# A Simple Model of Wealth Inequality

Gavin Engelstad

December 16, 2023

## 1 Introduction

Economic inequalities are a powerful and ever-present factor in people's lives. One's level of wealth determines where they live, who they know, and how healthy they are [12]. Therefore, understanding wealth dynamics can be vital to understanding how society as a whole functions.

In this paper, we explore the wealth distribution for individuals in the United States and build a remarkably simple model which approximates wealth inequality in the US. We also examine how wealth inequality can arise both at an individual level and at a more systemic group level.

### 1.1 Literature Review

Conventional Macroeconomic theory based on finding equilibriums in a world of representative agents struggles to model heterogeneity in agent behavior and conditions and far-from-equilibrium interactions which change the system without adding massive amounts of complexity [7, 3]. As an alternative, agent based models (ABMs) can be used to study the effects of these factors on the economy.

Although there are many robust published macro ABMs, very few are able to estimate inequality [5, 10, 6]. Instead, toy models like the Yard Sale Model (YSM), which rely on few parameters and simple agent interactions, do a better job [13].

In the YSM, there are many agents each of whom have some level of wealth. Each period, agents are paired up, and an agent in each pair is randomly selected to transfer wealth to the other. A transfer occurs based on a percentage of the net worth of the poorer agent, which varies based on whether the richer or poorer person is receiving the transfer [4].

This paper will attempt to build a model similar to the YSM that relies on fixed transfer amounts and varying probability of winning the transfer, instead of the YSM where transfer amounts vary and probabilities are equal, based on wealth that similarly approximates the real world wealth distribution.

## 2 Economic Background

Model accuracy will be measured via comparisons to real world Lorenz Curves and Gini Coefficients.

A Lorenz Curve (Figure 2.1) is a line that maps wealth percentiles of the population to the total wealth owned by those at or below that percentile. The 45° line represents perfect equality, and the farther a Lorenz Curve curves away from this Line of Equality, the more unequal the wealth distribution is.

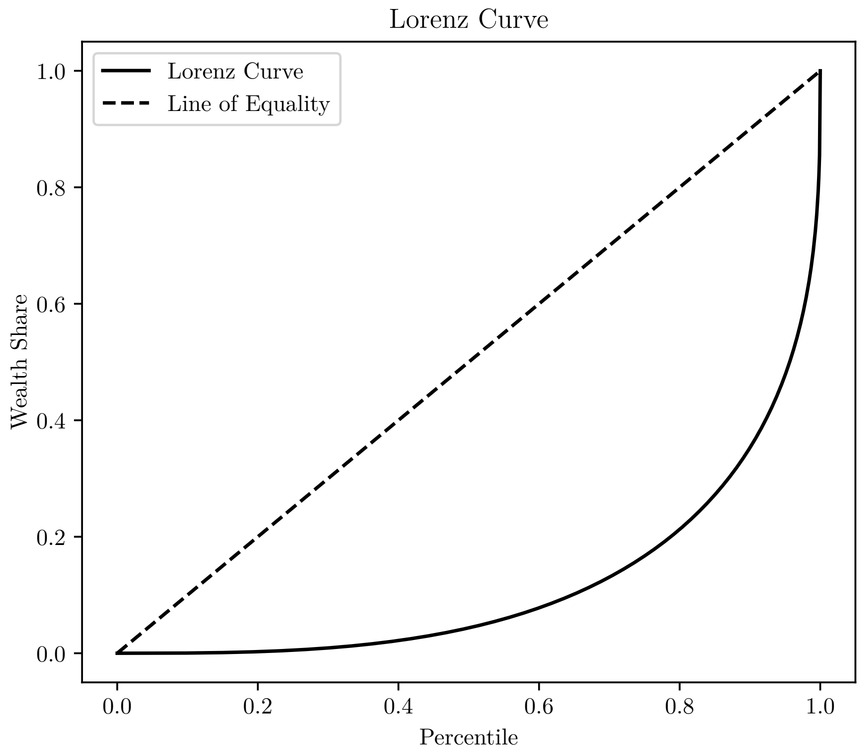



Figure 2.1: Example of a Lorenz Curve

A Gini Coefficient is the ratio of the area between the Lorenz Curve and the Line of Equality and the area under the Line of Equality. It ranges from 0 to 1, where 0 means perfect equality, or that wealth is equally distributed, and 1 means perfect inequality, or that all wealth is controlled by one person.[1] Table 2.1 shows a handful of Gini Coefficient values around the world.

# 3 A Basic Model of Wealth Inequality

## 3.1 Model

Similar to the YSM, the model will consist of many agents exchanging wealth.

Each period

1. Each agent will pair up with another one.

[1]A Gini Coefficient above 1 is possible if you allow for negative net worths. This, however, is irrelevant to the model.

| Country | Gini |
|---|---|
| United States | 0.850 |
| China | 0.701 |
| Uruguay | 0.774 |
| Argentina | 0.809 |
| United Kingdom | 0.706 |
| Germany | 0.788 |

| Region | Gini |
|---|---|
| Africa | 0.879 |
| Asia-Pacific | 0.885 |
| Europe | 0.816 |
| Latin America | 0.858 |
| North America | 0.842 |
| World | 0.889 |

Table 2.1: Gini coefficients around the world. Data from [11]

2. One agent in the pair will be randomly selected to transfer wealth to the other.

   Assuming $i$ and $j$ are paired agents with wealth levels $w_i$ and $w_j$, the probability of $i$ receiving a transfer from $j$ is

$$p_{ij} = \begin{cases} 0.5 + \frac{\alpha}{2} & \text{if } w_i > w_j \\ 0.5 & \text{if } w_i = w_j \\ 0.5 - \frac{\alpha}{2} & \text{if } w_i < w_j. \end{cases} \tag{3.1}$$

3. A transfer of amount $T$ will be initiated between the agents. If the agent selected to send the transfer has wealth less than $T$, the transfer will be for whatever amount of wealth the agent has to give.

$p_{ij}$ isn't affected by the magnitude of the difference between $w_i$ and $w_j$, only whether one is bigger than the other. In this case, the wealthier agent is $\alpha$ more likely to receive the transfer than the less wealthy one. Additionally, Step 3, where the transfer is completed, has a progressive redistributive mechanism built into how the model deals with agents who don't have enough wealth to afford a transfer. This gives a mechanism for wealth to trickle down, even when other factors in the model push wealth up.

The parameters in the model are described in Table 3.1.

| Parameter | Meaning |
|---|---|
| $\alpha$ | Difference in probability of richer and poorer agent receiving the transfer |
| $\overline{w}$ | Average wealth for all agents |
| $T$ | Transfer amount |
| $n$ | Number of Agents in the model |
| $t_f$ | Number of iterations |

Table 3.1: Model parameters

## 3.2 Model Behavior

When run, the model eventually reaches a steady state wealth distribution. Figure 3.1 shows the steady state Lorenz Curve for different $\alpha$ values. When $\alpha$ is higher, the richer person is more likely to receive the transfer, meaning the resulting distribution is more unequal. Similarly, a lower $\alpha$, means the distribution becomes more equal.

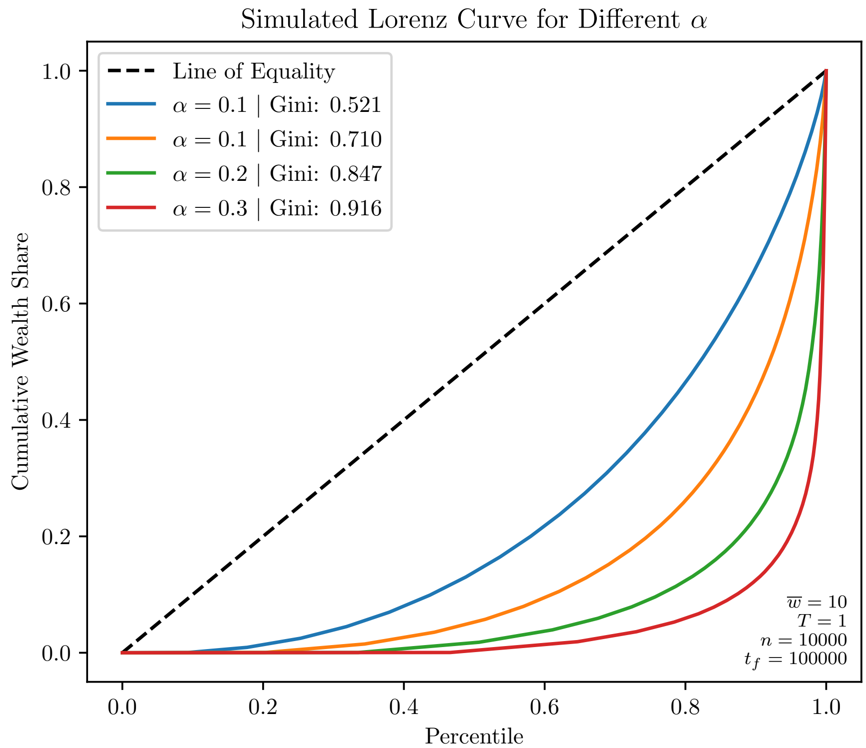
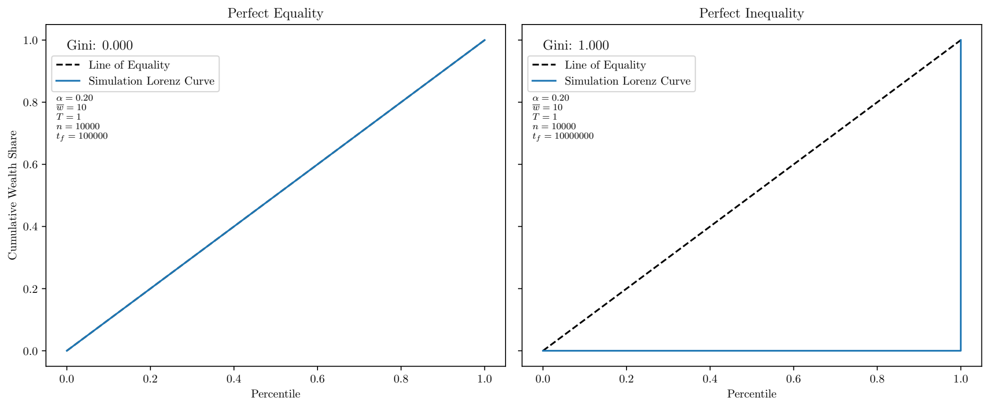



Figure 3.1: Model steady state Lorenz Curves for different $\alpha$ values.

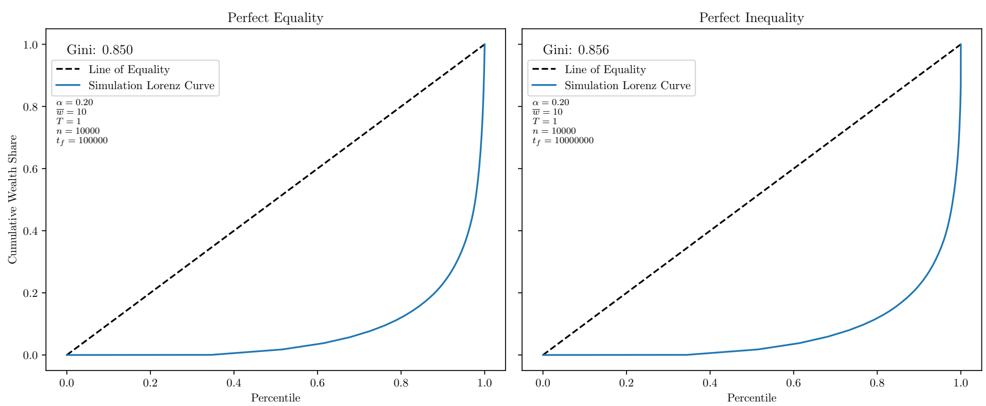
Furthermore, the resulting steady state is stable. Figure 3.2 shows the steady state that results from different initial conditions for the wealth distribution. Even though both simulations start at opposite distributions (Figure 3.2a), the resulting steady states converge (Figure 3.2b). Visually, the Lorenz Curves look identical and the Gini Coefficients in the two simulations are within 1% of each other. It did, however, take 100 times longer for the unequal start to converge than the equal start.

When the same simulation is run for larger $\alpha$ values, the unequal initialization begins to converge faster than the equal one. However, even when $\alpha = 1$, a perfectly equal start never takes 100,000,000 periods to converge like the perfectly unequal one does in Figure 3.2.

Also, social mobility is extremely high in the model. Figure 3.3 shows the transition matrix between quartiles over 10,000 periods. It shows negligible difference in one's chance of ending up in a certain wealth quartile given any start quartile. This demonstrates that even in the steady state, individual agents experience significant variation in their wealth over time. Even though the wealth share for a given percentile stays mostly constant in the Lorenz Curve, the agents that make up that percentile changes significantly as time progresses.




(a) Perfect equality and perfect inequality initial conditions




(b) Perfect equality and perfect inequality steady states

Figure 3.2: Steady states reached with different initial conditions

$$\begin{array}{c} \\ Q_1 \\ Q_2 \\ Q_3 \\ Q_4 \\ Q_5 \end{array} \begin{array}{c} \begin{array}{ccccc} Q_1 & Q_2 & Q_3 & Q_4 & Q_5 \end{array} \\ \begin{bmatrix} 0.2265 & 0.1665 & 0.2065 & 0.213 & 0.1875 \\ 0.177 & 0.2505 & 0.176 & 0.2095 & 0.187 \\ 0.199 & 0.1915 & 0.2285 & 0.201 & 0.18 \\ 0.1935 & 0.2125 & 0.204 & 0.187 & 0.203 \\ 0.204 & 0.179 & 0.185 & 0.1895 & 0.2425 \end{bmatrix} \end{array}$$

Figure 3.3: Transition matrix for probability of moving from the column quartile to the row quartile in 10,000 periods after reaching the steady state. Model parameters are the same as in Figure 3.2

## 3.3 Data Fit

The most surprising part of this model is that despite being such a simple model, it fits real world wealth distributions to a remarkable degree.

Figure 3.4 shows that when $\alpha = 0.2$, the resulting simulation Lorenz Curve nearly perfectly matches points along the real Lorenz Curve for the US wealth distribution.[2] The model has a calculated $R^2$ of 0.997, which again suggests the model has an extremely accurate fit.

The Gini Coefficient of the simulation, 0.849, is also very similar to that of the United States in Table 2.1, 0.850. In fact, even at the steady state, the simulation Gini Coefficient varies by up to around 0.005 period to period. The real Gini value is well within the simulation margin of error.

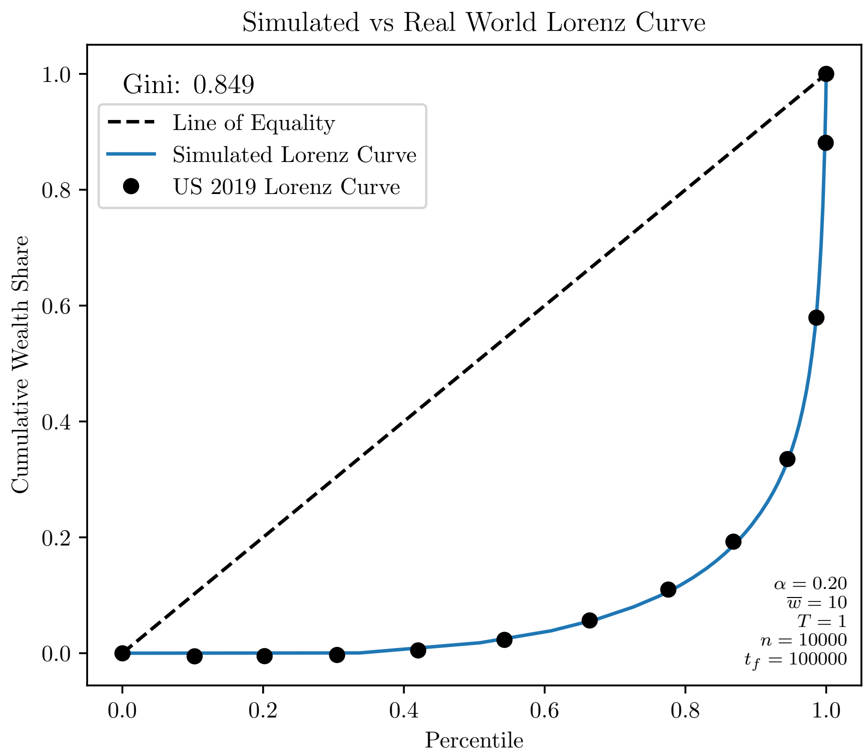



Figure 3.4: Simulated Lorenz Curve with real 2019 US Lorenz Curve points overlayed. Data from [1]

# 4 Modeling Group Wealth Dynamics

Individual wealth isn't the only factor that contributes to systematic wealth inequality. This second model attempts to integrate the advantage one gains from being in a wealthy group into the model presented in Section 3.

## 4.1 Model

The primary difference between this model and the basic model is that agents are now separated into groups. When deciding where to transfer wealth, the per capita average

[2] See Table A.1 for a numerical representation of how close these points are to the simulation curve.

wealth of the groups agents are a part of, $W_i$, is also taken into account, not just the wealth of the agents involved in the transfer, $w_i$.

Each time-step,

1. Agents are paired up

2. In each pair, one agent is selected to receive a transfer and the other is selected to give a transfer.

   Letting $i$ and $j$ be the paired agents, then

$$w_{ij} = \begin{cases} 1 & \text{if } w_i > w_j \\ 0 & \text{if } w_i = w_j \\ -1 & \text{if } w_i < w_j \end{cases} \tag{4.1}$$

   and

$$W_{ij} = \begin{cases} 1 & \text{if } W_i > W_j \\ 0 & \text{if } W_i = W_j \\ -1 & \text{if } W_i < W_j. \end{cases} \tag{4.2}$$

   The probability of $i$ receiving a transfer from $j$ given they're paired up is

$$p_{ij} = 0.5 + \frac{\alpha}{2} w_{ij} + \frac{\beta}{2} W_{ij}. \tag{4.3}$$

3. A transfer is initiated from the transferring agent to the receiving agent. Like in the basic model, if the transferring agent can't afford the transfer, the value of the transfer is capped at the wealth of the transferring agent.

The $\alpha$ parameter in Equation 4.3 represents the difference in probability of receiving the transfer based on your individual wealth while the $\beta$ parameter represents the difference in probability from being a part of the wealthier group. Also, Equation 4.1 and 4.2 only take into account whether an agent or the group an agent is a part of has more, less, or the same amount of wealth, not the relative magnitudes, like Equation 3.1 in the basic model.

The parameters in the model are described in Table 4.1.

## 4.2 Behavior

The grouped model yields a very similar steady state wealth distribution to the basic model. Figure 4.1 shows the steady state Lorenz Curve for the grouped model. Compared to the Lorenz Curve when $\alpha = 0.2$ in Figure 3.1, the shape and Gini Coefficient are very similar.

The groups in the model allow for comparison between wealth levels of each group. Figure 4.2 shows that when $\beta = 0$, the wealth shares held by each group approach the population shares for that group, even when the groups start with a disproportionate wealth distribution. Conversely, if $\beta > 0$, even by a small amount, the steady state wealth distribution leaves one group with a higher share of wealth than the others relative to their population shares.

| **Parameter** | **Meaning** |
|---|---|
| $\alpha$ | Difference in probability of richer and poorer agent receiving the transfer |
| $\beta$ | Difference in probability of agent in the richer group and agent in the poorer group receiving the transfer |
| $\overline{w}$ | Average wealth for all agents |
| $T$ | Transfer amount |
| $n_g$ | (For each group $g$) Number of agents in the group |
| $t_f$ | Number of iterations |

Table 4.1: Model parameters

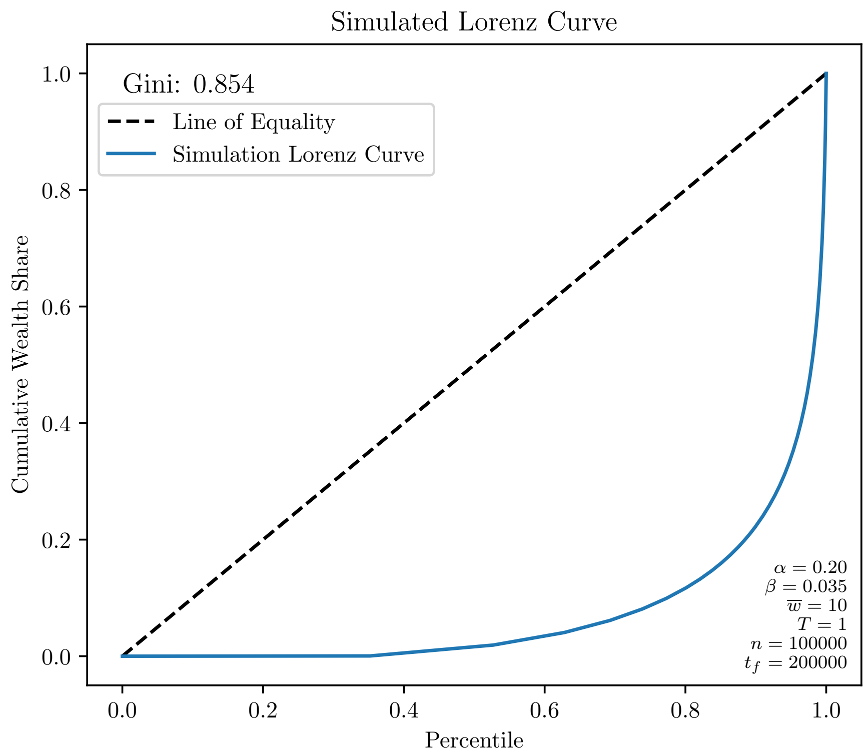



Figure 4.1: Lorenz Curve for the grouped model

By shading the Lorenz Curve, group disparities at each wealth share become apparent. Figure 4.3 shows that Group 1, the dominant group in the simulation, holds a very significant portion of the wealth, especially at higher percentiles. Interestingly, around the 95th percentile the share of wealth held by Group 2 and 3 stops increasing, suggesting that the top of the wealth distribution is only Group 1.[3]

[3]The other sudden changes in group slopes, namely at the 60th and 65th percentiles, I believe, are caused by the way the wealth share code deals with ties. In the case where wealth levels are equal, it puts Group

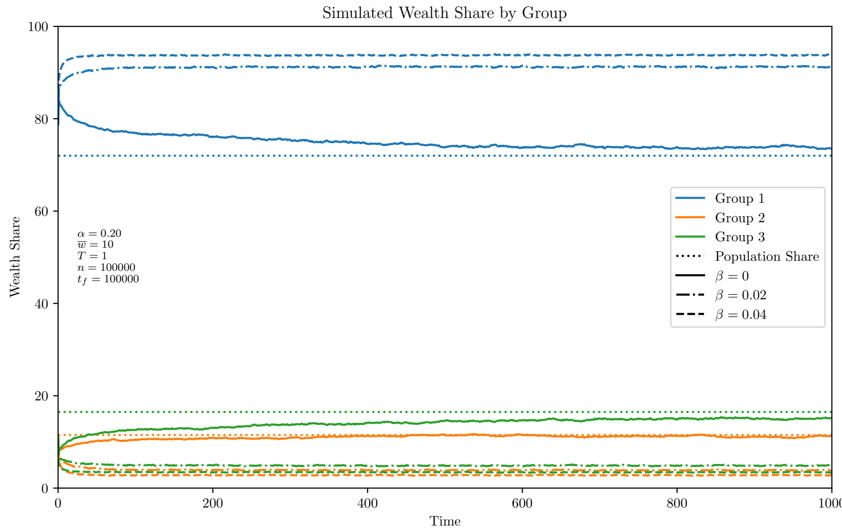



Figure 4.2: Simulation wealth shares for each group

*At $t = 0$, agents in Group A has 2 more wealth than those in Group B and agents in Group B have 2 more wealth than agents in Group C. Otherwise, whichever group ends up dominant in the steady state can be unpredictable.

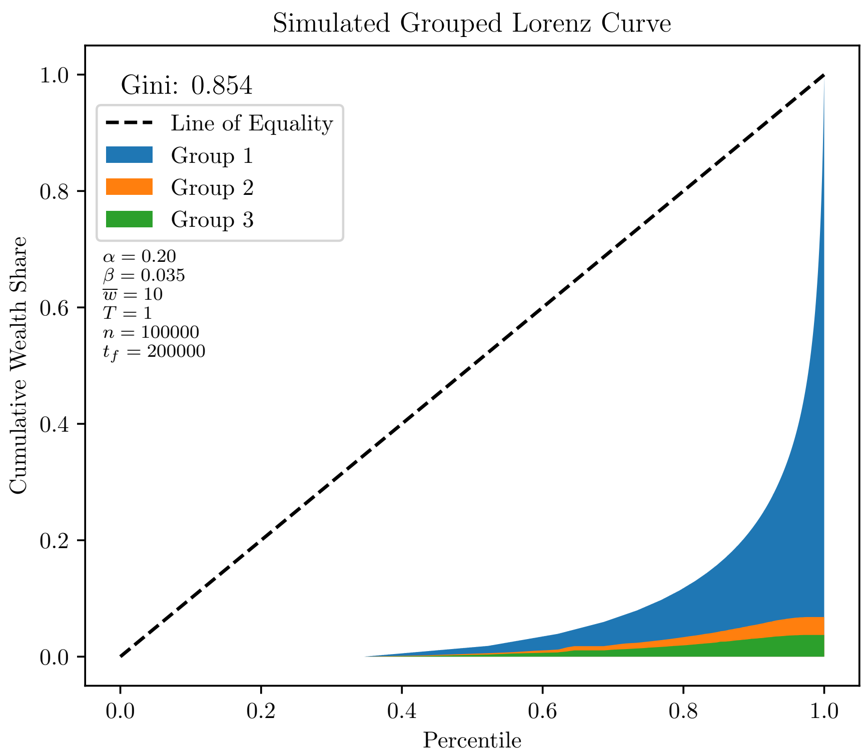



Figure 4.3: Lorenz Curve with group wealth shares overlayed. Shading represents the portion of that wealth share held by people in that group.

1 first, then 2, then 3. Weighting ties based on group shares at those values would likely get a smoother graph.

## 4.3 Data Fit

When group sizes are weighted to approximate the US racial distribution (Table 4.2),[4] $\alpha = 0.2$, and $\beta = 0.035$, the resulting distribution of wealth approximates that within the US.

| Racial Group | Simulation Group | Population Share |
|---|---|---|
| White (Non-Hispanic) | Group 1 | 0.720 |
| Hispanic | Group 2 | 0.115 |
| Black (Non-Hispanic) | Group 3 | 0.165 |

Table 4.2: Race and group population shares

At an individual level, the resulting Lorenz Curve and Gini Coefficient is similar to that of the model presented in Section 3, so it resembles the real Lorenz Curve in a similar manner (Figure 4.4).[5]

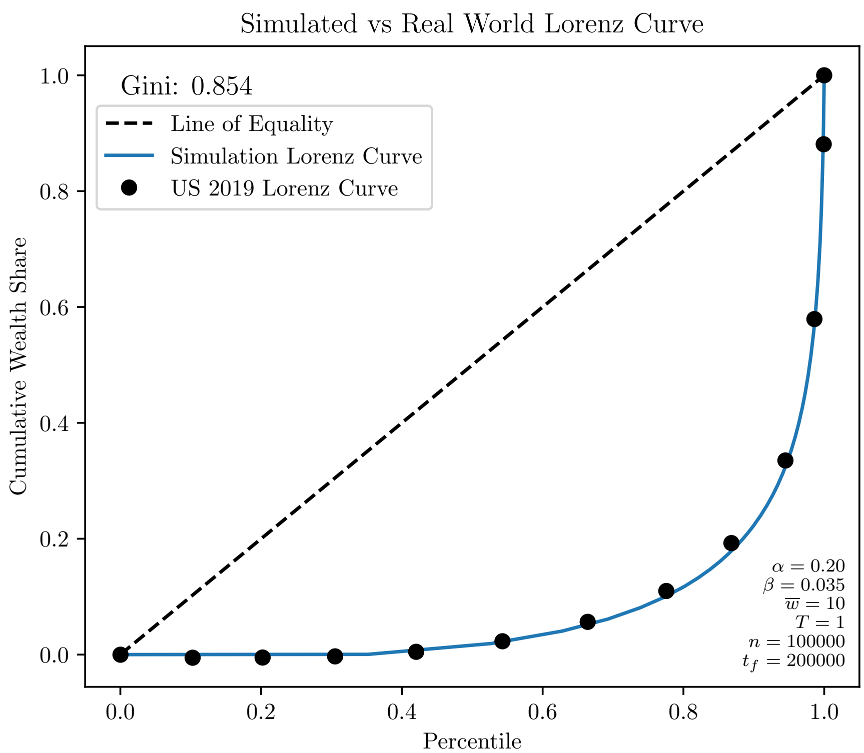



Figure 4.4: Simulated Lorenz Curve from the grouped model with real 2019 US points overlayed. Data from [1]

At a group level, the simulation wealth shares also approximate that of US racial groups. Figure 4.5 shows that as time goes forward in the simulation, the steady state wealth distri-

[4] Due to data constraints, all US data is normalized to only include White non-Hispanic, Black non-Hispanic, and Hispanic populations.

[5] See Table A.1 for the values and residuals numerically.

bution for groups approximates the real world.[6] This suggests that the addition of groups to the model allows the model to simultaneously approximate both real individual wealth dynamics and group wealth dynamics.

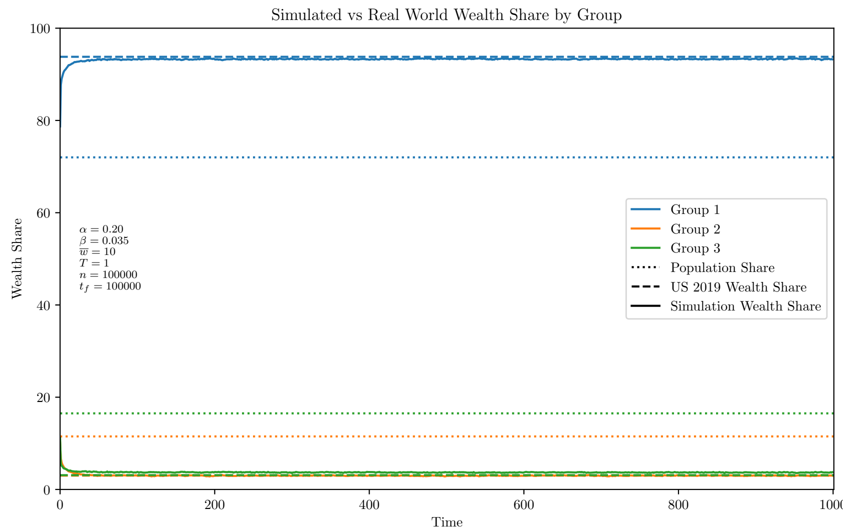



Figure 4.5: Simulated wealth shares over time with US 2019 data overlayed. Data from [1]

The ability to approximate grouped results is a direct result of the $\beta$ parameter. Figure 4.2 shows that when $\beta = 0$, which is functionally identical to the basic model, the model converges to a proportional wealth distribution, which isn't what's observed in the real world.

At an individual level, however, the model is only moderately effective at replicating real phenomena. Figure 4.6 demonstrates that the plateauing behavior at the 95th percentile observed in the model doesn't exist in the real data. Instead, the real world data continues to smoothly increase its slope upward to the top percentile.

This effect is further demonstrated in Table 4.3. The $R^2$ for the model as a whole is still very high, though this could also be due to the fact that multiple groups are represented in the data meaning the mean is a very poor predictor. By group, the $R^2$, though always fairly high, gets substantially lower for groups 2 and 3, suggesting the model is a worse fit for these individuals.

| **Group** | $\mathbf{R^2}$ |
|---|---|
| Group 1 | 0.998 |
| Group 2 | 0.877 |
| Group 3 | 0.805 |
| Total | 0.997 |

Table 4.3: $R^2$ values for the Figure 4.6. Each group is the $R^2$ just for that set of points, while the total $R^2$ is shown in the final row.

[6] See Table B.1 for a numerical representation of this.

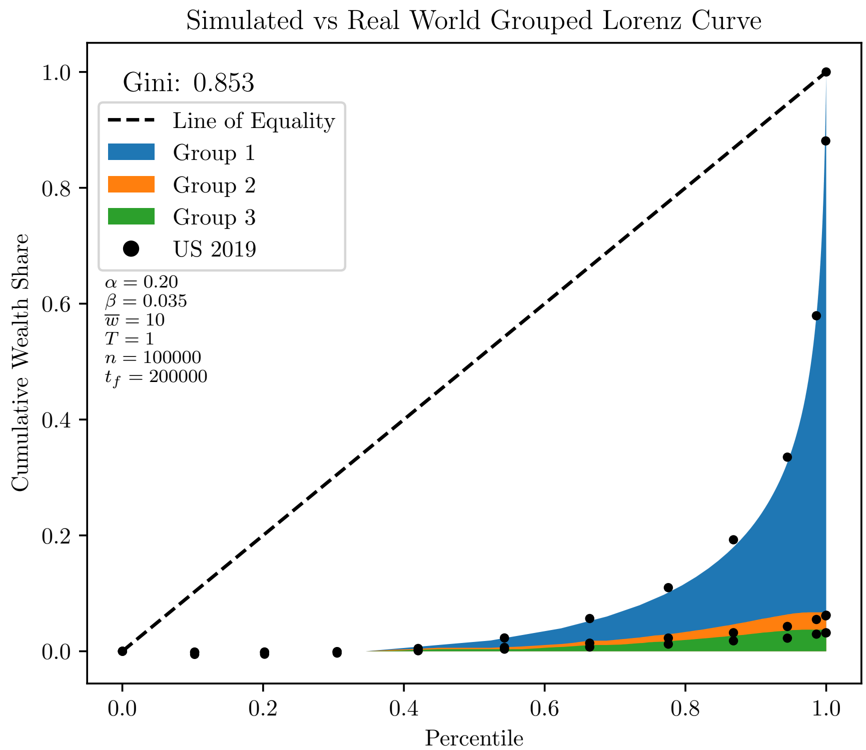



Figure 4.6: Lorenz Curve with group wealth shares shaded overlayed with US 2019 data. Data from [1]

Another way to evaluate the addition of parameters to a model is using the Akaike Information Criterion (AIC). The AIC value can be calculated using

$$\mathrm{AIC} = n \ln \left( \frac{\mathrm{RSS}}{n} \right) + 2k + 2$$

where RSS is the residual sum of squares, $n$ is the number of points, and $k$ is the number of parameters in the model. A lower AIC represents a better fitted model after accounting for the complexity added by a new parameter, so a model with a lower AIC is preferred [9].

Based on the grouped data, the AIC of the basic model is -213.90 and the grouped model is -336.97, suggesting that despite the weaker fit at an individual level, especially for groups 2 and 3, the added parameter does better explain the real world data.

# 5 Conclusion

Overall, the models presented in this paper are able to successfully replicate parts of the wealth distribution of the United States. The basic model in Section 3 closely approximates the wealth distribution for individuals, shown by the Lorenz Curve and Gini Coefficient, of the US in 2019 and the grouped model in Section 4 is able to approximate both the wealth distribution for individuals and racial groups, though struggles where those both interact.

As simple as they are, both models are remarkably accurate at simulating real world wealth dynamics, suggesting that more generally, wealth follows similar patterns where richer people experience advantages in the kinds of wealth transfers that occur every day.

The fact that the model didn't need to incorporate the relative wealth of agents[7] is especially interesting and may suggest that in the real world, class conflict can't be reduced to rich versus poor and includes complex interplays between every individual and the groups above and below them.

## 5.1 Limitations

This paper presents very simple models for very complex phenomena. This has the advantage that it allows for better analysis, since what happens is clearer, avoids risks of overfitting [12], and requires less computational power, but does mean that many factors at play in the real world don't exist in the model. The model is perhaps an oversimplification of what it tries to represent.

Also, data limitations restricted the level of analysis that could be performed. Only 13 points along the real Lorenz Curve were used since that's what's available [1] and constructing a Lorenz Curve for better comparison would require access to restricted Survey of Consumer Finances (SCF) data [8].

Finally, this paper's analysis is only in the US, primarily due to difficulty finding Lorenz curve data for other countries. This reduces the context in which the model can be interpreted.

## 5.2 Further Research

Further research could work to improve the accuracy of Figure 4.6, potentially by going back to the standard YSM and modifying that to add group privilege instead of creating a whole new basic model or fitting better parameters, since the parameters presented in their paper were found within a only a handful of trial-and-error steps.

With access to SCF data, the accuracy of the model could be tested against millions of points along the Lorenz Curve, instead of the 13 in this paper. This would get better results regarding the accuracy and predictiveness of the model.

Finally, the model could be applied to a global scope and compared to other countries. Doing this would require data regarding the net worth of individuals in a country, but would allow for a more comprehensive analysis of the model. Similar to the SCF, however, this data in other countries doesn't allow public access [2].

[7] and, in fact, has a worse fit when this is incorporated

# Appendices

## A Lorenz Curve Values Table

| Population Share | US 2019 | Basic Model | Grouped Model |
|---|---|---|---|
| 0.0002 | -0.0002 | 0.0000 | 0.0000 |
| | | (0.0002) | (0.0002) |
| 0.1027 | -0.0054 | 0.0000 | 0.0000 |
| | | (0.0054) | (0.0054) |
| 0.2021 | -0.0051 | 0.0000 | 0.0000 |
| | | (0.0051) | (0.0051) |
| 0.3051 | -0.0031 | 0.0000 | 0.0000 |
| | | (0.0031) | (0.0031) |
| 0.4203 | 0.0048 | 0.0088 | 0.0078 |
| | | (0.0040) | (0.0030) |
| 0.5429 | 0.0230 | 0.0249 | 0.0227 |
| | | (0.0019) | (-0.0003) |
| 0.6640 | 0.0565 | 0.0551 | 0.0523 |
| | | (-0.0014) | (-0.0042) |
| 0.7757 | 0.1098 | 0.1064 | 0.1015 |
| | | (-0.0034) | (-0.0084) |
| 0.8682 | 0.1925 | 0.1860 | 0.1796 |
| | | (-0.0065) | (-0.01282) |
| 0.9449 | 0.3351 | 0.3308 | 0.3258 |
| | | (-0.0043) | (-0.0093) |
| 0.9861 | 0.5792 | 0.5842 | 0.5792 |
| | | (0.0050) | (-0.0001) |
| 0.9993 | 0.8810 | 0.9439 | 0.9454 |
| | | (0.0629) | (0.0644) |
| 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| | | (0.0000) | (0.0000) |
| | | $\alpha = 0.2$ | $\alpha = 0.2$ |
| | | $\overline{w} = 10$ | $\beta = 0.035$ |
| | | $T = 1$ | $\overline{w} = 10$ |
| | | $n = 10000$ | $T = 1$ |
| | | $t_f = 100000$ | $n = 100000$ |
| | | | $t_f = 100000$ |

Model difference from US 2019 values in parenthesis

Table A.1: US, Basic Model, and Grouped Model Lorenz Curve Values. US data from [1]

# B Race Wealth Shares Table

| **Racial Group** | **US 2019** | **Grouped Model** |
|---|---|---|
| White (Non-Hispanic) | 0.938 | 0.93375 |
| | | (-0.00425) |
| Hispanic | 0.030 | 0.02995 |
| | | (-0.00005) |
| Black (Non-Hispanic) | 0.031 | 0.03629 |
| | | (0.00529) |
| | | $\alpha = 0.2$ |
| | | $\beta = 0.035$ |
| | | $\overline{w} = 10$ |
| | | $T = 1$ |
| | | $n = 100000$ |
| | | $t_f = 100000$ |

Model difference from US 2019 values in parenthesis

Table B.1: US 2019 and Grouped Model wealth shares by race. US data from [1]

* US values shown post-normalization to get rid of the "other" category

# References


[1] Aditya Aladangady and Akila Forde. "Wealth Inequality and the Racial Wealth Gap". In: *FEDS Notes. Washington: Board of Governors of the Federal Reserve System* (2021).

[2] Financial Conduct Authority. "Financial Lives 2022 survey". In: *Financial Conduct Authority* (2023).

[3] Alessandro Caiani et al. "Agent based-stock flow consistent macroeconomics: Towards a benchmark model". In: *Journal of Economic Dynamics and Control* 69 (2016), pp. 375–408.

[4] Christophe Chorro. "A simple probabilistic approach of the Yard-Sale model". In: *Statistics and Probability Letters* 112 (2016), pp. 35–40.

[5] Herbert Dawid and Domenico Delli Gatti. "Agent-based macroeconomics". In: *Handbook of computational economics* 4 (2018), pp. 63–156.

[6] Herbert Dawid et al. "A heterogeneous agent macroeconomic model for policy evaluation: Improving transparency and reproducibility". In: (2016), pgs.

[7] Giovanni Dosi and Andrea Roventini. "More is different ... and complex!: The case for agent-based macroeconomics". In: *LEM Working Paper Series* (2019).

[8] Board of Governors of the Federal Reserve Board. "2022 Survey of Consumer Finances". In: *Board of Governors of the Federal Reserve System* (2023).

[9] Glenn Ledder. *Mathematical Modeling for Epidemiology and Ecology.* Cham, Switzerland: Springer Undergraduate Texts in Mathematics and Technology, 2023.

[10] Pascal Seppecher. "Flexibility of wages and macroeconomic instability in an agent-based computational model with endogenous money". In: *Macroeconomic Dynamics* 16.S2 (2012), pp. 284–297.

[11] Credit Suisse. "Global Wealth Databook 2022". In: *Credit Suisse Research Institute* (2022).

[12] Hannah Thomas, Alexis Mann, and Tatjana Meschede. "Race and location: The role neighborhoods play in family wealth and well-being". In: *American Journal of Economics and Sociology* 77.3-4 (2018), pp. 1077–1111.

[13] Sheng Xu. "TIME-DEPENDENT ANALYSIS TO YARD-SALE MODEL OF WEALTH DISTRIBUTION". In: *Diss. Tufts University* (2022).