13 May 2020

A System Dynamics Approach for Study of Population Growth and the Residential Housing Market in the US

Gasser Galal Ali

Islam H. El-Adaway
Missouri University of Science and Technology, eladaway@mst.edu

Cihan H. Dagli
Missouri University of Science and Technology, dagli@mst.edu

Recommended Citation

G. G. Ali et al., "A System Dynamics Approach for Study of Population Growth and the Residential Housing Market in the US," *Procedia Computer Science*, vol. 168, pp. 154-160, Elsevier B.V., May 2020.
The definitive version is available at https://doi.org/10.1016/j.procs.2020.02.281

Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 168 (2020) 154–160

Procedia Computer Science

www.elsevier.com/locate/procedia

Complex Adaptive Systems Conference with Theme:
Leveraging AI and Machine Learning for Societal Challenges, CAS 2019

A System Dynamics Approach for Study of Population Growth and The Residential Housing Market in the US

Gasser Galal Ali[a]*, Islam H. El-Adaway[b], Cihan H. Dagli[c]

[a] *Graduate Student, Department of Civil, Architectural, and Environmental Engineering, Missouri University of Science and Technology, 218 Butler-Carlton Hall, 1401 N. Pine Street, Rolla, MO 65409, United States, e-mail: ggah4g@mst.edu,*

[b] *Hurst-McCarthy Professor of Construction Engineering and Management, Department of Civil, Architectural, and Environmental Engineering/Department of Engineering Management and Systems Engineering, Missouri University of Science and Technology, 228 Butler-Carlton Hall, 1401 N. Pine Street, Rolla, MO 65409, United States, e-mail: eladaway@mst.edu*

[c] *Professor, Department of Engineering Management and Systems Engineering, Missouri University of Science and Technology, Rolla, MO 65409, United States*

Abstract

The US Consensus bureau estimated the total construction spending at 1,320,305 Million Dollars, in February 2020, with an increase of 1.1% since last February. The construction market is large, and risky. Prediction of the market behavior, for several years ahead, is needed in order to take strategic investment decision for long and expensive projects. The goal of this research is to study the relationship between population growth and the housing market. To that end, a system dynamics model is developed. System dynamics is a top-down approach that starts with the high-level behavior of a complex system to simulate the behavior of that system over time. The developed model simulates the housing market by matching the population growth with the housing demand in monthly time steps. As such, the parameters of the developed model include birth rate, life expectancy, immigration, emigration, and construction seasonality. Using these parameters, the model simulates the population size and demand for housing. For validation, the outputs of the model are compared with real-life data for the US. When complete, the model should assist market researchers in simulating the housing market. This research benefits large real estate developers, construction companies, governmental and financial agencies.

Peer-review under responsibility of the scientific committee of the Complex Adaptive Systems Conference with Theme: Leveraging AI and Machine Learning for Societal Challenges

Keywords: Complex Systems; System Dynamics; Simulation; Housing; Construction.

1. Introduction

The construction industry is large and costly. In February 2020, the US Consensus Bureau estimated the total construction spending at 1,320,305 Million Dollars, with an increase of 1.1% since the previous February [1]. Of that number, the private

* Corresponding author. *E-mail address:* ggah4g@mst.edu

1877-0509
Peer-review under responsibility of the scientific committee of the Complex Adaptive Systems Conference with Theme: Leveraging AI and Machine Learning for Societal Challenges
10.1016/j.procs.2020.02.281

residential construction spending was 510,076 million dollars. The real estate market has several factors that affect its behavior as a market. Real estate is affected by supply and demand; houses become more expensive when the demand is high and supply is low. However, there is a delay between demand and supply because construction can take several months to several years, depending on the size of the project. Financial markets also affect construction and it is affected by the supply of land, material, and labor. In addition, buildings are durable and expected to survive for many decades. Even if a house is demolished, the land still exists and has a value. Finally, people who buy or rent houses can be users who live in it, or can be investors. Real estate as an investment is attractive to many buyers wishing to put their saving in such relatively low risk investment. However, real estate markets are subject to bubbles and crashes. In the US, the 2007 crisis affected the real estate markets. The speculation of housing bubbles is a major concern. However, real estate markets are very hard to predict, due to the aforementioned factors that influence them. Prediction of the housing market is essential for decision-makers, whether they are real estate developers, financers, or governmental agencies. This research studies the relationship between population growth and the housing market using System Dynamics (SD). Several researchers have successfully used SD to study the behavior, elasticity, and cyclic properties of the housing market. This research extends this field of study by following a holistic approach and dynamically simulating the relationship between high-level socio-economic parameters, the housing market, and construction spending in that sector.

2. Goal and Objectives

The goal of this research is to study the relationship between population growth and the housing market. To that end, a simulation is created using system dynamics. The simulation calculates the population growth in the US and matches it with the supply and demand of the housing market in order to simulate the total residential construction. In addition, an optional ABM alternative is presented and incorporated in the model.

3. Background

SD is a simulation technique that has a top-down approach. It has a high abstraction layer. It simulates the complex behavior of high-level parts of a system in order to simulate the emergent dynamic behavior. On the other hand, Agent-Based Modeling (ABM) takes a bottom up approach. ABM simulates the individual small agents in a complex system, using their simple rules and behavior, in order to simulate the emergent complex system behavior. [2]

Previous research has studied real estate economics in attempt to study the market behavior using SD. Wheaton [3] studied the cyclic properties of real estate. This was done by developing a stock-flow model driven by asset durability. In its essence, a stock-flow model of real estate assumes that rent will increase, or decrease, until the demand meets the current stock of real estate available for the market. This approach assumes that there is no vacancy (i.e. no vacant houses). Therefore, the market can be simulated, using that stock model and different scenarios of elasticities. The cycles of the market under economic shocks can be observed, until it becomes stable again [3]. Based on that idea, other research has developed models that take into consideration other factors [4] [5]. The vacancy assumption of the stock-flow model can be removed, in order to simulate the actual behavior of the market where houses can be vacant [5]. This research intends to build up on that approach for simulating real estate in order to simulate the total residential construction spending.

4. Methodology

4.1. Data Sources

Several data sources for the U.S. are used in this research to initialize and calibrate the model. The data is gathered from the CIA World Fact Book, U.S. Census bureau, U.S. Bureau of economic analysis, and the Federal Reserve Economic Data (FRED) at Saint Louis. The data includes: Birth Rate, Death Rate, Migration Rate, Workforce, Unemployment [6]; average income [7], Total Residential Construction Spending (TLRESCON) [8]; Population (POPTHM) [9]; and Number of Households [10]. The model is trained in this research using data for the entire US, however, it can be retrained using local or regional data so that it can output results in the local/regional scope.

4.2. Research workflow

In this research, SD is used to simulate population growth and the housing market. An alternative method to simulate population growth using ABM is also presented. In order to achieve that goal, a multistep approach, shown in Figure 1, is presented. Simultaneously, two models are developed; (1) an SD model, and (2) an ABM that can be used in the model only to simulate the population growth. Both models are calibrated using actual data from the US.

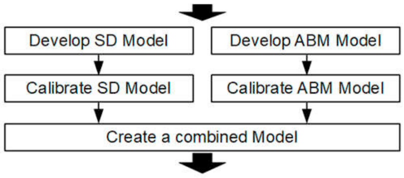

Figure 1: Research Steps

4.3. SD Model: Development

The outline of the SD model is shown in Figure 2. The outline can be logically separated into three parts. The first part, the part on the top simulates the population growth, which is affected by the birth rate, the migration rate, and the death rate. After calculating the population size, the number of payers can be calculated using the unemployment percentage and the labor force percentage. This research assumes that the demand for houses is connected to the number of people who have work and are therefore willing to buy or rent a house. This connection is affected by a factor called the Occupation Factor, which takes into consideration that not everyone capable will buy or rent one house, e.g. some people live together as a family or roommates, others may have multiple houses as an investment. Therefore, the number of payers and the occupation factor are used to calculate the household demand.

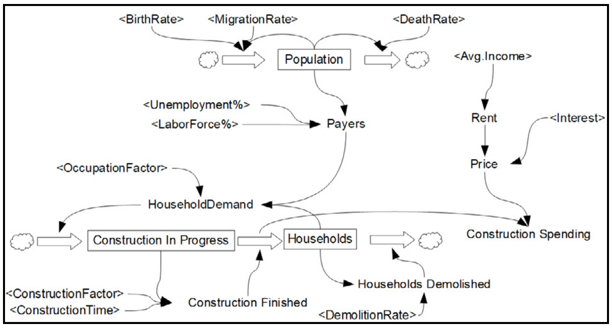

Figure 2: SD Model Outline

In the second part, on the bottom Figure 2, the household demand is used as the inflow for the occupation in progress stock. This translates that housing project have started. The number of construction projects (households) that are finished is calculated using a construction time, which translates to the average duration of a project, and the construction factor. The construction factor is used to induce the seasonality of construction in the US due to the effect of weather; in the winter, construction activity is slow, and in the summer, construction activity is higher. This factor is calculated using Equation 1. This results in a factor of 2/3 when t=0 (month 0), and a factor of 1 when t=6 (month 6).

$$ConstructionFactor_t = \frac{-\cos\left(\frac{time * \pi}{6}\right) + 5}{6} \quad (1)$$

Accordingly, the number of households constructed is calculated, and used as both the outflow of the construction in progress stock and the inflow of the households stock. Finally, the number of households demolished is calculated using the demolition rate, which is the average lifespan of a household.

Finally, in the third part, on the right part of Figure 2, the construction rate is used to calculate the total construction spending. This is calculated using a series of reasonable assumptions. First, using the average income of a person, it can be assumed that the rent of a house is a third of that income. Next according to an interest, calculated later through a calibration process, the price or cost of a house is calculated. This assumes that the price of a house is the present value of infinite monthly rent payments, according to Equation 2. By multiplying the price of a household by the construction rate (household construction per month), the total construction spending is calculated.

$$Price_t = \frac{Rent}{Interest_t} \quad where \quad Rent = \frac{averageIncome}{3} \tag{2}$$

4.4. SD Model: Calibration

There are two parameters in the SD that need to be calibrated. First, the initial number of construction in progress. Though the actual number of households can be set from actual data of the number of households [10], still the initial number of construction in progress needs to be calibrated by using the same data from the number of households. Another factor that needs to be calibrated is the interest rate. The interest rate is assumed, in this research to be time dependent, taking the form of Equation 3. The model is calibrated, to calculate the parameters a and b in Equation 3 by matching the calculated time series for the total residential construction spending with the actual data from the U.S. Census Bureau [8].

$$interest_t = a + time \times b \tag{3}$$

4.5. ABM: Development

ABM is used as an alternative to calculate the population growth. Cellular Automata (CA), which is an agent-based method, is used for that task. The model uses a modified CA based on John Conway's game of life [11], which is a two-dimensional cellular automaton where simple rules control when a cell is "alive" or "dead". The model used in this research uses modified rules, and introduces a stochastic behavior from the CA. The rules are as follows:

- The model has a defined number of cells, and each cell has one of two conditions: either "alive" or "dead".
- If a cell is alive, it has a 10% probability of dying
- If a cell is alive and the number of its surrounding neighbors is greater than or equal five, it has a 20% change of dying (starving due to low resources)
- If a cell is "dead" and the number of its surrounding "alive" neighbors is greater than or equal two, it has an 8% chance of getting born.

The combination of the rules stated above create an emergent behavior where the number of cells alive increase gradually with a small random effect. The number of cells alive can be used, after some factoring, as the population in the SD model.

4.6. ABM: Calibration

After getting the number of cells alive from the ABM, it is factored Equation 4, according to and used in the SD model as the population. Therefore, the model needs to be calibrated to adjust factors a and b in Equation 4, by matching the population calculated from the model with the actual data in the U.S. [9].

$$population_t = a + cellsAlive_t \times b \tag{4}$$

5. Results and Analysis

5.1. Results of SD Model

Figure 3 shows the results of the model for the population, compared to the actual data of the population from FRED [9]. As shown, the population calculated from the model is very close to the actual data. The population calculated is, however, linear because of the assumption that the birth rate, death rate, and migration, are constant numbers, although they are time dependent in reality.

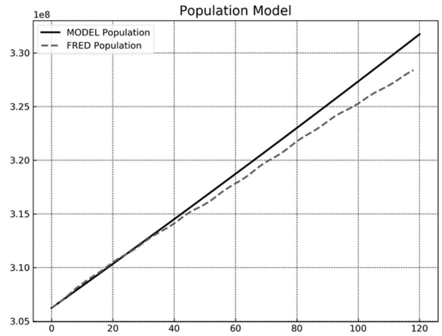

Figure 3: Population: Model Output and Actual Data from FRED [9]

Figure 4 shows the households calculated from the model vs the data from FRED [10]. It should be noted that, due to limitations of the data available, the data from the number of households is annual and therefore does not reflect the monthly variation in the time series. However, the number of households calculated by the model is still close to the actual data.

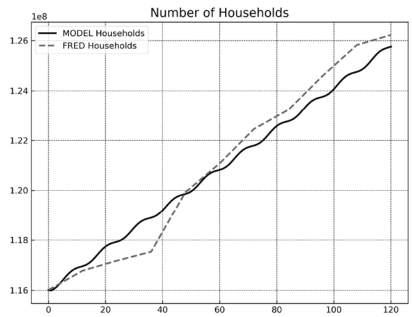

Figure 4: Households: Model Output vs Actual Data [10]

Finally, Figure 5 shows the model's output for the total residential construction spending, vs. the actual data from FRED [8]. The output of the model has a seasonality trend and growth rate similar to the actual data. There is some difference between the output and the data, like the drop in the models output between years 20 and 40, but still the difference is minimal.

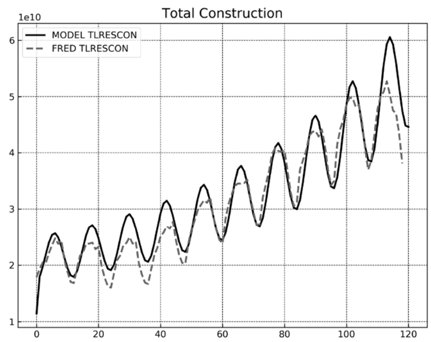

Figure 5: Total Residential Construction Spending: Model Output vs Actual Data [8]

5.2. Results of SD with ABM

The following section shows the results for the alternative configuration of the model using an ABM to simulate population growth. Figure 6 shows the output of the AMB model: Figure 6 (a), the ABM CA model is shown to have a growth pattern similar to Conway's Game of Life [11]; Figure 6 (b) shows a time series of the number of cells alive shown in Figure 6 (a). The ABM shows a steady growth of the number of cells alive.

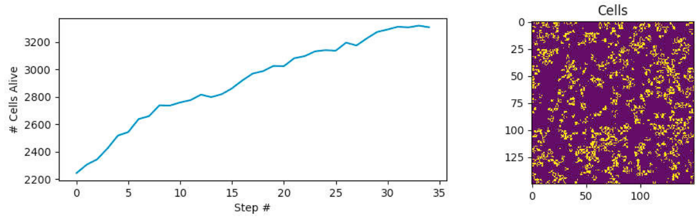

Figure 6: ABM CA Output. Left (a): Time Series of the Number of Cells Alive; Right (b): Grid of the CA

The ABM model was calibrated using the actual data for population [9], according to the previously shown Equation 4. Then, the ABM model was incorporated into the SD model. Figure 7 shows the population growth in the SD model after incorporating the ABM model, vs the actual results [9]. It should be noted that, due to data availability limitations, the plot for the actual population is annual. Therefore, the line for the actual data in Figure 7 does not show the monthly variations during each year. The output of the model using ABM shows the random variability of the population growth and at the same time has a long-term growth similar to the actual data.

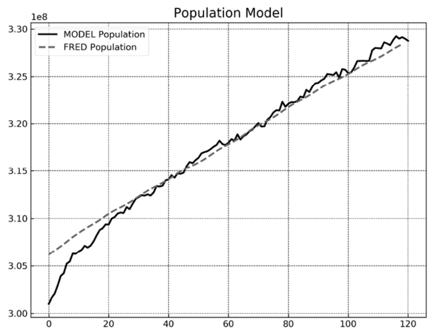

Figure 7: Population Growth in the SD Model using the ABM, vs Actual Data [9]

Figure 8 shows the number of households, using the ABM model. The figure shows small variations to the growth trend, compared to the results of the mode without the ABM, previously shown in Figure 4. The long-term growth is acceptable, but still there is some gap between it and the actual data from Fred [10].

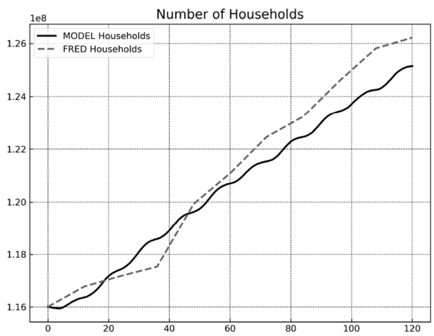

Figure 8: Number of Households in the SD using ABM, vs. Actual Data [10]

Finally, Figure 9 shows the total residential construction spending calculated using the model, vs the actual data from the Census Bureau [8]. The model shows a good random variability in the data that mimics the actual variability in the real data. In addition, the overall long-term growth rate of the sending is similar to the actual results. Overall, the trend of the graph has a realistic growth and stochastic behavior.

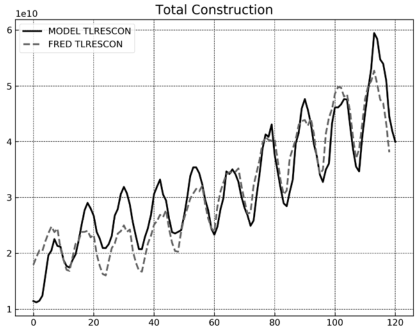

Figure 9: Total Residential Construction Spending in the SD model using ABM, vs Actual Data [12]

6. Concluding Remarks

This research investigates the relationship between population growth and the total residential construction spending in the US. The contribution of this research is that it takes a holistic approach by dynamically simulating the relationship between high-level socio-economic parameters. To that end, a SD model is developed. The model simulates the population growth, its relationship to supply and demand of residential construction, and finally the total residential construction spending. An optional alternative to simulating population growth, using ABM, is also incorporated into the model. The ABM part uses a modified version of Conway's Game of Life [11], with some stochastic behavior. The SD model and the ABM are calibrated using actual data from the US. Overall, the results of the model are acceptably close to the actual results of the US. This model can be used on a high-level by governmental agencies, or by researchers, to study the behavior of the market in the US. Though the model is created using data for the entire US, it can be trained using regional data on the state or city level. This was real estate developers can use to model to determine regional development plans. Using this framework, decision makers can prospectively anticipate the needs of the market and take decisions based on that. The model can also be used to test the stability of the market by introducing economic shocks into the model and studying the market's reaction.

7. References

[1] US Census Bureau. (2019). "US Census Bureau Construction Spending Survey," Construction Spending, 2019. [Online]. Available: https://www.census.gov/construction/c30/c30index.html. [Accessed: 03-May-2019].
[2] M. S. Eid and I. H. El-adaway. (2018). "Decision-Making Framework for Holistic Sustainable Disaster Recovery: Agent-Based Approach for Decreasing Vulnerabilities of the Associated Communities," Journal of Infrastructure Systems, vol. 24, no. 3, p. 04018009.
[3] W. C. Wheaton. (1999) "Real Estate 'Cycles': Some Fundamentals," Real Estate Economics, vol. 27, no. 2, pp. 209–230.
[4] Y. Barlas, B. Özba, and O. Özgün. (2007). "Modeling of Real Estate Price Oscillations in Istanbul," p. 23.
[5] X. Zhang, D. Geltner, and R. de Neufville. (2018). "System Dynamics Modeling of Chinese Urban Housing Markets for Pedagogical and Policy Analysis Purposes," J Real Estate Finan Econ, vol. 57, no. 3, pp. 476–501.
[6] Central Intelligence Agency. (2019). "North America: United States — The World Factbook - Central Intelligence Agency," The World factobook (North America: United States). [Online]. Available: https://www.cia.gov/library/publications/the-world-factbook/geos/us.html. [Accessed: 07-May-2019].
[7] U.S. Census Bureau. (2018). "Social Media Graphic: Median Household Income,". [Online]. Available: https://www.census.gov/library/visualizations/2018/comm/social-income.html. [Accessed: 07-May-2019].
[8] U.S. Census Bureau. (2019). "Total Construction Spending: Residential," FRED, Federal Reserve Bank of St. Louis. [Online]. Available: https://fred.stlouisfed.org/series/TLRESCON. [Accessed: 07-May-2019].
[9] U.S. Bureau of Economic Analysis. (2019). "Population," FRED, Federal Reserve Bank of St. Louis. [Online]. Available: https://fred.stlouisfed.org/series/POPTHM. [Accessed: 07-May-2019].
[10] U.S. Census Bureau. (2019). "Total Households," FRED, Federal Reserve Bank of St. Louis. [Online]. Available: https://fred.stlouisfed.org/series/TTLHH. [Accessed: 07-May-2019].
[11] A. Adamatzky, Ed.. (2010) "Game of life cellular automata". London ; New York: Springer.
[12] U.S. Bureau of the Census. (2019). "Total Construction Spending," FRED, Federal Reserve Bank of St. Louis. [Online]. Available: https://fred.stlouisfed.org/series/TTLCONS. [Accessed: 19-Nov-2018].