ends up making over-generalized assumptions about other sub-systems. While modeling complex systems such as demography, this approach will not yield desirable understanding of the target system. The reason is coupling, dependencies and interactions that exists between the sub-systems, all of which may evolve according to changes in other sub-systems. The final output of a complex systems is more/less than the simple addition of the outputs of its constituent sub-systems. Therefore, the proposed model incorporates the deterministic models of most of the important sub-systems of demography. Further, we make use of actual micro-level census data for initializing the simulation as well as the agent characteristics. Even though some of the other models make use of realistic data, the actual values are often assigned by making the use of statistical distributions around these values. However, this process ignores and alters the micro-characteristics of a population sample and results in non-optimal results.

In the proposed model, the micro-data, as detailed as the agent's household members and relationships, is assigned by directly using census data. This level of integration of micro-level data with the agent based modeling has enabled the model not only to deliver the analysis capabilities, but also the prediction capabilities. In addition, we designed every characteristic of the agent dynamic in nature, where the updated values are again a function of historical data at the start of the simulation and of newly generated data during the simulation. The exchange of information and interactions of the agents during the simulation result in changes in the individual

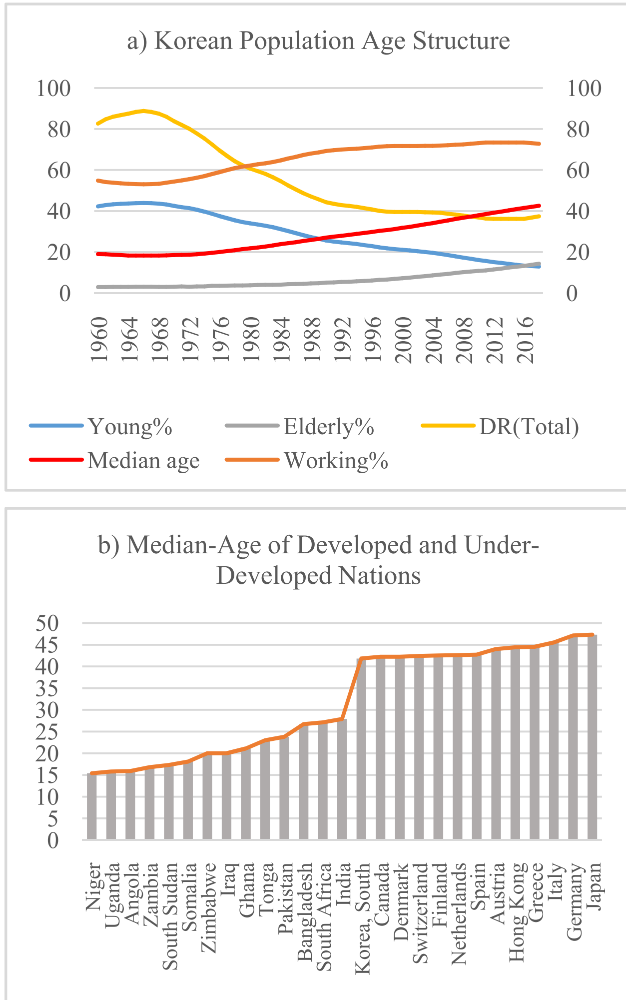

FIGURE 2. Historic data of Korean Population age structure, and Median age of some nations of world (Fig. a) and b) respectively). Fig. a) shows how Korean age structure is changing with time. DR stands for population dependency ratio. Fig. b) shows median ages of some developing and developed nations. The vertical axis in both the figures is age in years, horizontal axis in Fig. a) is calendar years, and in Fig. b) it is country names.

preferences of the agents. This leads to emergence of different demographic behaviors, making the simulation process complex in nature (like demography). Another argument that makes our proposed model holistic in nature is that it has a snapshot of real population as sampled agent population and is not a ceteris paribus. Various points discussed above deliver their benefits in the form of simulation results demonstrated in subsequent sections of the paper. Hence, the proposed model is unique and novel and is a step forward in simulating the demographic system using agent-based modeling.

The rest of the paper is organized as follows: the second section covers the Background and Motivation; the third section explains the Proposed Model; the fourth section discusses Implementation; the fifth section contains Experimental Results and Discussion; and the sixth section gives conclusion to the paper and discusses Future Directions.

II. BACKGROUND AND MOTIVATION

An ABM consists of autonomous agents with their individual rule-sets that are utilized for their decision processes. They consider the environment where they are situated, their own characteristic properties, and most important of all, the interactions with other agents [5]. Agents have properties like autonomy, perceptive, pro-active, learning, heterogeneity, bounded rationality, mobility, reactive, and especially interactions with other agents [6].

Billari *et al.* [5], [7] initiated the agent based social simulation for population dynamics. Preceding research in demographic contexts has been proposed in areas like household dynamics [9], residential mobility [8], from the interaction perspective, such as family formation [5], and migration [6]. These models take into account interactions but lack features such as usage of accepted social norms, usage of solid high quality data, and the scope of demographic activities with regards to model is limited.

The ABM models *often* take into account the micro-level or agent-specific characteristics but generally real data for simulation initialization is not considered. The rules and inferences used by agents are also often assumed analytically rather than deriving the rules based on actual circumstances of the population in consideration. This increases the chances of *outliers* in the simulations and results. Moreover, analytically assumed ABMs can be tools for limited analysis of the target system, but they cannot be used for prediction purposes.

The data-fed simulations were proposed in papers like [10] (which was extension of famous "Wedding Ring" [7]), [11], in which the usage of data into agent based simulation was proposed, [12]–[14] proposed the general population dynamics models, but the adequate ABM part in the models was lacking, and were largely governed by microsimulation-like processes. There are other examples in research which use real data for simulation initialization, and even take into account the interactions between the agents, but either are one sex models *ceteris paribus* [16], or use very low no. of agents (∼500) [15]. These approaches cannot be put to efficient use for demography, which includes large numbers of agents, as well dynamic updates of the majority of related events in simulation/ real time.

III. PROPOSED MODEL

A layman might take lower fertility rates and population decrease as good population behavior, as a decreasing population could mean the same resources would be shared by fewer people. However, that is not actually the case. Low fertility rates have a range of associated problems. Over time, it means a fewer children are being born, leading to fewer people growing up to be available as the work force. Further, this problem is generally found in developed countries like Japan, Korea, Spain, Germany, and many other European and Asian countries. Declining birth rates and increasing life expectancies in developed countries would eventually result in a decreased work force, and governments would be burdened to provide facilities catering to aging populations [17]. The median age in countries like Italy, Spain, Germany, Poland, is increasing, which is a sign of aging populations (also refer to Fig. 2b).

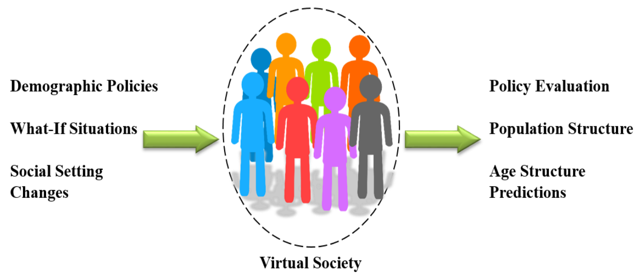

FIGURE 3. Conceptual diagram of virtual society.

Large elderly populations entail a large need for elderly care. This care is both social and economic in nature. To understand the context of this problem, the statistics of the elderly health in Korea can be looked at. Korean life expectancy at birth is 82.1 years, and about 22% of elderly persons in the age group of 80-84 are suffering from dementia, and that number is 33% for the above 85 years age group [18]. To take care of the elderly, individuals, families and societies on the whole, are required to invest social and economic resources (time and money). Countries like Korea, Japan already spend large amounts of money for the care of elderly populations. To overcome these problems in the future, governments have tried to put into practice the various policies for increasing the fertility rates, but these policies have not had adequate outcomes. As these policies have to be implemented at the national level, it becomes an expensive and time consuming process. With simulation and modeling, researchers and scientists design virtual societies and try to analyze the factors leading to behaviors such as lower fertility and delayed marriages. Based on this analysis, the right kind of policies can be designed and implemented. Our model hopes to serve this purpose. Fig. 3 is the very concise conceptual diagram of experimenting with virtual society.

The proposed model is developed to utilize the explanatory nature of the ABMs and in addition, also predict the population structures for future years [19]. In this analysis and prediction, we expect to determine the factors that are associated with the continuously ageing societies and declining birth rates. The decision process of the agents is driven by individual rule-sets rather than by randomness or transition probabilities. The agents' interactions with each other, and with the environment, induce changes in characteristics of the agents and eventually in the simulation output itself. We hope to discover general insights about all the socio-demographic activities implemented in this model. In the following subsections, we present details of our model which are mostly descriptive in nature. The available data in published research was utilized for various assumptions, and if such data was absent, plausible assumptions were arrived at after discussions with the domain experts.

A. SIMULATION PROCEDURE

Simulation starts with randomly distributed agents in the continuous space. The agents begin looking for suitable partners as simulation proceeds and dynamic processes start updating agents and the environment. The simulation advances as simulation ticks, with a single tick equivalent to one year in simulation, which starts from year 2005.

B. AGENTS IN SIMULATION

Agents in proposed model have numerous characteristic values. While most of these values are taken directly from census data, a few are generated by the simulation itself. Specifically, a spreadsheet is linked with model at simulation initialization, from where the census data is directly loaded. Agents have a unique (system generated) identification, gender, education, house id, age, marital status, number of children, education status, and income at simulation.

In addition, there are other characteristics that are derived by the simulation system. For accomplishing this, we make use of publically available census data from Statistics Korea (KOSTAT) [20] and Korea Statistical Information Service (KOSIS) [21]. These characteristics are income and education related data (explained in detail in later subsections), divorce rate and mortality rate.

C. ECONOMIC DIVISIONS

The virtual society in our simulation is allocated to four economic divisions, as per data on education, with better-educated agents having higher incomes. Available literature such as [21] and data available on KOSIS clearly indicates that higher education entails higher income. In addition to the agents who already have their income assigned from the census data, the income of other agents (who do not have any income yet) is assigned based on their education data. These divisions are as follows: if A_{edu} denotes the current education status of the agent A, and D_i denotes the i^{th} division:

$$\mathrm{A} \in D_1 \quad if\ (A_{edu} \leq 3) \tag{1}$$
$$\mathrm{A} \in D_2 \quad if\ (A_{edu} == 4) \tag{2}$$
$$\mathrm{A} \in D_3 \quad if\ (A_{edu} == 5) \tag{3}$$
$$\mathrm{A} \in D_4 \quad if\ (A_{edu} \geq 6) \tag{4}$$

The variable *edu* can take up eight values in our simulation according to the Korean education system:

$$edu \ \leq \ 3 \Rightarrow upto\ middle\ school \tag{5}$$
$$edu == 4 \Rightarrow high\ school \tag{6}$$
$$edu == 5 \Rightarrow junior\ college \tag{7}$$
$$edu \ \geq \ 6 \Rightarrow university\ and\ higher \tag{8}$$

D. AGENT AGE

Initial age (i.e., when the simulation starts) is an *int* value taken from the census data spreadsheet. In the simulation process, age is continuous in nature, updated continuously according to the simulation time. The simulation time is true to many decimal places, therefore, age becomes continuous in nature. The function *currentAge()* handles this, which returns the present age of the agent in that particular simulation time:

$$currentAge = time() - startTime() \tag{9}$$

E. AGENT EDUCATION

As previously stated, as per Korean education system, agents' education in simulation is divided into eight categories. The education status of the agents is updated at every simulation tick (in other words, on a yearly basis). This is done in accordance with the current age of the agent. The compulsory two year military training for Korean youth after high school is also taken into account and hence, the higher graduation age from college. The function *currEdu()* handles this in the simulation as follows:

$$currentAge == 15 \Rightarrow completed\ middle\ school \tag{10}$$
$$currentAge == 18 \Rightarrow completed\ high\ school \tag{11}$$
$$currentAge == 21 \Rightarrow completed\ junior\ college \tag{12}$$
$$currentAge == 25 \Rightarrow completed\ college \tag{13}$$
$$currentAge == 27 \Rightarrow completed\ masters\ level \tag{14}$$
$$currentAge == 32 \Rightarrow completed\ doctoral\ level \tag{15}$$

Further, the education of an agent in the simulation is bound by another variable called *desiredEducation*. This characteristic variable denotes the maximum level of education that an agent wants to achieve in his/her lifetime. At any time, $currEdu() \leq desiredEducation$. The desired education at the start of the simulation is assigned using education data for year 2005. Furthermore, to give them a real and dynamic nature, the desired education values themselves are also updated on a yearly basis. These updates are also done as per the regression developed from education related census data. Agents who are created (born) during the simulation process derive their desired education value from their parents. If $C_{desiredEducation}$ denotes the desired education of the newborn agent C to agents A and B, then always:

$$C_{desiredEducation} \geq \min\left(A_{desiredEducation}, B_{desiredEducation}\right) \tag{16}$$

This functionality is implemented by the Monte-Carlo method. Thus children's education decisions are shaped heavily by the education status of their parents. Generally, in the real world, children obtain the education that is equivalent to that of their parents. Additionally, we take into consideration the ABM interaction aspect by this setting. In this case, the parent-agents would affect the child agents.

F. AGENT INCOME

The population income data is obtained from KOSIS. We use this publically available income data conditioned and averaged for different education levels. As this data is not available at the micro or individual level, we utilized the available data and distributed the income values *normally* over the agent population in each of the divisions. This implies that agents belonging to separate divisions are assigned income based on the average income of their respective divisions. This is unlike other approaches whereby the income is normally distributed over the entire agent population. We understood that the major decision processes of the agents, for example, having a certain number of children, are governed by subtleties of respective income divisions of the agents. Sometimes it is understood that low income households have less disposable income to spend on children and consequentially could lead to the decision to have fewer children. This is of particular significance in countries like Korea where most of the students want to pursue higher education and expenditure on education itself is quite high [22]. This can imply that families with lower income should often be childless or have the fewest children. However, this assumption does not seem practical if we inspect the actual data. By experimenting with the setting of individual income divisions and incomes, we hope to analyze and explore the beliefs and assumptions beliefs like above.

Further, the individual incomes of all the agents, as well as the average income of the division, the value around which the agents' incomes were *normally* distributed, are updated at the start of each year in the simulation. In order to simplify and generalize the calculation of this value, we prepared a logarithmic model of progression of the agent income, based on KOSIS data. The simulation current *year* is the input to this model function.

Specifically, if $avgIncQuin_{year,D_1}$ is the average income of the division D_1 in *year*, then:

$$avgIncQuin_{year,D_1} = \alpha * \log(year + \beta) + \gamma \qquad (17)$$

Where α, β, and γ are the regression constants derived from the available data. We preferred logarithmic function over other functions as during the lengthier simulation runs, the value of *year* gets large and can severely affect the values obtained from this model function. The logarithmic functions tend to flatten over the large values and do not cause large variation in the output of this function.

The agents' incomes that are already assigned, are updated as per their respective updated division average incomes, and the *distance* of their incomes from that value. While doing that, we induce some randomness in this process so that these distances of the agents' incomes from the average values do not remain constant. This percentage of randomness lies between -10% and 10%.

The agents who do not have income yet, due to young age, or are created (born) during the simulation, have their incomes assigned dynamically during the simulation. The income division of the child agent, until he/she starts earning, is kept the same as the division of the father agent. The income of an agent is kept zero if the current education of the agent is less than the desired education. The Ph.D. and master's degree students are assigned some stipend before completing their education, which is equivalent to the often provided scholarship support. This stipend value too is taken from the Korean data.

G. AGENT MARRIAGE

When the simulation initializes, the agents begin looking for their life partners. They start communicating with random agents and these agents could be their potential spouses. For these communications to take place and lead to success in finding a spouse, agents need various conditions met. We do not assume the relationships before the marriage in our simulation model. The potential agent spouses must meet the following criteria:

- The agents' current age should be between fifteen and forty-five.
- One of the agents should have reached the desired education level (i.e. have completed studies).
- The agents should be of opposite sexes.
- The difference of the current ages of the agents should not be more than five years.
- Their house IDs should not be the same.
- They should not be married (present marital status must not be married).
- Their desired education difference should not be more than two levels.

The significance of these settings is manifold, as explained next.

The age limit set is the accepted biological human age for fertility related activities. In Korean culture, young men and women generally tend to complete their studies and secure a job before opting for marriage and having children. Same-sex marriages are rare in Korea [21]. Generally, when people get married, they prefer their spouses to be roughly of their age, hence the five-year age difference in the simulation. This is in line with [23]. The same house IDs criteria cover the case that agents who are related to each other cannot get married. Only agents who are not yet married, are widowed or widowed can get married. The two-level education difference reflects Korean societal behavior. Generally, people prefer spouses who roughly have the same level of education as themselves [23]. This setting suggests that a Ph.D. graduate would want to marry someone who is at least has bachelor's degree. By exposing agents to these conditions, we are trying to emulate a real-world setting in the simulation. If these conditions are satisfied for two communicating agents, they get married and the model logs some new information and updates the variables concerned: marital status of both the agents, first marriage time in simulation, spouse information, and a link is set up between the agents in the model which can be used by the agent couple for retrieving each other's characteristic properties. The location of the female agent is also changed to be that of the male agent.

H. AGENT'S FERTILITY

After the agents get married, they try to proceed toward family formation. In the simulation, there are two variables, *currentChildren* and *desiredChildren*, such that *currentChildren* always $\leq$ *desiredChildren*. The variable *desiredChildren* is the upper bound for having a certain number of children; this value is taken from [24]. After marriage, the agents' wish to have children is restricted by various factors. This can result in agents having fewer children than they desire.

For an agent couple to have children, the total income of husband and wife agents minus the expenditure on children is calculated and the net should be equal to or greater than the average income of the economic division they belong to. If in the *year*, an agent A in division D_i, whose income is denoted by A_{income}, whose spouse B's income is denoted by B_{income}, whose *currentChildren* is their current number of children, and whose $C_{expenditure}$ denotes expenditure on children, then in order to have a child:

$$A_{income} + B_{income} - C_{expenditure} \geq \delta * avgIncQuin_{year,D_1} \tag{18}$$

Where δ is a constant derived during the calibration process, $C_{expenditure}$ values are calculated from the children expenditure data of different income groups [25], [26], and is approximated as a concave function of total income and number of children *currentChildren*. As the *currentChildren* is zero initially, for the first child, (18) is the simple addition of both the agents' incomes.

In addition to income check, agents wait for a random time for having first (up to two years) and subsequent children (from one to four years). This is to mimic the age difference between the siblings in the real world. We always check that the age of the parent agents was between fifteen and forty-five years before they had children.

Thus, everything from age, education, personal income, division income, to expenditure on children is connected and affects various decision processes of the agents.

I. AGENT MORTALITY AND DIVORCE

The agent mortality and divorce are handled stochastically. To avoid unbeneficial complexity in light of the current work, deterministic models are not implemented for these processes. Thus, agents are exposed to the transition probabilities derived from the census data available at KOSIS at the start of each year. Widowed and/or divorced agents can remarry, if they satisfy all the marriage requirements as a normal agent, along with a probability check to incorporate the re-marriage rate. The divorce event occurs only in the married state, whereas the mortality event can affect agents in any state.

J. THE AGENT INTERACTIONS AND COMMUNICATION

The agents in the simulation live in the continuous space. The agent's housemates and agents living in up to ten model units, can be interacted with and characteristic information can be shared. Through a definition of thresholds and Monte-Carlo method, the characteristics of the interacting agents can be updated. Currently, we have implemented this functionality

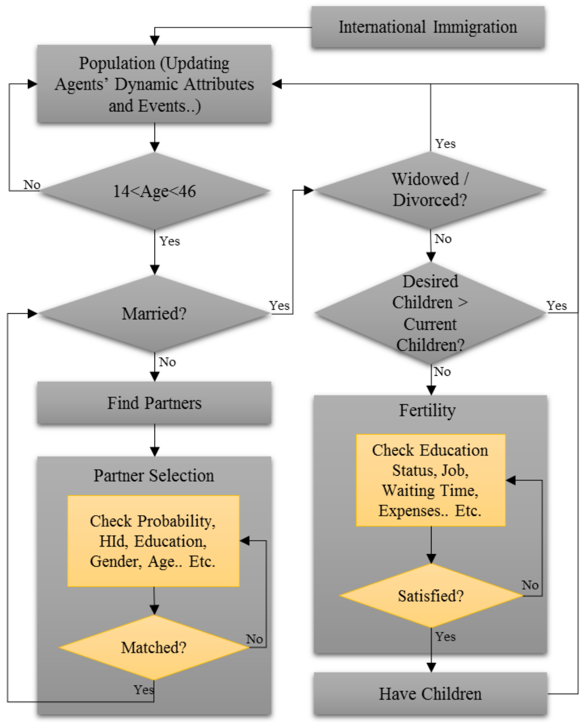

FIGURE 4. Simplified model flow.

for desired education and desired children factors. Only the agents that participate in the family formation activities update their *desiredChildren* factor. We have designed some experiments by taking this key feature into consideration.

The agent communication has been implemented such that it only takes place among the agents of the same income division, the general case in Korean society. The inspiration for this decision was the real world behavior, where people mostly communicate with their friends, families, and neighbors, which generally belong to same income group, and to some extent, are affected by their behavior. The neighbors of agents in the simulation can be updated at every simulation tick due to marriage, fertility and mortality events. The dynamic characteristics of other agents are ever changing as well. This signifies that the agent communication keeps on inducing subtle changes in agents and gradually this effect leads up to an emergent phenomenon in the simulation, which is evident in the results we demonstrate in Section V. Fig. 4 demonstrates the simplified model flow of the proposed model. The international immigration mentioned in this figure is utilized for experimental setup and is explained in the next section.

IV. IMPLEMENTATION

The proposed model is AnyLogic professional v. 7.1.2. The fertility event adds the agents to the simulation whereas the mortality event removes the agents from the simulation. After the simulation is initialized, agents starts interacting and looking for the potential spouses Due to diverse characteristics and the very flexible nature of the model, there is much potential for analyzing and exploring the demographic activities and behavior.

A. DATA USED IN SIMULATION

The census data used for the simulation is provided by KOSIS. We start the simulation with a data set of sampled 10,000 agents. The average of ten simulation runs is used for final analysis. Simulation starts from the year 2005 and is run up to the year 2050. The simulated data is collected and compared to the actual data, which serves as validation of the model.

B. VALIDATION, CALIBRATION AND EXPERIMENTAL ANALYSIS

Before any analysis with a model is pursued, it should undergo validation process. It is worth noting that as the decision process in our proposed model are not stochastic but deterministic, probabilities that are calculated directly from census data cannot be utilized in our model. The marriage rate probability for instance, can simply be computed by deriving it from census data, and then by removing matrimony requisites in the model, this probability value can handle the marriage process in simulation. This approach however, would attract some problems: the probability values calculated would represent the general marriage rates for the entire population (unless high quality data is available, and probabilities are derived from various instances of demographic behavior), and not the rates decided by the various behavioral-requirements put in place by a modeler in an ABM. Also, by direct usage of these rates, the end-product in model terms would only be a *quasi-ABM* or a *pseudo-ABM*, with limited usefulness in terms of obtaining newer insights, and understanding the functioning of the implemented system. Such a model becomes MiSim in behavior. As outlined in [27], such models might *implement* the desired functionality rather than *arriving* at or *deriving* desired functionality.

Fig. 5 presents the validation results of proposed model. The government of Korea provides statistical data for the population projections, available on KOSIS. This data provides high, middle, and lower statistical limits of these projections, which are employed for validation of our model. Fig. 5a presents the total population data. The population numbers in simulation are scaled to the actual population numbers. The model data lies well between the upper and lower limits of the statistical data. The close fit of statistical data with model data demonstrates the correct calibration and serves the purpose of validating our model.

To add further credence to the model validation, we also compared the young (0 to14 years), working (15 to 65 years), and elderly (above 65 years) population data independently. These results are presented in Fig. 5b, Fig. 5c, and Fig. 5d respectively. While the use of publicly available government provided data is straightforward, representing social norms in simulation is complicated and should only be mapped after discussions and agreement with appropriate domain experts [27]. The social norms in this model were derived after discussions with social scientists, and a common viewpoint in this regard was decided. While these standards cannot be called accurate, the concerned domain experts agreed them to be at a practical level of abstraction. As stated in [27],

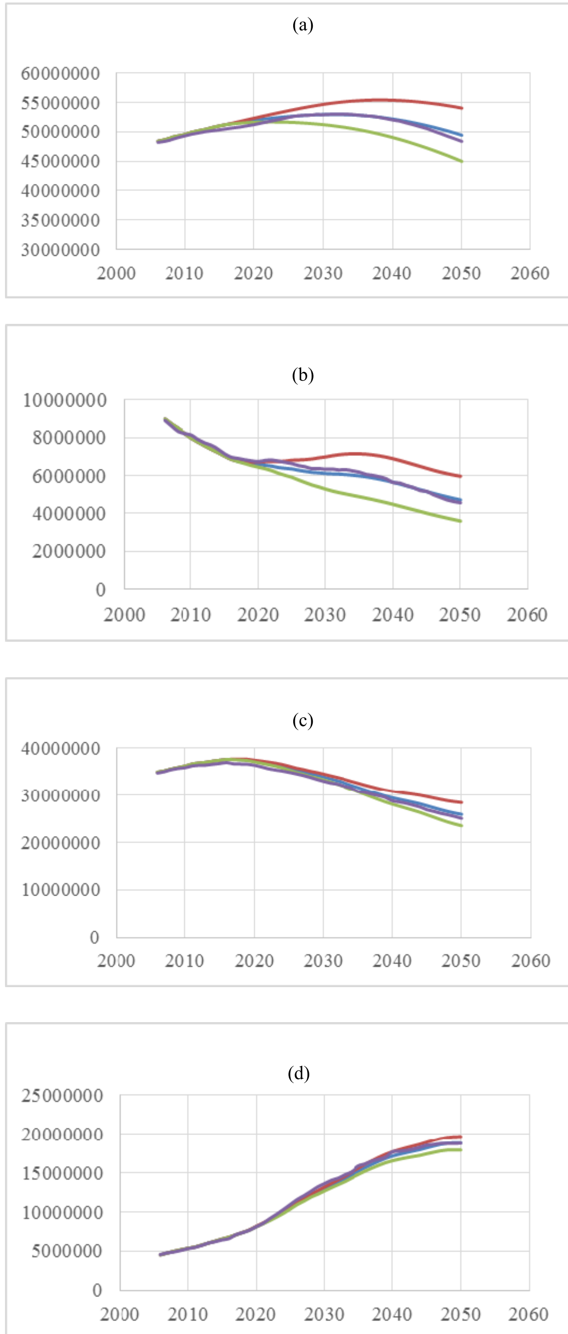

FIGURE 5. Validation Graphs. a) Total Population, b) Young Population, c) Working Population, d) Aged Population vs. respective statistical data from KOSIS. x-axis are years and y-axis are population numbers. The simulation data lies between the upper and lower bounds provided by the government, which demonstrates sound validation of the proposed model. Red, blue, green colors represent Statistical high, Medium, Low respectively and Purple color represents simulated data.

the mere comparison of simulation data of a complex system cannot be justified for its validation. The constituent components' functioning should also be validated so that they have desired functionalities comparable to the target system. Working on these lines, we presented the validation data not only for the total population, but also for the young, working and elderly populations, explained and justified the functionalities of underlying components of our model, and collectively, these points serve as the arguments for validation and fitness of purpose for our model.

C. EXPERIMENTAL ANALYSIS

To pursue analysis of the effect of various demographic activities, we implemented some virtual experimental scenarios. Factors such as immigration, desired children, magnitude of interaction, intra-quintile income comparison, and overall income. These values of these factors were changed and resultant effect was observed in the population numbers.

An increasing number of people have changed their views about children as well as marriage, and a rising percentage of people state that having children is not necessary for a happy marriage [29]. Thus, we figured that experimenting with a change in the desired number of children could give us some interesting insights. The experiment changes the distribution of desired children for the whole agent population in the simulation.

Korea has witnessed an increase in international immigration in recent times [30]. The census data provided by the government already includes the statistics for this. So international immigration was not handled independently in the simulation. Moreover, most of the immigration, whether skilled or non-skilled, is temporary in nature. Naturalization and permanent residency in Korea is quite low due to stringent Korean naturalization laws. The number of people seeking refuge in Korea is also quite low. However, in recent years, we have witnessed a refugee crisis unfolding in the European countries due to the exodus from war-torn countries in the Middle East. Germany, for example, received over a million refugees within a short span of time. Such an event can have many long-lasting implications for the host country, and altered population dynamics are one such implication. In order to simulate such an unforeseen and rapid influx of immigration to Korea, and to analyze its effect on population structure and numbers in the long run, we designed such a scenario in our model. Referring to data from the UN Refugee Agency and [31] and Asylum Information Database [32], we simulated an influx of one million immigrants in two years. The gender distribution, education, employment participation, and age distribution were taken from [31] and [32]. To avoid any unbeneficial complexities in the simulation, we assumed that immigrants would only pursue the demographic activities with fellow immigrants and not with the natives of the host country. We also reduced restrictions particularly those related to education, for finding spouses and having children. This assumption was made as data available reflects a crisis in refugee education. To take into account the real-world scenario in such cases, we introduced restrictions related to employment: whether a migrant is employed or not. Specifically, in the potential partner selection, at least a male should not be unemployed. Referring to [31] and [32], we distributed refugee migrants' preferences regarding the desired number of children. It is set to 50%, 30% and 20% for two, three, and four children respectively. If we look at the fertility rates of different countries, it is evident that under-developed and developing countries have high fertility rates [3]. The offspring of migrants have their number of desired children fixed as two children. This assumption is based on the fact that gradually, the migrants would change their outlook and make adjustments that are salient to the new country.

Further, we also decided to experiment with the effect of the magnitude of interactions among the agents in the simulation. Korea is a culturally homogeneous and connected country, so the impact of social trends can have dramatic outcomes and can become country-level phenomena rather quickly. The recent craze about cryptocurrencies is one such example. We have already explained in the previous sections how the agents would interact with and affect each other. Experimenting with interaction can give us interesting insights into the simulation and demographic behavior of the virtual society.

In this paper, we are proposing and experimenting with a rather new societal standard of economic-comparison within the same income division. We term it the intra-quintile comparison. As previously mentioned, agents' interactions are limited to only the agents in the same income division. Moving along these lines, we experimented with the comparison of agents' economic standards within the same income division. We stated in (18) that people's disposable household income must be greater than the average income of the quintile they are living in. This condition can be treated as a measure of *desired economic standard* within the income quintile. If a household is willing to drop this value, it would virtually pay less heed to others' *economic standards* and thus afford more expenditure. This would also mean that it would have more to spend on the children whilst satisfying the condition in (18). On the other hand, if a household decides to increase the value of this cut-off, the situation would exactly be the opposite of that explained above. The household would have less money to spend on children to clear the condition in (18). This whole scenario mimics the economic standard that a household wants to maintain. The more the value of this *standard*, the less would be the disposable income to spend. This scenario could have a direct impact on the child-bearing decisions of a family. Exploring the effect of this change can give us a new understanding about the economic-dynamics of the population.

For the purpose of experimentation, we also made some changes to the income of the agent population. If we have a look at the Korean GDP per capita in recent years, there has been a steady rise, but not without some abrupt changes.

TABLE 1. Experimental test case scenarios.

Factor		Cases	Implication
1. **Desired Children** (one, two, three, four and above children respectively) (%)		a. 5, 80, 10, 5	Change in attitude of people to have certain number of children.
		b. 0, 90, 5, 5	
		c. 0, 100, 0, 0	
2. **Immigration**		1 million immigrants in 2 years	Effect of sudden refugee migrants' influx.
3. **Interaction Effect**		a. Decrease to zero.	Effects of agents' interaction with other agents.
		b. Increase to 10%	
		c. Increase to 20%	
		d. Increase to 30%	
4. **Intra-Quintile Comparison**	**Increase**	a. By 10%	Effect of comparison of household's disposable income with the average income of respective quintiles.
		b. By 20%	
		c. By 30%	
	Decrease	d. By 10%	
		e. By 20%	
		f. By 30%	
5. **Income**	**Increase**	a. By 5-10%	Effect of increase and decrease in agents' income.
		b. By 15-20%	
		c. By 25-30%	
	Decrease	d. By 5-10%	
		e. By 15-20%	
		f. By 25-30%	

During the infamous 1997 Asian financial crisis, GDP per capita plummeted by about 33% [21]. Conversely, it also rose by about 20% in 2010 after the rebound from the global financial crisis. In the long run, these abrupt shocks to the economy could have an impact on the household economy and therefore, on the population dynamics as well.

TABLE 2. Population numbers of selected cases.

	Total Population	Young Population	Working Population	Aged Population
Original Simulated	48385651	4565521	25054945	18765186
Case 1a	49617388	5145728	25663103	18808557
Case 2	49641483	4695634	25628888	19316962
Case 5f	47457030	4392519	24223667	18840844

Taking a cue from this data, we reckoned it might be a good idea to experiment with this setting. This can enable us to assess the effect of sudden economic changes on population dynamics.

D. DESIGN OF EXPERIMENT

The details of experimental setup are summarized in Table 1. There are twenty test cases in total, which are run for ten runs each to average the readings. Each run takes about three minutes on average. Therefore, it took about 600 minutes (20*10*3) to generate the initial data for analysis.

The desired children are distributed in the simulation according to a survey that puts the number of desired children at 15% for one child, 70% for two children, 10% for three children, and 5% for four and above (all figures approximate). We made changes to this distribution by increasing percentage of two desired children. It is clear from this survey that the majority of the people want to have two children. If societal conditions are changed to favor and support child-bearing, more people could shift toward having more children. As two children is the most popular choice, we increased this value and observed the effect of this increase.

The agent interaction process is already explained in previous sections. The probability of this effect is about 3% which was derived during the calibration process. We changed the magnitude of this to 0%, 10%, 20% and 30%. A stronger effect signifies that more people are being influenced by other agents' behavior.

The quintile comparison was varied by 10%, 20%, and 30%, both increases as well as decreases. The income variation was also varied for both increase and decrease, simulating a sudden economic downturn and sudden economic jump. These values for these variations were decided based on historic GDP per capita fluctuations. The different values and nature of these factors would help us to understand their effect on the simulation behavior.

V. EXPERIMENTAL RESULTS AND DISCUSSION

In this section, we discuss the results obtained from the experimental setup. The objective of this experimental setup is to evaluate total population and its age structure. The experimental results are compared with the simulation numbers which are obtained in the original runs of the calibrated model. The starting population is sampled to be comparable to the actual population distribution. In case of immigration experiment, simulation is run until year 2075 to observe the long term impact of immigration, particularly the effect of the newer generation that is born in the host country.

We can roughly visualize how the change in initial conditions is affecting the population numbers. Some scenarios are yielding a big increase in the numbers, others yield subtle increases and some scenarios yield decreased numbers. The resultant numbers were subjected to *t-test* for evaluating whether the changes were statistically different or not (significant or insignificant). *t-test* was performed on all the test cases with a significance level of 0.05 and nine degrees of freedom. The *t-value* for such a setting is 2.262 for a two tailed test. Thus, if *t-values* are greater than 2.262 or less than −2.262, the population numbers are statistically different. The summary of this test for all the cases is presented in Table 3.

The summary of experimental simulation runs is shown in Fig. 6, Fig.7, and Fig. 8, which present the comparison of total population numbers for all the experimental setups throughout the simulation run. The graphs are labelled as case numbers which can be referred to in Table 1 or 3. In Fig. 9, apart from population numbers, we also present the population age structures for cases with the highest population numbers, which we term *best cases*, and lowest population numbers, termed *worst cases*. Two best cases are chosen for presentation, as the case of rapid influx of refugee migrants directly contributes to the population numbers and alters population age structures as well. Surprisingly, population numbers in the other best case are not very different from the immigration case. The population numbers for the young, working, and elderly population are presented in graphs in Fig. 9c to 9f.

If we take a look at outputs of cases 1a and 1b in Fig. 6a and 6b, and refer to *t-test* values in Table 3, we see a substantial increase in the population numbers. We had increased the values for two desired children by 10% and 20%. Simulation reflects that convincing people to change their attitudes and mindsets about child-bearing could result in constructive push to fertility rate, as results indicate that people can opt for having more children without facing any major economic challenges. Thus, encouraging people by means such as maternity and paternity leave, better facilities for child bearing, and change in job culture could help in giving a positive boost to fertility. Third case 1c did not yield any significant changes in the population numbers: rather, it resulted in a non-significant decline. The result is presented in Fig. 6c. In this case, agents' interaction effect on each other is nullified as all

TABLE 3. t-test values of different cases. The bold and italicized values are statistically same.

Factor		Cases	t-test value
1. Desired Children (one, two, three, four and above children respectively) (%)		5, 80, 10, 5	14.11969382
		0, 90, 5, 5	14.52406248
		0, 100, 0, 0	***-1.346338056***
2. Immigration		1 million immigrants in 2 years	7.356416865
3. Interaction Effect		Decrease to zero.	***0.099909235***
		Increase to 10%	-3.629549856
		Increase to 20%	-3.068057807
		Increase to 30%	-7.786693972
4. Intra-Quintile Comparison	**Increase**	By 10%	***-0.344529605***
		By 20%	-2.675989257
		By 30%	-6.532789026
	Decrease	By 10%	***2.203289404***
		By 20%	5.175533991
		By 30%	3.519867455
5. Income	**Increase**	By 5-10%	***1.912673512***
		By 15-20%	***0.879760702***
		By 25-30%	***0.313499713***
	Decrease	By 5-10%	***-1.837035896***
		By 15-20%	-8.174428147
		By 25-30%	-10.91703228

agents want to have two children. This change also reflects the effect of interactions among the agents. In cases 1a and 1b, there are 15% and 10% agents respectively who want to have more than three children. Their effect thus increased the values of desired children for some agents with lower desired number of children, which in turn must have influenced other

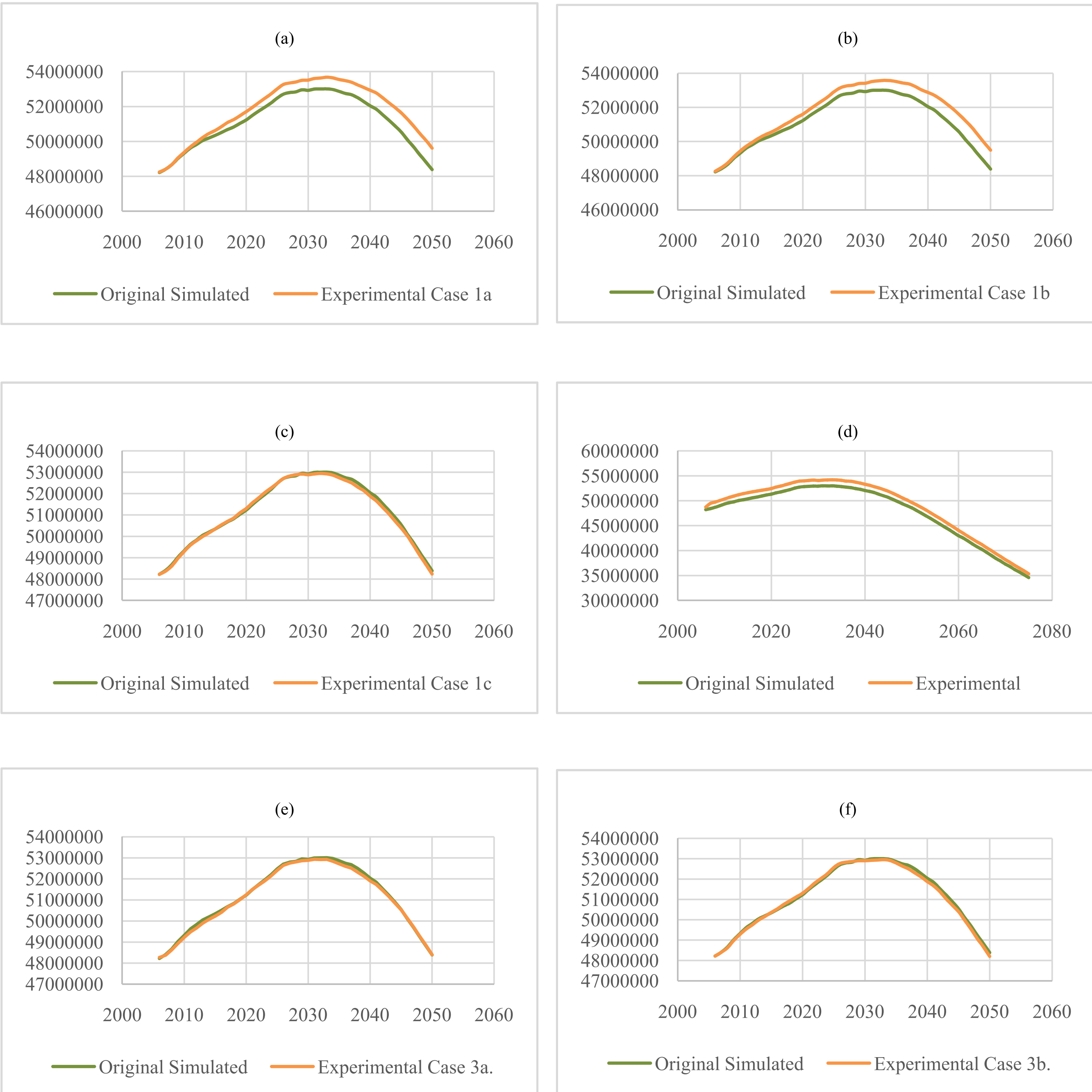

FIGURE 6. **Population summary of different test case scenarios. The graphs display the resultant total population numbers vs. the original simulated population numbers. Simulation is run from year 2005 to year 2050. For immigration case, it is run till year 2075. The experimental case numbers and corresponding test scenarios can be referred to in Table 1 or Table 3. There is a large, observable difference in population numbers in case 1a and 1b which involves changing preferences of couples for having certain number of children. x-axis are years and y-axis are population numbers. a) Desired children experiment case 1a. b) Desired children experiment case 1b. c) Desired children experiment case 1c. d) Immigration experiment. e) Interaction experiment case 3a. f) Interaction experiment case 3b.**

agents to do the same. This added effect resulted in emergence of agent behavior in the simulation and produced a significant increase in population over the years. It must be noted that agents with a higher number of desired children also are affected by agents with a lower number of desired children. These test cases assert the fact that interactions can, and do play an important role in the decision process of agents. The x-axis in all the graphs in Fig. 6, Fig. 7, Fig. 8, and Fig. 9 is years for which the simulation is run and the y-axis is the population numbers.

The immigration test case resulted in an obvious increase in population numbers due to direct contribution in

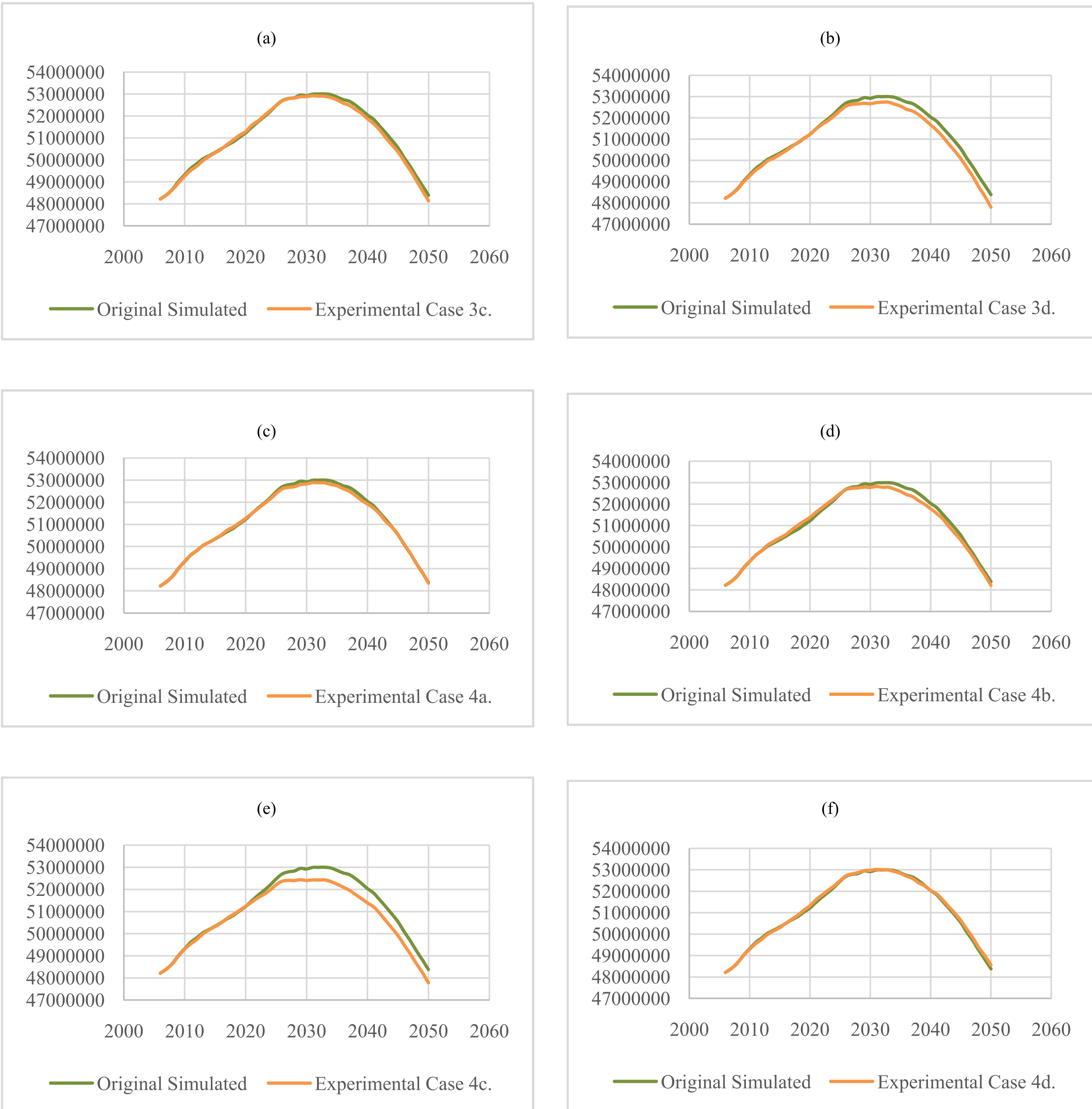

FIGURE 7. Population summary of different test case scenarios. The graphs display the resultant total population numbers vs. the original simulated population numbers. Simulation is run from year 2005 to year 2050. The experimental case numbers and corresponding test scenarios can be referred to in Table 1 or Table 3. There is a significant difference between the population numbers in case 4c and 3d. Case 4c involves experimentation with socio-economic preferences and case 3d covers magnitude of interaction-effect among the agents. x-axis are years and y-axis are population numbers. a) Interaction experiment case 3c. b) Interaction experiment case 3d. c) Intra quintile comparison experiment 4a. d) Intra quintile comparison experiment 4b. e) Intra quintile comparison experiment 4c. f) Intra quintile comparison experiment 4d.

population numbers. It is worth noting that we simulated the refugee influx for only two years and one million immigrants. Aside from the short-term increase in the population numbers, it is important to observe that in the long run, the population numbers begin to converge with the other *best case*. Fig. 6d and Fig. 9c can be referred to. In Fig. 9c, the population numbers of case 1a and case 2 seem to be almost similar. It can be concluded from this experiment that the short-term high-rate immigration would not produce any long term benefits in fertility rate and population numbers. Moreover, there are many social and economic implications of a high rate of immigration, especially in a country like Korea that is culturally homogeneous. Furthermore, in order to rely on immigration for a long-term boost in the fertility

FIGURE 8. Population summary of different test case scenarios. The graphs display the resultant total population numbers vs. the original simulated population numbers. Simulation is run from year 2005 to year 2050. The experimental case numbers and corresponding test scenarios can be referred to in Table 1 or Table 3. The test cases involving increase in agents' income have quite similar outputs to original simulation numbers. x-axis are years and y-axis are population numbers. a) Intra quintile comparison experiment 4e. b) Intra quintile comparison experiment 4f. c) Income increase experiment case 5a. d) Income increase experiment case 5b. e) Income increase experiment case 5c. f) Income decrease experiment case 5d.

rate, and better age distribution in the population, it should be sustained over a longer period of time.

In the experiment related to changes in the magnitude of the effect of interactions among the agents, some interesting results were obtained, presented in graphs in Fig. 6e, Fig. 6f, Fig. 7a, and Fig. 7b. While changing the interaction effect to zero resulted in a non-significant increase in the numbers, all other cases resulted in significant decrease in the numbers. This could be understood by paying attention to the initial distribution of the desired number of children in the virtual society. As a majority of the agents desire two children, and as the intensity of interaction grows, they overwhelm the other agents' decision with their preferences. As the second-largest majority of agents have three or more desired children, they must be shifting to two as their desired number. On the other hand, if we consider the discussion in the previous paragraph,

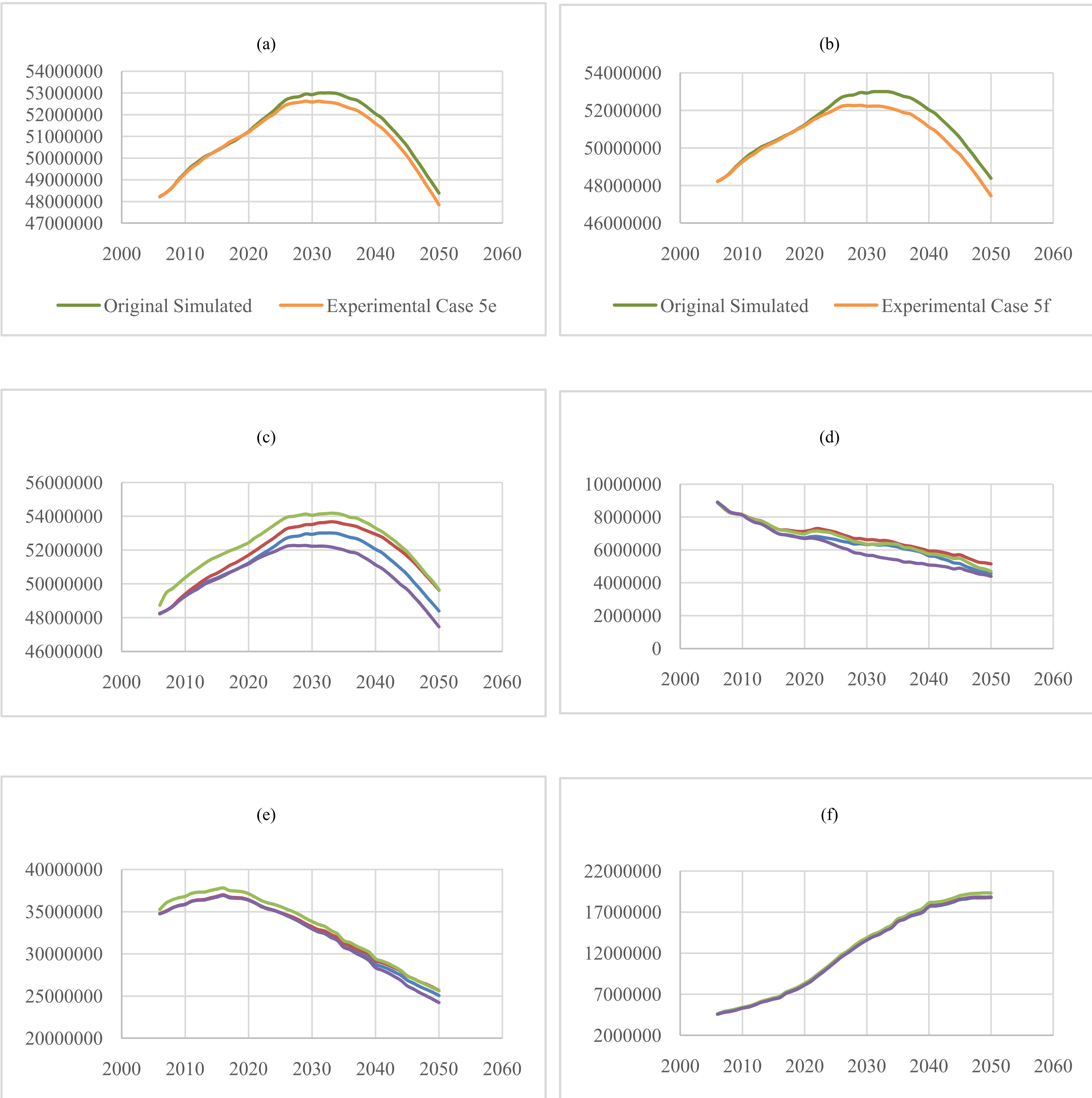

FIGURE 9. **Population summary of different test case scenarios. The graphs a) and b) display the resultant total population numbers vs. the original simulated population numbers. Simulation is run from year 2005 to year 2050. The graphs c) to f) demonstrate the original simulation population numbers vs. best cases (largest population numbers). Two best case shown here are case 1a (change in desired children) and case 2 (immigration). We also demonstrate comparison of young, working and aged population independently. The experimental case numbers and corresponding test scenarios can be referred to in Table 1 or Table 3. x-axis are years and y-axis are population numbers. For figures 8c to 8f, blue color represents original simulation, red color represents case 1a (desired children experiment), green color represents case 2 (immigration experiment) and purple color represents case 5f (income decrease by 25 to 30%).**

the population numbers would increase if somehow the effect of agents with more desired children toward those with fewer desired children is intensified. This understanding can aid government design and disseminate appropriate policies to encourage people to have more children.

The experimentation with intra-quintile comparison gave dramatic results. Out of the six experimental setups, four gave statistically significant different results. The fifth case also barely missed being statistically significant. The results show that people have a good deal of sensitivity toward their *economic-standard*, which we can attribute to societal and peer pressure. In both the scenarios, in other words, while i*ncreasing* or *decreasing* their self-realized *standard*, the results show a significant sensitivity. While dealing with this non-tangible factor *standard* is tricky in the real world, the other aspects that decide this *standard* can be worked upon. One of the major factors in deciding household disposable income is education expenditure of children. Korea spends large amounts of money in the education field, which is about 8% of national GDP, compared to the OECD average of 6%. The private expenditure on education is also highest, particularly at the tertiary level [22]. Various other studies show that expenditure on education and childcare is very high in Korea. One such study shows that 90% of parents feel burdened by childcare-related costs [29]. With that said and understood, such policies and education culture should be designed by government, which do not burden the overwhelming majority of children and parents. Such policies could help relieve intra-quintile societal-comparison pressure. The results of this experiment are presented in Fig. 7c to 7f, 8a, and 8b.

Lastly, we discuss the experiments run with changes in the income of the agents. Quite interestingly, even a 25-30% (case 5a, 5b and 5c) increase in income of agents did not result in significant increases in the population numbers. Although the increased income scenario did result in increased numbers, *t-test* data indicates that the increase was only non-significant. On the other hand, a decrease in the income did result in a significant decrease in cases 6e and 6f. The results are presented in Fig. 8c to 8f, 9a, and 9b. The outcome of this result set demonstrates that in many households, fertility related decisions are significantly affected by financial conditions. Further, the results from *increase* cases vs. *decrease* cases suggest that more people are *just above* the *desired economic situation* than *just below*. In fact, the lowest population numbers were obtained in case 5f, in which agents' incomes were decreased by 25-30%. Simulation results demonstrate that a decrease in income forced many households to delay child-bearing and they ultimately ended up having fewer children. This is an interesting insight into the socio-economic status of the population. Again, as we have already stated, one of the main factors deciding household disposable income is education and childcare expenditure. High education expenditure in Korea is one of the reasons why parents want to have fewer children [22], [29]. Through such understanding of the socio-economic situation of the population, the government can design and implement policies that benefit the population in general and *vulnerable* section of people (in the fertility context) in particular.

In Fig. 9c to 9f, we present a comparison of original simulation numbers vs. *best cases* and *worst case*. As already explained above, the experimental setups that give highest population numbers are termed *best cases* and vice-versa. If we examine Fig. 9d, we can see that the worst case also has the lowest youth population. The lowest youth population directly implies fewer people would grow up, and be available as the workforce. We can observe this in Fig. 9e in which working population numbers decrease substantially compared to the other cases. Thus, the primary reason behind the *worst performance* of this case is lower numbers for the young population, which is obviously due to the lower number of births. On the other hand, a *good* feature of best case 1a is how this case has the highest numbers for the young and working population among all the four cases presented in Fig. 9c to 9f, as well as the lowest elderly population among the three experimental cases. A higher youth and working population and a lower elderly population is an indicator of a good population age structure. Table 2 is the brief summary of population numbers of these cases when simulation stops.

VI. CONCLUSIONS AND FUTURE DIRECTIONS

In this paper, we proposed and implemented an agent based model for social demography that encompasses major activities related to socio-demographic dynamics of the population. We demonstrated validation, fitness of purpose, and an example analysis of the future population using our model. Our analysis threw light on some interesting insights and demonstrated the kinds of plans that can be designed for a positive push to fertility rates. The proposed model is very flexible and general in nature and can be reused by populating it with target-setting data. The type of analysis that can be pursued and addition of newer, complex details is virtually limitless, and the model can be expanded to accommodate the research questions being addressed. Important features of our model include the interactions among the agents, as well as the dynamic nature of the agents. Every characteristic of an agent in our model is dynamic in nature and is updated in accordance with real data. This is generally not the case in other such models.

This work is a step forward in answering and analyzing some key questions in social dynamics in general, and demography in particular. It motivates the modelers and researchers to pursue and analyze the social dynamics with techniques like agent-based modeling. Given the problem of ageing societies and declining birthrates, the understanding of social demography has come to be very important. Future generations should not be over-burdened with avoidable complications and providing a problem-free future for the next generation is the responsibility of the present generation.

We plan to continue our analysis with the model and introduce machine-learning (ML) and optimization techniques such as [33]–[36] for optimizing and automating the model for general usage. As per our understanding, ML and optimization techniques can benefit the simulation process by including them as a part of existing model. One of the better features of the agent-based modeling is its lower dependency on real data, which is generally contrary to ML techniques. New concepts like [37]–[40] can greatly benefit the working of our model. Incorporating such techniques can enable our model to handle bigger data sets more efficiently as we are currently using 10,000 agents in our simulation. These numbers can be improved and model can deliver even better results. In addition, the current work does not include land use and healthcare models, and we are working to add these features in the future extensions of our model. The inclusion of these models is needed to generalize the model and put it to practical use in these areas. Our next steps also include testing the government policies and eventually, help to design and streamline such policies.

REFERENCES

[1] D. Parisi, "Artificial life and human societies," in *Knowledge Management, Organizational Intelligence and Learning, and Complexity*, vol. 1. Oxford, U.K.: Eolss, 2009, pp. 270–287.
[2] N. Gilbert and K. G. Troitzsch, *Simulation for the Social Scientist*. Berkshire, U.K.: McGraw-Hill, 2005.
[3] (2018). *The World Factbook—Central Intelligence Agency*. [Online]. Available: https://www.cia.gov/library/publications/the-world-factbook/rankorder/2127rank.html
[4] United Nations DESA/Population Division. (2017). *World Population Prospects 2017*. [Online]. Available: https://esa.un.org/unpd/wpp/
[5] F. C. Billari, F. Ongaro, and A. Prskawetz, "Introduction: Agent-based computational demography," in *Agent-Based Computational Demography*. Heidelberg, Germany: Physica-Verlag, 2003, pp. 1–17.
[6] F. Heiland, "The collapse of the Berlin wall: Simulating state-level East to West German migration patterns," in *Agent-Based Computational Demography*. Heidelberg, Germany: Physica-Verlag, 2003, pp. 73–96.
[7] F. C. Billari, A. Prskawetz, B. A. Diaz, and T. Fent, "The 'wedding-ring': An agent-based marriage model based on social interaction," *Demogr. Res.*, vol. 17, no. 3, pp. 59–82, 2008.
[8] I. Benenson, I. Orner, and E. Hatna, "Agent-based modeling of householders' migration behavior and its consequences," in *Agent-Based Computational Demography*. Heidelberg, Germany: Physica-Verlag, 2003, pp. 97–115.
[9] N. Geard, J. M. McCaw, A. Dorin, K. B. Korb, and J. McVernon, "Synthetic population dynamics: A model of household demography," *J. Artif. Soc. Soc. Simul.*, vol. 16, no. 1, p. 8, 2013.
[10] E. Silverman, J. Bijak, J. Hilton, V. D. Cao, and J. Noble, "When demography met social simulation: A tale of two modelling approaches," *J. Artif. Soc. Soc. Simul.*, vol. 16, no. 4, p. 9, 2013.
[11] S. Hassan, J. Pavón, L. Antunes, and N. Gilbert, "Injecting data into agent-based simulation," in *Simulating Interacting Agents and Social Phenomena*. Heidelberg, Germany: Physica-Verlag, 2010, pp. 177–191.
[12] J. W. Bae, E. Paik, K. Kim, K. Singh, and M. Sajjad, "Combining microsimulation and agent-based model for micro-level population dynamics," in *Proc. Int. Conf. Comput. Sci. (ICCS)*, San Diego, CA, USA, vol. 80, Jun. 2016, pp. 507–517.
[13] M. Sajjad, K. Singh, E. Paik, and C.-W. Ahn, "A data-driven approach for agent-based modeling: Simulating the dynamics of family formation," *J. Artif. Soc. Soc. Simul.*, vol. 19, no. 1, p. 9, 2016.
[14] K. Singh, M. Sajjad, and C.-W. Ahn, "Towards full scale population dynamics modelling with an agent based and micro-simulation based framework," in *Proc. 17th Int. Conf. Adv. Commun. Technol. (ICACT)*, 2015, pp. 495–501.
[15] A. Grow and J. Van Bavel, "Assortative mating and the reversal of gender inequality in education in Europe: An agent-based model," *PLoS ONE*, vol. 10, no. 6, p. e0127806, 2015.
[16] S. Klüsener, F. Scalone, and M. Dribe, "Exploring the role of communication in shaping fertility transition patterns in space and time," in *Agent-Based Modelling in Population Studies*. Cham, Switzerland: Springer, 2017, pp. 369–403.
[17] M. Myrskylä, H. Kohler, and F. C. Billari, "Advances in development reverse fertility declines," *Nature*, vol. 460, pp. 741–743, Aug. 2009.
[18] Y. J. Kim, J. W. Han, Y. S. So, J. Y. Seo, K. Y. Kim, and K. W. Kim, "Prevalence and trends of dementia in Korea: A systematic review and meta-analysis," *J. Korean Med. Sci.*, vol. 29, no. 7, pp. 903–912, Jul. 2014.
[19] K. Singh and C.-W. Ahn, "A holistic agent based model for demography," in *Proc. IEEE Symp. Ser. Comput. Intell. (SSCI)*, Dec. 2017, pp. 1–8.
[20] KOSTAT. (2018). *Statistics Korea*. [Online]. Available: http://kostat.go.kr/portal/english/index.action
[21] KOSIS. (2018). *Korean Statistical Information Service*. [Online]. Available: http://kosis.kr/eng/
[22] OECD. (2014). *Education at a Glance 2014: Country Notes*. [Online]. Available: http://www.oecd.org/edu/education-at-a-glance-2014-country-notes.htm
[23] M. McPherson, L. Smith-Lovin, and J. M. Cook, "Birds of a feather: Homophily in social networks," *Annu. Rev. Sociol.*, vol. 27, no. 1, pp. 415–444, Aug. 2001.
[24] "The solution to dissolve the gap between the desired number of children and the number of children born," Hyundai Res. Inst., Seoul, South Korea, Tech. Rep. 10-24, Aug. 2010. [Online]. Available: http://hri.co.kr/upload/publication/201085144343%5B1%5D.pdf
[25] J.-K. Kang, "The high price of raising a well-educated child," Korea Joongang Daily, Seoul, South Korea, Apr. 2013. [Online]. Available: http://koreajoongangdaily.joins.com/news/article/article.aspx?aid=2969998
[26] L. Claire, "Low-income households cut education spending in Korea," Korea Herald, Seoul, South Korea, Sep. 2015. [Online]. Available: http://www.koreaherald.kr/view.php?ud=20150916001169
[27] F. Polack, "Filling gaps in simulation of complex systems: The background and motivation for CoSMoS," *Natural Comput.*, vol. 14, no. 1, pp. 49–62, 2015.
[28] F. A. C. Polack *et al.*, "Reflections on the simulation of complex systems for science," in *Proc. 15th IEEE Int. Conf. Eng. Complex Comput. Syst. (ICECCS)*, Mar. 2010, pp. 276–285.
[29] B.-E. Kim, "Korea's birthrate lowest among OECD members," Korea Times, Seoul, South Korea, Mar. 2017. [Online]. Available: http://www.koreatimes.co.kr/www/nation/2017/03/119_225996.html
[30] H.-J. Ock, "Number of foreigners in Korea exceeds 2 million," Korea Herald, Seoul, South Korea, Jul. 2016. [Online]. Available: http://www.koreaherald.com/view.php?ud=20160727000493
[31] UNHCR. (2017). *The UN Refugee Agency*. [Online]. Available: http://www.unhcr.org/
[32] AIDA. (2017). *Asylum Information Database*. [Online]. Available: http://www.asylumineurope.org/
[33] X. Zhao, L. Zhang, P. Shi, and M. Liu, "Stability of switched positive linear systems with average dwell time switching," *Automatica*, vol. 48, no. 6, pp. 1132–1137, 2012.
[34] X. Zhao, Y. Yin, B. Niu, and X. Zheng, "Stabilization for a class of switched nonlinear systems with novel average dwell time switching by T–S fuzzy modeling," *IEEE Trans. Cybern.*, vol. 46, no. 8, pp. 1952–1957, Aug. 2016.
[35] H. Wang, P. X. Liu, and B. Niu, "Robust fuzzy adaptive tracking control for nonaffine stochastic nonlinear switching systems," *IEEE Trans. Cybern.*, to be published.
[36] P.-Y. Duan, J.-Q. Li, Y. Wang, H.-Y. Sang, and B.-X. Jia, "Solving chiller loading optimization problems using an improved teaching-learning-based optimization algorithm," *Optim. Control Appl. Methods*, vol. 39, no. 1, pp. 65–77, Jun. 2017.
[37] X. D. Zhao, P. Shi, X. L. Zheng, and J. H. Zhang, "Intelligent tracking control for a class of uncertain high-order nonlinear systems," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 9, pp. 1976–1982, Sep. 2016.
[38] J.-Q. Li, H.-Y. Sang, Y.-Y. Han, C.-G. Wang, and K.-Z. Gao, "Efficient multi-objective optimization algorithm for hybrid flow shop scheduling problems with setup energy consumptions," *J. Clean. Prod.*, vol. 181, pp. 584–598, Apr. 2018.
[39] B. Niu, H. Li, T. Qin, and H. R. Karimi, "Adaptive NN dynamic surface controller design for nonlinear pure-feedback switched systems with time-delays and quantized input," *IEEE Trans. Syst., Man, Cybern., Syst.*, to be published.
[40] H. Wang, W. Liu, J. Qiu, and P. X. Liu, "Adaptive fuzzy decentralized control for a class of strong interconnected nonlinear systems with unmodeled dynamics," *IEEE Trans. Fuzzy Syst.*, vol. 26, no. 2, pp. 836–846, Apr. 2018.

KARANDEEP SINGH received the bachelor's degree in computer science and engineering from Punjabi University, Punjab, India, and the master's degree in computer engineering from the PEC University of Technology, Chandigarh, India. He is currently pursuing the Ph.D. degree with the Korea University of Science and Technology (UST), at the Cognitive Informatics Research Section, Electronics and Telecommunication Research Institute (ETRI), South Korea. His research interests include modeling and simulation, big data, and cognitive machine learning.

CHANG-WON AHN received the Ph.D. degree in industrial engineering with specialization in stochastic processes and queuing theory from the Korea Advanced Institute of Science and Technology in 1998. He is currently a Principal Researcher at the Cognitive Informatics Research Section, Electronics and Telecommunication Research Institute (ETRI), and a Professor with the Department of Computer Software, Korea University of Science and Technology (UST). His research interests include big data, modeling and simulation, and social digital twin.