Problem Set 2

Part I: Modeling Households in Vietnam

The optimization problem is

$$\max_{\{c_t\}_{t=0}^{T-1},\{a_{t+1}\}_{t=0}^{T-1}} U = \sum_{t=0}^{T-1} \beta^t \frac{c_t^{1-\gamma}}{1-\gamma} \ ,$$
$$\text{s.t. } a_{t+1} = (1+r)(a_t + y_t - c_t) \ ,$$
$$y_t = \begin{cases} G_t e^{(\rho \ \log \ y_{t-1}+\epsilon_t)} & \text{if } t < t_r \\ \kappa y_{t_r-1} & \text{if } t \geq t_r \end{cases} \ ,$$
$$a_t \geq 0 \ ,$$
$$c_t > 0 \ ,$$
$$a_T = 0 \ ,$$

where $|\rho| < 1, G_t$ is age-specific average of income, $\epsilon_t \sim \mathcal{N}(0, \sigma^2)$, and the initial wealth endowment, $a_0 \geq 0$, is given.

1. **Data Preparation**

   From VHLSS 2018 Data, there will be 10 data files that can be used to solve this firm problem: muc4a, muc5a1, muc5a2, muc5b1, muc5b2, muc5b3, muc6a, muc6b, muc7, muc123a. The variables that will be used in this exercise are ***household expenditure, household consumption, household income, household fixed assets, household wealth.*** The process to construct the new data from the VHLSS 2018 Data set is shown below.

   To begin with, we used file "muc123a" as the base to merge the labels of the other files to create a new data frame for new data file. In "muc123a", filter to keep the columns of "tinh", "huyen", "xa", "diaban", "ho so", "matv", "hsize", "m1ac2", "m1ac3", "m1ac5". Then from the "muc4a" file, we used columns “m4ac7: average days work per month”, “m4ac8: average hours work per day”, “m4ac11: cash received from main job”, “ m4ac12f: other salary”, “m4ac21: cash received from secondary job”, “m4ac22f: other salary 1”, “m4ac25: other salary 2”, “m4ac27: average hours per day for secondary job” to define the average individual income (only used the income of the first and secondary job), and from this results, we compute the household income.

We run on a data segment part123a to fetch information about heads of households who are male and aged 25 years and older. It begins by defining a new column hsize, which gives us the size of the household. This is done by grouping the data on the household identifiers (tinh, huyen, xa, diaban, hoso) and looking at the max value of matv, which informs us about the number of individuals in the household. Then, it drops the data so that only rows where m1ac3 is 1 (household head) and m1ac2 is 1 (presumably male) are kept. It then takes only the male household heads who are at least 25 years old from the data set, according to the m1ac5 variable (presumably age). A list of given columns is then specified, including identifiers, household size, and personal characteristics such as gender, role, and age. The dataframe is truncated to retain only these required columns, and a final copy of filtered data is stored in a new dataframe df for further use or analysis. Next, we import income data from'muc4a.csv' and select relevant columns representing household identifiers and income variables. Then calculate each individual's income by summing their earnings from various sources. Using these individual incomes, it calculates the total household income by grouping individuals within the same household. The calculated household income is merged back into individual-level data and the main dataset ('df'), using household and individual identifiers as the merging keys. Finally, it fills any missing values with 0 to ensure data completeness for further analysis. This process integrates income data into the main dataset, enabling analysis of income patterns and relationships with other household characteristics.

Then we process household expenditure data from the 'muc5a1.csv' file. It first loads the data into a panda DataFrame and removes suitable columns, including household identifiers and expense variables. It then aggregates the data by household, calculates the total holiday food and drink expenditure by summing the 'expense bought' and 'expense self-supplied/received' columns, and adds this total as a new column called 'HH_exp1'. This total cost information is aggregated back into the original DataFrame. Finally, there are duplicate household records removed so that each household will only exist once in the final DataFrame and this is then printed. This action prepares the data for analysis by providing an aggregated set of holiday food and drink expenses by household.

About the household daily expenditure on nonfood and others, the code reads household expenditure data is processed from 'muc5b1.csv'. It first loads the data into a pandas DataFrame and only loads specific columns relating to household IDs and expenses. It groups the data by household, calculates the total expense per household by summing up matching expense columns, and adds the total as a new column to the DataFrame. To condense each house into a single one, it recombines this aggregate cost data with the original and removes duplicate entries for households. The final resulting DataFrame, then containing household-level aggregate expenses on non-food, is printed out. This derived data can then be used in other analysis as a foundation, such as with other household-level data in order to construct a general picture of a home's spending. About the household annual consumption expenditure, we use expenditure data from the 'muc5b2.csv' file. It begins by loading the data and selecting useful columns representing household identifiers and expense information. It thereafter calculates total household expenditure by grouping the data by household identifiers and adding 'expense bought' and 'expense self-supplied or received' columns. This total cost is then added back to the original data such that each household record now has its total cost. Finally, the code removes duplicate household records such that each household is only once in the dataset. The resulting part5b2 DataFrame has processed and cleaned household expenditure data ready to be analyzed. About the other spending that is considered as household expenditure, we use household expense data from the 'muc5b3.csv' file. It begins by importing the data and choosing specific columns that involve household identifiers and year expenses. It proceeds to compute the total expense of each household by grouping data based on household identifiers and then summing up the column for annual expense. The total expense is then added as a new column in the DataFrame. To ensure data cleanliness, the code removes duplicate household records and keeps only the first occurrence of each unique household. Finally, the processed DataFrame, with details on other household expenditure and an independent total expenditure column, is displayed. This prepares data for analysis and merging with other household-level data. Finally, about the household's accommodation expenditure, we use data from a CSV file 'muc7.csv' as a pandas DataFrame part7. It then selects some columns that have household identifiers and accommodation costs such as leasing costs, water, electricity, and garbage disposal. The DataFrame is filtered to retain only these chosen columns. Next, a new column 'HH_exp6' is created to store the total housing cost of each household, which is calculated by summing the water, electricity, and garbage collection costs. Finally, the modified part7 DataFrame with the selected columns and the calculated total housing cost is printed. The cleaned data can be utilized for additional analysis or merged with other datasets to obtain a complete view of household finances.

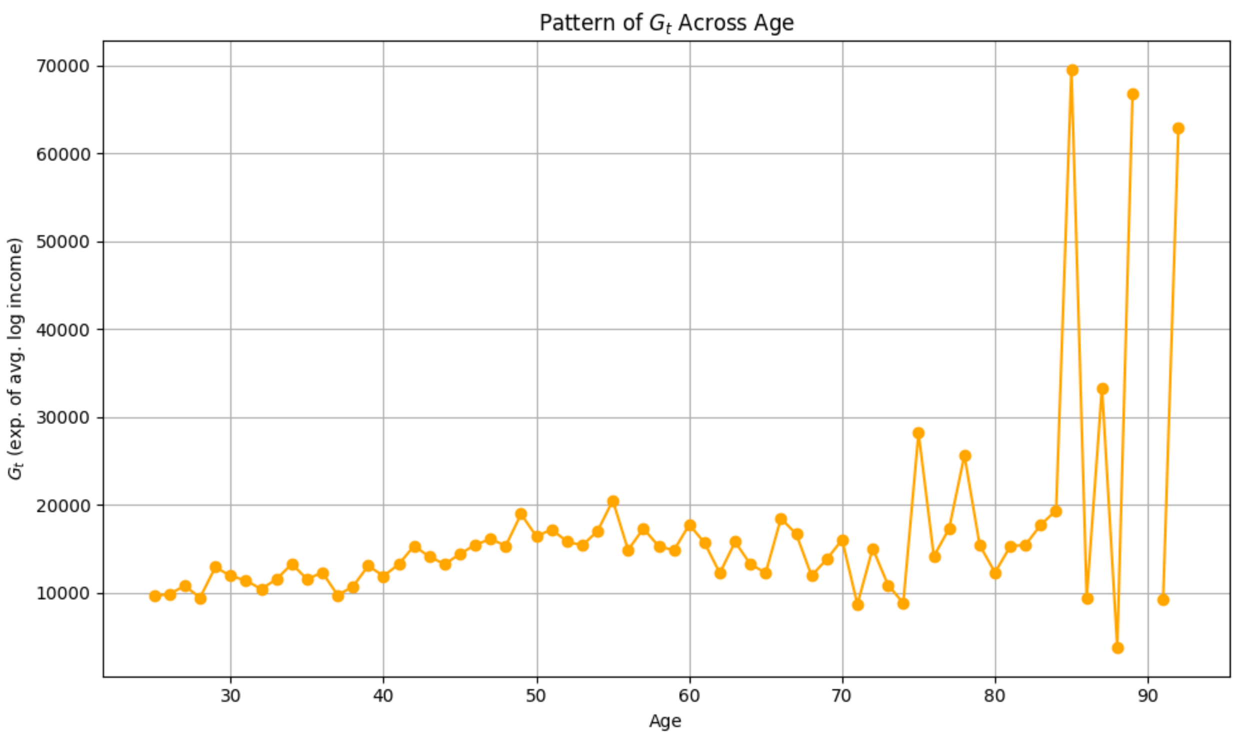
Next, we consider the household fixed assets by using asset data from a file named 'muc6a.csv' and calculating the aggregate value of these assets per household by multiplying quantity, current price, and percentage ownership. Then it reads durable goods data from 'muc6b.csv' and calculates the total value of the durable goods for each household by multiplying quantity and current price. It then merges the calculated wealth values with the main data set df, presumably containing other household variables. Finally, the code calculates total household wealth by summing the values of fixed assets and durable goods and making a new column 'HH_Wealth' in the df dataset. This provides a clear view of the wealth of every household in terms of their properties. Then we focus on processing fixed asset data of a household. It starts by importing data from "muc6a" and converting the column of ownership percentage to decimals. Then it chooses specific columns that are of concern to household identification and asset data, replacing missing values with 0. The core functionality lies in calculating total household asset wealth by multiplying quantity, value, and ownership percentage. This wealth is subsequently aggregated by household via a group-by operation and remerged with the original dataset. Redundant household records are finally eliminated to achieve uniqueness, and the resulting DataFrame of calculated and aggregated household asset wealth is illustrated. The code cleans and analyzes household fixed asset data in order to make conclusions regarding household wealth.

About the household durable goods, we use date from file "muc6b" and select suitable columns that identify households and their durable goods. It imputes missing values with 0 and calculates the sum of value of durable goods per household by multiplying value and quantity. Next, it calculates this value by household and forms a new DataFrame with the total durable goods wealth per household. The aggregated wealth is joined back into the original DataFrame and duplicate household records are dropped. The resulting DataFrame provides a complete view of household durable goods and their proportion in total household wealth. Next, we calculate and show the total household wealth. The line df['HH_Wealth'] = df['HH_wealth1'] + df['HH_wealth2'] creates a new column named 'HH_Wealth' in the dataframe df. This column represents the total wealth of each household, calculated by summing the values from two existing columns: 'HH_wealth1', representing wealth from fixed assets, and 'HH_wealth2', representing wealth from durable goods. The '+' operator performs the addition for each row (household) in the dataframe. Following this calculation, the line df simply displays the updated dataframe, allowing the user to view the newly computed 'HH_Wealth' column alongside the other household data. This is a typical way to confirm the results of a data manipulation operation within a Jupyter Notebook environment. The variables in this dataset relate to individuals' employment and income details. Specifically, m4ac6, m4ac7, and m4ac8 represent employment characteristics from the primary job: the number of months worked annually, the average number of days worked per month, and the average number of hours worked per day, respectively. Similarly, m4ac16, m4ac17, and m4ac18 capture the same information for a secondary job. Income from these jobs is recorded in several fields: m4ac11 indicates cash received from the main job, while m4ac12f captures additional salary from other sources related to the main job. For the secondary job, m4ac21 records direct cash earnings, and m4ac22f and m4ac25 include other supplementary salary components.

We rename column for clarity. The code begins by renaming the 'm1ac5' column to 'age' for clarity and computes the natural logarithm of household income ('HH_Income'), replacing any 0 values with 'NaN' to handle undefined logarithms. It then groups the data by age and calculates the average log income for each age group. To obtain average income on the original scale, it applies the exponential function to reverse the logarithm transformation. Finally, it creates a new DataFrame called 'process_df' containing the age and the corresponding average income ('Gt') for each age group, which is then displayed. Essentially, this code segment prepares, transforms, and analyzes household income data to understand the relationship between age and average income.

2. Pattern of G_t

After computing the average of log income for each age group we get Gt in the model for each age t



The figure plots the pattern of age of Gt, the exponential of mean log income, to capture the normal life-cycle trajectory of income. Gt is flat with a modest rising trend between the ages of 25 and 40, reflecting gradual income rises during early adulthood, which can be explained by human capital accumulation and career development. Between ages 40 and 60, a steady increase is observed, peaking somewhat in mid-50s, which is consistent with the empirical facts that individuals are shown to peak in income at mid-career. From age 60 and onwards, the pattern begins to decrease and become more volatile, possibly due to retirement, lower labor-force participation, or transitioning to non-wage sources of income. For the age group 80 and above, the trend is highly irregular with high peaks and dips, reflecting the impact of small sample sizes or the effects of outliers on the average. Typically, the Gt pattern displays a normal income life-cycle with highly unstable behavior in the extreme older age groups due to lack of data.

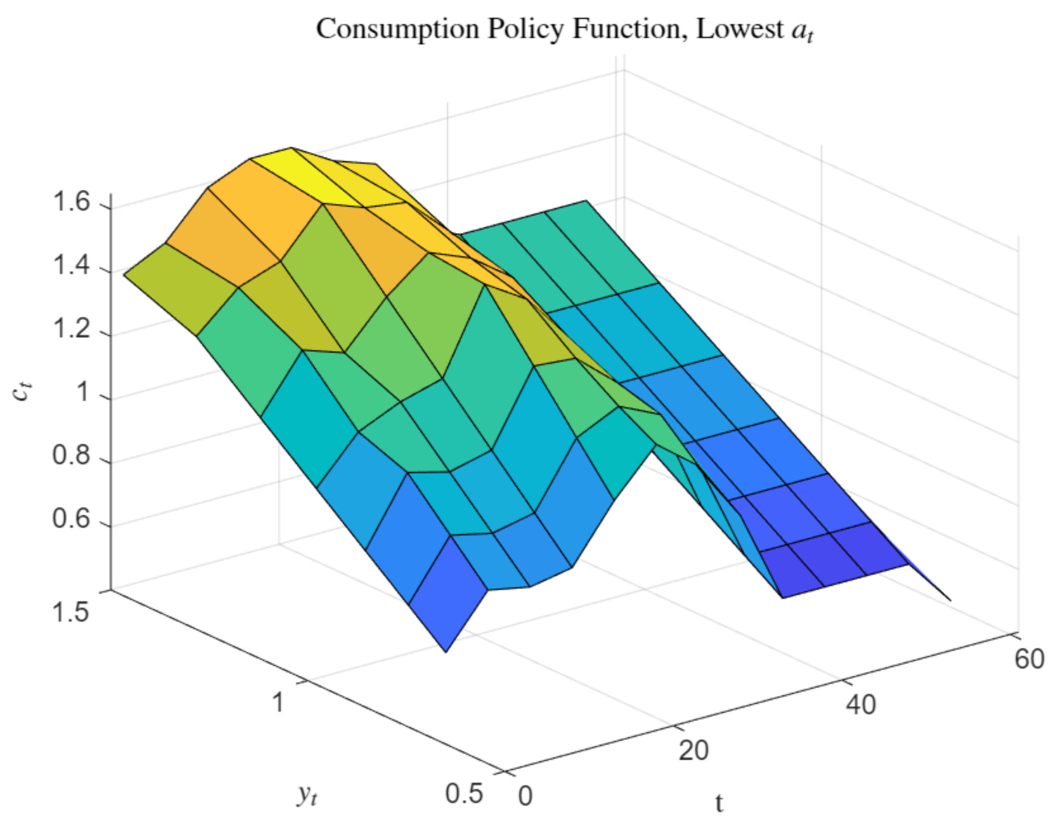
According to the graph, from the age of 45, the data fluctuates dramatically since people at this age mostly passed away; therefore, the data observations become less than the previous age. That leads to the fluctuation at the end of this Gt graph.

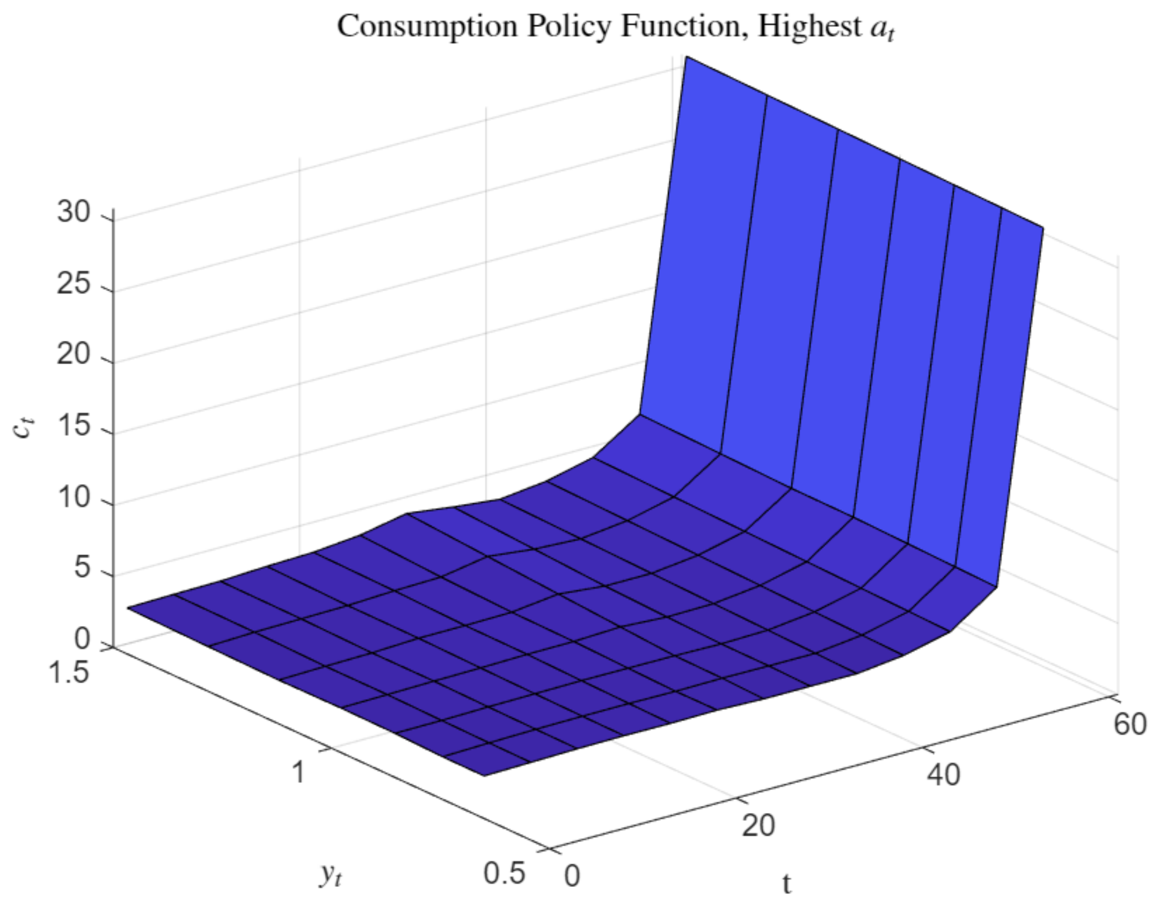
3. Policy function

These are the two consumption policy function illustrations when the agent has low and high level of wealth respectively. The x-axis is time t (age), y-axis is income level y_t, z-axis is consumption c_t.

- In the lowest level:
  - The 3D graph shows how individuals with minimal wealth level choose consumption over time and across income states
  - At younger ages and higher incomes, consumption is higher
  - Over time, consumption drops significantly
  - Tight liquidity constraints: with low assets, consumption must closely track income
  - There is less consumption smoothing, and the individual is highly sensitive to income shocks

- In the highest level:
  - This is the case for individuals with high assets
  - These individuals consume much more, especially as they age
  - Toward later life stages, there's a sharp increase in consumption, suggesting a decumulation of assets

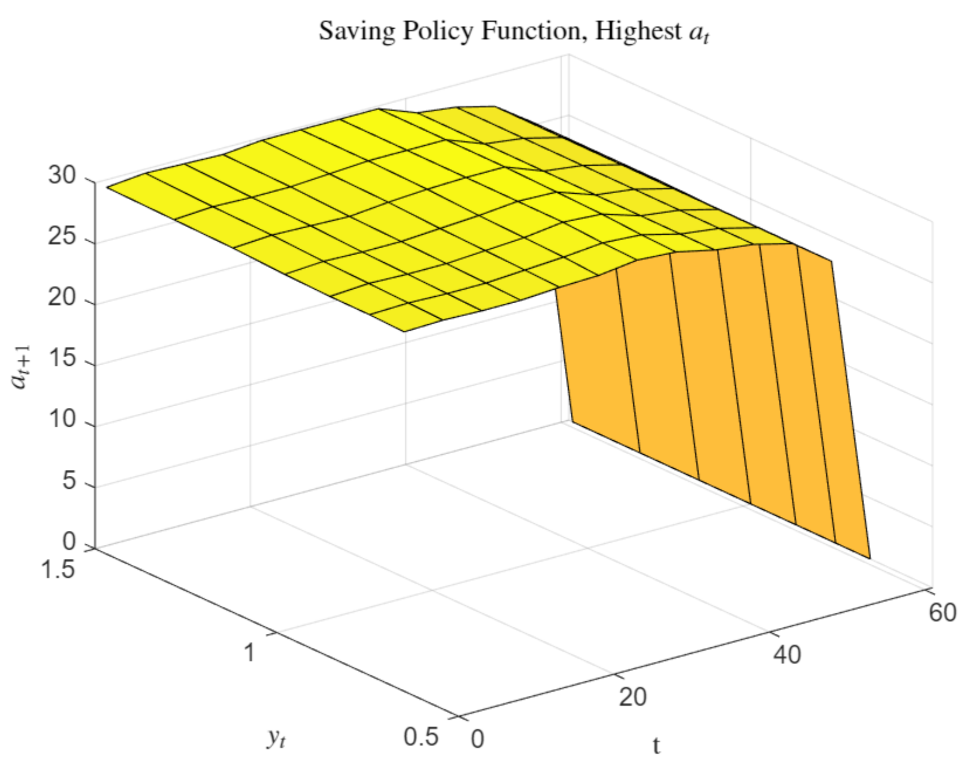




These are the saving policy function of the agent in low and high levels of wealth.

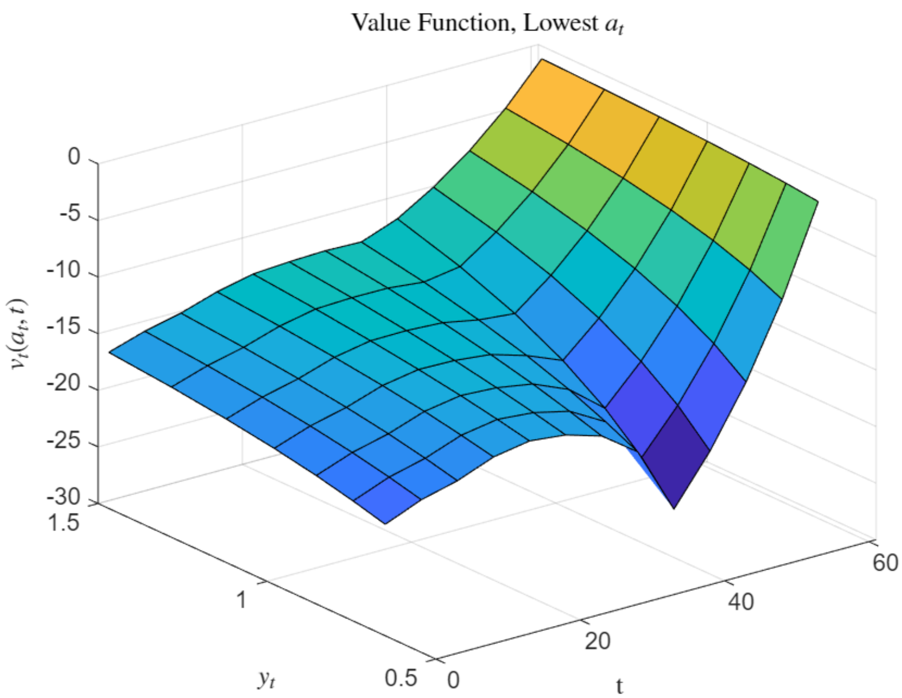
- At the lowest level
    - The savings are low or zero in most cases
    - High sensitivity to income early in life
    - It reflects precautionary saving behavior when young
- At the highest level
    - At high initial asset levels, agents continue saving heavily through most of life
    - Indicates strong bequest motive or limited incentive to increase consumption

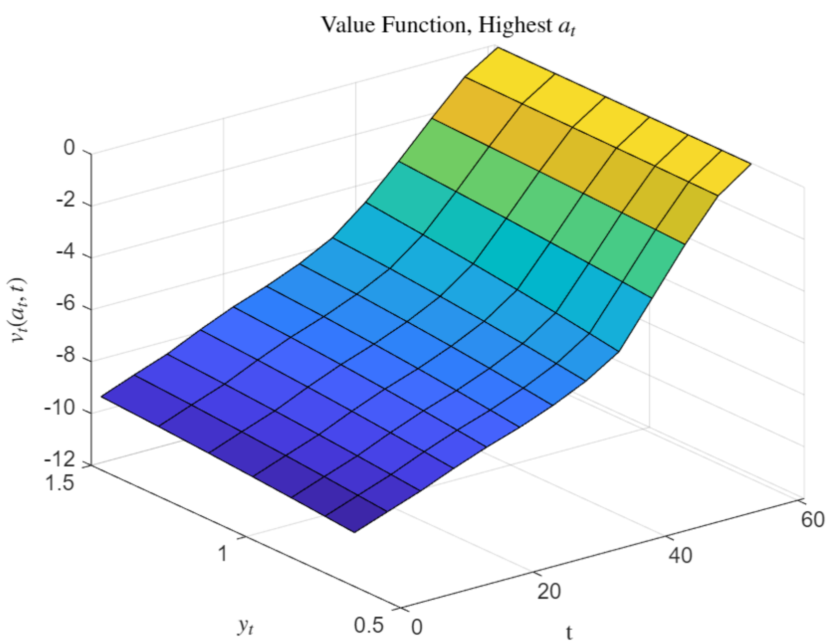
- Precautionary saving is evident for low-asset individuals and at high a_t behavior is less sensitive to shocks which indicates **self-insurance**



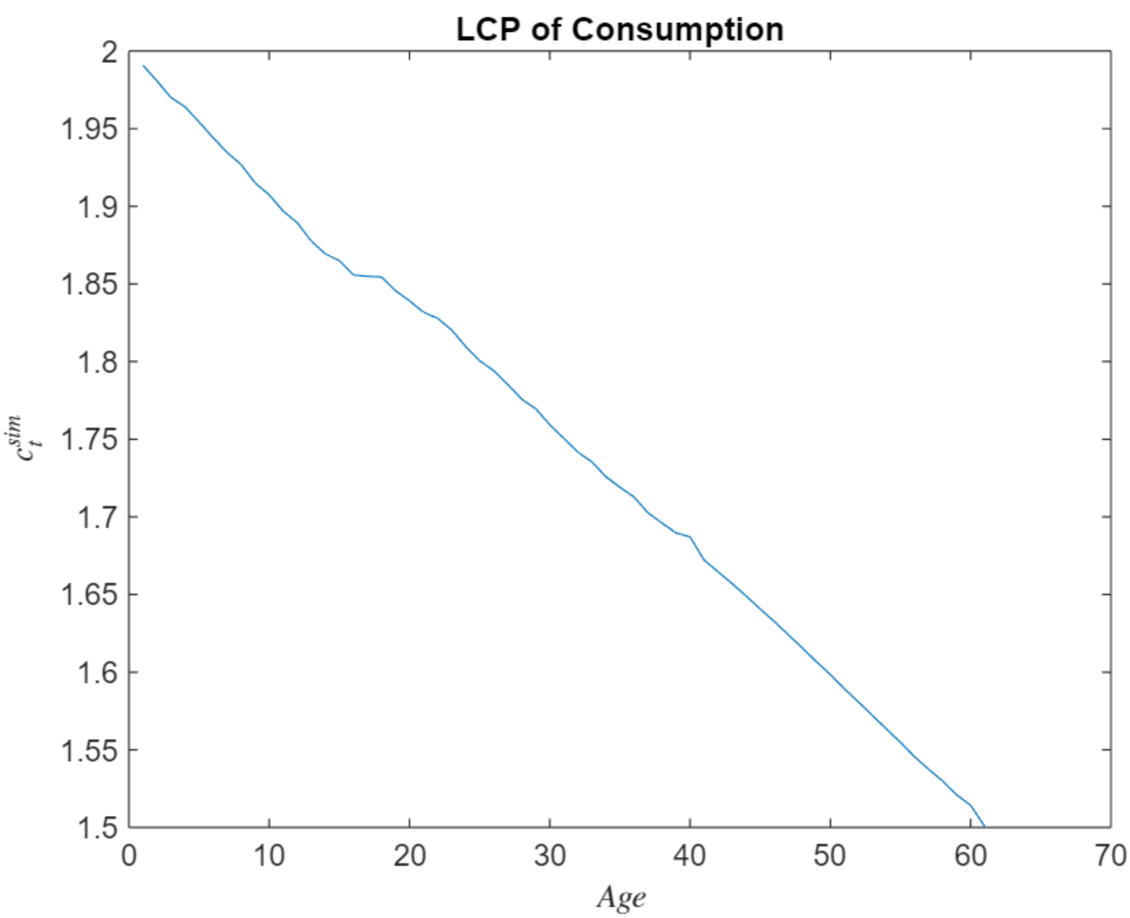
These are the value function of the agent in low and high levels of wealth. The axis includes time t, income y_t, and the value function a_{t+1}.

- At lowest level:
    - For poor agents, the value function is very negative, especially at older ages and low income, reflecting low utility due to tight budget constraints
    - The value improves with higher income
- At highest level:
    - Wealthy agents enjoy much higher utility
    - The value function increases steadily with time and income, indicating greater ability to smooth consumption and handle uncertainty
    - The curve is relatively flat over income => diminishing marginal utility of income at high wealth levels

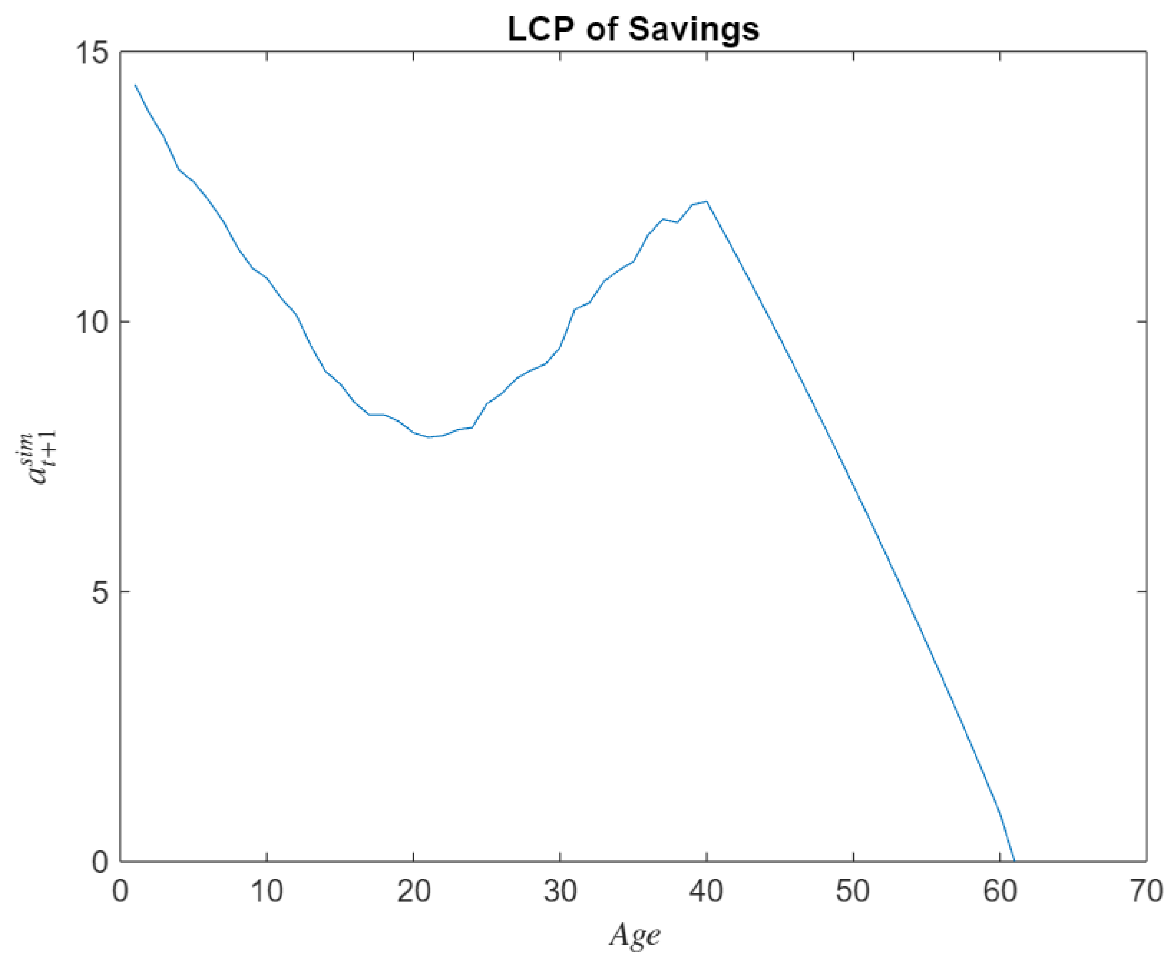


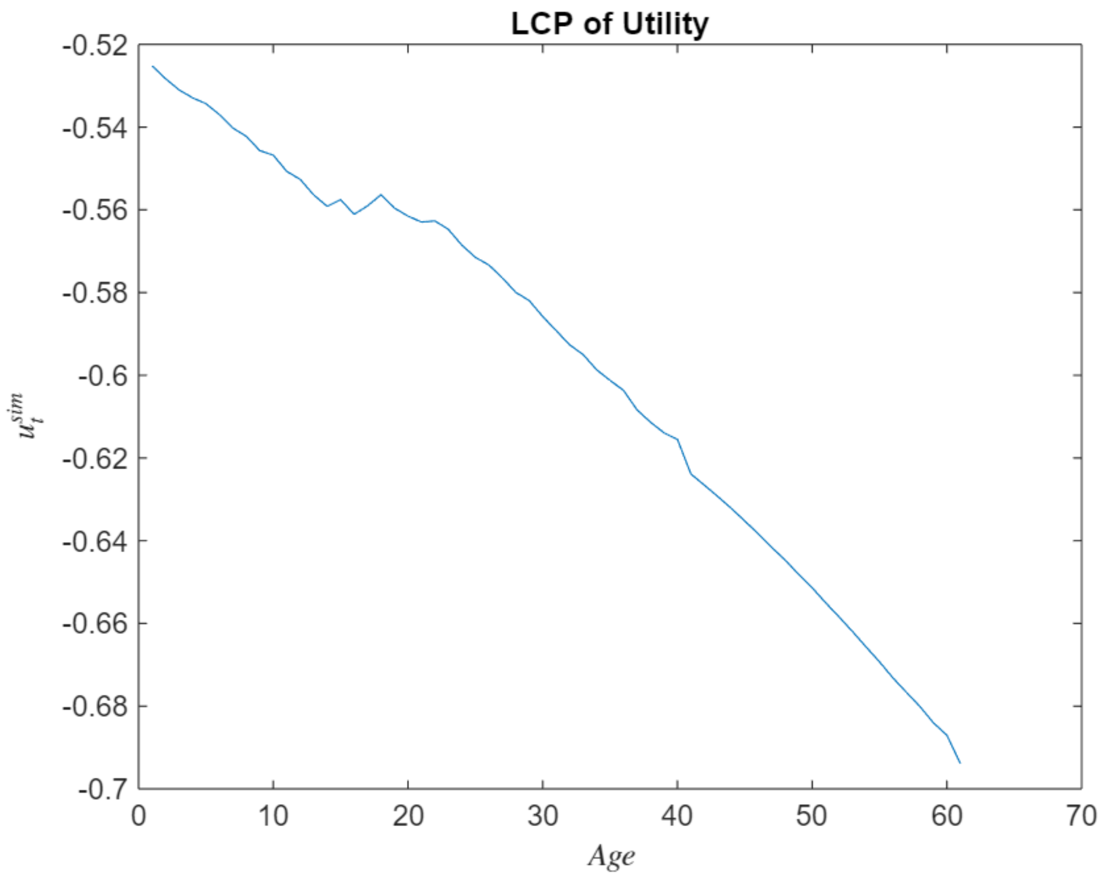


In the LCP consumption graph presents the simulated life-cycle profile of consumption, c_t^{sim}, plotted against age. The profile is monotonically decreasing in form, with consumption initially starting at a fairly high level in young ages and then steadily declining subsequently. This form suggests that model members front-load consumption, perhaps due to an intrinsic time preference (low discount factor) to present over future utility. The inability to locate hump-shaped shape - the typical characteristic of empirical data - implies that the model may be lacking features such as earnings growth, retirement, or precautionary saving motives that typically generate such curvature. The lack of a kink or slope drop at the age of retirement also implies either omission of retirement from the model or a calibration where labor earnings continue through life. Overall, the declining consumption path may be a combination of declining or level income profiles, lack of terminal utility or bequest issues, and zero borrowing constraints, and all of these lead agents to steadily reduce consumption as they age.

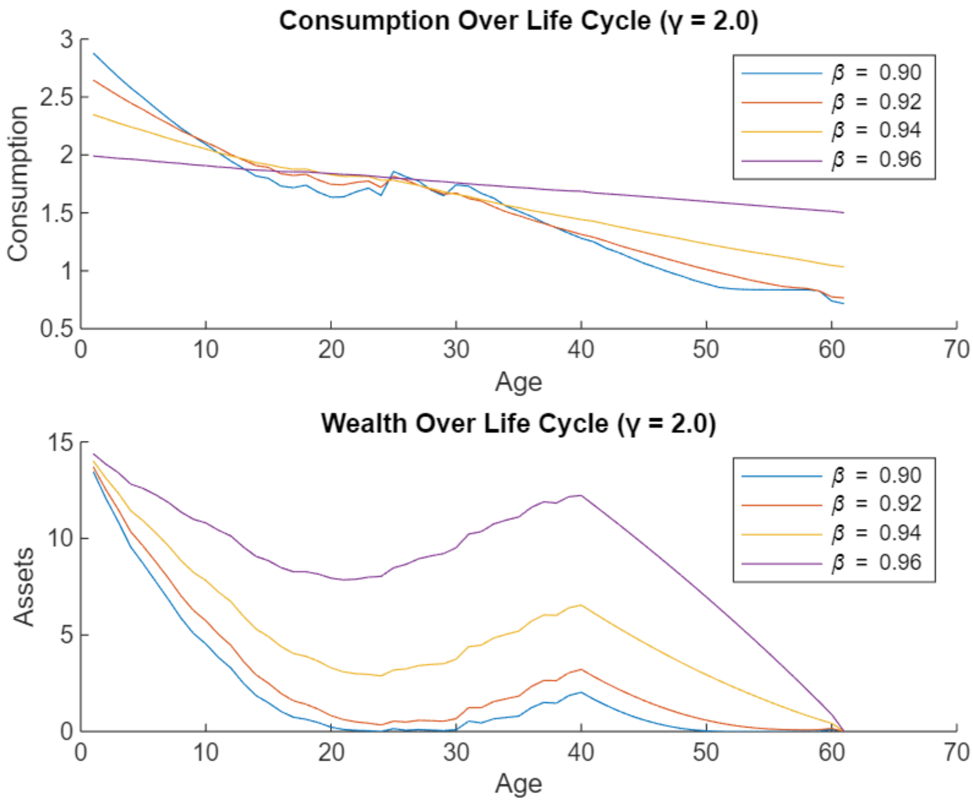


This LCP of Savings presents the simulated life-cycle profile of savings, a_{t+1}^{sim}, which exhibits a U-shape early in life, rising through midlife, and declining sharply before retirement. This profile replicates low saving early in life due to borrowing constraints or low income, followed by asset accumulation during prime working years and retirement savings decumulation late in life. The absence of precautionary saving late in life suggests minimal risk or uncertainty in the model specification.




LCP of Utility
u_t^{sim}
Age
-0.52
-0.54
-0.56
-0.58
-0.6
-0.62
-0.64
-0.66
-0.68
-0.7
0
10
20
30
40
50
60
70

Simulating the model for different values of β and γ



The figure reveals life-cycle patterns in consumption and wealth (asset) for a range of discount factor β levels with a relative risk aversion coefficient γ=2.0 fixed. Top panel shows age-specific consumption and bottom panel, age-specific accumulation and decumulation of wealth from age 0 to 65.

All lines in the consumption panel slope downwards over age, but the rate of decline depends drastically on the value of β. The lower discount rates (e.g., β=0.90) have more steep declines, showing that more impatient people consume more at the start of their life and plummet consumption significantly later. On the other hand, higher β values (β=0.96, for example) mean less life-cycle variation in consumption because patient consumers prefer more smooth consumption throughout their life. These kinds of consumers save more and consume less in earlier ages, producing flatter consumption paths.

In the wealth panel, initial wealth is similar for all possible values of β, but life-cycle paths vary considerably. Those with lower β (less patient) dissipate their wealth quickly, experiencing asset levels near zero in early adulthood and low wealth levels through most of their lives until retirement. Those with higher β accumulate and maintain higher wealth levels throughout their lives, especially in middle and late adulthood. For β=0.96, the wealth curve does exhibit a well-defined hump shape, rising in late middle age before falling towards retirement, as would be predicted if precautionary saving and consumption smoothing both exist.

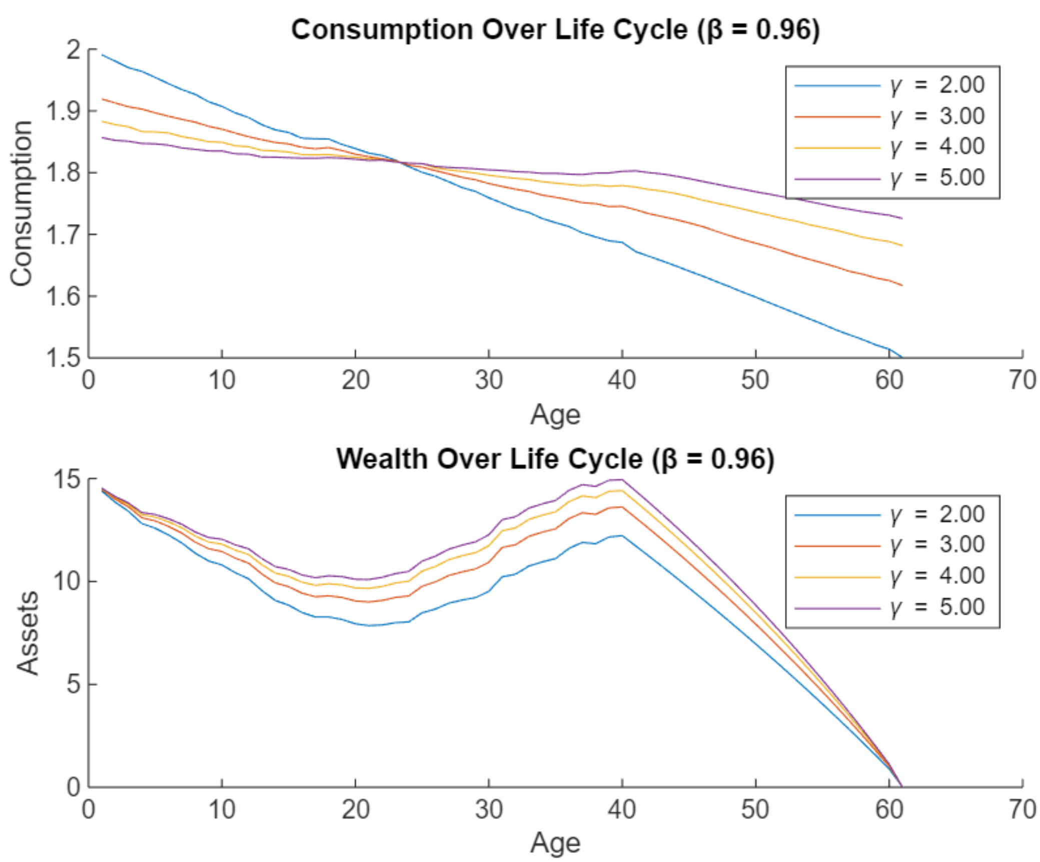
The graph as a whole shows how the degree of patience (represented by β) has a significant effect both on consumption and on saving behavior across the life cycle, with higher patience individuals better smoothing consumption and accumulating more wealth.

The figure provides life-cycle profiles in consumption and wealth (assets) for different values of the discount factor β and with a given relative risk aversion parameter γ=2.0. The upper panel is age profiles in consumption, and the lower panel is accumulation and decumulation of wealth between ages 0 and 65.

For each line on the consumption panel, there is a fairly declining slope with respect to age, but the decline rate is highly sensitive to the value of β. Declining discount rates (e.g., β=0.90) mean steeper decline, i.e., more impatient agents consume more in early life and dramatically reduce consumption in old age. Contrarily, greater values of β (e.g., β=0.96) are in accordance with smoother consumption paths since patient families prefer smoother consumption over the life cycle. Patient families have larger stocks of future time, and this generates paths of consumption that are less curvaceous in shape.

Early-life levels of wealth are equal across all β values in the panel but vary significantly over the life cycle. Lower-β (least patient) individuals spend wealth quickly, holding extremely low values of assets in early adulthood and low wealth levels until old age. In contrast, higher-β individuals accumulate and maintain more wealth over their lifetimes, especially middle and late adulthood. At β=0.96, the wealth curve is a small hump rising in late middle age before declining into retirement, as would be predicted by a precautionary saving motive and consumption smoothing behavior.

Overall, the figure illustrates how the degree of patience (measured in β) is a key determinant of life-cycle consumption and saving behavior, with patients smoothing consumption more and accumulating more wealth.



The graph shows life cycle trends in consumption and wealth (asset) for various levels of discount factor β with constant relative risk aversion coefficient γ=2.0.

Top panel graphs age-specific consumption and bottom panel, age-specific savings and dissaving of wealth from age 0 to age 65.

Lines for each of the levels of consumption decrease with age, but the gradient changes tremendously with the level of β. The lower discount rates (e.g., β=0.90) have steeper slopes, meaning that more impatient people consume more at the start of their life and significantly fall off consumption much later. Larger values of β (β=0.96, e.g.) mean less life-cycle variation in consumption since patient consumers enjoy smoother consumption throughout their life. Such consumers save more and consume less at younger ages, producing flatter consumption profiles.

For the wealth panel, initial wealth is the same for all possible values of β, whereas life-cycle profiles vary considerably. Less patient ones with lower β consume wealth quickly, with levels of assets almost zero in early adulthood and low wealth levels in most parts of their lives before retirement. They tend to build up and maintain higher wealth levels throughout their lives, especially in middle and old age. For β=0.96, the curve does have a clear hump shape, rising in late middle age before falling towards retirement, as would happen if precautionary saving and consumption smoothing happen.

The life-cycle graph shows how the degree of patience (represented by β) has significant impact both on consumption and saving behavior across the life cycle, with more patient people smoother consumption and accumulating more wealth.

The figure displays consumption life-cycle profiles and wealth (assets) profiles at different discount factor β values and relative risk aversion parameter γ=2.0. The top figure is consumption age profiles and the bottom one is accumulation and decumulation of wealth between ages 0 and 65.

For every line on the consumption panel, there is a rather declining slope in relation to age, but the speed of decline is very sensitive to the level of β. Falling

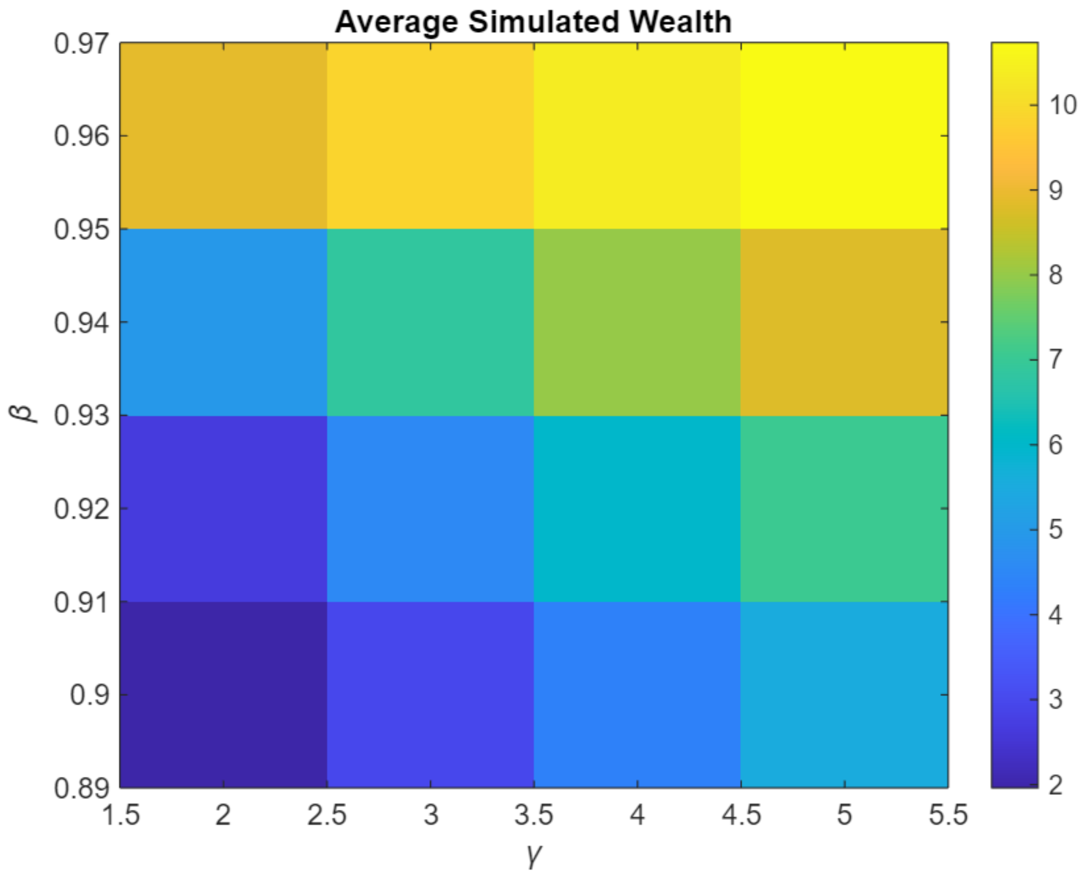
discount rates (e.g., β=0.90) mean steeper decline, i.e., impatient agents consume more at early ages and aggressively reduce consumption in advanced age. By contrast, larger values of β (say, β=0.96) are in line with smoother consumption profiles since patient families prefer smoother consumption throughout the life cycle. Patient families have larger stocks of future time, and this generates less curvaceous shapes of consumption paths. Early-life levels of wealth are the same for all values of β in the panel but vary quite strikingly across the life cycle. Lower-β (least patient) individuals waste wealth early on, with very low asset levels in early adulthood and low wealth levels for most of their lives. Those with higher βs accumulate and maintain more wealth over their lifetimes, especially middle and late adulthood. At β=0.96, the wealth curve is a small hump climbing in late middle age before declining into retirement, as would be predicted by a precautionary saving motive and consumption smoothing behavior.
More generally, the figure illustrates how the degree of patience (measured in β) is a key determinant of life-cycle consumption and saving behavior, with patients smoothing consumption more and accumulating more wealth.

The figure illustrates the impact of varying degrees of risk aversion (γ) on consumption and wealth trajectories over the life cycle with a constant discount factor (0.96). The upper panel depicts consumption paths, and the lower panel presents asset accumulation.

On the consumption panel, all trajectories depict a monotonic decline in consumption over age, but with a greater rate of decline for lower levels of risk aversion. Since γ is small (e.g., γ = 2.0), individuals have higher early-life consumption and allow consumption to fall sharply thereafter. Alternatively, with large values of γ (e.g., γ = 5.0), individuals are flatter in their consumption path, indicating higher taste for consumption smoothing due to increased risk aversion. That is, risk-averse individuals are more concerned with smoothing consumption over time.

In the wealth panel, all individuals begin life with the same levels of assets and show hump-shaped behavior: assets decline at early life, rise in middle age, and then decline sharply as individuals dissave during their retirement. However, those with higher values of γ accumulate and hold more assets over their life. For example, the $\gamma = 5.0$ path consistently has higher asset holdings than the $\gamma = 2.0$. This identifies a precautionary saving motive—risk-averse families build up more buffer stock savings to protect themselves against future income risk or consumption shock.

In conclusion, these panels demonstrate that higher risk aversion (γ) produces more smoothed consumption paths and more wealth accumulation during the life cycle, as economic theory predicts relates precautionary saving and consumption smoothing to individuals' attitudes towards risk.



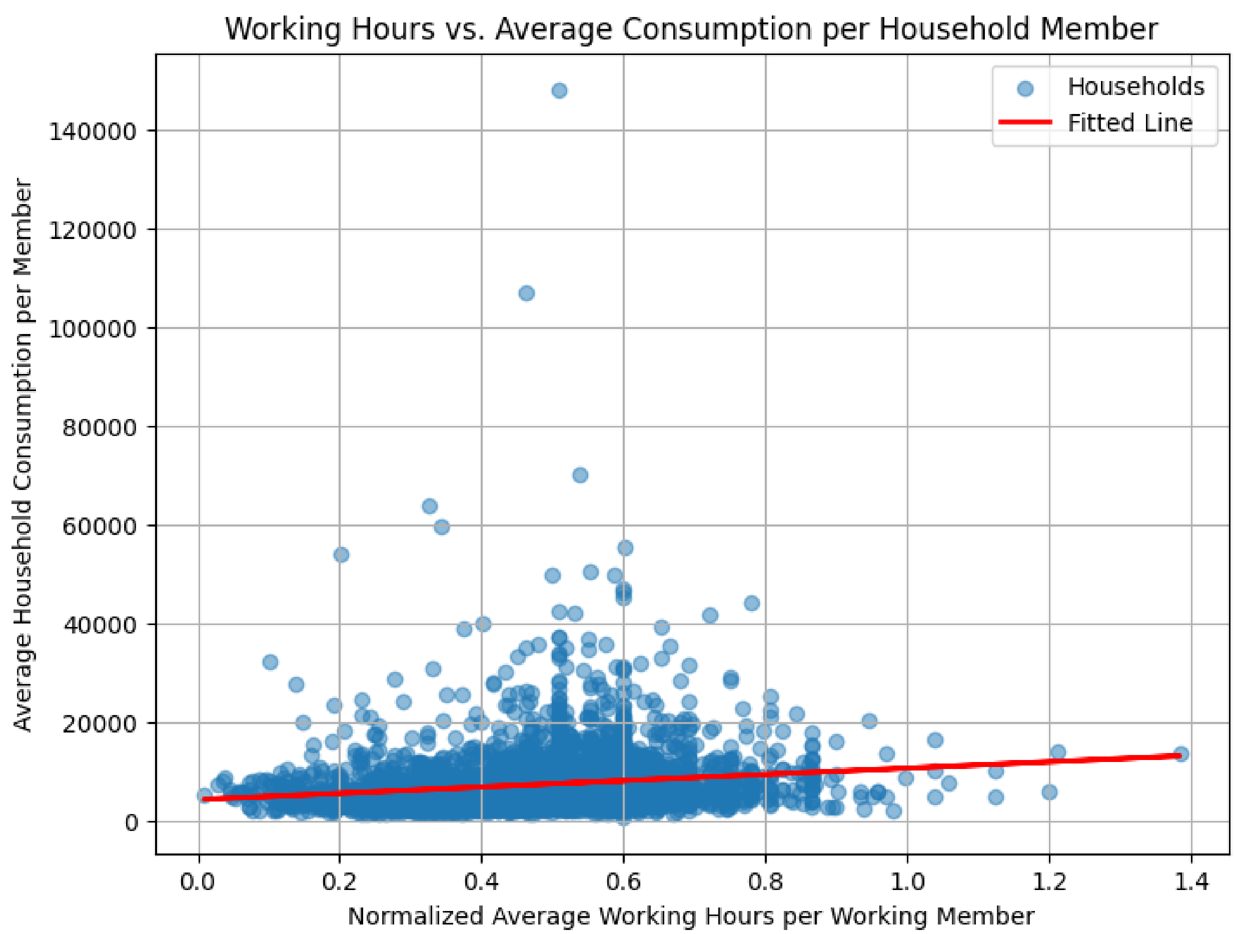
This heatmap shows how average wealth outcomes vary and the correlation between variables of this model after simulating. The lowest correlation is 2 and the highest correlation is 10 corresponding to each value of β and γ. As β **increases** (moving upward on the graph), average simulated wealth also increases. This suggests that more patient agents (who value future consumption more) tend to accumulate higher levels of wealth over time. As γ increases, average wealth also tends to increase although the effect is more gradual than for β. Higher γ implies greater risk aversion that can lead to more precautionary saving and, thus, higher wealth accumulation.

4. Extension

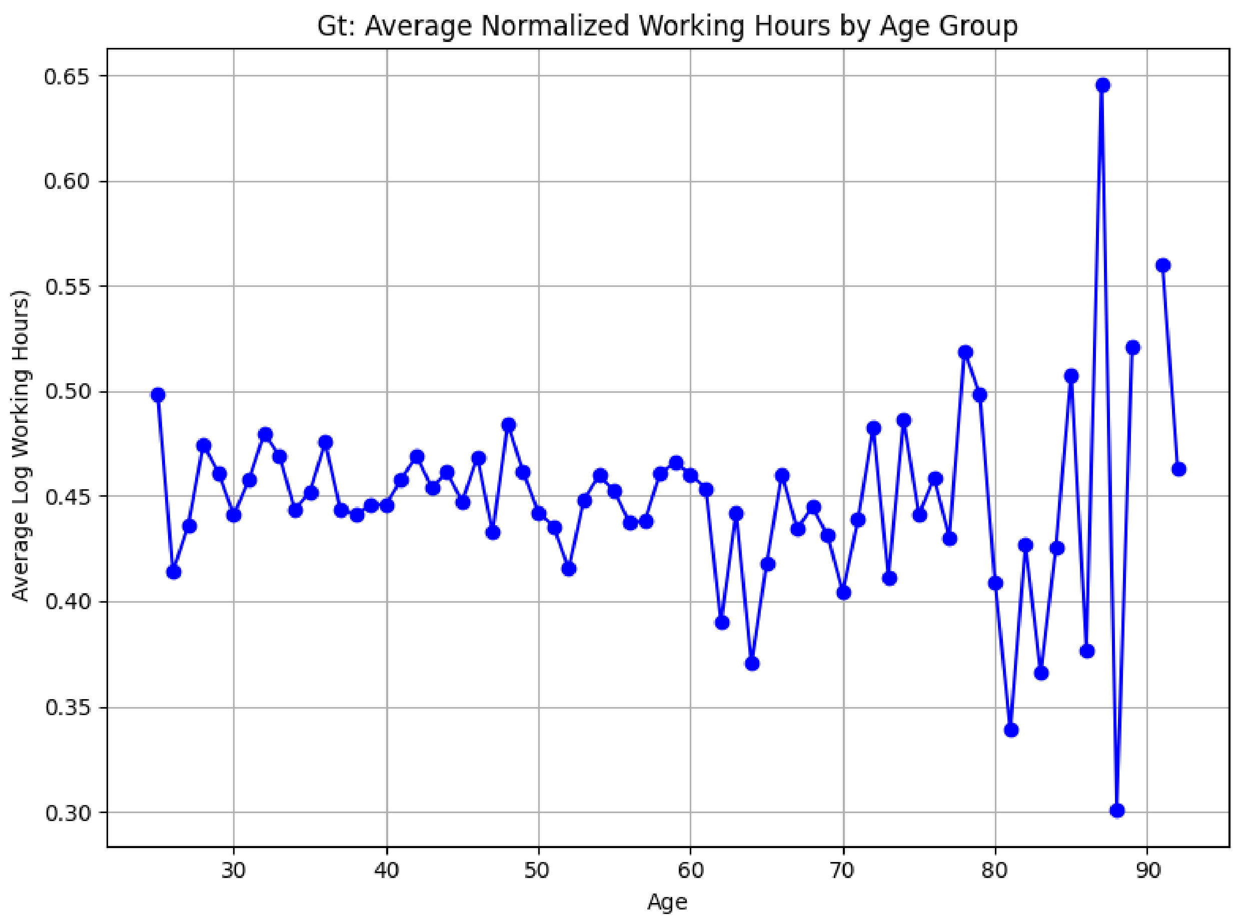
For this extension part, I will use "working hours" as an added determinant to this model. From observation, there is a problem in Vietnamese working hours is that the working hours distribution among labors is not equal, the more hours the workers commit, the higher income the workers receive, and the workers with higher income will consume more.

The hypothesis is that workers who have higher income will have higher consumption. The variables will be used is working hours of worker, and in order to compute number of working hours, we use data file "muc4a" and select columns "m4ac6, m4ac7, m4ac8" then calculate the sum of them to compute working hours of the first job and columns "m4ac16, m4ac17, m4ac18" then calculate the sum of them to compute working hours of the second job. Next, we compute the total income calculating the sum of the columns "m4ac11, m4ac12f, m4ac21, m4ac22f, m4ac25". This data contains many mistakes in collecting process, so we have to eliminate these mistakes to have the correct and clean data. Consider the conditions that working hours are more than zero and income equals to zero => remove these observations; and remove the members with income equals to zero. Then we aggregate this clean data at the household level. For each household, total working hours were adjusted by dividing by the number of active workers, yielding the average working hours per household. To standardize the working hours, normalization was performed based on the assumption of each individual has 260 working days per year and 16 active hours per day (which is assumed by letting 24 hours minus 8 hours per day for sleeping). In terms of consumption, total household expenditure (HH_exp) was calculated and averaged across all household members to obtain per capita consumption. An OLS regression analysis was conducted, resulting in an R-squared value of 0.026, indicating that only 2.6% of the variance in the dependent variable could be explained by the model, which indicates a significantly low explanatory variable.

Below is the scatter plot graph of working hours versus average consumption per household member. Each point represents a household, while the red line indicates the fitted linear regression line. This shows a dense clustering of households with normalized working hours between 0.2 and 0.8, and consumption levels predominantly below 20,000. A small number of households exhibit exceptionally high consumption values, indicating the presence of significant outliers. The low R-squared value of 0.026 obtained from the regression analysis, indicating that only 2.6% of the variation in consumption is explained by working hours.



Finally, we compute the G_t value (average normalized working hours by age group) which is illustrated as below. The same with the G_t graph in the initial model, the extension one also has a dramatic fluctuation from the age of 80 to 90 due to the lack of observations in reality.



The new recursive formulation

In order to analyze the extension version of this model, based on the initial formulation, we have the new formulation as below

$$V_t(a_t, t, y_t) = \max_{c_t, n_t, a_{t+1}} \frac{c_t^{1-\sigma}}{1-\sigma} + \gamma \frac{(1-n_t)^{1+\frac{1}{v}}}{1+\frac{1}{v}} + \beta V_{t+1}(a_{t+1}, t+1, y_{t+1})$$

$$s.t.\, a_{t+1} = (1+r)(a_t + y_t - c_t),$$

$$y_t = \begin{cases} G_t n_t e^{(\rho log y_{t-1} + e_t)} & if\ t < t_r \\ \kappa y_{t_r - 1} & if\ t \ \geq t_r \end{cases},$$

$$a_t \geq 0,$$

$$a_0 \geq 0\ given,$$

$$a_T = 0,$$

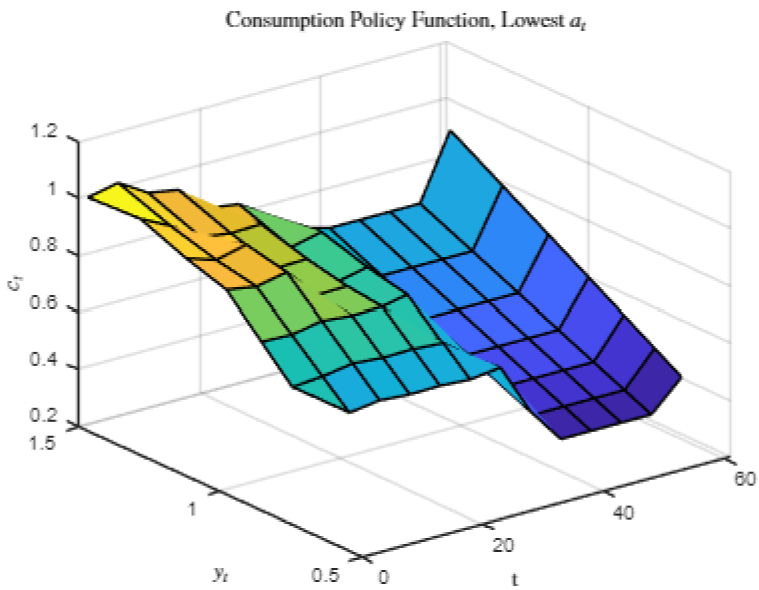
$$c_t > 0,$$

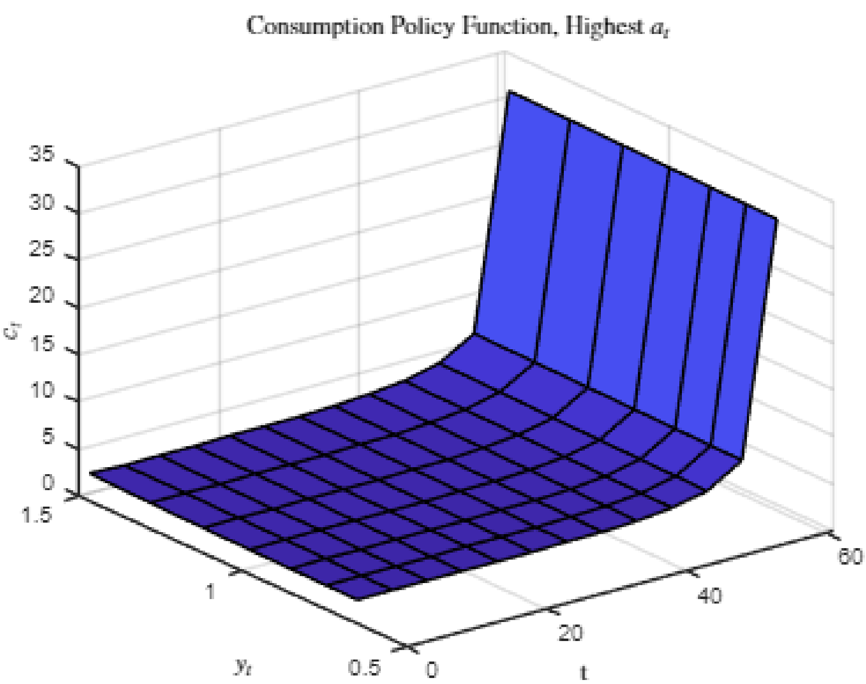
$$1 = n_t + l_t$$

Policy function

To begin with, below are two consumption policy function graphs that show how consumption c_t varies with income y_t, and working hours t.

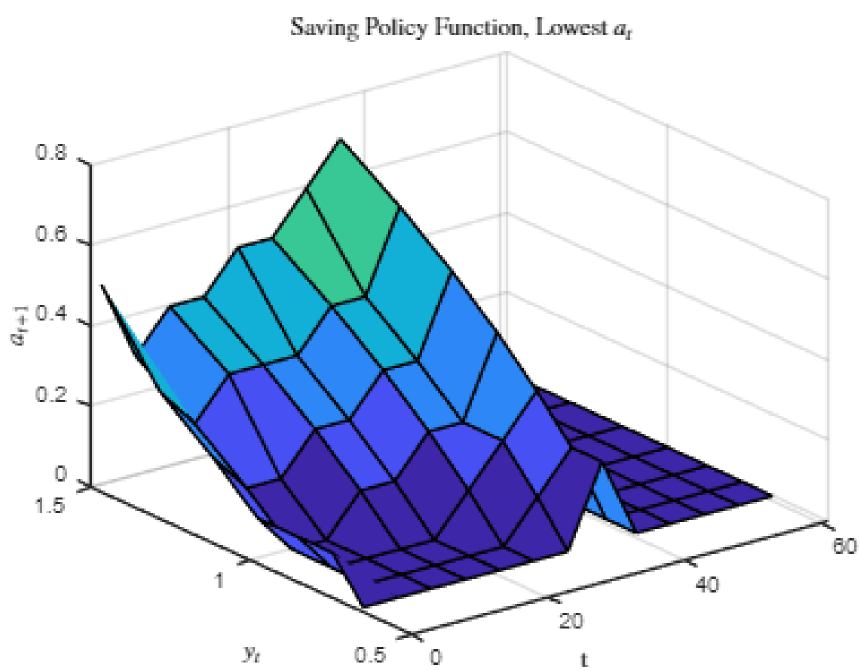
- At the lowest level:
    - Consumption declines over increasing working hours
    - Income increases, consumption tends to increase, but the relationship is not linear
- At the highest level:
    - Consumption values are higher as working hours increase
    - Consumption increases rapidly with more working hours since optimizing labor boost strong financial positions
    - Relatively flat behavior across y_t, especially at higher t

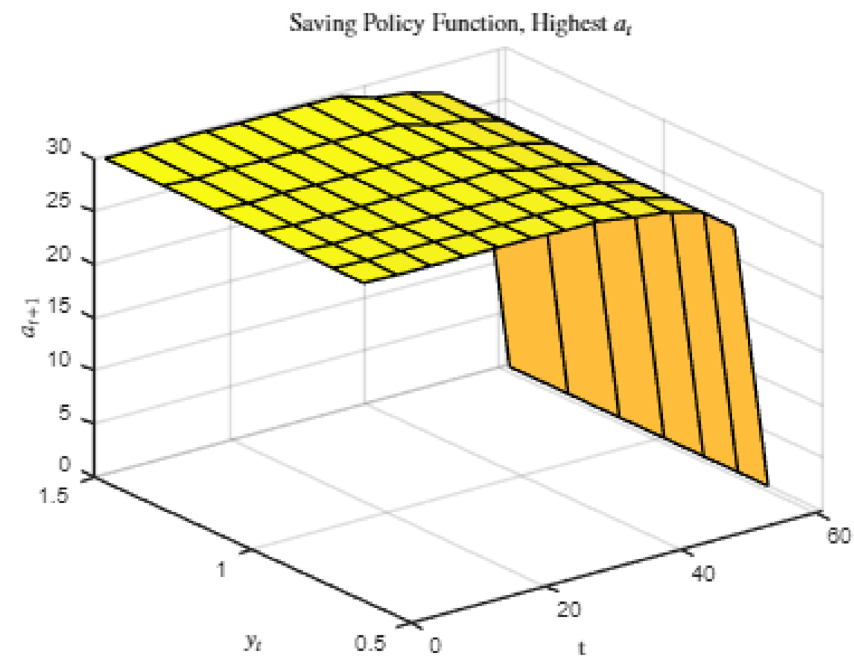

Consumption Policy Function, Lowest a_t
1.2
1
0.8
0.6
0.4
0.2
c_t
1.5
1
0.5
y_t
0
20
40
60
t


Consumption Policy Function, Highest a_t
35
30
25
20
15
10
5
0
c_t
1.5
1
0.5
y_t
0
20
40
60
t

Below are two saving policy function graphs

- At the lowest level:
    - Limited saving capacity, especially early in life
    - The graph shows nonlinear and irregular patterns
    - Low across the board when t is small (early life) and income is low, but increase slightly with rising y_t and t
- At the highest level:
    - Very high and stable levels
    - As t approaches retirement (around t=50 - 60), savings drop sharply, indicating the beginning of dissaving since in the age of retirement, people have no or less working hours, which leads to less income and less saving. Even, some individuals have no income after retirement.

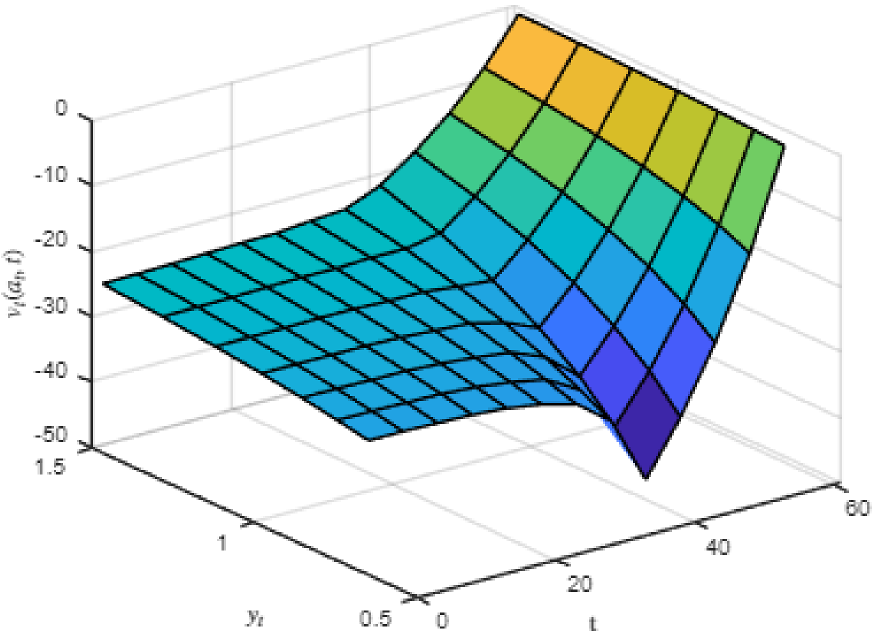




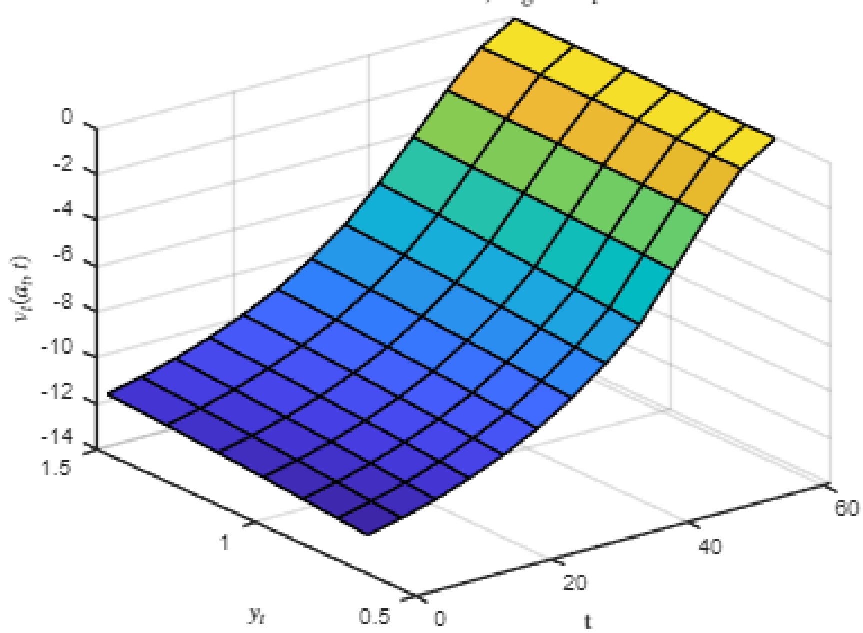
Below are value function graphs:

- At the lowest level:
    - Lifetime utility (value) is much lower for individuals with low asset holdings
    - Value increases with both income and working hours, but drops toward retirement
- At the highest level:
    - Value function is significantly higher, and increases steadily over time and with income => stable and optimized
    - At lower value of y_t, value remains high

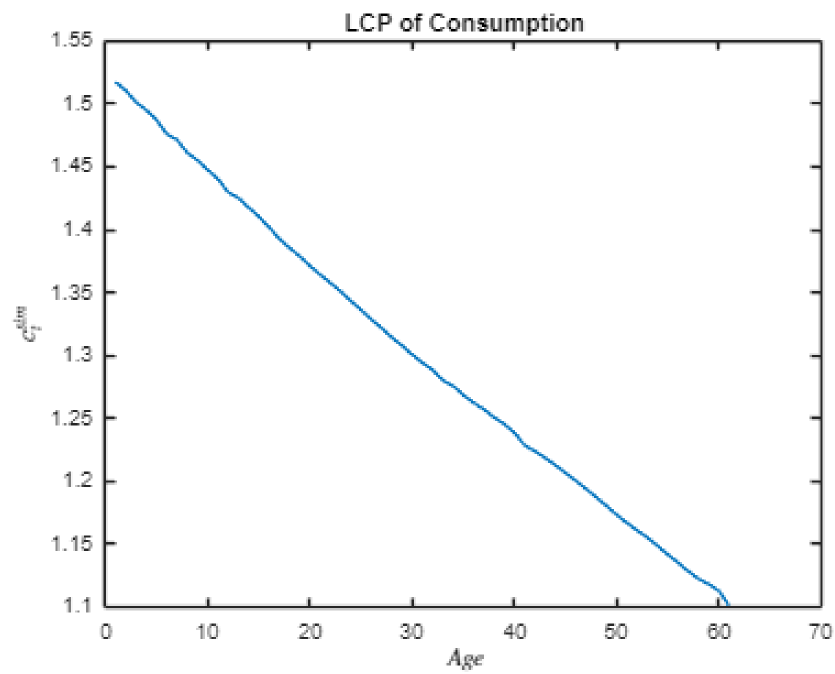
Value Function, Lowest a_t



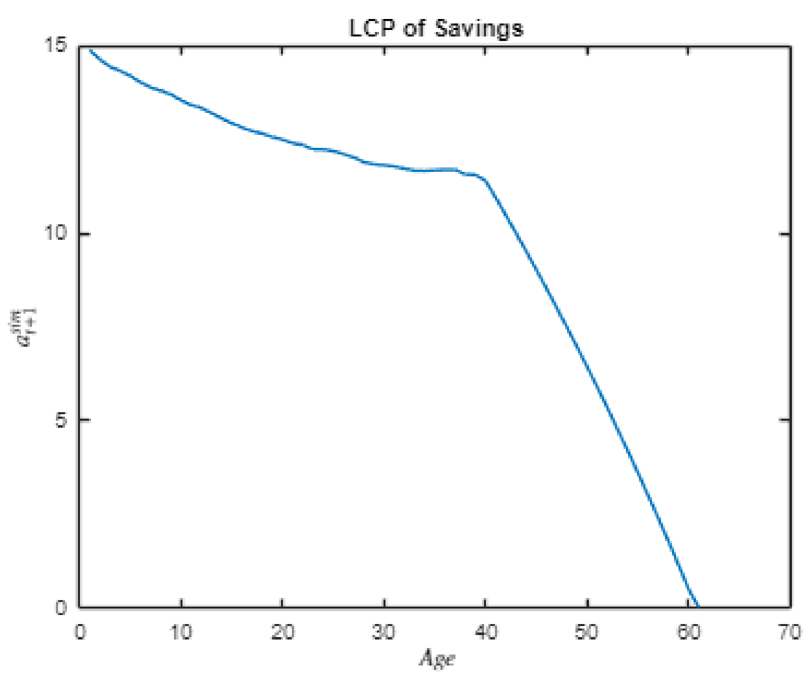
Value Function, Highest a_t



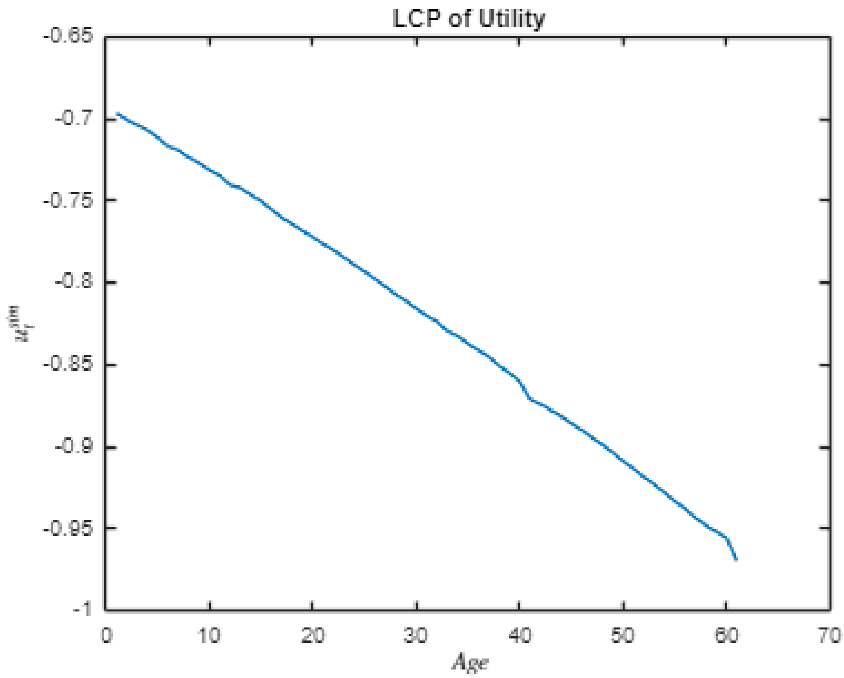
In the LCP of Consumption, consumption decreases gradually with age. More specifically,



The LCP of Savings graph shows that savings are high in youth, peak around mid-life, and decline steeply thereafter, reaching zero around retirement age.

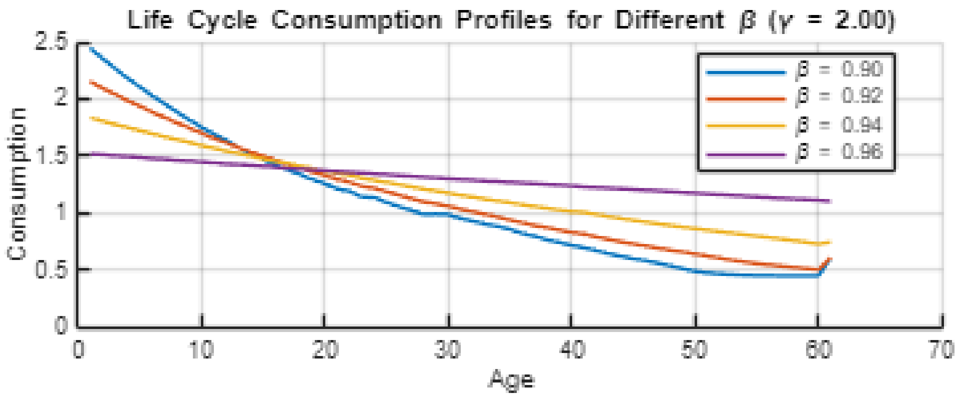


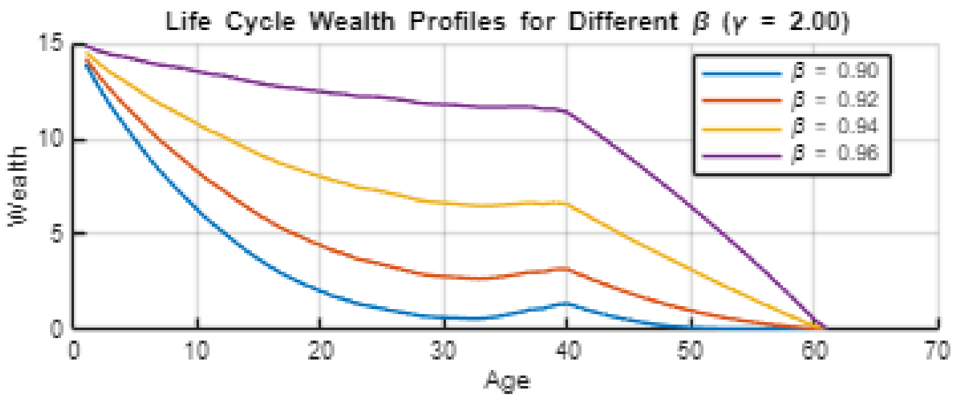
The LCP of Utility shows the negative and decline steadily with age which indicates decreasing marginal utility of consumption with age.

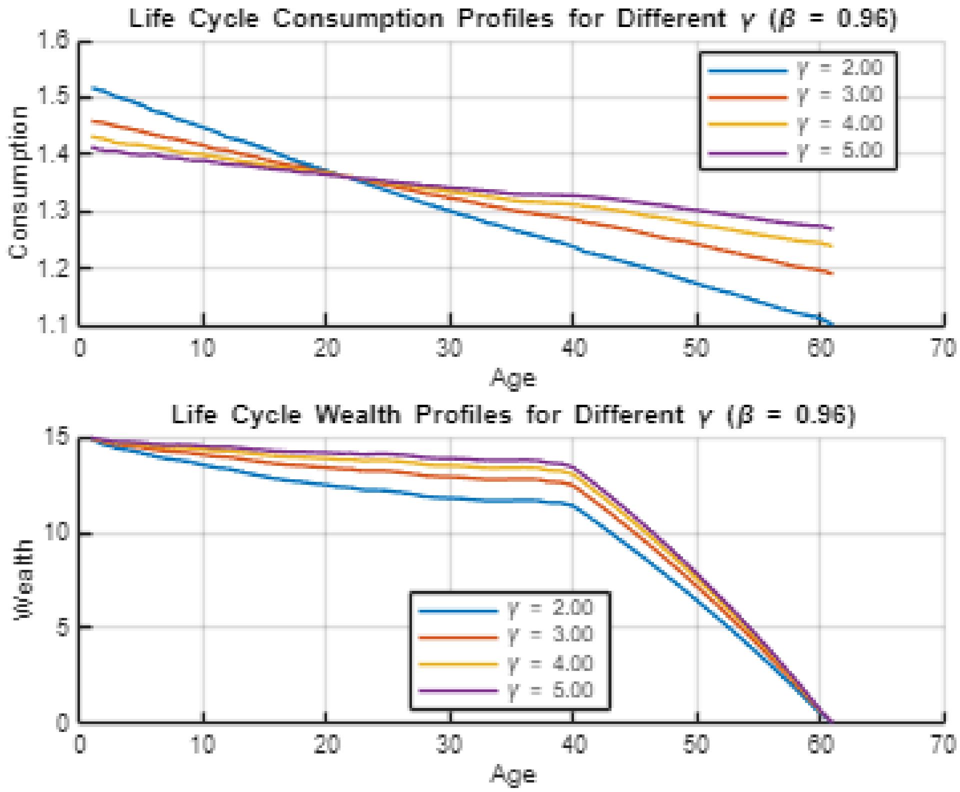


In general, the model illustrates that at the young age, the individuals have more working hours will consume more, and by the time, the consumption level declines as they reach retirement age since after retirement, they have less or even no income.

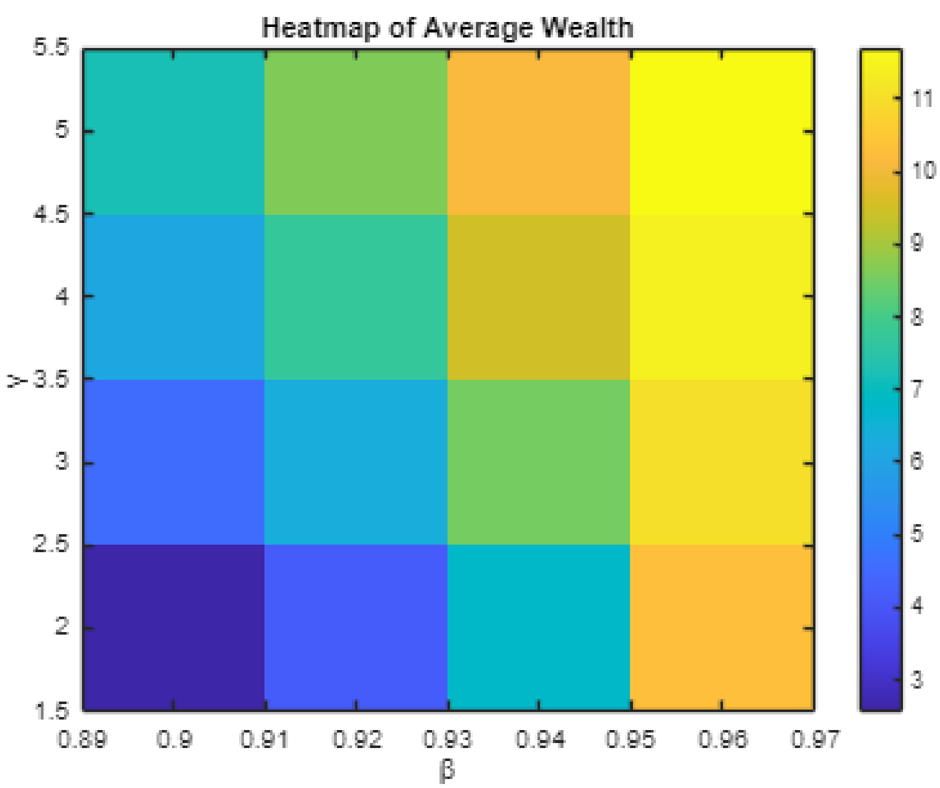
Then we consider the life cycle consumption profile over different values of β and γ. Overall, the high β indicates the smoothness consumption over life cycle. In contrast, the low β indicates the decline.


Life Cycle Consumption Profiles for Different β (γ = 2.00)
Consumption
Age
β = 0.90
β = 0.92
β = 0.94
β = 0.96
0
0.5
1
1.5
2
2.5
0
10
20
30
40
50
60
70


Life Cycle Wealth Profiles for Different β (γ = 2.00)
Wealth
Age
β = 0.90
β = 0.92
β = 0.94
β = 0.96
0
5
10
15
0
10
20
30
40
50
60
70



Life Cycle Consumption Profiles for Different γ (β = 0.96)
Consumption
Age
γ = 2.00
γ = 3.00
γ = 4.00
γ = 5.00
Life Cycle Wealth Profiles for Different γ (β = 0.96)
Wealth
Age
γ = 2.00
γ = 3.00
γ = 4.00
γ = 5.00



Heatmap of Average Wealth
γ
β

This model (using the same parameters as the initial model) and the life cycle profile aligns with the hypothesis.

The comparison of the real-world and simulated data exposes very significant variations in life cycle consumption and wealth. Consistent with conventional life-cycle models, the simulated data reveals a smoother, more theoretical drop in both consumption and wealth over time. By contrast, the actual data shows more variation - especially in spending patterns - probably reflecting real-world shocks, heterogeneity, and imperfect markets. While the real data shows wider dispersion, the simulated data exhibits typically higher and more consistent average consumption-to-income and wealth-to-income ratios. Besides, the age group consumption variance is much greater in the simulated data to show more unexplained unpredictability in spending behavior. These observations indicate the suitability of the simulated model in capturing the entirety of real economic behavior.

```
           consumption_to_income  wealth_to_income  consumption_variance
age_group
20-24                        NaN               NaN                   NaN
25-29                   2.898996          3.382270          2.199277e+08
30-34                   3.161397          7.143595          3.291132e+08
35-39                   3.862334          7.226388          2.096600e+08
40-44                   3.829098          7.529010          3.057302e+08
45-49                   3.205159          8.504305          4.145537e+08
50-54                   3.810067          8.693933          9.561165e+08
55-59                   3.585104          6.276802          6.534557e+08
60-64                   3.418057          8.580961          9.980919e+08
65-69                   3.153631         10.544835          5.419577e+08
70-74                   3.084113          4.527235          3.821556e+08
75-79                   2.465252          3.184701          5.341640e+08
<ipython-input-2-7d3925ae793c>:19: FutureWarning: The default of observed=False is deprecated and will
  grouped_stats = df.groupby("age_group").agg({
```

The combined data of consumption-to-income ratio, wealth-to-income ratio, and consumption dispersion by age have varying life-cycle patterns. The consumption-to-income ratio increases from the age group 25 - 29, peaks between ages 35 - 44, and declines gradually with advanced age, the lowest being in the age group 75 -79. This pattern suggests that individuals will consume a larger share of their income during their years of maximum working life, and decrease their consumption ever more as they approach and enter retirement. On the contrary, the wealth-to-income ratio climbs sharply with age, particularly between 30 and 69 years, and peaks for the 65 - 69 group. This accord with conventional savings conduct, as individuals save throughout working life for old age. Also importantly, the ratio declines after age 70, marking decumulation of assets in later life. The variance in consumption is strongest in the 50 - 64 age group, as would be expected for increased heterogeneity of financial behavior in late career and early retirement ages. Variance then declines across higher ages, as would be expected for more homogeneous consumption behavior and maybe with fixed incomes or constrained resources in retirement. In combination, these patterns validate theoretical life-cycle model predictions of consumption smoothing, wealth accumulation, and eventual depletion in subsequent years.

From my observations, **this data set is so noisy**, perhaps it has too many mistakes in the collecting process since it is aggregated data from many resources, so we cannot have the exact conclusion. Therefore, it is difficult to have the best fit parameters in model in order to have the exact conclusion for all hypotheses. Via preparing data to apply to the model, it is obvious that it is really hard to have the finalized data and requires many steps to clean data to have the most suitable data. Hence, if we want to have the best fit, we need to combine with another suitable data set or find another superior and clean data set.