Using Markov Chains to Model Economic Mobility in the United States

Authors: Alan Estrada, Daniel Kang, Jordan Rogers
Applied Math 115: Mathematical Modeling

Abstract

The study of economic mobility is essential to understanding how economic experience is affected by present economic position and public policy. It also provides data against which public policies intended to facilitate economic mobility can be evaluated. In this paper we leverage Markov Chains and Individual Agent Modeling to assess the state of economic mobility in America today, ranging from the long term impact of your ancestor's wealth bracket on your income and accumulated wealth to direct analysis of social policy proposals like racial equity and school quality standards.

1 Introduction

We often hear sayings such as "the rich get rich, the poor get poorer", "the middle class/American Dream is dead", and even seemingly data-driven sayings such as the claim that 70% of wealthy families lose their wealth by the 2nd generation, and 90% by the 3rd generation [1]. While it is easy to blindly accept or dismiss such claims and base our assumptions about the economic system we live in upon such sayings, it is important to collect and analyze the data to confirm or deny such statements and use that information to assess the state and direction of the American economy and how it is affected by social policy.

Fortunately, the collection part has been done numerous times across the globe, making analytics and assessment much more accessible. In 2006, a group of researchers from Forschungsinstitut zur Zukunft der Arbeit Institute for the Study of Labor in Bonn, Germany published a paper titled *American Exceptionalism in a New Light: A Comparison of Intergenerational Earnings Mobility in the Nordic Countries, the United Kingdom and the United States* [2]. This study published income quintile transition matrices from parent to child. The U.S. Father-Son transition matrix, which is the backbone of much of our work, is included below with the father's quintile (1 being the bottom 20% and 5 being the top 20%) given by the row, and the son's by the column.

$$\begin{bmatrix} 0.421 & 0.245 & 0.153 & 0.102 & 0.079 \\ 0.194 & 0.284 & 0.208 & 0.174 & 0.140 \\ 0.194 & 0.186 & 0.256 & 0.202 & 0.162 \\ 0.125 & 0.182 & 0.198 & 0.252 & 0.243 \\ 0.095 & 0.122 & 0.189 & 0.234 & 0.360 \end{bmatrix}$$

The Stanford Center on Poverty and Inequality also reports that segregation and school quality are two of the other main factors that are strongly correlated with economic mobility [3]. The institute proves a strong negative correlation between standard measures of racial segregation and upward mobility, as well as a strong positive correlation between areas with higher test scores (controlling for income levels), lower dropout rates, as well as smaller class sizes, and higher rates of economic mobility. By simulating the passing of multiple generations, we hope to investigate the impact of combinations of racial integration and school quality on the potential mobility of individuals in a given population.

2 Description of Model

2.1 Prior Generational Wealth & Income Quintile Model

In order to assess the impact of prior generational wealth and income quintile on the economic mobility of individuals in the United States, we ran an individual agent based markov chain simulation to evolve and track a population of 50,000 individuals (initialized as 10,000 individuals in each of the income quintiles) and their descendents both on their income bracket and their accumulated wealth (see "RogersEconomicMobility.m").

Accurate simulation of the latter requires accurate information on income, individual savings rate, and investment returns by quintile. Fortunately, data exists for all, with very detailed data for the former [4] and somewhat more sparse data for the latter [5] from which we must make reasonable extrapolations to model wealth accumulation. For investment returns we assume an average returns equal to that of the US equity market. From these sources we derive the following:

Quintile	[1]	[2]	[3]	[4]	[5]
Income ($1000s)	15	35	55	90	160
Savings Rates (%)	0	0	2	6	12
ROI (%)	7	7	7	7	7

2.2 Racial Integration & School Quality Economic Mobility Model

Then, in order to assess the impact of racial integration and school quality on the economic mobility of individuals in the United States, we parameterize these characteristics into a probabilistic transition matrix with states *1* (lower income) and *2* (higher income), using *r* and *s* to represent a number within an index of [0,5] for racial integration and school quality, respectively. The probabilities for the transition matrix are constructed, where *p* is the probability of transitioning from a lower income level to a higher income level, and *q* is the probability that an individual in a high income level will stay at that level in the next generation:

$$P(A_{t+1} = B) = \frac{(rs)^2 + r^2}{50} = p \qquad P(B_{t+1} = B) = \frac{r^2 + s^2}{50} = q$$

The probability of transitioning from *State 1* (lower income) to *State 2* (higher income) is determined primarily by racial integration [3] where the quality of schools, despite a given level, will ultimately be affected by whether or not segregation exists (since the segregated minorities will always get the lower quality education), thus the *r*s* relation in the equation. The probability of staying at the high income level (*State 2*) equally weighs the importance of racial integration and school quality in economic mobility. This populates the transition matrix below:

$$\begin{bmatrix} 1-p & p \\ 1-q & q \end{bmatrix}$$

In order to better understand how these parameters (racial integration and school quality) impact economic mobility for future generations, we characterize four different regions with different combinations of racial integration and school quality on a scale of [0,5].

Region Type	Racial Integration r	School Quality s
Highly integrated population with good schools	4.5	5
Highly integrated population with bad schools	4.2	1.15
Poorly integrated population with good schools	1.1	4.3
Poorly integrated population with bad schools	1.3	1.8

2.2 Sensitivity Model

In order to assess the impact of each individual markov chain component to social mobility, we change each markov element by a finitely small number over many iterations, and distribute the small change to retain a valid markov chain.

For each finitely small change in one element per time, we calculate two social mobility metrics for each markov chain: the Bartholomew index and trace index [6]. Below is the mathematical formulation for each metric with mb representing the Bartholomew index and mt representing the trace index. Furthermore in the equations below k represents the number of categories that exist within the matrix which in our case is 5.

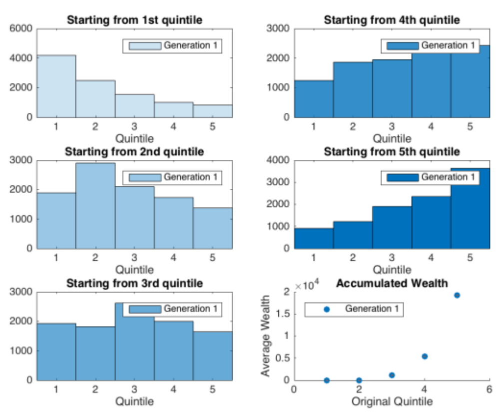
$$m_T = \frac{k - trace(P)}{k-1} \qquad\qquad m_B = (1/k)\sum_i \sum_j P(i,j)|i-j|$$

Each of these indices measure social mobility in a different way. The trace index calculates a metric that increases as the percentage of people that stay in the same class increases and decreases in the other direction. Furthermore this index only has values that lie from [0,1] where 0 indicates no social mobility and 1 indicates perfect social mobility. Moreover, the Bartholomew index weighs transition probabilities by the number of categories traversed by the matrix (see “TransitionSensitivityAnalysis.py” for the code).

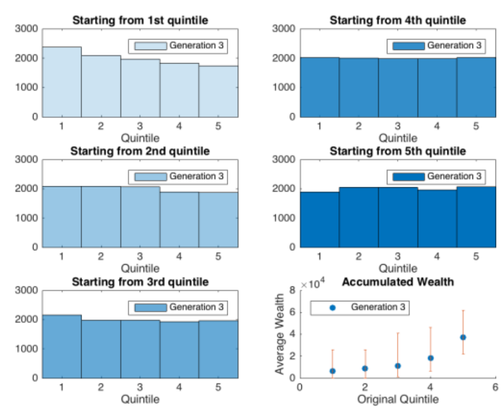
3 Analysis

3.1 Prior Generational Wealth & Income Quintile Analysis

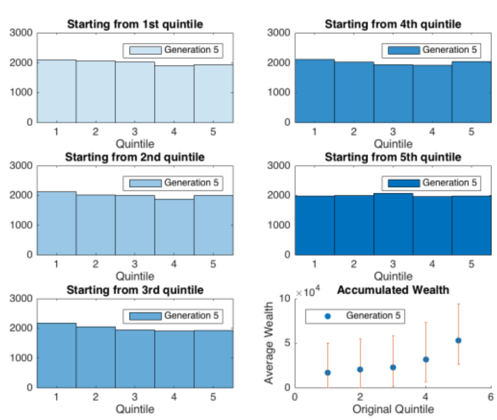
The results from the investigation on the impact of prior generations’ income on the long-term income and wealth of their descendents speak for themselves in the charts below. Each chart shows the distribution of descendents incomes based on the starting quintile, as well as the accumulated wealth (with the 95% coverage interval show with the vertical red bars).



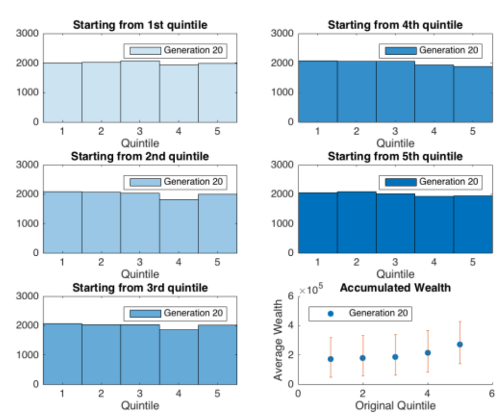
Generation 1 Income and Wealth



Generation 3



Generation 5



Generation 20

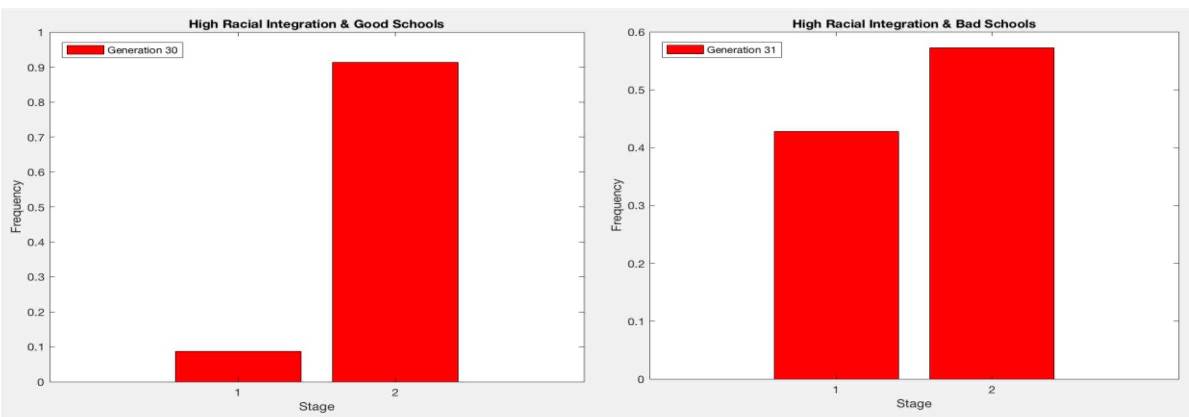
While the first generation of descendents do show a noticeable skew in income distribution, by the third generation the field is virtually even with the notable exception of the lowest quintile, which still displays a bias to the lower quintiles that cannot be exclusively attributed to the stochastic nature of the simulation. Referring back to the transition matrix we observe that the diagonal elements (the probabilities of a son being in the same quintile as his father) are the largest probabilities for each row, and this element in the 1st quintile is by far the largest, about 60% larger than the 2,3,4 quintiles, and 17% larger than the 5th quintile. The 2nd quintile also notably has the highest probability of remaining stationary of the 2,3,4 quintiles, thus it makes intuitive sense that the 1st quintile will take longer to reach equilibrium.

From the fifth generation onwards, no notable differences exist when looking strictly at income. This makes intuitive sense because if we take the transition matrix to a reasonably high power and multiply it by the uniform, five element column vector $[.2\ .2\ .2\ .2\ .2]^T$ we arrive at the steady state distribution of $[.2\ .2\ .2\ .2\ .2]^T$; the simulation merely confirms this is distributed evenly regardless of initial family income quintile.

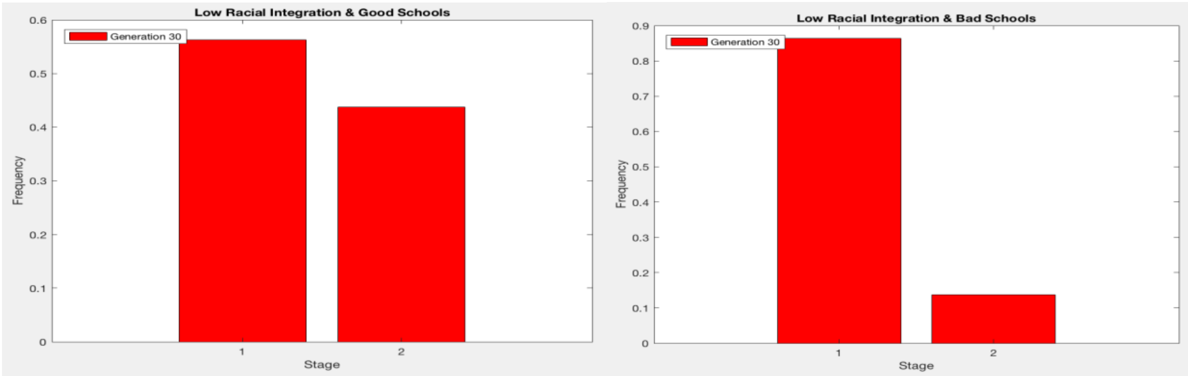
While long term income quintile is unaffected, accumulated wealth bears a different verdict. Looking at the bottom right graph of each of the above illustrations, we see the starting quintile has a significant impact on accumulated wealth, even though all but the lowest income quintiles reach equilibrium in about 3 generations. In our simulation there is only one investment/return cycle per generation, so the wealth spread is lower than in the real world where 20-30 investment cycles occur before the next generation is in the workforce.

3.2 Racial Integration & School Quality Economic Mobility Analysis

The results of the stochastic simulation (averaging 500 trials) from the investigation of racial integration and school quality on economic mobility for each region type, with an initial population half split between High Income (*State 2*) and Low Income (*State 1*), and after 30 generations, are shown below (see ‘MobilityMarkov.m’):



Regions with good schools and high racial integration enable individuals to more freely move to higher income levels with the passing of each generation since, intuitively, there are more opportunities for the majority of the population. The model for this region yields a population with ~10% of individuals in *State 1* (lower income) and ~90% of individuals in *State 2* after 30 generations (above). Regions with high racial integration but bad schools would intuitively provide less opportunity than the prior example, but still allow for all individuals in the population to have the same opportunity to transition income levels. As a result, the model for this region yields a population with ~43% of individuals still in *State 1* and ~57% of individuals in *State 2* (also above).



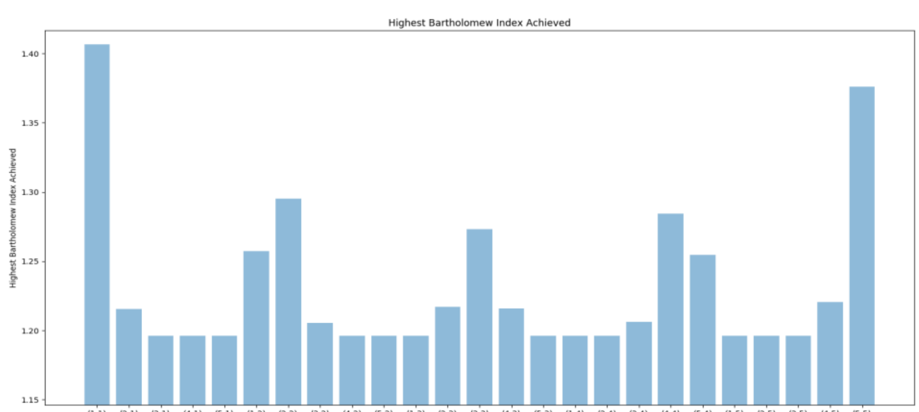
Regions with good schools and low racial integration experience higher population transitions to *Stage 1* (lower income), despite the positive effects of beneficial schooling. Roughly 57% of the population finds itself in a lower income level and ~43% in a higher income level, after 30 generations (above). The intuition behind this stems from the interaction between racial segregation and schooling, where even in places with good schooling, poor integration invariably leaves a part of the population with diminished opportunities. Correspondingly, regions with bad schools and poor racial integration, experience even more transition into the lower income level since populations are crippled by both racial inequity and poor education. In this case, the model yields that ~87% of the population will be in the lower income level after 30 generations (also above).

3.3 Sensitivity Analysis

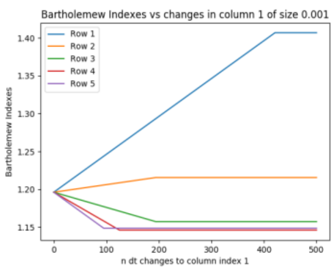
To analyze the results of the sensitivity analysis for the Markov chain representing the US. population that doesn't account for unemployment, we will have two sections, with each section analyzing which transition probabilities increase each respective social mobility indicator the most. Each section will start of by analyzing for the highest achieved social mobility index that occurred when the specific index of choice was decreased. Afterwards, each section will get the top ten percent of the highest performing indices in the Markov chain and analyze how rapidly they converge to their maximum value to determine the highest achieving indices. The first section we will be analyzing the Bartholomew index and the second section will analyze changes to the trace index.

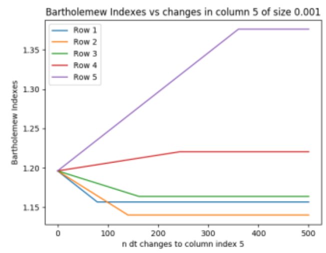
Bartholomew Indexes:

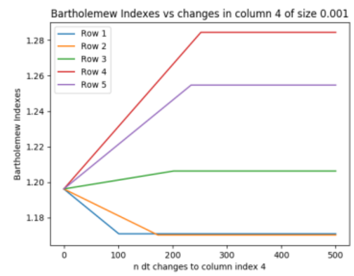
Below are the results for the highest achieved Bartholomew index when each index was reduced by a step size of .001 for 500 steps.



Furthermore, as can be seen in the graph above, the top performing transition probabilities that achieved the highest bartholomew index when reduced in stepsizes of .001 where transition probabilities with indexes of (1,1), (5,5), (2,2), (4,4), (3,3). Therefore decreasing these probabilities have the highest potential in the US of impacting social mobility. Below it is shown how fast the changing each the above transition probabilities are in achieving their respective maximum Bartholomew Index. Furthermore because each index changes linearly only two graphs are shown however all graphs are generated in the code "Sensitivity Analysis." Also the graphs below are representative of the rest of the graphs in the sense that all rates of change are the same.

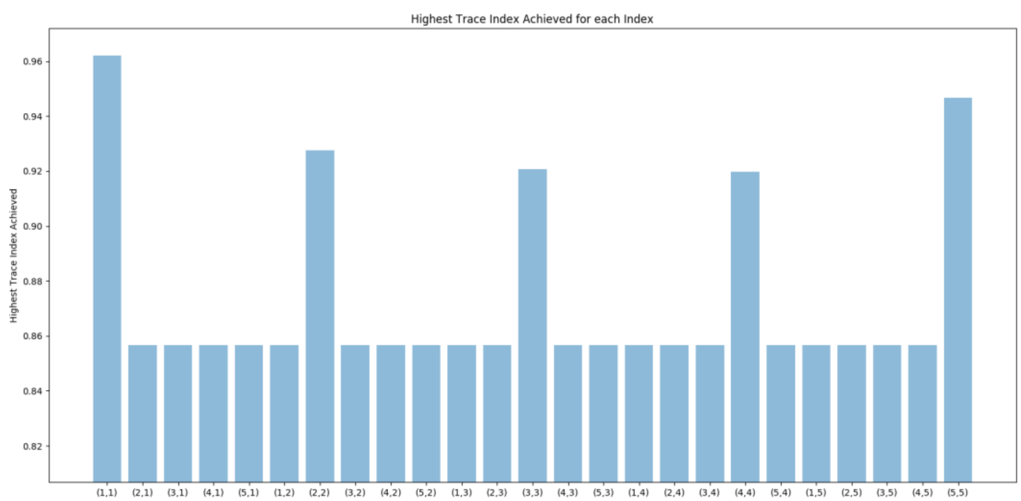




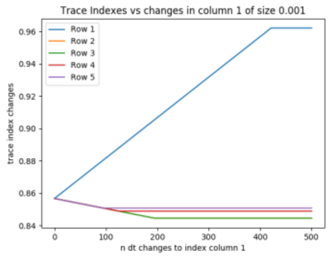


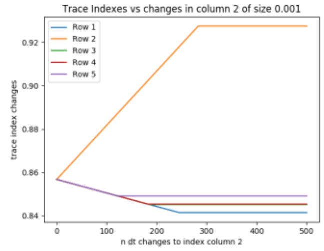
Trace indexes:

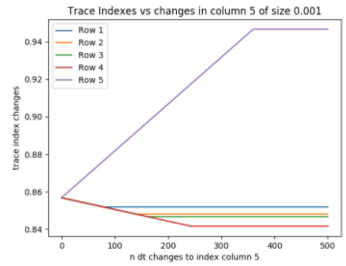
Below are the results for the highest achieved trace indexes when each transition was reduced by a step size of .001 for 500 iterations.



Furthermore, as can be seen in the graph above, the top performing transition probabilities that achieved the highest trace indexes when reduced in stepsizes of .001 where transition probabilities with indexes of (1,1), (2,1), (2,2), (3,2). Therefore changing these probabilities have the highest potential in the US of impacting social mobility. Below its shown how fast the changing each the above transition probabilities are in achieving their respective maximum trace indexes.







Therefore the transition probabilities at indices at all indexes of interest converge to their respective max trace index once they have been completely decreased.

4 Discussion

4.1 Prior Generational Wealth & Income Quintile Discussion

Our results from the study of the original generation's income on long term income and accumulated wealth provide evidence on all of the aforementioned aphorisms.

First and foremost we can all rest a bit easier because there is no evidence to suggest that the American Dream is dead and with it the middle class. We see the same exponential distribution of wealth that always exists, and while your position on that distribution *is* influenced by our ancestors, when looking at income after 5 generations it is virtually impossible to distinguish between individuals based on their ancestor's income (without looking at their bank accounts). This balancing process does occur a bit slower for those in the lowest quintile, but that bottom 20% is far from an absorbing state.

Second, there is some evidence to support that the rich get rich and the poor get poorer, but this all comes down to saving rates and investment returns. No wealth accumulates until you reach the middle quintile and have enough disposable income to actually invest, and even then the saving rates are low until you break into to the 4th or 5th quintile. In the long run everyone gets richer, but the curve is set early by the first few generations; beyond that the wealth distribution continues to expand in magnitude but stays fairly static in ratio.

While the simulation does not directly address where the 70%/90% wealth lost number comes from, it is easily explained by the transition matrix. The probability of the child (2nd generation) of a 5th quintile parent remaining in the 5th quintile is 36%, meaning 64% of children will have income lower than their parents. The third generation leaves us with only 13% of the descendants of the original high earners still in the top tier (another 11.6% drop to a lower quintile and then return to the 5th quintile). This 64%/87% is quite close the oft cited 70%/90%.

This model does make several assumptions that can impact the applicability of its findings

- Income quintiles stay constant over whole simulation
    - This will affect wealth accumulation but not transitions
- Transition probabilities and savings rates do not fluctuate as the do in the real world of long periods of time
    - This does not reflect changes in social policy to address income inequality
- Using averages for income, savings rates, and returns instead of a second random process does not show the true spread of wealth accumulation

4.2 Racial Integration & School Quality Economic Mobility Discussion

Our results when modeling racial integration and school quality as parameters of economic mobility [3] reveal that the full benefits of a good schooling system can only be realized if the population is well integrated together. While the results of the two extreme cases (good schools and high integration; bad schools and high segregation) are quite intuitive, the intermediate cases reveal that even with good schools, a largely segregated population will force more of the population to a lower income level, since by definition, segregated communities provide inequitable opportunities for one demographic over another. On the other hand, we find that even with bad schools, a region with high racial integration may possibly enable more of the population to still enter a higher income level class. This is because despite a worse school system, a population that is equitable in its opportunities provides more of a probability for success in the future.

This model also makes several assumptions, which may impact the applicability of its findings:

- There are only two income states (lower and higher)
  - This will prevent a more nuanced understanding of income transitions
- The scaling of *r* and *s*, and the corresponding indices chosen for region characterization in the probabilistic model, limits analytical capability
  - Slightly arbitrary choices for parameter magnitudes in different region types limits the model to be only directionally relevant

4.3 Sensitivity Discussion

Our results for the sensitivity analysis indicate that regardless of which metric we use for determining social mobility, either the Bartholomew Index of Trace Index, the transition probabilities that will affect the respective social mobility metric will be the same.

When using the Bartholomew Index, which weighs great changes in states more heavily than small changes in states, e.g. the transition probability from going from quintile one to quintile five is weighed more heavily than going from quintile one to quintile two, as the social mobility metric, or when using the Trace Index, which in effect treats movement between any quintiles as equal since it only takes into account how likely each generation is in staying in their father's generation, the results show that the greatest potential positive impacts for social mobility would occur by decreasing the transition probability of the kids of the people in the first quintile of society. The second greatest potential positive impact for social mobility, if using either metric mentioned above, is by decreasing the transition probability of the richest people in the US.

However while these are the greatest potential changes, as can be seen in the graphs that show the change in the social mobility metrics through a change in transition probabilities, regardless of which index that is in the trace of the transition matrix that you choose to decrease in the beginning, the social mobility index will increase at the same rate. More formally social mobility metrics increase at the same rate when you decrease any transition probability that is in the index (1,1), (2,2),(3,3),(4,4), and (5,5), however decreasing a transition probability to be less than 0 is invalid and therefore maximum potentials are only reached with the highest initial transition probabilities. Therefore, as long as you decrease the transition probabilities of someone staying in the same class as their parents for any class social mobility will increase by the same amount.

This model also makes many assumptions as well, which may impact the applicability of this model.

- Decreasing the probability of one specific quantile class into moving into the same or another quantile class will cause for all the other probabilities in the row of the markov chain to increase equally.
- Changing a specific probability of one class transitioning to another class does not affect the transition probabilities of the new class.

5 Conclusion

Especially in the recent context of more polarizing economic class issues, the questions of economic mobility and the determinants of income transitions remain urgent in today's society. As our models suggest, there are many moving parts in the economic mobility of individuals and families. We find that the American Dream is still live and well and economic mobility is not only achievable from any starting point, in the long-term it does not matter where your family started. From a perspective of wealth, the rich do indeed get richer (assuming their descendents with lower incomes don't overdraw their inheritances), but the poor get richer as well once they move up to the income brackets where surplus income makes investing more feasible.

Moreover, we find that school quality and societal integration are important and interactive determinants of future income class. As our model suggests, the capability of a population to derive the full benefits of a good schooling system is highly dependent on the ability of the community's to integrate all members of its population.

Lastly, we find from the sensitivity analysis that regardless of which metric we use to measure social mobility, the greatest impacts on social mobility occur when we change the stability of either the poorest members of society and second decrease the rate at which the richest people stay rich.

References

[1] Taylor, Chris. "70% of Rich Families Lose Their Wealth by the Second Generation". Retrieved March 31, 2018, from http://time.com/money/3925308/rich-families-lose-wealth/

[2] Jantti, Markus et al. "American Exceptionalism in a New Light: A Comparison of Intergenerational Earnings Mobility in the Nordic Countries, the United Kingdom and the United States". Retrieved March 18, 2018, from http://ftp.iza.org/dp1938.pdf

[3] Chetty, R., Hendren, N., Kline, P., & Saez, E. (2015). Economic Mobility. *Pathways: The Poverty and Inequality Report. The Stanford Center on Poverty and Inequality*.

[4] Wikimedia Foundation. "Household Incomes in the United States". Retrieved March 25, 2018, from https://en.wikipedia.org/wiki/Household_income_in_the_United_States#Household_income_and_demographics

[5] Saez, Emmanuel, and Zucman, Gabriel. "The Distribution of US Wealth, Capital Income and Returns since 1913", slide 44. Retrieved March 25, 2018, from https://eml.berkeley.edu//~saez/course/SaezZucman2014Slides.pdf

[6] Montgomery, James. "Social Mobility". Retrieved March 19, 2018, from https://www.ssc.wisc.edu/~jmontgom/socialmobility.pdf

Contributions:
Jordan Rogers: Abstract, introduction through transition matrix, effect of 0th generation's income quintile on long term income and accumulated wealth (code, description, analysis, and discussion), slides 2,3,4, and parts of assumptions/limitations and conclusion
Daniel Kang: Introduction after transition matrix, Racial Integration & School Quality Economic Mobility (code, model description, analysis, discussion, conclusion), corresponding slides
Alan Estrada: Sensitivity Analysis (code, model description, analysis, discussion and conclusion)