

Working Paper

Measuring lifetime sales tax progressivity: A simulation-based approach

Jorge Barro, Ph.D.

Fellow in Public Finance, Rice University's Baker Institute for Public Policy

Clint Hamilton

University of California, Berkeley

© 2018 by the James A. Baker III Institute for Public Policy of Rice University
This material may be quoted or reproduced without prior permission, provided appropriate credit is given to the author and the James A. Baker III Institute for Public Policy.

Wherever feasible, papers are reviewed by outside experts before they are released. However, the research and views expressed in this paper are those of the individual researcher(s) and do not necessarily represent the views of the James A. Baker III Institute for Public Policy.

This paper is a work in progress and has not been submitted for editorial review.

Introduction

Consumption taxes are the third largest source of government revenue in the U.S., generating nearly as much revenue as corporate income and property taxes combined and exceeded only by personal income and Social Security taxes. In 2016, consumption taxes accounted for 17 percent of all federal, state, and local government taxes collected, providing roughly 3 percent of total U.S. GDP in revenue.¹ Each day consumption taxes are paid millions of times with the purchase of most consumer goods. Though its impact on equity has been extensively disputed and discussed, much is left to discover regarding the progressivity of consumption taxes.

One difficulty in evaluating consumption taxes, such as sales and excise taxes, results from the lack of data matching the taxpayer with the tax payment. Usually, buyers pay the tax, while sellers collect it for the government. Although the incidence of the tax may fall more heavily on either the buyer or the seller, equity of the buyer-borne tax burden is difficult to assess because very little is ever known about the buyer. In contrast, when individuals pay income taxes to the federal government, extensive paperwork documenting individuals' unique economic circumstances are documented in submissions to the IRS, allowing researchers to measure certain features, including tax progressivity.²

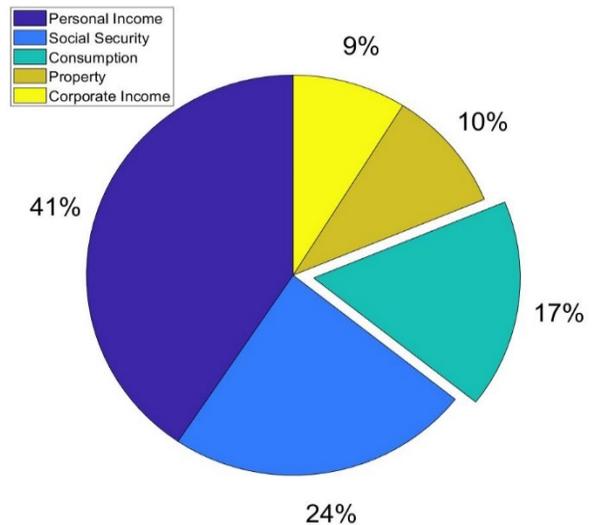


Figure 1: Composition of 2016 Tax Revenue in the U.S.

¹ OECD Revenue Statistics: <https://www.oecd.org/tax/revenue-statistics-united-states.pdf>

² Progressivity of a tax is defined here as the change in taxes paid divided by income as income increases. If this value is increasing, then the tax is generally *progressive*, and if this value is decreasing, then the tax is generally *regressive*. Without loss of generality, any measurement of this value is referenced as a tax's level of progressivity.

Since the progressivity of the sales tax is difficult to directly measure, this paper introduces an indirect approach combining simulated household income with realizations of consumption behavior from survey data. This approach circumvents two issues regarding the measurement of consumption tax progressivity. First, the simulated approach overcomes the anonymity of buyer information by using granular consumption expenditure data. The data pairs uniquely identified individuals with a constructed consumption tax base that can be modified to replicate the composition of any existing tax jurisdiction's specific consumption tax base. Second, as discussed below, annual consumption tax progressivity provides a poor indication of an individual's actual tax burden. To overcome this limitation, the quantitative strategy involves simulating both lifetime income and lifetime consumption to provide a measure of lifetime consumption tax progressivity.

Economist James Poterba first proposed measuring consumption tax progressivity from a lifetime perspective as he evaluated lifetime excise tax burden.³ This paper contributes to that concept by presenting a lifetime sales tax progressivity simulator, which is applied to a sales tax base constructed for illustrative purposes.⁴ The results of the simulation applied to this tax base show that the total lifetime sales tax bill increases with lifetime income, but lifetime sales taxes are slightly regressive.

Theory

Consumption and Income

In the 1930's, British economist John Maynard Keynes provided a general rule of thumb relating consumption and income: as income increases, the percentage of income allocated to consumption declines. Estimates from the Consumer Expenditure Survey (CEX) shown in Figure 2 confirm this simple relationship, which Keynes (1936)

³ Poterba, J. (1989). Lifetime incidence and the distributional burden of excise. *The American Economic Review*, 79(2), 325.

⁴ Given the variability in state and municipal tax bases across the U.S., any given tax base may be more or less progressive than the example herein.

described as “decreasing marginal propensity to consume.”⁵ Because the allocation of consumption declines as income increases, many opponents of sales taxation argue that consumption taxes are inherently regressive. From the narrow lenses of annual consumption and annual income, very little argument to the contrary exists. A closer look at consumption, savings, and income patterns throughout a lifetime, however, reveal the potential that a better measure of consumption tax progressivity is based on longer-term income and consumption behavior. To that extent, this paper argues in favor of evaluating *lifetime* consumption tax progressivity, rather than an arbitrary periodic measure, and presents a lifetime consumption tax progressivity simulator based on multiple data sources and a tax base constructed for illustration.

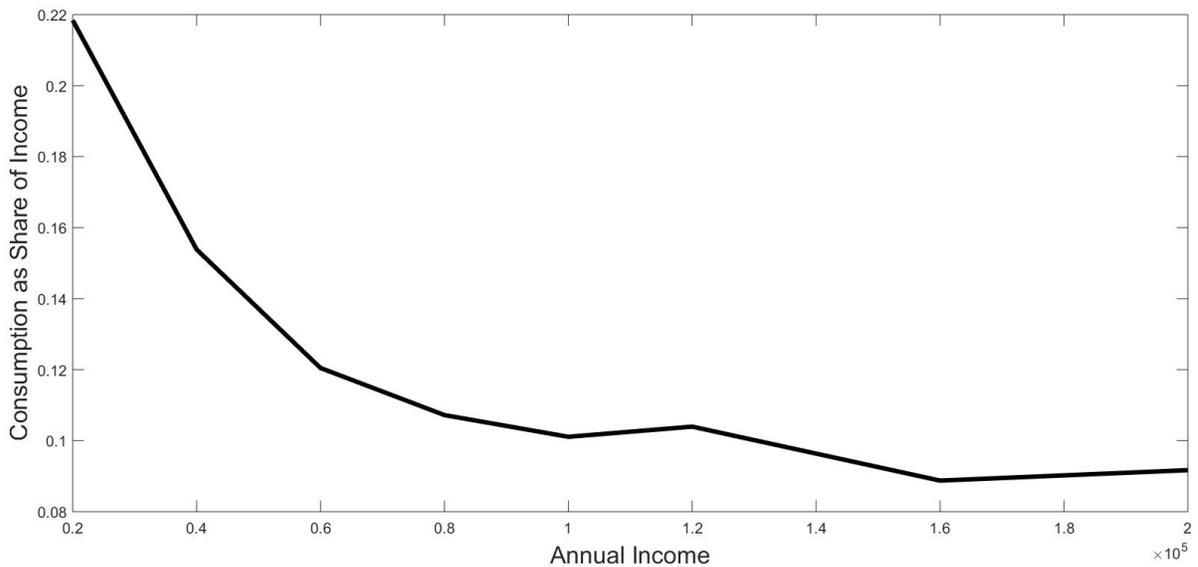


Figure 2: Marginal Propensity to Consume, as measured by the Consumer Expenditure Survey

Life-cycle Economic Theory

To understand why lifetime measurements provide a better understanding of consumption tax progressivity, consider the Life-cycle Hypothesis (LCH) popularized in the 1960’s by Nobel Laureate Franco Modigliani.⁶ The LCH, as represented in Figure 3, claims that because income increases early in an individual’s lifetime and decreases later on,

⁵ Keynes, John Maynard (1936). *General theory of employment, interest and money*. New York: Harcourt Brace Jovanovich.

⁶ Modigliani, Franco (1966). "The Life Cycle Hypothesis of Saving, the Demand for Wealth and the Supply of Capital". *Social Research*. **33** (2): 160–217.

individuals tend to borrow early in life, pay off debt and save for retirement in their working years, then dis-save later in life. This economic behavior is motivated by the inherent nature of individuals to smooth consumption throughout their lifetime.

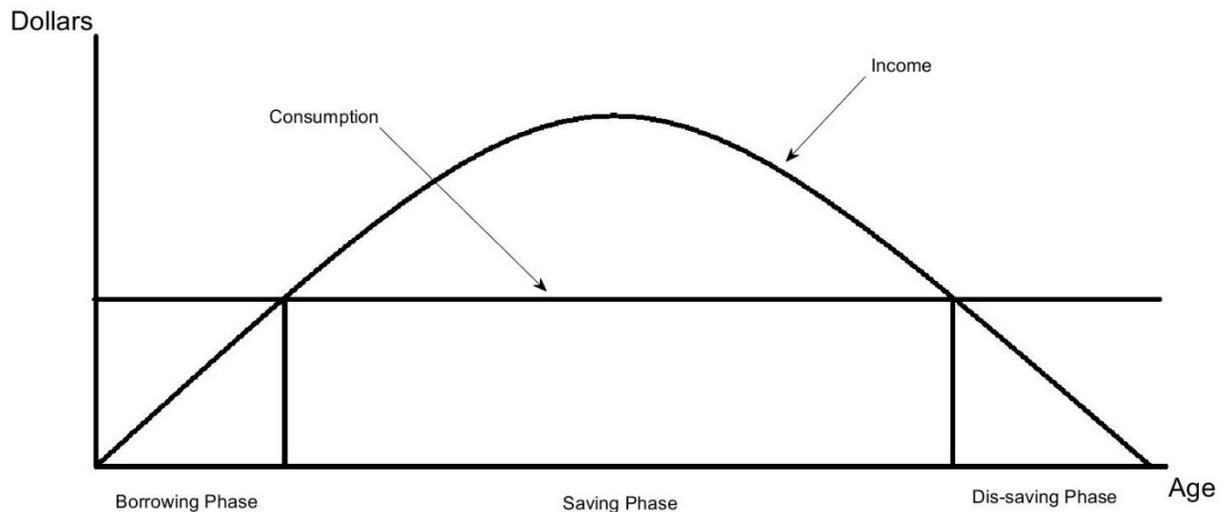


Figure 3: Representation of the Life-cycle Hypothesis.

First, note that income can be very low and consumption still be positive. Consider the extreme case where income is zero and the individual consumes a positive amount. If the individual pays sales taxes on any of that consumption, then the effective sales tax is infinite. Following the strict measurement of progressivity, this individual would have an infinite effective sales tax rate – the most regressive tax possibly achievable! This has nothing to do with an individual’s ability to pay the tax; it’s just a feature of life-cycle economic behavior. Although it is an extreme case numerically, this example could easily apply to people who are economically disadvantaged or students simply trying to smooth consumption over the life-cycle.

Second, note that if an individual’s income were to rise in a single period and fall again to the usual income profile in the graph, then consumption would hardly be affected. An individual seeking to smooth consumption over the life-cycle would simply save most of the residual income and spread the transitory gain over the remaining years. However, if the individual’s entire income profile were to rise, then we should expect consumption to

rise considerably more by some proportion. Another Nobel laureate, Milton Friedman, popularized these concepts earlier in the 1950's, and the theory is called the Permanent Income Hypothesis (PIH).⁷ The PIH states that individuals make consumption decisions based on total expected lifetime income, rather than income earned in a particular year. Since income variability throughout an individual's lifetime may reflect persistent changes or idiosyncratic instances, observed consumption may say more about income than income says about consumption. To that extent, progressivity measurements by periodic income are an arbitrary measure of equity.

Elements of both the LCH and the PIH are represented in Keynes' broader theory of decreasing marginal propensity to consume. Both theories show why studying consumption and income in a single period to determine an individual's consumption tax burden is an incomplete exercise in public economics. Instead, what matters for measuring consumption tax progressivity involves a comparison between lifetime income and lifetime taxes paid. To that end, the goal of this study is to project lifetime income profiles using panel data to estimate income variability and pair the income profiles with the consumption patterns of households with similar incomes at each age. These quantities can then be used to measure lifetime sales tax progressivity.

Quantitative Analysis

Simulating Lifetime Income, Lifetime Consumption, and Lifetime Taxes

The process of measuring lifetime sales tax progressivity starts with simulating income over a lifetime for a large number of households. Details of this process and estimation are included in the appendix. Each point in the simulated panel includes the household's age and annual income. Average simulated income is plotted over the life-cycle in Figure 4 and generates the shape premised by the LCH.

For each of these data points, the next step is finding an actual household in the CEX which is an exact age match and an approximate income match. For an exact age and an

⁷ Friedman, Milton. "The permanent income hypothesis." In *A theory of the consumption function*, pp. 20-37. Princeton University Press, 1957.

approximate income interval (i.e., some defined interval around the simulated income point), many observations may match the criteria, and, indeed, this is the hope. In that case, an observation is chosen randomly. If no observations match the criteria, the interval expands iteratively until a match is found.

Once this process is complete, the computation has generated a simulated income profile over each age of the lifecycle, as well as a consumption bundle corresponding to each age-income pair for a large number of simulated households. The next step is to parse a set of expenditure items included in the tax base and to apply a sales tax rate on these items.⁸ This step generates a sales tax bill for each age and income in a simulated agent’s lifetime. In the last step, incomes are summed across all ages for each simulated household, and the same is done for tax bills. This results in a lifetime income and lifetime sales tax bill for each simulated household, as shown in Figure 5. To generate a lifetime effective sales tax rate, we simply divide the lifetime sales tax bill by income, as shown in Figure 6.

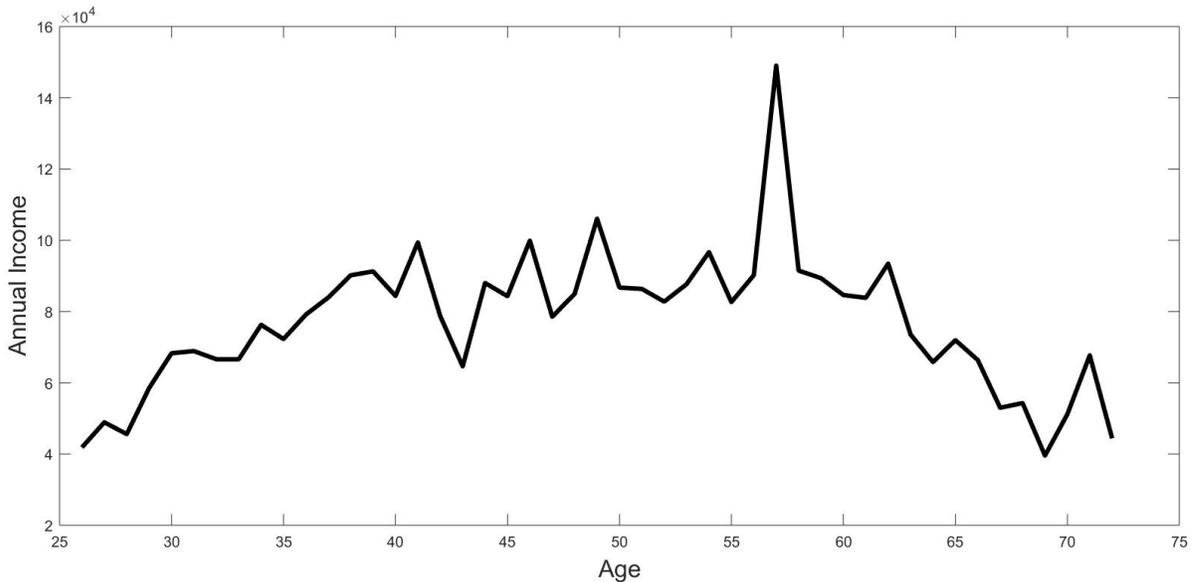


Figure 4: Average simulated income over the life-cycle.

⁸ We apply a sales tax of 8 percent.

Results

Figure 5 presents the results of the simulation by relating total lifetime income and total lifetime sales taxes. The graph shows positive relationship, implying that higher lifetime income households pay more sales taxes over their lifetimes. A simple linear regression suggests that for every \$10 increase in household lifetime income, the household pays 1.7 cents in lifetime sales tax.⁹

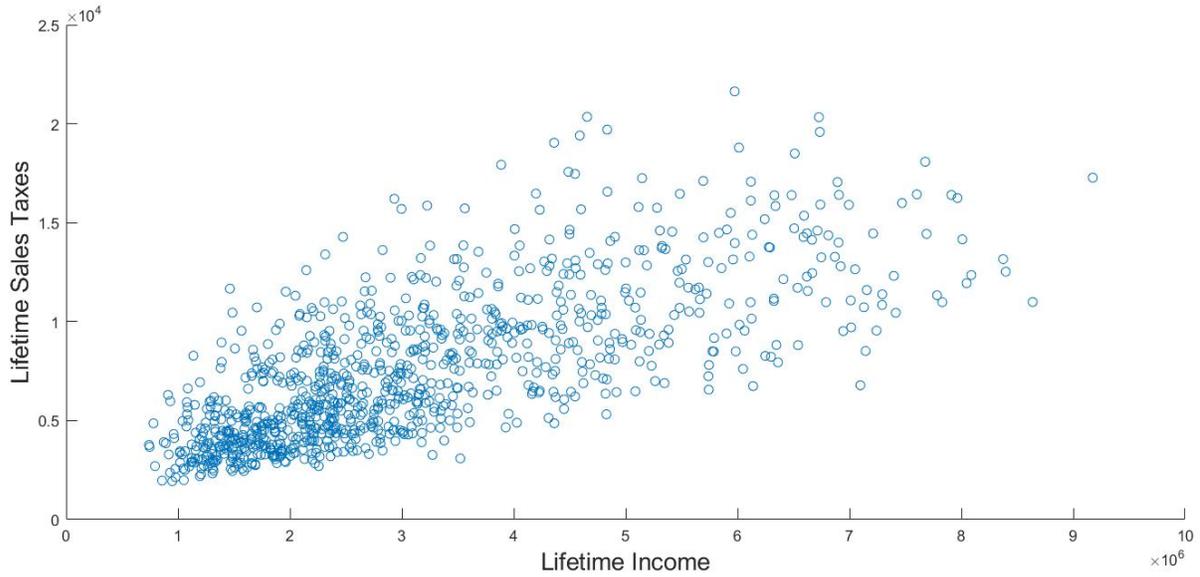


Figure 5: Total lifetime sales taxes paid and total lifetime income.

Although Figure 5 shows total sales tax payments increasing with lifetime income, this comparison does not imply progressivity. To measure progressivity, lifetime sales taxes paid must be divided by lifetime income and plotted against lifetime income; in other words, we must consider effective sales taxes. Figure 6 shows the relationship between simulated lifetime income and simulated lifetime effective sales taxes, which measures lifetime sales tax progressivity. The coefficient on the linear term of the estimated ordinary least square regression equation is small and negative, suggesting that the lifetime sales tax in this example is slightly regressive.¹⁰

⁹ The estimated linear equation is: $y = 0.0017x + 2.3e3$.

¹⁰ The estimated linear equation is: $y = -2.3e-10x + 0.0033$.

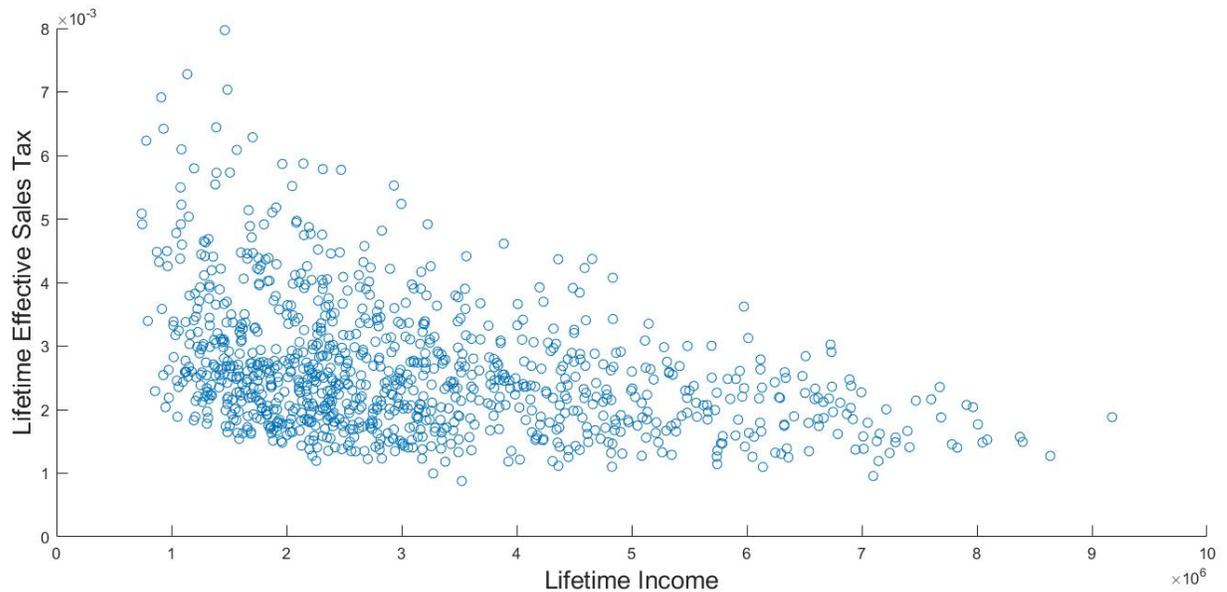


Figure 6: Lifetime sales tax progressivity.

Limitations

While the consumption tax progressivity simulator presented in this study overcomes many of the limitations that inhibit scholars from evaluating sales tax equity, certain other limitations remain. An important limitation is the possibility that deeper dynamic relationships between income and consumption over the life-cycle play a larger role in shaping lifetime sales tax progressivity. For example, we assumed income risk is only dependent on the previous period's income. In reality, some individuals may experience more persistent economic circumstances relative to the ones assumed in the computation. Without richer longitudinal data, this effect is difficult to capture. However, the approach presented in this paper reflects current values and only requires a few years of data.

Conclusion

This paper presents a lifetime consumption tax simulator that overcomes analytical problems caused by consumption tax payer anonymity and theoretical misrepresentations corresponding to periodic consumption tax progressivity measurements. The methodology combines longitudinal income data from the PSID with consumer expenditure data from the CEX to project lifetime consumption taxes paid as it relates to lifetime income. Although the simulator was introduced using a hypothetical tax base for

illustration, the tax base applied could reflect any existing U.S. tax base to measure its lifetime progressivity.

Appendix

Simulating Lifetime Income

To simulate lifetime income, we generate set of Markov transition matrices corresponding to income quintiles for each age cohort. Because we use quintiles, the initial distribution is generated by dividing the unit interval into five disjoint and equally-spaced bins of length 0.20 and separating random draws from a uniform distribution accordingly. Elements of each Markov matrix give the probability of being in a particular quintile in the initial observation and transitioning into a particular quintile in the subsequent observation.¹¹ Because of the computational intensity of the simulation, the number of simulated agents is limited to 1,000.

Following Storesletten et al. (2004), we use the PSID data and define household income in a similar manner.¹² We use the 2013 and 2015 waves of the PSID.¹³ Household earnings are defined by labor income of the household head and spouse plus transfers. To ensure the data is most representative of the national population we use the longitudinal weights provided by the PSID.¹⁴ Similar to Storesletten et al. (2004), we do not include household-year observations that have negative income or that have a change in income by more than a factor of 20.¹⁵ Ultimately, we are able use about 3,000 household observations from the PSID to calculate our Markov matrices.¹⁶

¹¹ A period in this case is two years, corresponding to the frequency of the PSID. Income is assumed to remain constant throughout the period at the 2015 level, and survival is certain.

¹² Storesletten, Kjetil, Christopher I. Telmer, and Amir Yaron. "Consumption and risk sharing over the life cycle." *Journal of monetary Economics* 51, no. 3 (2004): 609-633.

¹³ After 1997, the PSID is conducted only every other year.

¹⁴ Storesletten et al. (2004) corrects for a number of data issues many, if not all, of which are remedied or largely mitigated by using the sample weights.

¹⁵ This would be households whose income increases at least 20 times from the previous panel or whose income decreases to 1/20 or less of its value in the previous panel.

¹⁶ As noted in a later footnote, a larger number of observations are used to calculate income cutoffs between quintiles.

For the purposes of this analysis, quintiles are used (i.e., five quantiles). Let the number of quantiles be represented by N . Let people be indexed in order of income from lowest to highest collectively within the set Z^+ . Let w_i thus represent the weight of the person with the i^{th} highest income. Then the person with the highest income in quantile q is given by p_q where

$$p_q = \max_p \{ p \in Z^+ \mid p = \sum_{i=1}^p w_i \leq \frac{q}{N} \}.$$

We define p_0 as 1 and p_5 is naturally equal to the total number of people. To calculate the income brackets (i.e. quintiles) with maximum accuracy, we use all available households when calculating income cutoffs for a given cohort in a given year. We then apply these income cutoffs to the subset of households available in both years to determine which income bracket these households fall into in a given year for their respective cohort.¹⁷ We use the data to calculate Markov transition matrices for income between the years between 2013 and 2015. Age cohorts are defined by reported age of the household head in 2015. For each cohort ages 26 to 72, we create a Markov matrix, M_a , where a is the age in 2015.¹⁸ An element of the matrix, $M_a(i,j)$, is the weighted percentage of individuals in the i^{th} quantile of cohort a in 2013 who are in the j^{th} quantile of cohort a in 2015. Note that the matrices represent within cohort, not within population, transition percentages.

Simulating Lifetime Consumption

After generating a panel of simulated lifetime income profiles, the next step is to match each age-income realization with a randomly selected individual with the same age and approximately the same income in the Consumer Expenditure Survey (CEX). This approach captures variation in individuals' consumption bundles as they vary with age and income. To match a simulated age-income realization with a corresponding individual in the CEX data, the simulator first attempts to find a randomly-selected observation with an exact age match and income within a narrow range. If the match is unsuccessful, the income range increases iteratively until match is found.

¹⁷ In the PSID, some individuals are only in the 2013 or 2015 sample. We include these individuals when calculating income brackets; however, we cannot use them when calculating the Markov matrices.

¹⁸ These age bounds were chosen, in part, due to data availability.

The CEX dataset contains detailed consumer expenditure data, which allows specification of a subset of consumption items contained within a particular tax base. For illustrative purposes, the sales tax base used in this study is the same set of expenditure items constructed in Barro (2017),¹⁹ which was designed to represent a typical U.S. state or municipal consumption tax base. A proportional tax is applied uniformly to the items in the respective taxable consumption bundle of the individual, and the total tax bill for the year is summed across all items.

In the final step, sales taxes and incomes are summed across each simulated individuals' ages to arrive at lifetime sales tax and lifetime income. For each simulated individual, total lifetime sales tax is divided by total lifetime income, which generates the effective lifetime consumption tax rate. If an individual has consumed zero taxable items throughout the entire lifetime, the observation is treated as survey error and discarded.

¹⁹ This tax base excludes most services, as well as food purchased for home consumption and any items reserved for excise taxation.