Problem Set 2: Modeling Households in Vietnam

Do Binh Kiet - 220251

May 19, 2025

Contents

1 Data selection

1.1 Variable Selection from VHLSS 2008

- **Income (y_t):**
  - m4ca11: Total wage/salary received in the past 12 months from the main job.
  - m4ac12f: Other wage payments received from the main job aside from salary in the past 12 months.
  - m4ac22f: Total wage/salary received in the past 12 months from the second job.

  These variables are sourced from the dataset `muc4a.csv`. The sum represents the household's gross labor income from its primary and secondary employment activities over the year.

- **Consumption (C_t):** We use total household expenditure as a proxy for consumption. The variable selected is:
  - hhex1nom: Total nominal household expenditure in 2008.

  This variable is sourced from the dataset `hhexpe08.csv`. It provides a comprehensive measure of the household's spending across various categories.

- **Wealth (a_t):** Household wealth is approximated by the value of physical assets owned by the household. We construct this measure as the sum of:
  - m6ac6: Value of fixed assets at current prices (2008).
  - m6bc6: Value of durable appliances at current prices (2008).

  These variables are likely sourced from a dataset related to assets, potentially named `muc6ab.csv` or similar (please verify the exact filename). This measure captures tangible wealth but may exclude financial assets or liabilities, which are often harder to measure accurately in surveys.

1.2 Calculating the Age-Specific Income Profile (G_t)

The model specifies an age-dependent deterministic component of income, G_t. This profile is estimated directly from the VHLSS data using the following steps:

1. **Identify Household Head:** Information on the household head's age and gender is extracted, typically found in demographic files like `muc123a.csv`.

2. **Filter Sample:** As per the problem set instructions, the sample is restricted to households where the head is male.

3. **Merge Data:** The constructed household income variable (sum of m4ca11, m4ac12f, m4ac22f) is merged with the age of the male household head.

4. **Calculate Log Income:** For each household in the filtered sample, the natural logarithm of the total income is computed. Note: Households with zero or non-positive reported income might require specific treatment (e.g., exclusion or adding a small constant before taking the logarithm) to avoid computational issues.

5. **Average Log Income by Age:** Households are grouped by the age of the male head (t). Within each age group, the average of the log incomes is calculated. Let this average log income at age t be denoted by $\bar{g}_t$.

6. **Compute G_t:** The age-specific income profile G_t is obtained by exponentiating the average log income for each age group:
$$G_t = \exp(\bar{g}_t)$$
   This G_t represents the average income level for a household with a male head of age t, forming the baseline income level around which the persistent and stochastic shocks operate in the model's income process: $y_t = G_t e^{(\rho \log y_{t-1} + \epsilon_t)}$.

1.3 G_t discussion

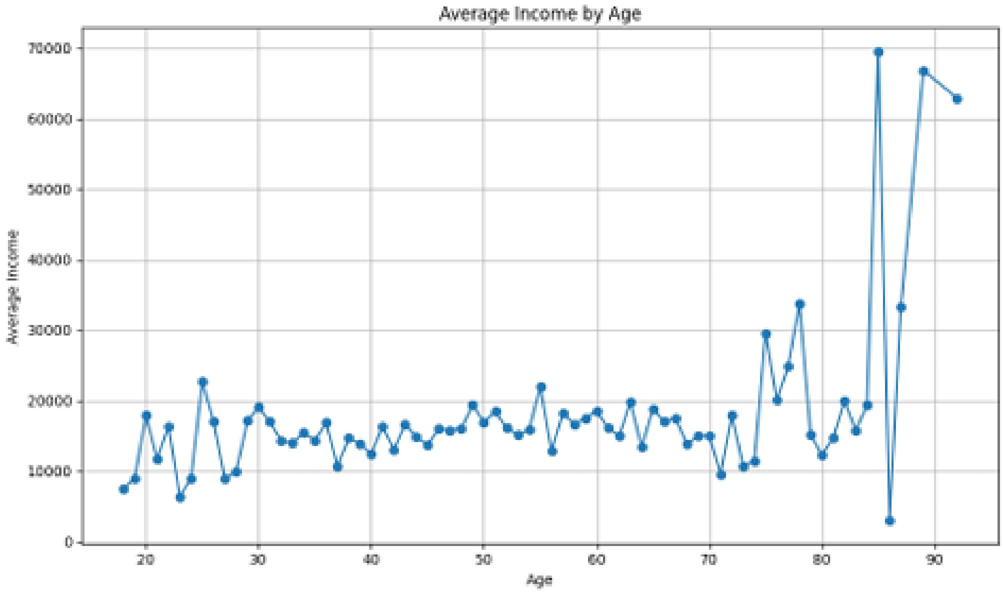

Figure 1: Age-specific income profile

We could see that this graph diverges from the hump shape that tends to be associated with the standard model while showing a lot more volatility and unclear upward or downward trends. This might be a result of the small sample size.

2 Baseline Model Simulation and Results

The baseline life-cycle model is solved and simulated with parameters $\beta = 0.94$ and $\gamma = 2.00$. The following analysis is based on the graphical results, which use a terminal age $T-1 = 60$ and a retirement age $t_r \approx 40$.

2.1 Policy Functions

The policy functions dictate optimal consumption and savings at each age and asset level. For the saving function, each line represents the optimal next-period assets a_{t+1} for a given current asset level a_t across age.

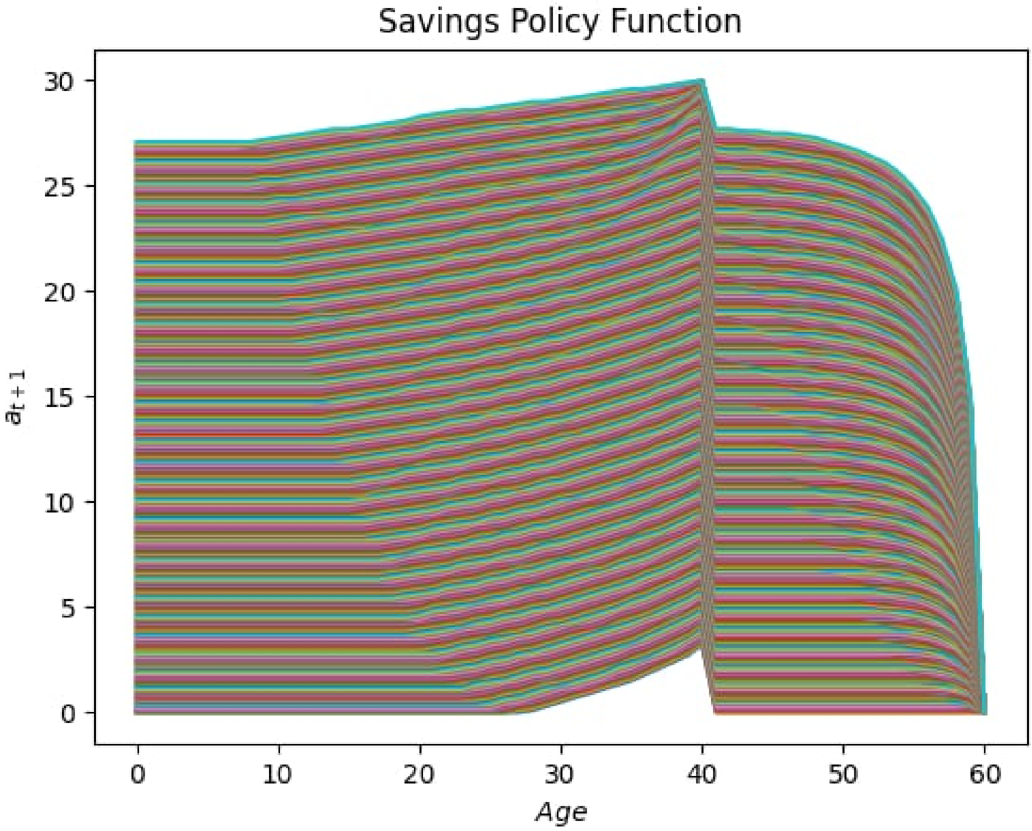

Figure 2: Savings Function

- Savings (a_{t+1}) generally increase with age during the working life (up to age 40). Households accumulate assets for retirement.
- A distinct kink is visible around age 40, likely corresponding to the retirement age t_r. After this point, savings decrease sharply as households begin to dissave to finance retirement consumption.
- The borrowing constraint $a_{t+1} \geq 0$ is active, particularly for households with low asset levels and at older ages, where many policy lines bunch at zero.
- At very young ages, savings are low, often at the zero bound, indicating that consumption needs are prioritized or the borrowing constraint is binding.

As for consumption, each line represents the optimal consumption C_t for a given asset level a_t across age.

- Households with higher asset levels consume more at any given age.
- For a given asset level, consumption is relatively flat or slightly declining during the early to middle working years (up to age 40, the presumed $t_r - 1$).
- After age 40 (retirement), consumption generally decreases for most asset levels, as households start dissaving.
- Towards the very end of life (approaching age 60), there's a sharp upturn in consumption for households with remaining assets. This is consistent with the absence of a bequest motive and the terminal condition $a_T = 0$, leading households to consume all remaining wealth.

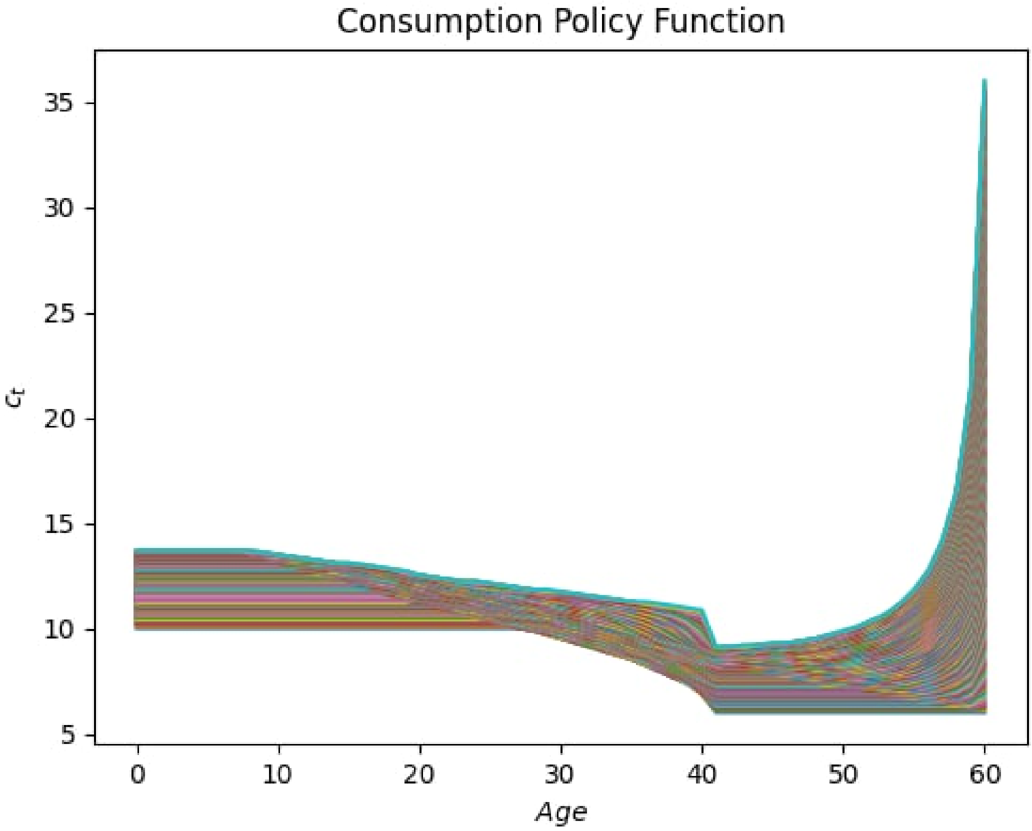

Figure 3: Consumption function

The value function $V_t(a_t)$ represents the maximum lifetime utility from age t onwards, given assets a_t.

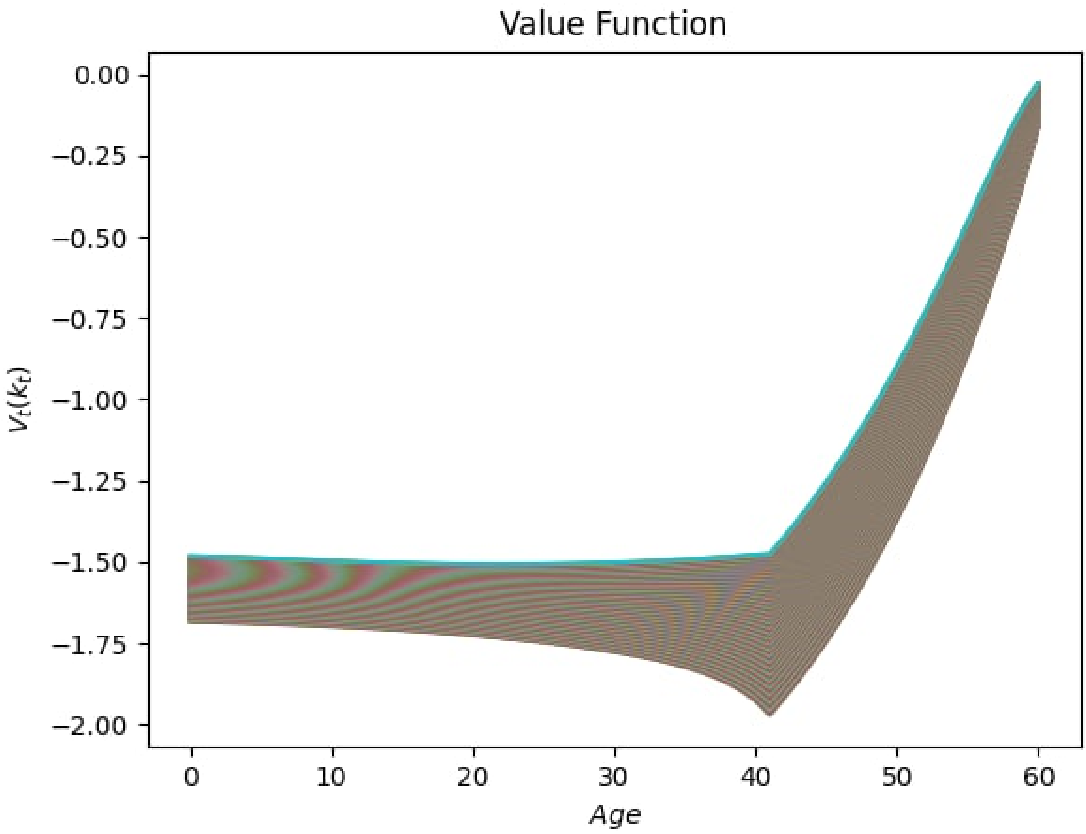

Figure 4: Value function

- The value function is increasing in assets, as more assets allow for higher lifetime consumption and thus higher utility.
- For a given asset level, the value function generally decreases with age during the working life, reflecting a shorter remaining time horizon to derive utility.
- There's a noticeable change in slope around age 40 (retirement).
- As households approach the terminal age (60), the value function converges towards zero (or a value reflecting utility from terminal consumption if $C_{T-1} > 0$), especially for low asset levels, consistent with the finite horizon. The CRRA utility function $C^{1-\gamma}/(1-\gamma)$ with $\gamma = 2$ means utility values are negative if C is not extremely large or $1 - \gamma < 0$.

2.2 Simulated Life-Cycle Profiles of Consumption and Wealth

The simulation returns the following graphical results:

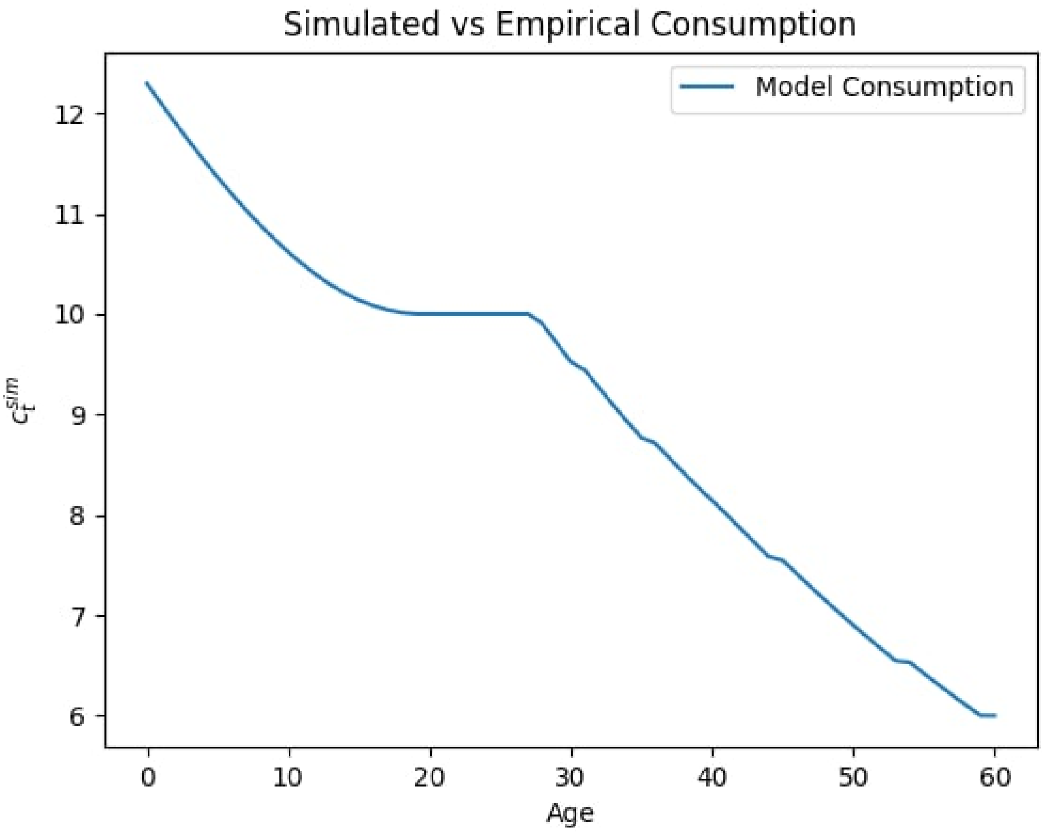

Figure 5: Simulated consumption

Average Simulated Consumption

- Consumption starts relatively high at young ages, then declines throughout the life cycle.
- There's a slight plateau or slower decline between approximately age 15 and 28, after which the decline continues more steadily.
- The profile does not exhibit the empirical "hump shape" often seen in data. Instead, it shows a generally decreasing trend.

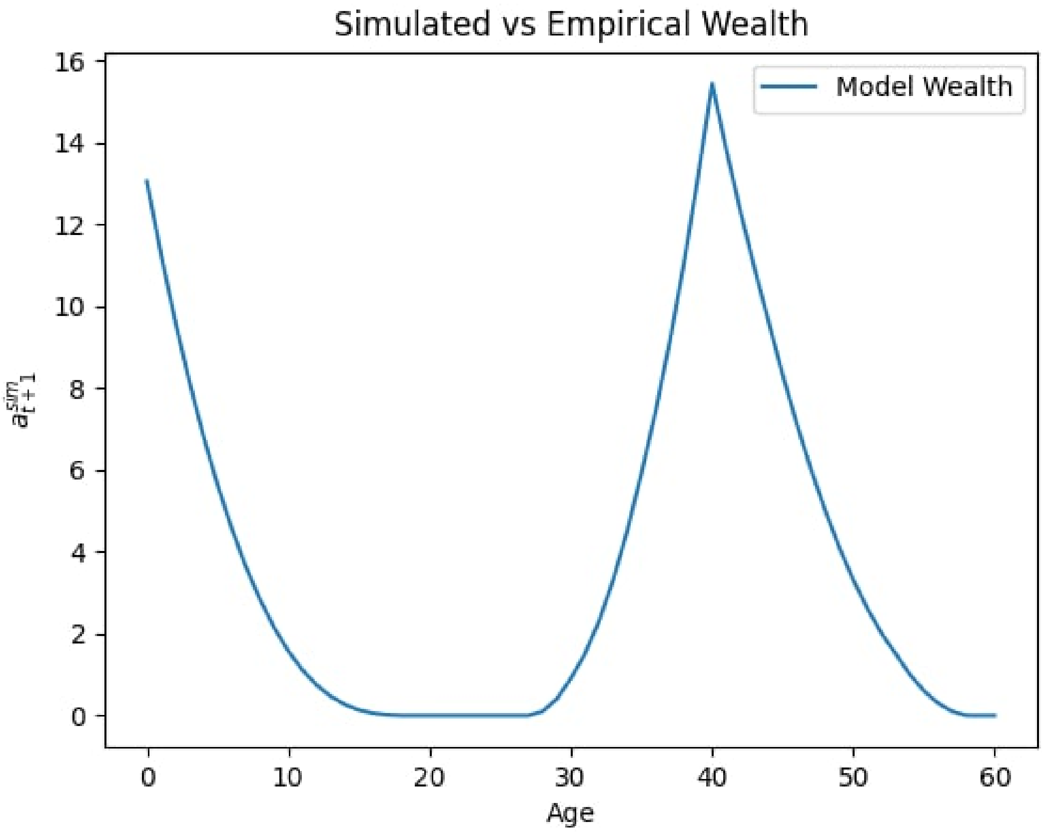

Figure 6: Simulated wealth

Average Simulated Wealth

- Wealth starts at an initial level (if $a_0 > 0$) and is rapidly decumulated in the early years of life.
- Average wealth remains at or very close to zero for a significant portion of the working life (approx. ages 15 to 28-30), indicating the borrowing constraint is binding for many.
- Wealth accumulation begins later in working life, peaking around age 40 (coinciding with t_r). This peak signifies maximum retirement savings.
- After age 40, wealth is rapidly depleted to finance retirement consumption, returning to zero by the terminal age (60).
- This profile shows the three phases: early dissaving/constraint, mid-to-late working life accumulation, and retirement dissaving.

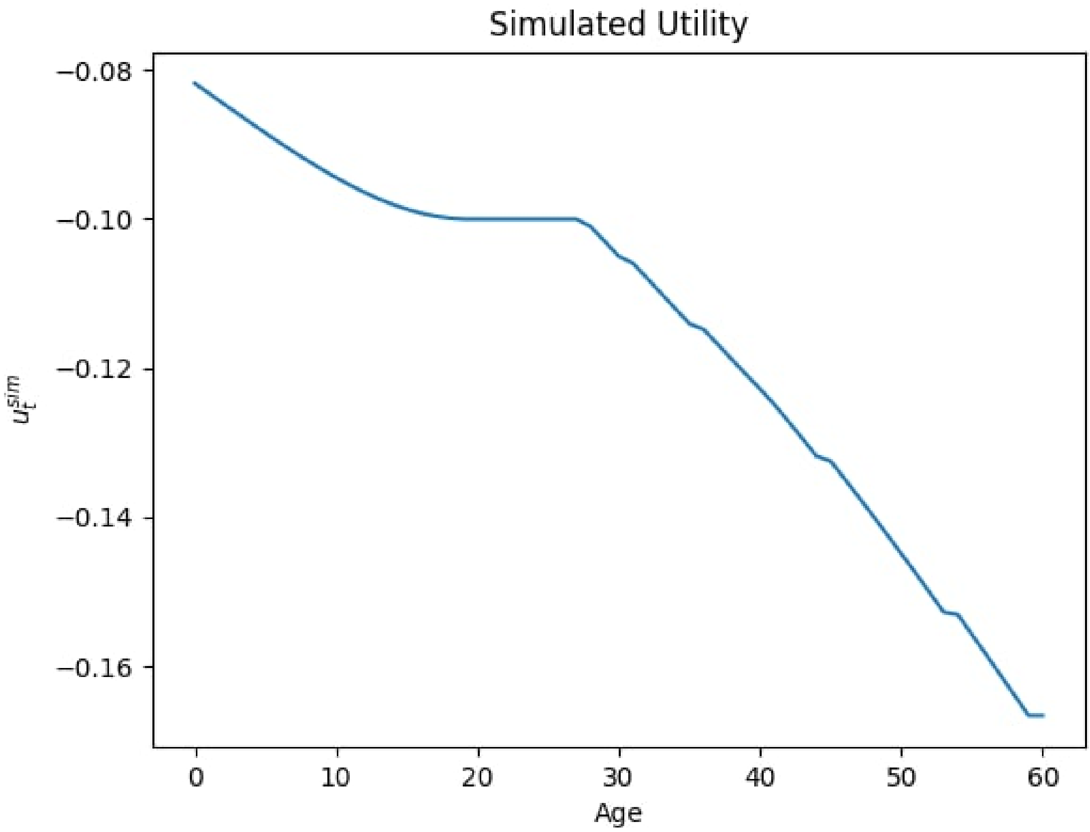

Figure 7: Simulated utility

Average Simulated Utility

- Average period utility (u_t^{sim}) starts higher and declines over the life cycle, mirroring the consumption profile.
- This is expected given that period utility is a direct, increasing function of period consumption.

2.3 Sensitivity Analysis: Varying β and γ

The analysis examines how consumption and wealth profiles change with the discount factor β and the coefficient of relative risk aversion γ.

2.3.1 Impact of Varying the Discount Factor (β)

Here, $\gamma = 2.00$ is held constant, and β varies within $\{0.90, 0.92, 0.94, 0.96\}$.

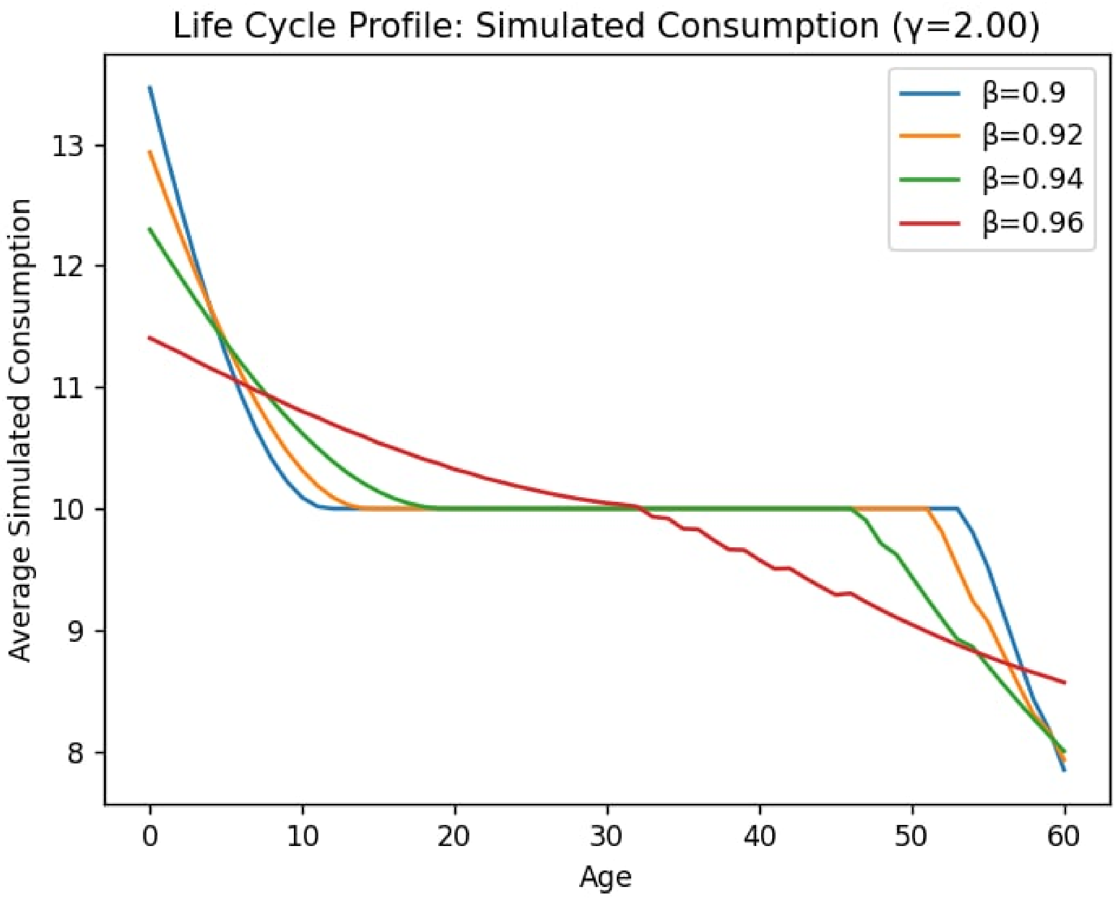

Figure 8: Consumption sensitivity to different discount factors

- A higher β (greater patience) leads to lower consumption early in life and slightly higher consumption later in life. This results in a flatter consumption profile over the working years for more patient households.
- Households with lower β (less patient) consume more when young and less when old.
- The overall declining trend in consumption over the life cycle persists across different β values.

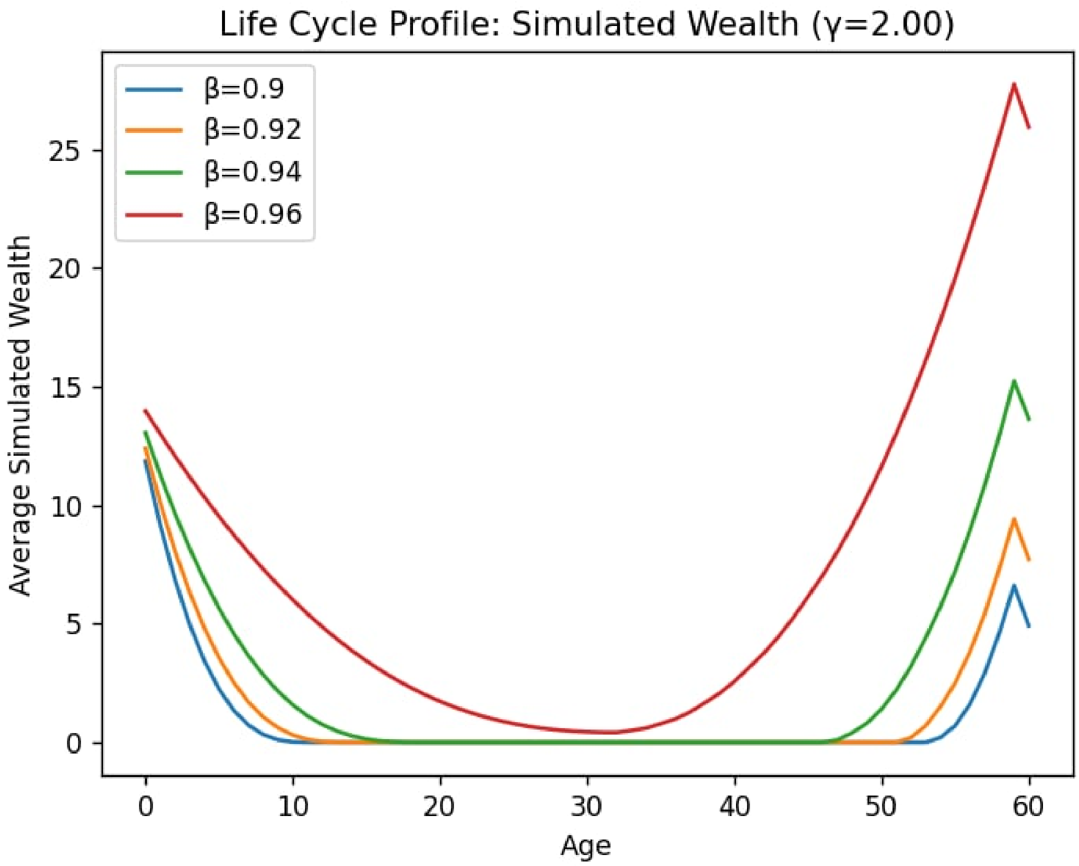

Figure 9: Wealth sensitivity to different discount factors

- Higher β (more patient) households accumulate significantly more wealth. The peak wealth achieved before retirement (around age 40) is substantially higher for $\beta = 0.96$ compared to $\beta = 0.90$.
- More patient households save more vigorously, leading to a larger stock of assets for retirement.
- The period during which wealth is near zero (due to the borrowing constraint) is shorter for more patient households, as they begin saving earlier and more.

2.3.2 Impact of Varying Risk Aversion (γ)

Here, $\beta = 0.96$ is held constant, and γ varies within $\{2.0, 3.0, 4.0, 5.0\}$.

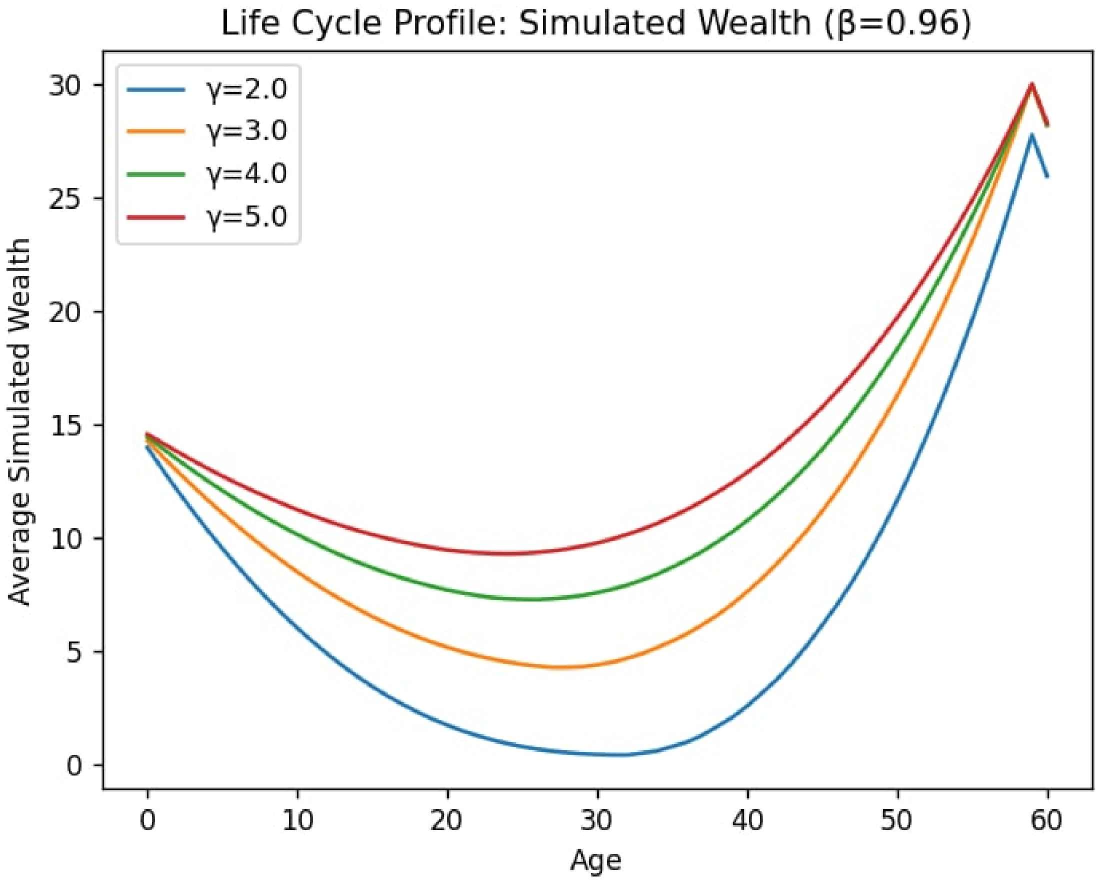

Figure 10: Consumption sensitivity to risks

- Higher risk aversion (higher γ, meaning lower Intertemporal Elasticity of Substitution - IES) leads to a lower and flatter consumption path throughout most of the life cycle.
- Households with higher γ are less willing to substitute consumption intertemporally and are more prudent, leading to reduced consumption, especially early on, to save more.
- The overall declining trend remains, but the level is lower for higher γ. The sharp drop at the end (age 60) is more pronounced for lower γ.

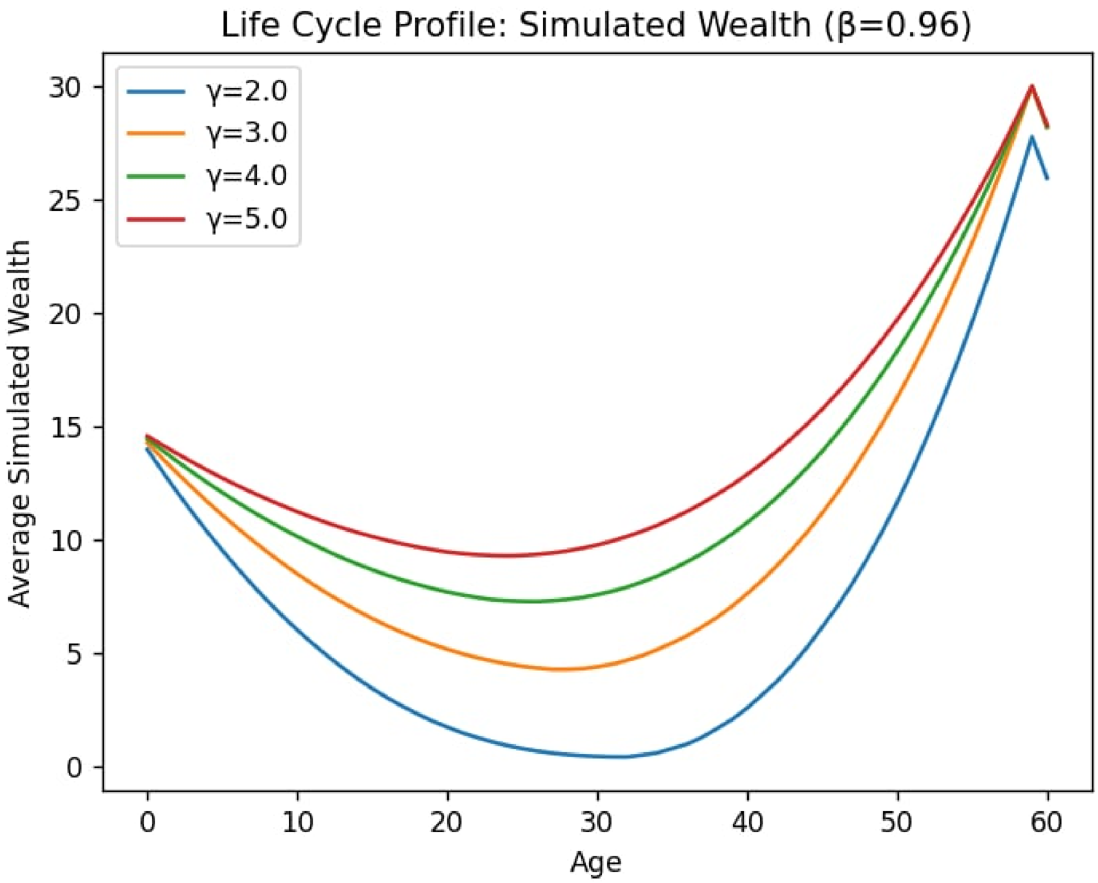

Figure 11: Wealth sensitivity to risks

- Higher risk aversion (γ) generally leads to higher wealth accumulation. This is driven by a stronger precautionary saving motive: more risk-averse households save more to buffer against future income uncertainty.
- Consequently, peak wealth before retirement is higher for larger values of γ.
- Even for patient households ($\beta = 0.96$), risk aversion plays a crucial role in determining asset buildup.

2.3.3 Combined effect of parameters

This heatmap summarizes the combined effects of β and γ on average simulated wealth.

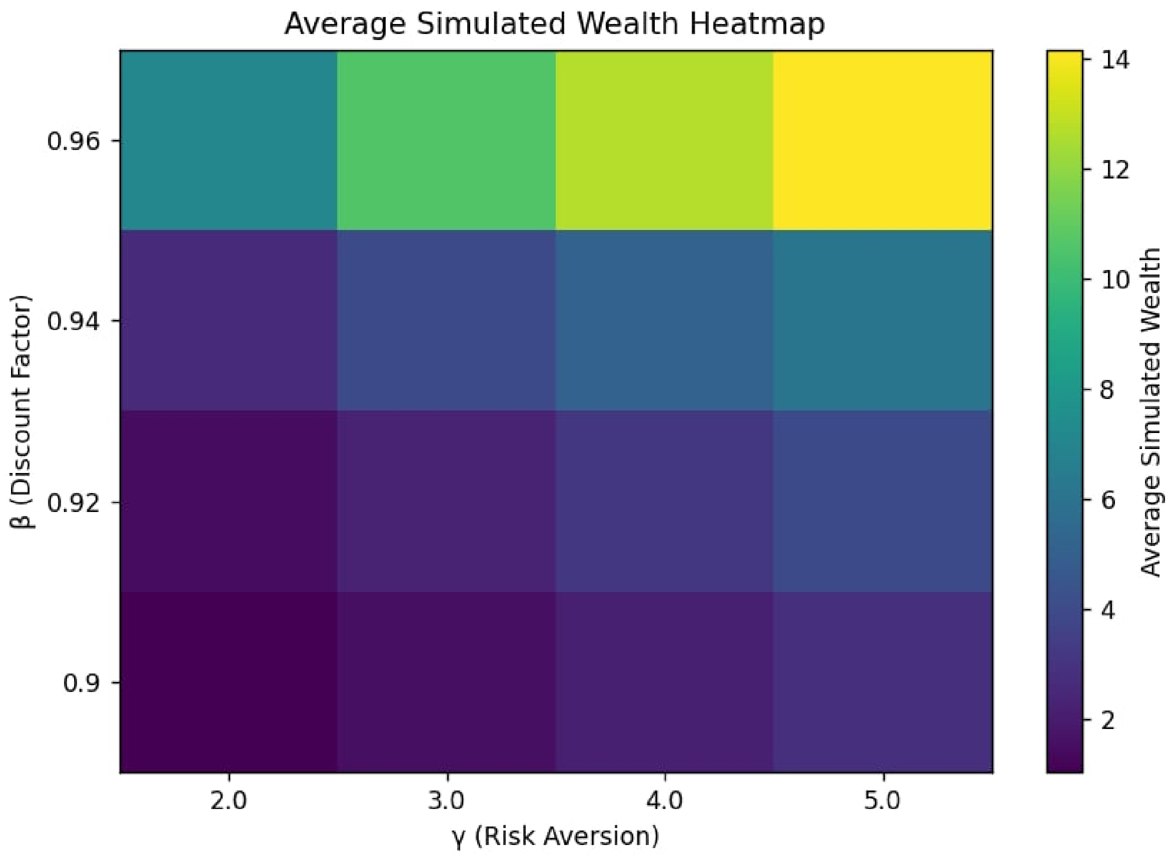

Figure 12: Heatmap of average simulated wealth

- Average simulated wealth is highest (brighter yellow colors, ≈ 14 units) for combinations of high patience ($\beta = 0.96$) and high risk aversion ($\gamma = 5.0$).
- Conversely, wealth accumulation is lowest (darker purple/blue colors, $\approx 1 - 2$ units) for impatient households (low $\beta = 0.90$) with low risk aversion ($\gamma = 2.0$).
- The gradient is strong along both axes:
  - Increasing β (moving up) generally increases wealth for any given γ.
  - Increasing γ (moving right) generally increases wealth for any given β.
- This visually confirms that both a greater desire to defer consumption (higher β) and a stronger motive to self-insure against risk (higher γ) contribute positively to wealth accumulation.

2.4 Chosen Determinant to Add: Total Living Area and Associated Committed Expenditures

The model is extended to include committed expenditures related to total living area. The rationale is that larger dwellings incur higher non-discretionary costs (utilities, maintenance) and can proxy for permanent income/wealth.

2.4.1 Empirical Findings

- The OLS regression shows a statistically significant positive coefficient for $total_leaving_area$ (coef = 282.91, p-value ¡ 0.001) on total nominal household expenditure.
- An R-squared of 0.172 indicates that total living area explains about 17.2% of the variation in household consumption.

```
Regression: Household Consumption on Total Living Area
                            OLS Regression Results
==============================================================================
Dep. Variable:               hhex1nom   R-squared:                       0.172
Model:                            OLS   Adj. R-squared:                  0.172
Method:                 Least Squares   F-statistic:                     1439.
Date:                Sun, 18 May 2025   Prob (F-statistic):          3.05e-286
Time:                        23:14:33   Log-Likelihood:                -79954.
No. Observations:                6934   AIC:                         1.599e+05
Df Residuals:                    6932   BIC:                         1.599e+05
Df Model:                           1
Covariance Type:            nonrobust
=====================================================================================
                        coef    std err          t      P>|t|      [0.025      0.975]
-------------------------------------------------------------------------------------
const              1.584e+04    595.116     26.609      0.000    1.47e+04     1.7e+04
total_living_area   282.9127      7.459     37.930      0.000     268.291     297.534
==============================================================================
Omnibus:                     5033.138   Durbin-Watson:                   1.333
Prob(Omnibus):                  0.000   Jarque-Bera (JB):           139608.352
Skew:                           3.167   Prob(JB):                         0.00
Kurtosis:                      24.050   Cond. No.                         161.
==============================================================================
```

Figure 13: Empirical relationship between living area and expenditures

These findings support the idea that living area is related to consumption, partly through committed expenditures. The model incorporates this as $C_{commit}(H_i) = \theta \times H_i$, which reduces income available for discretionary consumption $y'_t = y_t - C_{commit}$. For the simulation, a C_{commit} value of 0.450 units is used.

3 Simulation Results of the Modified Model with Committed Housing Expenditures

The modified model includes $C_{commit} = 0.450$. Results are for $\beta = 0.94, \gamma = 2.0$. The terminal age appears to be $T - 1 = 70$.

3.1 Life-Cycle Profiles of Consumption and Wealth (Modified Model)

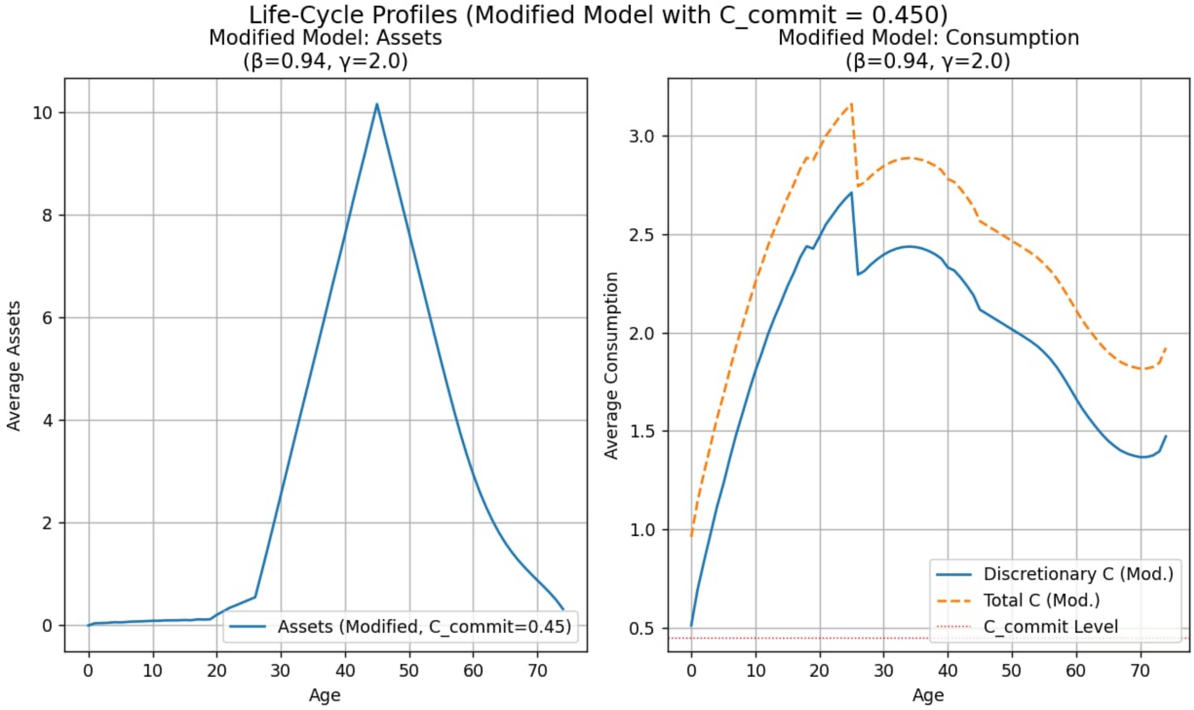

Figure 14: Life-Cycle Profiles

Wealth profile:

- The wealth accumulation path shows a pronounced hump shape.
- Households start with zero assets; accumulation begins relatively early.
- Peak wealth of approximately 10 units is reached around age 45-50.
- After the peak, wealth is rapidly decumulated for retirement, approaching zero by age 70.
- Compared to a baseline without C_{commit} (and potentially different T), this profile might show lower peak wealth if C_{commit} significantly impacts savings capacity. The PDF's Figure 13 (left) is identical.

Consumption profile:

- **Discretionary Consumption (C_t, solid blue line):** Starts low, rises steeply to a peak of around 2.7 units at age 30-35, then gradually declines. This reflects smoothing of discretionary income.
- **Total Consumption ($C_t + C_{commit}$, dashed orange line):** Shifted upwards by the C_{commit} level (0.450, dotted red line). It exhibits a similar hump shape, peaking earlier (age 30-35) at about 3.2 units.
- Even when discretionary consumption is low, total consumption remains above the C_{commit} floor.
- The introduction of C_{commit} makes households effectively "poorer" in terms of discretionary resources, which should generally lower discretionary consumption and savings compared to a model with the same gross income but no committed expenditure.

These patterns are consistent with the PDF's Section 4.1.

3.2 Sensitivity Analysis

This heatmap shows average wealth for different β and γ in the modified model with $C_{commit} = 0.450$.

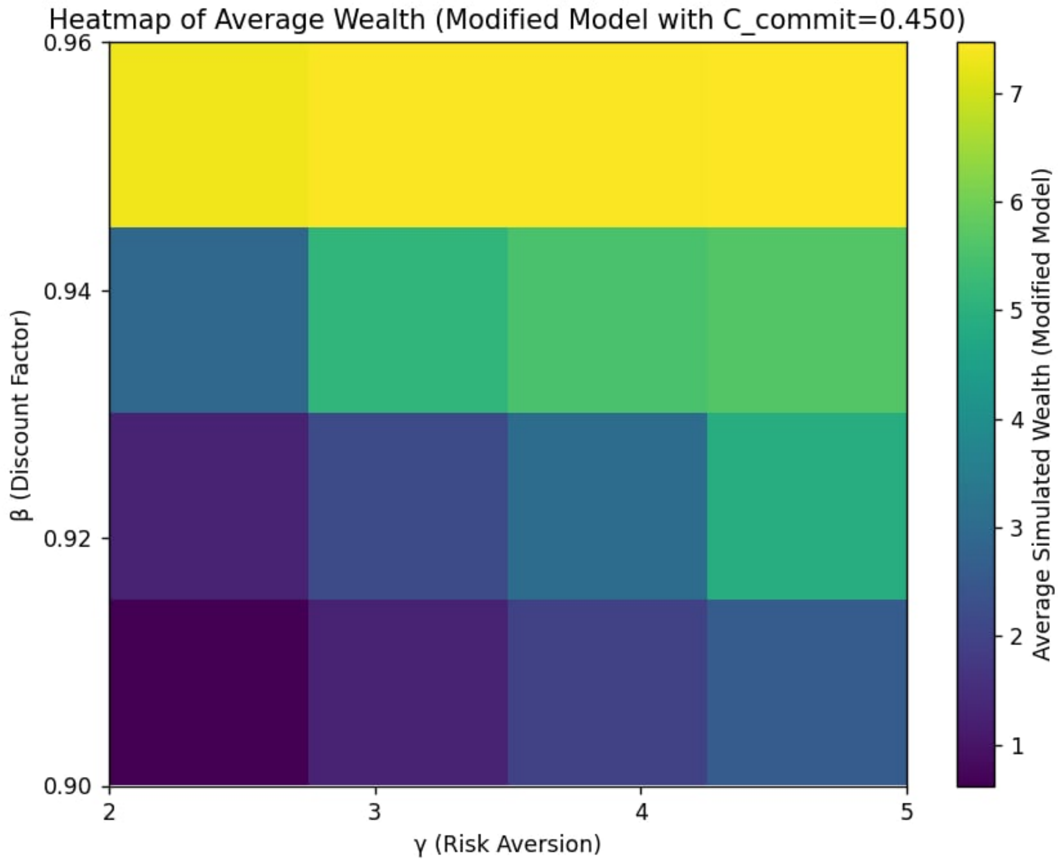

Figure 15: Heatmap of Average Wealth for the Modified Model

- The qualitative patterns are similar to the baseline model's heatmap (Figure ??).
- Highest average wealth (brighter yellow, $\approx 7+$ units) occurs for high patience ($\beta = 0.96$) across all displayed risk aversion levels ($\gamma = 2$ to 5), and also for high risk aversion ($\gamma = 5$) with $\beta = 0.94$.
- Lowest wealth (darker purple, ≈ 1 unit) is for low patience ($\beta = 0.90$) and low risk aversion ($\gamma = 2.0$).
- The overall levels of wealth are lower than in the baseline heatmap (Figure ?? showed peaks around 14). This is expected, as C_{commit} reduces resources available for saving. For instance, at $\beta = 0.96, \gamma = 5.0$, baseline wealth was ~ 14, now it's $\sim 7+$.
- This confirms that C_{commit} acts as a persistent drain on resources, reducing the capacity for wealth accumulation across all preference parameters.

3.3 Discussion and Comparison with Baseline Model

- **Reduced Discretionary Resources:** C_{commit} directly lowers income available for discretionary choices.
- **Impact on Wealth Accumulation:** Overall capacity to save is diminished, leading to systematically lower wealth levels as seen in the heatmap comparison.
- **Consumption Smoothing:** Households still smooth discretionary consumption, but total consumption has a floor.

4 Sensitivity Analysis of Modified Model Findings

This section focuses on the sensitivity of the modified model's outcomes to the level of C_{commit}, holding $\beta = 0.94, \gamma = 2.0$. Three levels are considered: Low (0.250), Medium (0.450 - baseline modified), and High (0.650). Terminal age is $T - 1 = 70$.

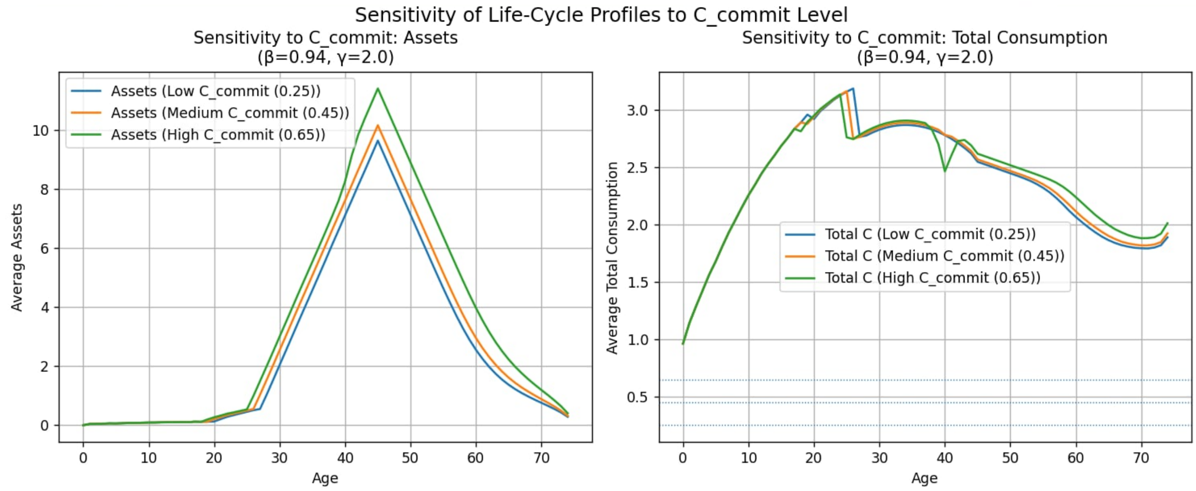

Figure 16: Sensitivity of Life Cycle Profiles

4.1 Impact on Wealth Accumulation

- A higher C_{commit} level leads to substantially lower wealth accumulation.
- **Low C_{commit} (0.250):** Households accumulate the most wealth, peaking above 10 units.
- **Medium C_{commit} (0.450):** Peak wealth is around 10 units.
- **High C_{commit} (0.650):** Peak wealth is lowest, around 9.5 units, and the period of positive asset holding might be shorter or accumulation slower.
- This is intuitive: diverting more income to non-discretionary expenses directly reduces saving capacity. The graph shows a clearer difference: Low C_{commit} peaks near 11, Medium near 10, High near 9.

4.2 Impact on Consumption Profiles

- The right panel shows *total* consumption. The dotted lines represent the different C_{commit} levels (0.25, 0.45, 0.65).
- With a **Low C_{commit} (blue line)**, total consumption is lower initially but allows for higher discretionary consumption (the difference between the solid line and its respective dotted line). The peak total consumption is around 3.1.
- With a **High C_{commit} (green line)**, total consumption starts higher (due to the higher floor) but discretionary consumption is squeezed. Peak total consumption is around 3.25.
- The overall hump shape of total consumption persists. Higher C_{commit} shifts the entire profile upwards but also means a larger fraction of total consumption is non-discretionary, reducing flexibility and likely welfare from discretionary goods.
- The peak of total consumption appears to occur at roughly the same age (around 30-35) across different C_{commit} levels, but the peak value is higher for higher C_{commit}. This implies that discretionary consumption is more heavily suppressed by higher C_{commit} to maintain the hump shape of total consumption.

5 Comparison of Modified Model Predictions with Empirical Data

5.1 Life-Cycle Profiles and Ratios

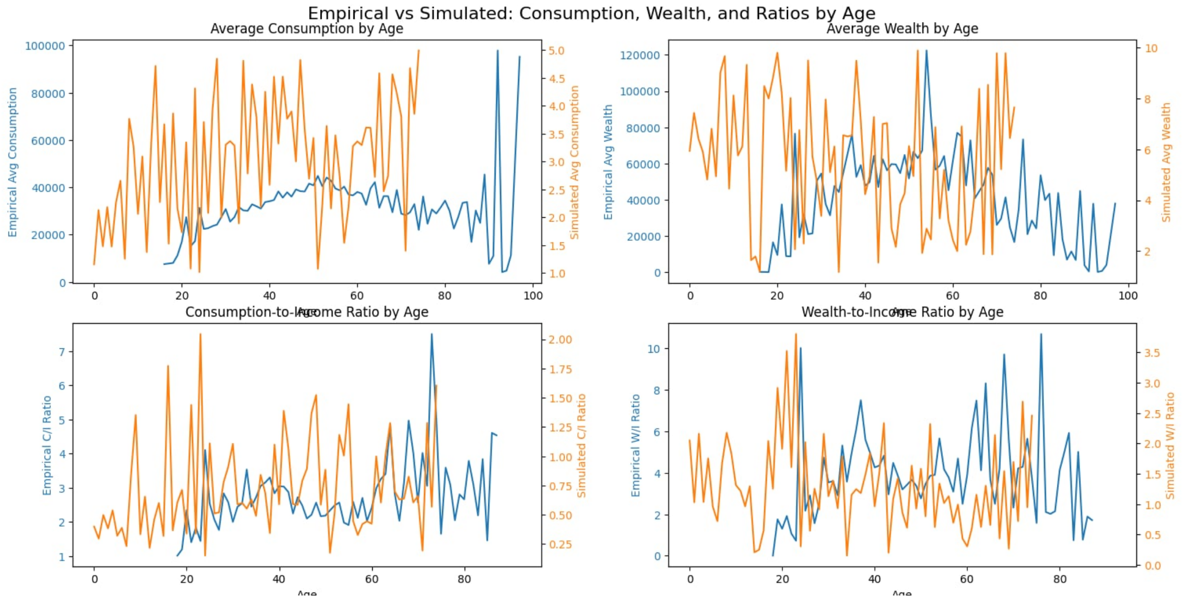

Figure 17: Comparing empirical and simulated profiles

This ratio is unit-less and thus more directly comparable.

- **Empirical Pattern:** The empirical C/Y ratio (blue line) starts high for young households (often ¿1), declines during prime working years (though remaining relatively high, often between 2-4), and can become more volatile or rise again at older ages.
- **Simulated Pattern:** The simulated C/Y ratio (orange line, right y-axis) starts relatively low (as households save from initial income), rises slightly, and then typically declines as income grows faster than consumption during peak earning years, before potentially rising in retirement as income falls. In the provided plot, the simulated C/Y is generally much lower (mostly below 1.5) and flatter than the empirical C/Y.
- **Mismatch:** The model significantly underpredicts the C/Y ratio compared to the VHLSS data across most ages. The empirical data suggests households consume a much larger fraction of their income than the model predicts. This could be due to:
  - Underestimation of true consumption or overestimation of income in the VHLSS data (or vice-versa).
  - The model households being overly prudent or having too strong a savings motive (parameters β, γ).
  - Omission of factors like informal support, access to non-asset based smoothing, or different types of income shocks in the model.
  - The model's definition of income (Y_{gross}) versus how it's measured in VHLSS for the denominator.

Wealth-to-Income Ratio (A/Y) by Age (Figure ??, Bottom-Right):

- **Empirical Pattern:** The empirical A/Y ratio (blue line) is volatile but generally shows an increasing trend during working life, peaking in late middle age.
- **Simulated Pattern:** The simulated A/Y ratio (orange line, right y-axis) also increases during working life as assets are accumulated, and then declines in retirement. The simulated peak A/Y appears to be [e.g., lower than the peaks observed empirically / roughly in the same range but with less volatility].

- **Mismatch/Match:** The model captures the general trend of A/Y increasing with age during working years. However, the levels and volatility differ. The model's A/Y seems [e.g., smoother and potentially lower on average] than the empirical data. This is a common finding, as standard models often struggle to generate the high levels of wealth heterogeneity and A/Y ratios seen in data without features like bequests or very high income risk.