Modeling Residential Segregation Based on Schelling's Model: An Analysis Using ENIGH 2022 Data

Jonathan Arturo Preciado Reyes

June 2024

Abstract

Residential segregation is a phenomenon where population groups are distributed unequally across urban spaces. This project aims to model residential segregation patterns based on income levels using data from the 2022 National Survey of Household Income and Expenditure (ENIGH), which divides the population into ten equal parts (deciles) based on their average income. By using Schelling's segregation model, this study introduces a new dimension: positions with dynamic costs based on demand. The model simulates the segregation of different socioeconomic groups, showing clear disparities in housing affordability. Results indicate that lower-income deciles allocate a higher percentage of their income to housing compared to higher-income deciles, reflecting real-world inequalities. This highlights the importance of government regulation to ensure equitable access to housing.

1 Introduction

Residential segregation can be defined as the unequal distribution of population groups across urban spaces. These groups can be determined by income, religion, or migration status. This project focuses on modeling residential segregation patterns according to income levels. Data from the 2022 National Survey of Household Income and Expenditure (ENIGH) will be used, dividing the population into ten equal parts (deciles) based on their average income.

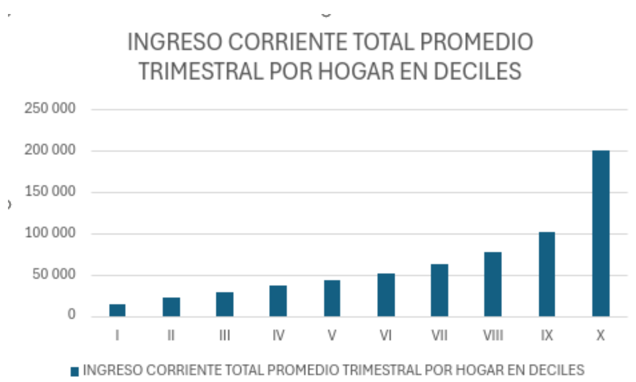

Figure 1: Average quarterly household income by decile

2 Introduction

2.1 On the Importance of the Topic

Mexico is the second most income-unequal country in the Organization for Economic Cooperation and Development (OECD) (Solís 2016). This inequality is physically visible: the spatial location of different socioeconomic groups has significant implications in various areas such as access to education and health. Condominiums, urban slums, and unregulated settlements are characteristic of the country's urban areas (Monkkonen, P. 2012) and are a sign of residential segregation. Thus, this project seeks to simulate these characteristic clusters or areas, based on Schelling's segregation model.

2.2 On the Base Model: Schelling's Segregation

Schelling's segregation model is a concept developed by the economist Thomas Schelling to explain how even small individual preferences can lead to significant segregation in a society. Agents or individuals decide to move to an empty position or stay in their position according to a certain measure of satisfaction, which depends on whether there is a certain measure of similar individuals (in this case, those belonging to the same group) in their neighborhood (adjacent agents). This model allows the observation of segregation patterns according to these groups. Clusters of agents of the same type form emergently.

In the proposed new model, a new dimension will be presented: positions with cost. The cells or positions that agents will occupy will have a dynamic cost that depends on demand. The affordability of these cells depends on the agent's income, so their movement depends not only on their satisfaction but also on the economic possibilities of each group.

3 Statement

3.1 Theoretical Framework

3.1.1 Context and Definitions

Residential Segregation.
Population groups unequally distributed across urban space.
Housing Affordability.
The cost of housing in relation to a person's or family's income. Housing is considered affordable if it generally does not exceed 30
Displacement.
Migration of people from one area to another.
Supply and Demand.
A fundamental principle of market economics. It describes how the price of a good determines the quantity of that good that producers are willing to sell and consumers are willing to buy.

3.1.2 Literature Review

Schelling's Segregation Model.
As explained in the introduction, this agent-based model was proposed by economist Thomas Schelling. This model gives rise to emergent properties of segregation based on the individual preferences of each agent.

Gentrification Model by Niloofar Bagheri-Jebelli, Andrew Crooks, and William G. Kennedy. In 2021, Bagheri-Jebelli et al. presented a dissertation at the *Conference of the Computational Social*

Science Society of the Americas addressing the problem and effects of gentrification by proposing an agent-based model. Cities undergo various processes such as urban growth, expansion, contraction, and gentrification. The model proposed by the authors simulates these dynamics to capture the effect of agents' choices and actions on the city's structure, emphasizing the rise in prices according to the previously mentioned factors.

3.1.3 Key Concepts:

Segregation, displacement, migration, housing affordability, price regulation, supply and demand, gentrification, urban formations.

3.1.4 Research Question.

> **How does the cost of housing in certain areas affect residential segregation patterns in a model based on Schelling's segregation model?**

A model based on Schelling's segregation model is proposed, aiming to capture housing price adjustments and their relationship with different social groups. Specific objectives include measuring the historical prices and how they differ according to each emerging sector. It will also analyze how quickly each population group stabilizes according to different parameters. **Hypothesis:** The model will stabilize after several iterations, where clear segregation will be observed. The prices of the areas or clusters will be close to the agents' budgets.

4 Development

4.1 Parameters

Parameter	Range	Description
population_density	$[0, 1]$	A normalized measure of how populated the grid will be.
cost_update	$[1, 24]$	A measure that sets how often the cost of the cells will be updated in ticks.
housing_budget	$[0, 1]$	A normalized measure indicating how much of the agent's salary will be allocated to housing payment.
required_similarity	$[0, 1]$	A normalized measure indicating the level of similarity agents require to be satisfied.
alpha	$[0, 1]$	A measure or weight that affects how drastic the price adjustment will be.
initial_price	$[1000, 10000]$	A measure that sets the initial cost of all cells. It also serves as the lower bound for this price.

Table 1: Model Parameters

4.2 Agents

The agent population is divided into ten equal parts, called deciles. Each decile has a fixed income or salary, which will be used to afford a cell or position in the grid. As in the traditional Schelling model, agents have a satisfaction variable that depends on their surroundings. An agent is considered satisfied if the similarity with its surroundings (the ratio between the number of neighbors from the same decile and the total number of neighbors) is greater than or equal to the 'required_similarity' parameter. If the agent has no neighbors, it will not be satisfied. Agents will move if they are not satisfied with their neighborhood or if the cell they occupy is not affordable. A cell is considered affordable for an agent if it does not exceed their 'budget', a variable calculated based on their salary and the 'housing_budget' parameter.

Visually, the deciles are represented as follows:

Decile	Color
I	Red
II	Orange
III	Brown
IV	Yellow
V	Green
VI	Lime Green
VII	Turquoise
VIII	Cyan
IX	Sky Blue
X	Dark Blue

4.3 Cells

Each cell or patch will have a dynamic cost that adjusts to demand. The demand for a cell is modeled based on whether a neighboring agent (regardless of whether they are occupying a cell) finds the cell affordable. The price adjustment will be explained later.

4.4 Initialization

According to the population density set by the user, ten equally sized groups will be created and randomly distributed. All cells receive the same initial price, which is manageable through the interface, will be uninhabited, and will have no demand.

4.5 Update

With each iteration, agents will update their parameters and act accordingly. The costs of the cells will be updated depending on the 'cost_update' variable. Below is the flowchart of the agents and the cost update formula.

4.5.1 Agent Update

Agent update is governed by the following logic expressed in a flowchart:

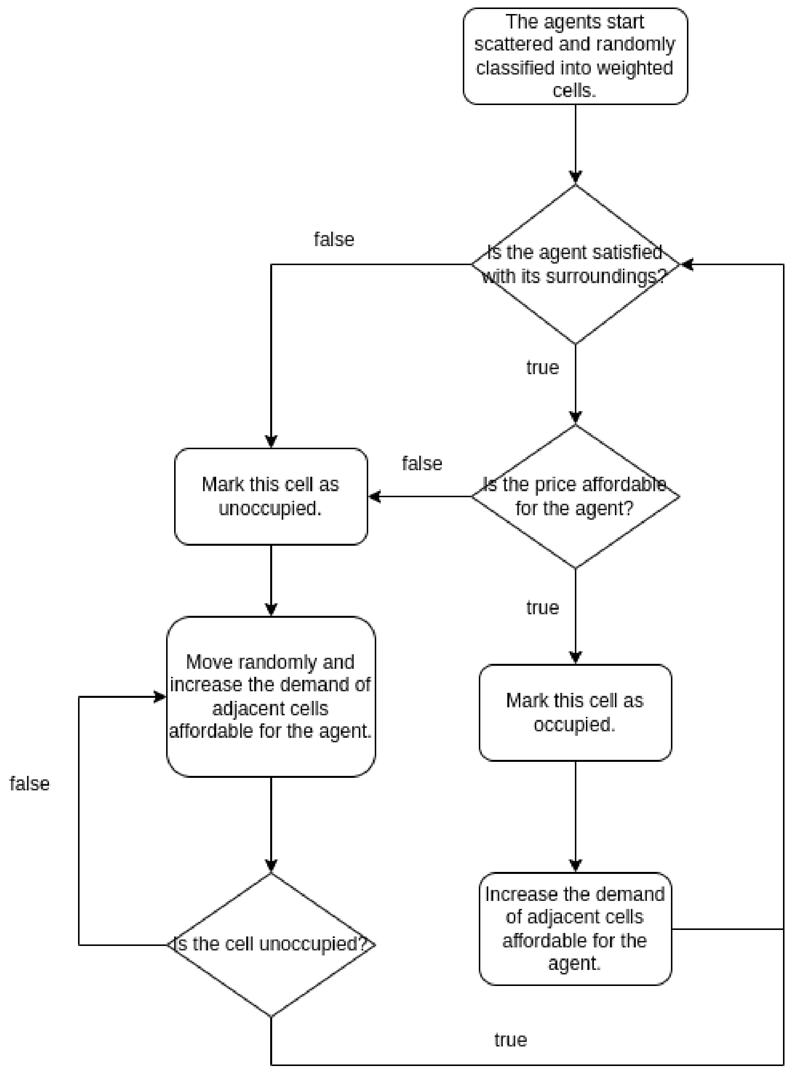

Figure 2: Agent Flowchart

4.5.2 Price Adjustment

The following is the price adjustment formula:

$$\text{new\_price} = \begin{cases} \max\left(\left(1 + \alpha\left(\frac{\text{demand}-\text{avg}}{\text{avg}}\right)\right) \times \text{price}, \text{initial\_price}\right) & \text{if avg} > 0 \\ \text{price} & \text{if avg} \leq 0 \end{cases}$$

Where:
'avg' is the average demand of the area.
'demand' is the current demand for this cell.
'price' is the price variable of this cell.
'new_price' will be the new value of 'price' for this iteration.
'alpha' is a user-defined variable that increases or decreases the weight of the adjustment.

This function is piecewise defined to avoid division by zero. The maximum between the calculated price and the initial price is taken to prevent prices from falling below the lower bound.

If the demand for this cell is greater than the average, the new price will rise, as the product of α by the ratio will be positive. Conversely, if the demand is lower than the average, the ratio becomes negative, and the price decreases to maintain competitiveness.

It should be noted that potential demand is reset with each tick to keep the data updated and prevent this value from increasing with each iteration. It is also important to remember that this price update varies according to the 'cost_update' variable set by the user.

4.6 Termination

The model will terminate when all agents are satisfied.

5 Results

5.1 Suggested Parameters

To better appreciate segregation, it is suggested to have a population density greater than 50% but less than 75%, so that the model can converge (note that the higher the density, the longer it takes for the model to converge). It is also suggested to set a required similarity of 50%, allowing the model to consider the influence of cell prices rather than just the traditional similarity of the Schelling segregation model. Regarding cost updates, this can vary and affects the model's convergence time. Take a value of 12 to simulate that each tick represents a month, and the model's convergence represents an urban area. For the housing budget, a value of 30% is recommended as it is a value suggested by various real-world sources. Consider a high parameter for α so that prices can adjust more quickly. For the model analysis, a value of 0.8 was used. Finally, consider an initial price less than or equal to the product of $13411 * hb$, where hb is the housing budget, to allow agents in the lowest decile to settle.

5.2 On Segregation

With the above in mind, consider the following simulation with a population density of 75%:

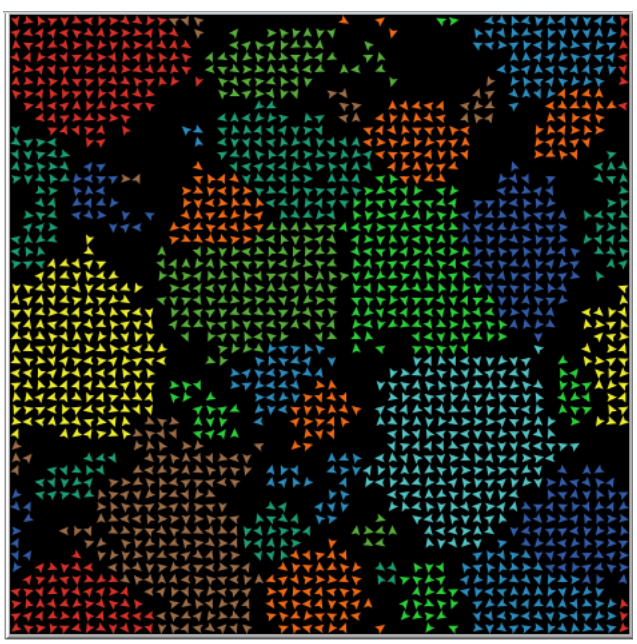

Figure 3: After 3028 iterations, the model converges with a clear distinction between deciles

It is possible to observe segregation in the model, noting that the lower deciles are not typically near clusters of higher deciles due to the higher housing prices.

5.3 On Housing Prices

For quantitative data, consider the average housing expenditure per decile in the final iterations:

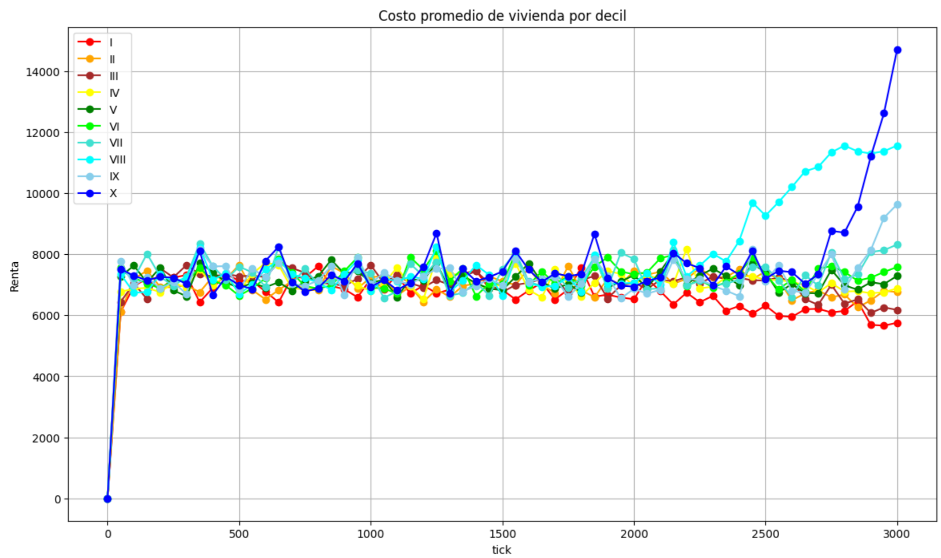

Figure 4: Average housing expenditure per decile

In general, the higher the decile, the more is spent on housing. However, this can be misleading as it does not consider how significant this expenditure is for each class. Thus, consider the following graph representing the percentage of income each decile allocates to housing payments:

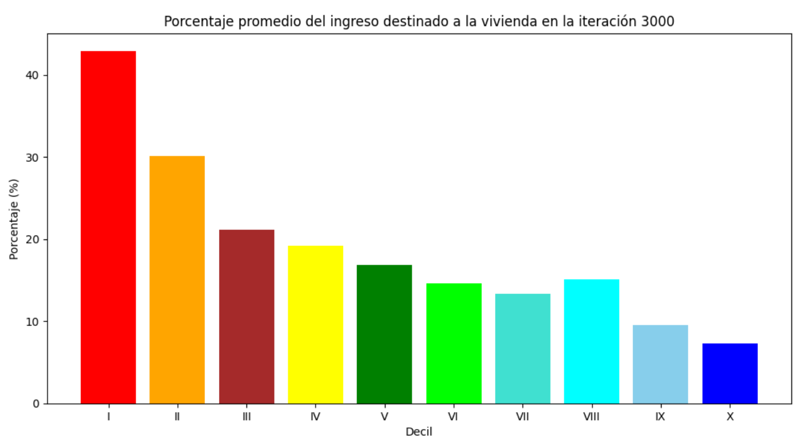

Figure 5: Average percentage of income allocated to housing in one of the final iterations

One of the final iterations (not the last) was taken to provide a more realistic sample of the model since there are no clusters in the initial iterations, and in the last iteration, all agents are satisfied.

Paradoxically, the lower deciles allocate more than 40% of their income to housing, while in the higher deciles, this figure is barely close to 10%.

6 Conclusions

According to Villegas (2022) and data from ENIGH and ENVI (National Housing Survey),

> "...people in decile I must allocate 101.46% and 99.16% of their income to cover the expenses of an average household according to ENVI and ENIGH, respectively. This means that if they want to have a home, they would have to use all their income, thus making access to decent housing impossible. At the other extreme, we find that the percentage of income that the highest decile (X) must allocate is only 6.18% and 6.04% according to ENVI and ENIGH, respectively, which reflects a serious inequality in access to decent housing in Mexico."

Although the model results differ from the reality where decile I would need to allocate more than 100% of their income to housing (largely because this is difficult to model), this model does reflect the disparity in housing access. In particular, decile X does approach this 6.18% expressed by the author, which highlights a significant disparity in housing access. While the research hypothesis is not entirely fulfilled, it is possible to observe segregation in the model, and the results are close to reality.

It is the responsibility of government agencies to regulate real estate speculation and ensure access to housing for all. It is interesting to observe how this problem has been dealt with in other countries (such as China, where 90% of the population owns their homes[1]), as this problem is observable internationally.

6.1 Possible Extensions

Personally, from the beginning, I considered the possibility of modeling social mobility, so that, with a certain probability, an agent could move from one decile to another. The model could easily be extended to this, but in reality, social mobility is very low in Mexico. Regarding gentrification, it would be possible to model it by spontaneously introducing, say, more agents from deciles IX and X in areas normally inhabited by lower or middle deciles.

[1]BBC News Mundo. (2021, October 1). Evergrande: what has led 90% of Chinese citizens to own their homes (and how it influences the crisis of the country's second-largest real estate company). https://www.bbc.com/mundo/noticias-internacional-58670064

7 Bibliography

Sánchez Peña, L. (2012). Cambios en la segregación residencial socioeconómica en México. Realidad, Datos y Espacio: Revista Internacional de Estadística y Geografía, Mayo - Agosto.

Monkkonen, P. (2012). La segregación residencial en el México urbano: niveles y patrones. EURE (Santiago), 38(114), 125-146 e - ISSN 2395-8537.

Instituto Nacional de Estadística y Geografía. (2022). Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH) 2022: Tabulados. https://www.inegi.org.mx/programas/enigh/nc/2022/#tabulados

Solís, A. (2016, 7 octubre). México ocupa el segundo lugar en desigualdad entre los países de la OCDE. Forbes México. https://www.forbes.com.mx/mexico-ocupa-el-segundo-lugar-en-desigualdad-entre-los-paises-de-la-ocde/

Onu-Habitat. (s. f.). Elementos de una vivienda adecuada. https://onuhabitat.org.mx/index.php/elementos-de-una-vivienda-adecuada

Bagheri-Jebelli, N., Crooks, A., & Kennedy, W. G. (2021). Capturing the effects of gentrification on property values: An agent-based modeling approach. En Proceedings of the 2019 International Conference of The Computational Social Science Society of the Americas (CSSSA 2020) (pp. 245–264). Springer. https://doi.org/10.1007/978-3-030-86692-1_14

Villegas Lomelí, C. A. (2022, septiembre 13). Desigualdad en el ingreso y asequibilidad de vivienda en México. Nexos. https://economia.nexos.com.mx/desigualdad-en-el-ingreso-y-asequibilidad-de-vivienda-en-mexico/