

Mass**INC** Policy Center

Middle Class Cost Model Data and Methods

2025

1) What We're Estimating and Why the Numbers Can Seem "High"

This project estimates the minimum annual wages required for households to meet a middle-class participation budget in a given state and year. We deliberately do not define "middle class" as an income percentile. Instead, we use a budget-constraint definition: A household is "securely middle class" if it can cover the costs of items that are widely associated with a middle-class lifestyle and maintain basic financial resilience without relying on debt, unpaid bills, or skipping essential care.

This approach aligns with the residual income tradition in affordability measurement: What matters is not just how much income you have, but whether you have enough left over after unavoidable obligations to maintain stability.

Why the results can look high:

- We include cost categories that many "cost of living" tools omit—notably retirement savings, emergency savings, and postsecondary education—because these are core to stability and economic mobility. We treat "middle class" as avoiding vulnerability, not merely avoiding poverty.
- We model taxes and income-linked obligations consistently across states and years.

2) Scenario Structures and Units of Analysis

We built a multi-dimensional dataset indexed by:

- State
- Year
- Age cohort (26 to 34, 35 to 44, 45 to 54, 55 to 64, 65 and older)
- Family type (1 adult; 2 adults; 1 adult, 2 kids; 2 adults, 2 kids (2 earners); and 2 adults, 2 kids (1 earner))

To avoid errors due to small sample sizes, households with children are not modeled for age cohorts 55 to 64 and 65 plus.

3) Budget Architecture: Fixed Participation Costs vs Income-Linked Obligations

We separate fixed costs and income-linked obligations. This separation is computationally necessary to solve a circular budget constraint.

A. Fixed Costs: Essential Living and Economic Participation Costs

Costs that vary by place, household type, and year, but are treated as not directly increasing with modeled income within the middle-class range. These include:

- Housing
- Health insurance premiums
- Medical & dental out-of-pocket costs

- Food
- Transportation
- Clothing
- Cleaning & personal care supplies
- Other household supplies
- Media & communication
- Childcare (if applicable)

B. Income-Linked Obligations

Costs that explicitly depend on income, requiring an optimization algorithm to solve for income and costs. These include:

- Taxes
- Retirement savings (target rate varying by age)
- Emergency savings (target rate varying by age)
- Student loan payments (for adults)
- Family higher education contributions (for families with children)

4) Fixed Participation Cost Components: Justification and Operationalization

4.1 Housing and Utilities

Data and methods: US Department of Housing and Urban Development Fair Market Rents (FMRs) provide a stable signal of geographic/time variation, while American Community Survey (ACS)-based housing costs show variation in costs by age and family type, but can have small sample sizes. We therefore estimate housing costs using a hierarchical Bayesian model that:

- treats ACS housing costs as the estimand
- uses log(FMR) as an offset to anchor state-year market levels
- partially pools across states and years to stabilize estimates

This improves precision without replacing ACS with FMR.

Caveats:

- FMR is a rent benchmark, not a direct measure of household spending; it provides a statistical anchor, not the concept being measured.
- Small-cell volatility is reduced by partial pooling.
- ACS data were only available through 2023. FMRs for 2024 were inputted into the statistical model to estimate data for 2024.

4.2 Food

Data and methods: The US Department of Agriculture (USDA) Food Plans; the Moderate-Cost Plan is used as the baseline budget. Values are national (replicated across states) and vary by year and household composition. USDA describes the plans as nutritionally

adequate diets at successively higher budgets. The Moderate-Cost Plan reflects a realistic consumption pattern for working households facing time constraints. Lower-cost food plans implicitly assume substantial time available for meal planning, bulk preparation, and price-minimizing shopping strategies. Because most middle-class households rely on full-time employment and face binding time constraints, the model assumes a food basket that includes a mix of home-prepared meals and convenience-oriented items consistent with observed consumption patterns.

Caveat:

- No state-level variation in food costs. This will lead to conservative estimates in high-cost regions and will overstate costs in others.

4.3 Transportation and Clothing

Data and methods: These estimates are drawn from Bureau of Labor Statistics (BLS) Consumer Expenditure Survey (CE) tabulations (national, by year and family type). They are treated as baseline annual spending levels and replicated across states.

Caveats:

- CE is expenditure-based, not a required minimum measure. Using the mean of an "essential" category implicitly embeds current norms.

4.4 Cleaning and Personal Care Supplies, Media and Communications, and Other Household Supplies

Data and methods: These components are constructed from CE microdata using custom aggregations of Universal Classification Codes (UCCs). Categories are aggregated to annual household expenditures by year and family type and applied uniformly across states. The custom aggregation includes baseline communication services and modest media access but excludes high-cost discretionary entertainment categories.

Caveats:

- These expenditures reflect average baseline consumption patterns and may include modest discretionary elements.
- Because CE data capture observed spending rather than a strict minimum bundle, results should be interpreted as representing a societal norm rather than a survival minimum.

4.5 Health Insurance Premiums

Data and methods: For non-Medicare households, premiums are drawn from the Medical Expenditure Panel Survey Insurance Component (MEPS-IC) annual state tables. MEPS-IC is an employer survey covering private employers and state and local governments that produces state-level estimates of employer-sponsored insurance premiums and employee contributions. For Medicare-eligible households, premiums are based on Medicare Part B standard premiums.

Caveats:

- MEPS-IC reflects employer-sponsored insurance premiums. It does not capture the full marketplace/individual market premium distribution, and it is not a measure of Medicare premiums (which follow a different pricing and eligibility structure).
- Supplemental Medicare coverage premiums are not explicitly modeled and may cause total healthcare costs for older households to be understated.

4.6 Medical, Vision, and Dental Out-of-Pocket Spending

Data and methods: Out-of-pocket (OOP) spending is estimated using the Medical Expenditure Panel Survey Household Component (MEPS-HC), accessed through the Integrated Public Use Microdata Series (IPUMS). IPUMS harmonizes and integrates MEPS microdata across years and provides consistent variable definitions and documentation. We measure OOP using EXPSELPAY, which captures direct payments made by individuals and their families for medical care during the year. This measure includes OOP payments for physician services, hospital care, prescription drugs, and other healthcare services, as well as dental and vision services when paid directly by the household. MEPS captures actual payments made under real-world insurance arrangements; as a result, the estimates reflect deductibles, copayments, coinsurance, uncovered services, and the distribution of utilization across households.

Caveats:

- Estimates are produced at the Census region level (Northeast, Midwest, South, and West) due to MEPS public geography sample size limitations.
- Because MEPS-HC microdata are only available through 2023, 2024 OOP values are projected by inflating the 2023 amounts using the Consumer Price Index for all Urban Consumers (CPI-U) Medical Care index (annual average growth factor).

4.7 Childcare

Data and methods: Childcare costs are constructed using multiple sources.

- Primary source: The baseline estimates are drawn from the National Database of Childcare Prices (NDCP), a federally produced dataset providing state-level childcare price estimates derived from administrative and survey sources. The NDCP provides the most comprehensive nationally consistent estimates, but it is available only through 2023 and contains missing observations for some states and years.
- Supplemental source: For states or years lacking NDCP observations, additional data points were obtained from the University of Washington Self-Sufficiency Standard data repository, which publishes state-specific childcare cost estimates constructed using comparable market-price methodologies.
- Fill missing state-year combinations within the NDCP time window (2008 to 2023): Complete-case observations were used to estimate year-to-year growth rates in childcare costs. Growth patterns were calculated using states with continuous data coverage. These estimated annual changes were applied to interpolate missing values within the observed NDCP timeline.

- Projection outside observed years (2005 to 2007 and 2024): Because NDCP coverage begins after 2005 and ends at 2023, childcare costs were extended using the CPI-U for education-related services.

Caveats:

- The assumption of one infant and one toddler represents a maximum-cost phase that most households experience only temporarily, if at all. Results should therefore be interpreted as measuring financial resilience during peak childcare years rather than long-run average expenses.
- Childcare needs depend on the ages of children rather than the age of the householder; age buckets serve as a life-cycle proxy and cannot perfectly capture household timing differences.
- Some households rely on informal or unpaid caregiving arrangements (help from family members, flexible work schedules, or staggered employment). These arrangements are not modeled directly because they are not consistently observable across datasets and vary widely by circumstance.
- The one-earner household scenario should not be interpreted as a costless alternative; it implicitly involves foregone earnings rather than reduced economic burden.
- Childcare prices vary by provider type and quality; modeled values represent average market rates for center-based care.

4.6 Dining Out & Recreation

Spending on dining out and recreational activities is derived from the Ludwig Institute for Shared Economic Prosperity's Minimal Quality of Life (MQL) Index methodology. Specifically, expenditures include youth sports equipment, dining out, celebratory meals, travel-related food and lodging, and weekend leisure activities.

The MQL framework defines these costs based on observed spending patterns among households in the middle of the income distribution, recognizing that time constraints, social participation, and modest leisure activities are widely treated as necessary components of contemporary economic participation rather than optional consumption. Full methodological details, data sources, and calculation procedures are documented in the [MQL methodology report](#).

5) Income-Linked Obligations: Justification and Operationalization

5.1 Taxes

Data and Methods: Tax liabilities are modeled using the National Bureau of Economic Research TAXSIM-35 microsimulation engine. For each state, year, and family type, taxes are simulated across a dense grid of wage levels. The resulting tax schedule includes federal income taxes, state income taxes, and payroll taxes.

For any modeled wage level, total taxes are calculated by interpolating the effective tax rate implied by the simulated wage grid. This approach preserves nonlinear features of tax systems, including bracket transitions and credit phase-ins and phase-outs, while allowing efficient integration into the income solver.

TAXSIM is a widely used academic microsimulation tool designed to produce consistent, comparable tax liabilities across jurisdictions and years.

- Payroll taxes: TAXSIM reports total Federal Insurance Contributions Act (FICA) taxes, including both employer and employee shares. Because the model is constructed from a household cash-flow perspective, only the employee portion is included, calculated as one-half of total FICA.
- Treatment of years without statutory coverage: TAXSIM provides complete federal tax law through 2023, while state coverage varies slightly by module and year. For years without statutory coverage, effective tax schedules are held constant at the most recent available law year (e.g., 2023 schedules applied to 2024). Tax liabilities therefore continue to vary with wages, but the structure of the tax system is assumed to be unchanged.

Caveats:

- Post-coverage years assume no major tax law changes until updated statutory data become available.
- Refundable credits (e.g., Earned Income Tax Credit and Child Tax Credit) can generate negative net income tax at lower wage levels; these negative values reflect statutory policy design and are intentionally preserved in the model.

5.2 Retirement Savings

Method: Retirement savings are modeled as a percentage of income that varies by age to reflect lifecycle patterns of earnings and financial obligations:

- Ages 26 to 34: 8 percent
- Ages 35 to 44: 10 percent
- Ages 45 to 54: 10 percent
- Ages 55 to 64: 12 percent to 15 percent
- Ages 65 and older: 0 percent

The model assumes that households in their 20s and early 30s save at a meaningful but somewhat lower rate than mid-career households, reflecting early-career earnings constraints. Mid-career households save at or near the canonical 10 percent benchmark. Pre-retirement households increase saving rates modestly to reflect catch-up savings and the completion of major child-related expenditures.

Financial planning research frequently recommends total lifetime saving rates of approximately 12 percent to 15 percent of earnings, inclusive of employer contributions. Employer-sponsored plans commonly provide contributions averaging roughly 3 percent to 5 percent of pay. Because employer contributions do not reduce household disposable

income, the model includes only the employee-side contribution, calibrated to approximate a 15 percent total saving rate when employer matches are considered.

Caveats:

- Retirement adequacy benchmarks vary across financial planning frameworks and depend on assumptions about wage growth, investment returns, retirement age, and desired replacement rates. The modeled rates represent a normative adequacy benchmark rather than a universal prescription.
- Actual saving behavior varies widely across households; many households save substantially less than the recommended levels. The model reflects a standard of financial resilience consistent with maintaining middle-class living standards over the life course.
- The age-varying schedule simplifies complex lifecycle behavior; in practice, saving patterns may respond to debt, childcare costs, housing transitions, and macroeconomic shocks.

5.3 Emergency Savings

Method: Emergency savings are modeled as an age-varying contribution rate applied to wages. Contributions are intended to represent the ongoing maintenance of a liquid financial buffer equivalent to several months of essential expenses, recognizing that reserves must be periodically replenished after use.

The contribution rate varies across life stages to reflect both saving capacity and risk exposure over the lifecycle:

- **Ages 26 to 34 (2 percent):** Early-career households face constrained cash flow due to student loans, childcare, and housing formation costs. Contributions are therefore modest, representing gradual buffer accumulation rather than full reserve building.
- **Ages 35 to 44 (5 percent):** Rising household responsibilities increase exposure to income and expenditure shocks. Higher contributions reflect the need to actively build and maintain liquidity during years of elevated financial risk.
- **Ages 45 to 54 (4 percent):** Households typically have higher earnings but also face competing obligations such as postsecondary education expenses for children. Contributions remain substantial but slightly reduced as buffers are assumed to be partially established.
- **Ages 55 to 64 (4 percent):** Liquidity remains important prior to Medicare eligibility, when medical and employment risks remain meaningful.
- **Ages 65 and older (2 percent):** Households are assumed to maintain smaller ongoing contributions primarily to preserve liquidity rather than accumulate new reserves during retirement years.

Caveats:

- The emergency savings component represents a normative resilience standard, not observed household saving behavior.
- Actual liquidity needs vary widely across households depending on housing type, health status, and access to informal support networks.

5.4 Student Loans

Data and methods: Public expenditure surveys do not reliably observe total student loan payments. Instead, payments are modeled using statutory Income-Driven Repayment (IDR) formulas in effect during each year of analysis. IDR programs determine required payments as a function of income and household size, allowing payments to be derived directly from modeled wages.

Annual payments are calculated as:

$$\text{Payment}_t = r_t \times \max(0, \text{AGI} - (m_t \times \text{Poverty Guideline})) \times a_t$$

where:

- r_t = statutory payment rate in year t
- m_t = poverty-income protection multiplier
- a_t = fraction of the year that payments were active
- Poverty guidelines vary by household size and geography (CONUS, Alaska, Hawaii).

Adjusted Gross Income (AGI) is defined as wages net of retirement contributions within the model framework.

Poverty thresholds are obtained from the US Department of Health and Human Services (HHS) poverty guidelines via the Office of the Assistant Secretary for Planning Evaluation application programming interface, including historical series and Alaska and Hawaii adjustments.

Payments follow the most widely available IDR structure in each period:

Period	Policy regime	Payment rule
2005 to 2008	Early IBR-style assumption	20% of income above 150% poverty
2009 to 2011	IBR	15% above 150% poverty guideline
2012 to 2019	PAYE / REPAYE	10% above 150% poverty guideline
2020	COVID pause begins	Payments active for partial year only

Period	Policy regime	Payment rule
2021 to 2022	Federal payment pause	Payments set to \$0
2023	SAVE transition	10% above 225% poverty guideline; partial-year activation
2024	Administrative forbearance	Payments set to \$0

Partial-year activation parameters explicitly capture months in which payments were suspended or newly implemented.

Student loan payments are applied to the 26 to 34 age cohort, reflecting the typical repayment window following undergraduate education.

5.5 Higher Education Contribution for Children

Method: Higher-education contributions are based on College Scorecard net price estimates for public institutions, which measure average cost of attendance after grant aid by state, year, and family income bracket. Net price reflects the amount households are expected to finance through income, savings, or borrowing under the US financial aid system.

State-year averages are constructed using enrollment-weighted net prices. The model uses the upper income brackets (\$75,001 to \$110,000 and above \$110,001), which most closely align with the income range represented by middle-class households in this framework.

Because bracket-level data are available beginning in 2009, earlier years (2005 to 2008) are backcast and later years (2024) extended using the CPI-U for education to maintain a consistent time series.

These costs only apply to households in the 45 to 54 age cohort. The estimates are only modeled for households with two children. We assume both children attend college at the same time and parents cover half of the expected family contribution (the student will utilize loans to cover the other half).

Caveats:

- Net price represents average aid-adjusted costs and may not reflect the experience of any specific institution or household.
- Restricting analysis to higher income brackets reflects the study's focus on middle-class households rather than the full income distribution.
- Age of householder is used as a proxy for having college-age children and should be interpreted as a lifecycle exposure rather than a universal expense.

6) Solving for Required Income

For each combination of state, year, age group, and family type, the model calculates the level of earnings required for a household to fully cover its financial obligations. Required income is defined as the level at which total resources exactly equal total costs.

In practice, this means identifying the wage level at which a household can pay for:

- fixed costs (such as housing, food, transportation, and healthcare)
- obligations that rise with income, including taxes, retirement saving, emergency saving, student loan payments, and higher-education contributions

A key complication is that several of these costs depend on income itself. For example, higher wages increase taxes and savings contributions, which in turn change how much income is required. Because income appears on both sides of the calculation, the required wage cannot be computed through simple addition.

Instead, the model uses a numerical optimization procedure to identify the minimum wage level that satisfies the household budget constraint. Starting from an initial income range, the algorithm evaluates total costs—including all income-dependent obligations—and systematically adjusts income until both sides of the equation are equal within a very small tolerance.

Because taxes and several obligations increase automatically as income rises, this procedure finds the lowest income at which total earnings are sufficient to cover all modeled costs simultaneously. In technical terms, the algorithm solves for the income level at which the difference between income and total obligations equals zero (a root-finding problem).

Childcare expenditures were excluded from the wage base used to calculate retirement savings, emergency savings, and student loan repayment obligations. This prevents a mechanical feedback loop in which higher wages raise proportional savings requirements, which then further increase required wages solely because childcare expenses are present.

For tax calculations, the model incorporates the tax-preferred treatment of childcare available through Dependent Care Flexible Spending Accounts. A capped portion of childcare spending is treated as paid with pre-tax income, reducing taxable earnings up to the statutory limit, while any childcare expenses above that cap remain part of taxable income. Taxes are therefore calculated on taxable wages rather than total gross earnings.

The resulting value represents the income required for a household to meet modeled obligations without running a deficit.