

Has Consumption Inequality Mirrored Income Inequality?

Mark Aguiar and Mark Bilal

August 24, 2020

1 Summary and Introduction

The aim of this paper is to examine to what extent the increase in income inequality since 1980 was mirrored by consumption inequality. They find that consumption inequality has tracked income inequality while most of the previous literature argues the opposite; that is, previous to this paper, the literature addressed that consumption inequality has not kept pace with income inequality. For instance, using Consumer Expenditure Survey, Krueger and Perri (2006) find around 33 percent increase between 1980 and 2010 in income and only 11 percent increase in expenditure. There has been a concern about mis-measurement in analysis using Consumption Expenditure Survey. Specifically, discrepancy between aggregate consumption reported in the CE and national income and product account (NIPA) personal consumption expenditures. The mis-measurement is particularly problematic when it is systematic (e.g. mis-measurement tend to occur in high-income household). To mitigate the problem of systematic mis-measurement, this paper consider measuring consumption inequality based on how high- versus low-income households allocate spending toward luxuries versus necessities (double difference).

2 Demand System Estimation

The object of interest in this paper is the inequality measure. Taking income inequality as an example, the inequality measure is defined as follows:

dividing households into 5 bins, the 5–20, 20–40, 40–60, 60–80, and 80–95 percentile groups, compute the ratio of the mean of the top income group to the mean of the bottom income group. The inequality measure is computed from the difference between total expenditure in low-income household and that in high-income household. To estimate this difference, they conduct two-step estimation procedure: they first estimate expenditure elasticities using 1994–1996 data and, then, back up total expenditure from estimated elasticities.

For $h = 1, \dots, H$, $i = 1, \dots, I$, $j = 1, \dots, J$, and t , x_{hjt} and X_{ht} denote reported expenditure of good j and reported total expenditure at time t by household h .

Reported expenditure is represented by multiplication of the true expenditure and measurement errors:

$$x_{hjt} = x_{hjt}^* e^{\zeta_{hjt}}, \quad (1)$$

where x_{hjt}^* denotes the true expenditure and ζ_{hjt} consists of three measurement errors: good-time specific, income-time specific, and residual good-household specific errors. Specifically, ζ_{hjt} is represented by

$$\zeta_{hjt} = \psi_t^j + \phi_t^i + \nu_{hjt} \quad (2)$$

- ψ_t^j : e.g. food may be under-reported for all household.
- ϕ_t^i : e.g. the rich may under-report all expenditures.
- ν_{hjt} : the residual errors, which is assumed to be an non-systematic (independent) measurement error.

The first-order (linear) approximation of the Engel curve:

$$\ln x_{hjt}^* - \ln \bar{x}_{jt}^* = \alpha_{jt}^* + \beta_j \ln X_{ht}^* + \Gamma_j Z_h + \varphi_{hjt} \quad (3)$$

In terms of reported variables,

$$\ln x_{hjt} - \ln \bar{x}_{jt} = \alpha_{jt} + \beta_j \ln X_{ht} + \Gamma_j Z_h + u_{hjt}, \quad (4)$$

where $\alpha_{jt} = \alpha_{jt}^* + \beta_j (\ln X_{ht}^* - \ln X_{ht})$ and

$$u_{hjt} = \phi_t^i + \nu_{hjt} + \varphi_{hjt} \quad (5)$$

The first-stage is to estimate equation (4) using the 1994-1996 Consumer Expenditure Survey. Note that β_j is not estimated consistently if ϕ_t^i is correlated with total expenditure X_{ht} . Therefore, results from estimation procedure is conditional on expenditure elasticities in 1994-1996 (checked in the section discussing robustness).

In the second-stage, total per-household expenditure of each income group is backed up by estimating the inversion of equation (4). From the first-stage regression, two predicted value can be obtained: $\hat{\beta}_j$ and $\hat{x}_{hjt} = \ln x_{hjt} - \ln \bar{x}_{jt} - \hat{\Gamma}_j Z_h$, where \hat{x}_{hjt} is represented by

$$\begin{aligned}\hat{x}_{hjt} &= \alpha_{jt} + \phi_t^i + \beta_j \ln X_{ht}^* + \varphi_{hjt} + \nu_{hjt} \\ &= \alpha_{jt} + \phi_t^i + \beta_j \ln X_{it}^* + \beta_j (\ln X_{ht}^* - \ln X_{it}^*) + \varphi_{hjt} + \nu_{hjt} \\ &= \alpha_{jt} + \phi_t^i + \beta_j \ln X_{it}^* + \varepsilon_{hjt},\end{aligned}\tag{6}$$

where $\varepsilon_{hjt} = \beta_j (\ln X_{ht}^* - \ln X_{it}^*) + \varphi_{hjt} + \nu_{hjt}$. They implement this estimation with further linearity assumption. By regressing \hat{x}_{hjt} on $\hat{\beta}_j \times D_{it}$ and dummies, differences of total expenditure between income group i and income group 1, $\delta_{it} = \ln X_{it}^* - \ln X_{1t}^*$, can be estimated. Note that the second-stage is conducted with the sample without 1994-1996 to avoid correlated sample error.

3 Results

TABLE 1—TRENDS IN INEQUALITY: RATIO OF HIGH-INCOME TO LOW-INCOME RESPONDENTS

	1980–1982	1991–1993	2005–2007	2008–2010	log change 1980–1982/ 2005–2007	log change 2005–2007/ 2008–2010
Labor earnings	6.41	8.47	7.88	8.59	0.21	0.09
Before-tax income	4.75	5.80	6.40	6.50	0.30	0.02
After-tax income	4.21	5.12	5.87	5.92	0.33	0.01
Consumption expenditures	2.47	2.77	2.93	2.77	0.17	-0.06
Non-durable expenditures	2.31	2.58	2.76	2.62	0.18	-0.05

Notes: High income refers to respondents who report before-tax household income in the eightieth through ninety-fifth percentiles. Low income refers to respondents in the fifth through twentieth percentiles. The elements of the first three columns are the ratio of the average of high-income respondents to the average for low-income respondents, where the averages are taken over the pooled years indicated at the head of the respective column. The last two columns are the log difference of the first and third columns and the third and fourth columns, respectively. All variables are converted into constant dollars before averaging. Definitions of each series and sample construction are given in the data section.

TABLE 3—TRENDS IN CONSUMPTION INEQUALITY BASED ON RELATIVE EXPENDITURE PATTERNS

	(1)	(2)	(3)	(4)	(5)
log inequality, 1980–1982	0.85 (0.07)	0.90 (0.06)	0.82 (0.08)	0.71 (0.05)	0.91 (0.06)
log change, 1980–1982/1991–1993	0.27 (0.08)	0.17 (0.06)	0.20 (0.07)	0.27 (0.06)	0.15 (0.07)
log change, 1980–1982/2005–2007	0.48 (0.08)	0.35 (0.07)	0.43 (0.08)	0.46 (0.06)	0.30 (0.07)
log change, 2005–2007/2008–2010	-0.06 (0.08)	-0.04 (0.06)	-0.05 (0.08)	-0.05 (0.06)	-0.04 (0.06)
Categories included	All	All	All	Without durables	Without tobacco
Specification	OLS	WLS	WLS	WLS	WLS
First-stage instrument	Income	Income	Lagged expenditure	Income	Income

Notes: This table reports the estimated change in consumption inequality for top versus bottom income quintiles obtained from the second-stage regressions. Column 3 uses the first-stage estimated expenditure elasticities reported in column II of Table 2, while all other specifications use the column I estimates. The estimated parameters in the first row represent log inequality between the high-income and low-income households in 1980–1982. The next three rows represent the relative growth in total expenditure for high-income households relative to low-income households for the period specified. See the specification in the text for full details. The first column implements the second stage by OLS while the remaining columns implement weighted least squares, using the average NIPA shares for 1980–2010 as weights. For column 4, the weights are adjusted by multiplying the NIPA shares by the average share of each category's expenditure that is nondurable in our CE sample. The standard errors are calculated using a bootstrap with 100 replications.