SPURIOUS CORRELATIONS IN QUANTILE-BASED CONSUMPTION SPILLOVER TESTS

by

Han Wang

A dissertation submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Economics.

Spring 2021

SPURIOUS CORRELATIONS IN QUANTILE-BASED CONSUMPTION SPILLOVER TESTS

by

Han Wang

Approved: ____
Michael A. Arnold, Ph.D.
Chair of the Department of Economics

Approved: ____
Bruce W. Weber, Ph.D.
Dean of the Lerner College of Business and Economics

Approved: ____
Louis F. Rossi, Ph.D.
Vice Provost for Graduate and Professional Education and
Dean of the Graduate College

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: ____

Olga Gorbachev, Ph.D.
Professor in charge of dissertation

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: ____

Jeremy Tobacman, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: ____

Desmond J. Toohey, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: ____

Terry Campbell, Ph.D.
Member of dissertation committee

ACKNOWLEDGMENTS

I am extremely grateful to my doctoral dissertation committee, Drs. Olga Gorbachev, Jeremy Tobacman, Desmond Toohey, and Terry Campbell. I am also grateful to Drs. David Stockman and Matthew White, who guided me through the process of finding a dissertation topic. I also thank the tutors at the Writing Center at the University of Delaware, the participants of Fall Econ Graduate Research Symposium of the Department of Economics at the University of Delaware in 2017 and 2018, the participants of the 9th Annual Graduate Students' Forum at the University of Delaware in 2019, and the participants of the Econ Seminar of Department of Economics at University of Delaware in 2020. Finally, I would like to acknowledge the financial support from the Graduate College at the University of Delaware.

TABLE OF CONTENTS

Appendix

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Consumption spillovers are difficult to estimate. Many tests in the literature argue that spillovers cause positive correlations between individual consumption levels and aggregate income quantiles. This paper develops simulation-based procedures for evaluating reduced-form tests for consumption spillovers. I find that the correlation found in prior tests may be spurious, arising from the mechanical relationship between a household's income in a given period and a quantile of the income distribution in that period. This paper also explores the mechanical correlation's determinants and proposes strategies for estimating unbiased consumption spillover effects.

Chapter 1

INTRODUCTION

Income inequality and household debt-to-income ratio have been rising since before the 2008 financial crisis. The income share of the top 10% of the U.S. income distribution rose from 33% to 50% from 1978 to 2007 (Piketty and Saez 2003). Notably, there was a six percentage-point jump during the five-year economic expansion before the 2008 financial crisis[1] . At the same time, the household debt-to-income ratio was also increasing. The ratio increased from 80% to over 170% from 1978 to 2007 (Mian and Sufi 2010), and there was also a jump by nearly 50 percentage points between 2001 and 2007. The consensus in the literature is that a larger credit-to-GDP ratio, especially a larger household debt-to-income ratio, can lead to a larger probability of a financial crisis[2].

Consumption spillover may help explain how the rising income inequality caused higher household leverage and thus the 2008 financial crisis. For instance, according to the expenditure cascade hypothesis (Frank 2014), households have a preference for comparing themselves with the people right above them in the income distribution ("comparing upwards"). This preference can cause over-consumption and financial distress of the non-rich as the income concentration at the top of the income distribution rises[3], but the empirical evidence for this theory is mixed[4]. Nevertheless, there is clearly evidence that other types of consumption spillovers have negative externalities. For example, people may compare themselves with the lottery winner in their neighborhood. Empirical evidence shows that this preference can increase one's consumption (Kuhn et al. 2011) and debts (Agarwal, Mikhed, and Scholnick 2020). Consumption spillovers may also happen between an individual and the group to which this individual belongs. The group may consist of the people with the same geographic location, employer, or any other demographic characteristics as the individual. Empirical evidence shows that this comparison may increase one's consumption (De Giorgi, Frederiksen, and Pistaferri 2020; Charles, Hurst, and Roussanov 2009) but lowers one's well-being (Luttmer 2005; Guven and Sørensen 2012; Daly, Wilson, and Johnson 2013).

Many tests in the literature argue that spillovers cause positive correlations between individual consumption levels and aggregate income quantiles (and their transformations). I will call these tests "quantile-based consumption spillover tests." Drechsel-Grau and Schmid (2014) use the German Socio-Economic Panel (SOEP) data over 2002-2011 and divide the sample into eleven classes according to consumption for each year. The reference group of an observation is the observations in all the higher classes than the class of the observation. The authors find a positive correlation between the growth in a household's consumption and the growth in the consumption of its reference group. Quintana-Domeque and Wohlfart (2016) use data from British Household Panel Survey over 1998-2008 and find that a household's food consumption away from home is positively correlated with the average food consumption of the top 20% of the income distribution of the household's county in any given year. Bertrand and Morse (2016), using the data from Consumer Expenditure Survey over 1980-2008, show that the household consumption of the bottom 80 percent of the income distribution in a state-year cell is positively correlated with the average of the 80th percentile of the income distribution in that state in years t, t – 1, and t - 2.

Are these correlations spurious? After all, household income is mechanically correlated with aggregate income quantiles. Manski (1993) and Angrist (2014) show that an individual variable y_{it} in a certain group t is mechanically correlated with the group mean μ_t in various specifications. However, if the individual variable y_{it} follows normal distributions $N(\mu_t, \sigma_t^2)$ in each group t, the p^{th} percentile of y_{it}, such that $0 < p < 100$, in the group is a function of the group mean and variance:

$$p^{th}\ percentile\ of\ y_{it} = \mu_t + c_p \sigma_t \qquad (1)$$

where the group mean μ_t and group standard deviation σ_t are random variables, and $c_p = \sqrt{2}\,\mathrm{erf}^{-1}\left(\frac{2p}{100} - 1\right)$ is a constant determined by the given percentage rank p. For example, if $p = 80$, the formula above can be simplified as

$$80^{th}\ percentile\ of\ y_{it} \approx \mu_t + 0.842 \cdot \sigma_t$$

Therefore, as long as the variance of the second term in equation (1), $var(c_p \sigma_t)$, is not extremely large, the individual variable y_{it} will be mechanically correlated with its p^{th} percentile in the group t given that the variable y_{it} is mechanically correlated with its group mean μ_t.

I construct a simulation to show that this mechanical correlation exists, and thus the correlation between household consumption levels and aggregate income quantiles can be spurious. This simulation does not feature consumption spillover, but it can be shown that the individual consumption is still statistically correlated with the aggregate income quantiles, indicating that the correlation must be spurious. The simulation is constructed as follows:

$$lnC_{ist} = 0.7 lnY_{ist} + 0.3\epsilon_{ist}, \quad lnY_{ist} \sim N(0,1)\ and\ \epsilon_{ist} \sim N(0,1) \qquad (2)$$

All the households i in a period t are divided into different groups according to the household's location s. There are 50 periods, 50 locations, and 1,000 observations in each location-period cell. The logarithm of income lnY of a household i in a location-period cell (s, t) is drawn from a standard normal distribution, and log consumption of the household lnC is a linear combination of the log income and a shock ϵ_{ist} drawn from another standard normal distribution, independent of the income. The income elasticity of consumption is set to be 0.7. The aggregate income quantile of interest is the 80th percentile of the income distribution of the location-period cell to which the observation belongs. These percentiles are calculated using the same method in Stata[5]. In a group with 1,000 observations, the 80th percentile of the income distribution is equal to the average income of the observations whose percentage ranks are 80% and 80.1%.

Figure 1 shows that there is a mechanical correlation between household incomes (the left panel) and aggregate income percentiles and this mechanical correlation biases the consumption spillover tests (the right panel). The figure plots the logarithm of the cell-level 80th percentiles against the observation-level log incomes (the left panel) and the log consumption (the right panel), respectively. The blue crosses represent all the observations in both panels. The red rhombuses represent the observations whose incomes are between the 70th and 90th percentiles of the income distributions of location-period cells to which the observation belongs. The green straight line in the left panel is a 45-degree line, on which the log income of an observation is equal to the log 80th percentile of income distribution of the location-period cell to which the observation belongs. As it is shown, the maroon area in the left panel indicates a strong correlation between the logarithm of the cell-level 80th

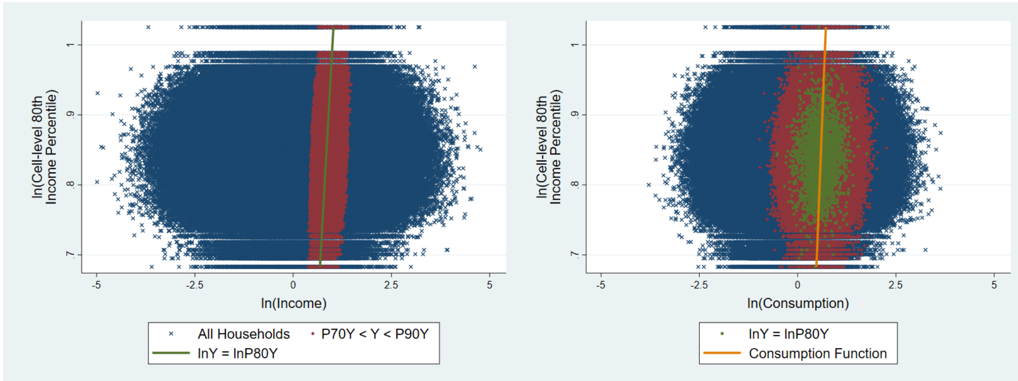

Figure 1: The Mechanical Correlation. The left panel illustrates the mechanical correlation between the log household income and the log 80th percentile of the income distribution in a location-period cell. The right panel shows the mechanical correlation bias in the quantile-based consumption spillover tests.

percentile and the observation-level log income for the observations whose incomes fall in the 20-percentile neighborhood around the percentile. It is reasonable to expect that the estimated coefficient is equal to one.

Similarly, the right panel illustrates a strong correlation between the logarithm of the cell-level 80th percentile of income distribution and the observation-level log consumption. The green squares consist of the data points whose incomes are the 80th percentile of the income distribution in each cell. The yellow straight line in the right panel represents the consumption function. On this line, the log consumption of every data point is equal to 70% of the log 80th percentile of the cell-level income distribution to which the data point belongs. As it is shown, the maroon area and the green area indicate a strong correlation between the observation-level log consumption and the logarithm of the cell-level 80th percentile of the income distribution in the neighborhood of this percentile, with an estimated coefficient expected to be 0.7.

The corresponding estimation results of Figure 1, presented in Table 1, confirm the mechanical correlation. Columns 1 to 3 reports the estimation results for all the regressions of household income, while columns 5 to 7 the regressions of consumption. The specifications read:

$$lnY_{ist} = \phi + \psi ln\bar{Y}(80thIncomePercentile)_{st} + \epsilon_{ist}$$
$$lnC_{ist} = \alpha + \beta\ ln\bar{Y}(80thIncomePercentile)_{st} + v_{ist} \quad (3)$$

where $ln\bar{Y}(80thIncomePercentile)_{st}$ denotes the reference variable, i.e., the log 80th percentile of the income distribution of the location-period cell to which the observation belongs. The regressions in each block of Table 1 differ in their samples. The sample of the first column in each block is all the observations. The observations whose incomes are between their corresponding cell-level 70th and 90th percentiles comprise the sample in the second column of each block. The sample of the third column is the set of all the 80th percentiles of the cell-level income distribution. As shown in each block of Table 1, the estimated coefficients of the percentile approach the expected value as the range of household income narrows down the 80th percentile. In columns 1 to 3, the estimated coefficient increases from 0.491 to 1.000. In columns 5 to 7, the estimated coefficient increases from 0.344 to 0.719.

So far, I have shown that the individual income levels are mechanically correlated with aggregate income quantiles and that the estimate of the consumption spillover effects using an aggregate quantile as the independent variable may be spurious. I will call the effect of this mechanical correlation on the estimate of consumption spillover effects in the quantile-based regression the "mechanical correlation bias in the quantile-based consumption spillover tests." Now the questions are 1) what factors can affect this bias? 2) can the transformations of the quantile and

Table 1: OLS Estimation of the Mechanical Correlation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent Variable		*ln(Income)*				*ln(Consumption)*		
Sample:	All Incomes	P70 < Y < P90	Y = P80	Y < P80	All Incomes	P70 < Y < P90	Y = P80	Y < P80
lnY(The 80th Income Percentile)	0.482***	0.871***	1.000***	0.407***	0.339***	0.618***	0.667***	0.287***
	(0.0141)	(0.00676)	(1.74e-08)	(0.0120)	(0.0107)	(0.0106)	(0.136)	(0.00966)
Constant	-0.405***	0.127***	2.52e-08*	-0.693***	-0.285***	0.0825***	0.0331	-0.487***
	(0.0119)	(0.00569)	(1.47e-08)	(0.0101)	(0.00903)	(0.00892)	(0.115)	(0.00813)
R-squared	0.000467	0.0323	1.000	0.000573	0.000400	0.00678	0.00951	0.000441
Observations	2502500	497851	2500	2000544	2502500	497851	2500	2000544

Notes: The dependent variables in columns 1 – 4 and 5 – 8 are log household income and log household consumption in a period, respectively. The independent variable is the log 80th percentile of the income distribution of the location-period cell to which the observation belongs. Regression samples are defined using these cell-level percentiles. "P80 Neighbors" represent the set of observations used to calculate the 80th percentile. Standard errors are in parenthesis. ***, **, * denote statistical significance at the 1, 5, and 10% level, respectively.

other variants to the simple univariate regression applied by the three quantile-based consumption spillover tests above eliminate this bias from estimating the consumption spillover effects? 3) can any single technique embedded in a full specification eliminate the bias if the full specification cannot? 4) can a new combination of those single techniques eliminate the bias if none of the single transformations can?

The discussion on the mechanical correlation bias in the econometrics literature focuses on the relationship between an individual outcome and the average outcome or other average characteristics of the group to which the individual belongs (Manski 1993; Angrist 2014). While Manski (1993) is mainly focused on the difficulty in identifying peer effects from other effects, the author points out one case where the estimated peer effect is a mechanical phenomenon. Built on Manski (1993), Angrist (2014) lists more cases and show that the estimated coefficients in these cases are simply functions of these two parameters: 1) the estimated coefficient in an OLS regression of individual outcome on individual characteristics, and 2) the estimated coefficient in a two-stage least squares (2SLS) regression of individual outcome on average group characteristics instrumented by group dummies. Angrist (2014) suggests that a clear distinction between the subjects (i.e., the people affected by the peer effects) and the peers (i.e., the people providing the peer effects) is important to avoid the mechanical correlation. This distinction implies two conditions: the exogenous condition and the separation condition. The exogenous condition requires that the peer characteristics are independent of the characteristics of the subjects. The separation condition requires that the econometrician include only the subjects into the estimation sample, leaving the peers out of the sample. Among the three quantile-based consumption spillover tests that I mentioned above, the method in Drechsel-Grau and Schmid (2014) fails both conditions. The methods in Quintana-Domeque and Wohlfart (2016) and Bertrand and Morse (2016) satisfy the two conditions theoretically, but the simulations below show that these two methods still cannot reduce the mechanical correlation bias to an economically insignificant level in some cases.

To the best of my knowledge, this is the first paper discussing the mechanical correlation bias in *quantile-based* consumption spillover tests, and I find the factors that affect this bias. I show that the technique of lagged reference variable proposed by Quintana-Domeque and Wohlfart (2016) and the technique of out-of-regression sample embedded in Bertrand and Morse (2016) do not eliminate the bias in some cases. I evaluate the performance of the three quantile-based consumption spillover tests in Drechsel-Grau and Schmid (2014), Quintana-Domeque and Wohlfart (2016), and Bertrand and Morse (2016) using the realistic simulated data that matches the CEX data with respect to the key moments that affect the size of the mechanical correlation bias. The test in Bertrand and Morse (2016) is the best among the three at minimizing the mechanical correlation bias. Having analyzed all the techniques embedded in the three tests, I propose a new method, which reduces the bias to the level of one-tenth of a percent. Using the new test, I find no evidence of consumption spillovers in the CEX data.

Chapter 2 discusses the four determinants of the mechanical correlation bias in quantile-based consumption spillover tests. Chapter 3 summarizes the differences among the three quantile-based consumption spillover tests in Drechsel-Grau and Schmid (2014), Quintana-Domeque and Wohlfart (2016), and Bertrand and Morse (2016). Chapter 4 examines the performance of the techniques embedded in the three tests in a simple simulation where two important parameters of the simulation are set to their extreme values. Chapter 5 re-examines the performance of these techniques in a complex simulation where the simulated data is calibrated to match the CEX data. Eventually, I propose my methods to minimize the bias and apply the new test to the CEX data. Chapter 6 concludes.

ENDNOTES

[1] Interestingly, the level of this share in 2007 was about the same as the level was in 1928, right before the Great Depression.

[2] Schularick and Taylor (2012), Bordo and Meissner (2012), and Perugini, Hölscher, and Collie (2015) show that debt-to-GDP ratio is the best predictor of a financial crisis. Jordà, Schularick, and Taylor (2016), Büyükkarabacak and Valev (2010), Mian and Sufi (2010) and Mian, Sufi, and Verner (2017) show that household credit is more important than business credit or public debt to predict a financial crisis..

[3] See Alvarez-Cuadrado and Japaridze (2017), Belabed, Theobald, and van Treeck (2018), and Cardaci (2018) for theoretical models.

[4] For international data, see Bordo and Meissner (2012), Perugini, Hölscher, and Collie (2015), and Stockhammer and Wildauer (2018); for U.S. household data, see Wildauer (2016), Thompson (2018), Coibion et al. (2016), and Georgarakos, Haliassos, and Pasini (2014).

[5] In Stata, to calculate the 80th percentile of the income distribution, observations are ranked according to their incomes first. Then, the 80th percentile is the income of the first observation whose cumulative weight exceeds 80 if there is no such an observation whose income rank is strictly equal to 80. In the other case, where there *is* such an observation, the 80th percentile is the average of the income of this observation and the income of this observation's next-higher neighbor.

Chapter 2

DETERMINANTS OF THE "BETA HAT"

In the previous chapter, I showed that household income is mechanically correlated with aggregate income quantiles. Also, this mechanical correlation can bias the estimate of consumption spillover effects. In this chapter, I will explore which factors determine the size of this bias. With the determinants of the bias held constant, I will examine which technique can eliminate the bias from a quantile-based consumption spillover test in later chapters.

I find that four factors can affect the estimated coefficient on the percentile of the cell-level income distribution in the regression of household consumption: the range of household income that defines the estimation sample, the weight of measurement error in the log income, the between-group variance of group-level means of the log income, and the between-group variance of group-level variances of the log income.

Now I will use simulations to illustrate the effects of these determinants. First, I need to construct a baseline simulation and estimate a baseline regression. The baseline simulation here is the same as the previous one, as described by equation (2). This baseline regression is also the same as the previous regression of consumption, as described by equation (3), except that the regression sample here is the set of the observations whose incomes are less than the 80^{th} percentiles of the income distributions of their cells. The baseline regression reads:

$$lnC_{ist} = \alpha + \beta\, ln\bar{Y}(80thIncomePercentile)_{st} + \epsilon_{ist},$$

$$if\ Y_{ist} < \bar{Y}(80thIncomePercentile)_{st} \tag{4}$$

I will call the estimated coefficient on the variable of interest, i.e., the percentile or its transformation, the "beta hat" in the rest of the paper. The estimation results of the baseline regression are presented in column 8 of Table 1. The beta hat is about 0.3. To explore the effects of the four determinants, I will perform four experiments, each being a different variation to the baseline simulation or the baseline regression.

The first experiment shows that the beta hat approaches the income elasticity of consumption as the range of household income used to define the regression sample narrows down to the percentile of the independent variable. In this experiment, the simulation is the same as the baseline. Multiple regressions are estimated using different samples, which are defined as ranges of household incomes. The generic specification reads:

$$lnC_{ist} = \alpha + \beta\, ln\bar{Y}(80thIncomePercentile)_{st} + \epsilon_{ist},$$

$$if\ \bar{Y}\left(X^{th}IncomePercentile\right)_{st} \leq Y_{ist} \leq \bar{Y}\left(Z^{th}IncomePercentile\right)_{st} \tag{5}$$

The 80th percentile of the cell-level income distribution is always a boundary of all these ranges (with only one exception). The other boundary increases from the minimum income to the maximum income of the cell ten percentiles at a time. At the minimum (or the "0th percentile"), the range is the incomes less than the 80th percentile. At the maximum (or the "100th percentile"), the range is all the incomes higher than the 80th percentile. At the 80th percentile, the sample is the set of observations used to calculate the 80th percentiles in all the cells. At any other percentile, the range is the incomes in-between the other boundary and the 80th percentile. The regression at the 0th percentile is the same as the baseline regression.

The estimated coefficients and their 95% confidence intervals are displayed in Figure 2 for the regressions in this experiment. The beta hat declines from 0.7 as the

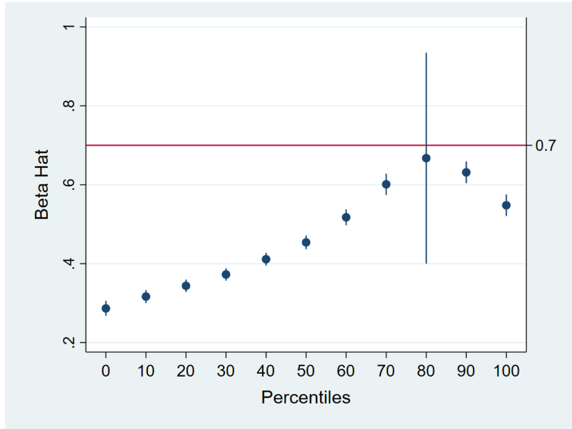

Figure 2: Range of Household Income and The Beta Hat. This figure shows how the beta hat approaches 0.7 (the income elasticity of consumption) as the regression sample, defined as a range of household income, narrows down to the 80th percentile of the cell-level income distribution. A percentile on the horizontal axis represents a certain range that is defined using the percentile

boundaries of the range move away from the 80^{th} percentile. One needs to go back to Figure 1 to see the underlying reason. In the left panel, the two boundaries of the maroon area seem to be clear-cut straight lines, parallel to the 45-degree line, while the two boundaries of the blue area seem to be noisy, curved, and parallel to the y-axis. These boundaries indicate that a percentile will become less correlated with the 80^{th} percentile as it moves away from the 80^{th} percentile. The same pattern applies to the scatter plot of consumption in the right panel. In conclusion, researchers must make sure that the range that is used to define the regression sample is symmetric relative to the median in order to avoid this kind of mechanical correlation. I will choose the 20th and 80th percentiles as the boundaries of the sample for the regression that I will propose in chapter 5.

The second experiment shows that the beta hat declines as the weight of measurement error in household income increases. In this experiment, a noise term is added to the original income as follows:

$$lnY_{ist}^{*} = \sqrt{1 - b^2} \cdot lnY_{ist} + b \cdot u_{ist},$$

$$u_{ist} \sim N\,(0,1)\ and\ cor(lnY_{ist}, u_{ist}) = 0 \quad (6)$$

The original income Y is considered as the true income, according to which the consumption is calculated in the same way as the baseline simulation. The noise term u_{ist} is considered the measurement error, following the same standard normal distribution as the true log income lnY, but the two variables are independent. The observed log income lnY^{*} is a weighted sum of the true log income and the noise term, and the sum of squares of the two weights is equal to one. I call the coefficient on the error term b "the weight of measurement error." This setup ensures that the observed log income has the same between-group variance of group-level means and between-group variance of the group-level variance as the true log income. A series of simulations are performed, using different weights for the measurement error. The weight runs from 0 to 1 with an increment of 0.1. The regression for this experiment is the same as the baseline regression, except that the true income is replaced by the observed income everywhere in the regression.

As the weight of the measurement error in log income goes to one, the independent variable in the baseline regression, i.e., a percentile of the income distribution, will approach a complete noise. Therefore, the beta hat will approach zero. The estimation results are displayed in Figure 3. The estimated coefficients and their 95% confidence intervals are on the vertical axis, and the weight of measurement error labels the horizontal axis. In the case of weight equal to 0, the regression is

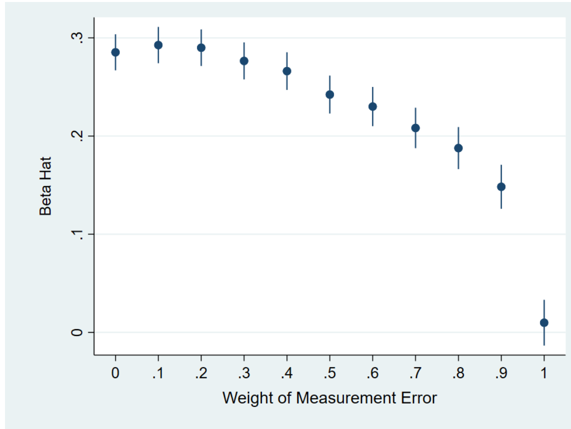

Figure 3: Weight of Measurement Error and The Beta Hat. The coefficient b on the horizontal axis represents the weight of measurement error, as defined in equation (6). This figure shows that the beta hat decreases as the weight increases.

identical to the baseline regression. As it is shown in the figure, beta hat increases at first but immediately begins to decline as the weight of the measurement error increases.

To show why the beta hat eventually declines, I compare the scatter plots for consumption between the baseline simulation and the simulation in which the weight of measurement error is equal to 0.9 in Figure 4. In the left panel, the green area is symmetric relative to the consumption function. However, in the right panel, due to the measurement error, the green area becomes less symmetric relative to the consumption function but more symmetric relative to the vertical axis. The same situation that occurs to the consumption function and the green dots, which are only associated with the 80^{th} percentile, can be applied to all the quantiles. Therefore, the maroon area, which represents the regression sample, is more symmetric relative to the vertical axis in the right panel than that in the left panel. Therefore, the fitted line is more vertical in the right panel than the one in the left panel. In conclusion, an increase in the weight of measurement error will make the axis of symmetry of the regression sample closer to the vertical axis, thus lowering the beta hat. Therefore, the weight of measurement error will be an important consideration when I construct the simulations in chapters 4 and 5.

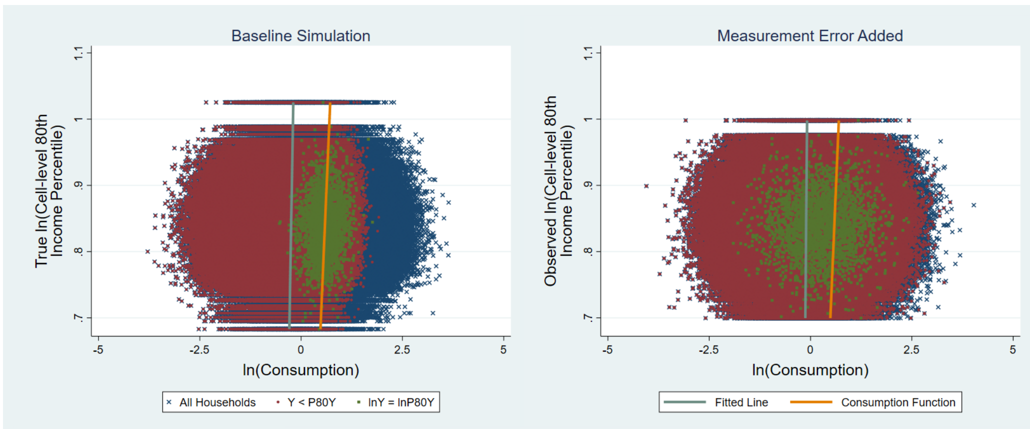

Figure 4: Weight of Measurement Error – An Example. This figure illustrates the measurement error in log household income lowers the beta hat. The left panel shows the case in which there is no measurement error, while the right panel shows the case in which measurement error added in the way described by equation 6 and the weight is 0.9.

The third experiment shows that the beta hat approaches the income elasticity of consumption as the between-group variance of group-level means of the log household income goes to infinity. To show that, I construct twelve simulations using different levels of the between-group variance of group-level means. The data generating process is:

$$lnY_{istj} \sim N(\mu_{stj},\ 1\ ),\quad \mu_{st} \sim N\left(0,\ m_j^2\right)\ and\ m_j^2 = 0,\ 10^{-4} \times 2^0,\ \ldots,\ 10^{-4} \times 2^{10}$$

In each simulation j, the cell-level means of log incomes μ differ across location-time cells (s, t), while the cell-level variance is set to one for all the cells. Specifically, the mean of a cell is a random variable whose value is drawn from a normal distribution. The mean of the distribution is set to zero for all the simulations, but the variance m^2 can take different values across these simulations. Twelve simulations are constructed, with the between-group variance of group-level means running from 0 to 0.1024 exponentially. In the case where the between-group variance is equal to zero, the simulation reduces to the baseline simulation where there is no variation in group-level means. Each simulation of the series has a corresponding regression, whose specification is the same as the baseline regression.

The between-cell variance of cell-level means here m_j^2 is equivalent to the between-group variance of group-level means $var(\mu_t)$ in equation (1). An increase in this type of variance, with the variance of the second term of the equation held constant, enhances the correlation between the mean and the percentile, and thus enhances the correlation between the individual variable lnY_{ist} and its cell-level percentile for the whole sample. Since the individual log income lnY_{ist} is correlated with the individual log consumption lnC_{ist}, the correlation between the individual log consumption in the whole sample and the log 80th percentile of the cell-level income distribution will also be strengthened. As we will see, this also applies to the observations below the 80th percentile of the cell-level income distribution if there is no between-group variation in group-level variances.

The estimation results of the entire series of this experiment are plotted in Figure 5. As predicted, the confidence interval becomes smaller and smaller as the between-group variance of group-level means of log household income increases.

More importantly, the figure shows that the beta hat approaches 0.7, which is equal to the income elasticity of consumption chosen for the consumption function. Next, I illustrate how this variance affects the beta hat and its confidence interval using an example.

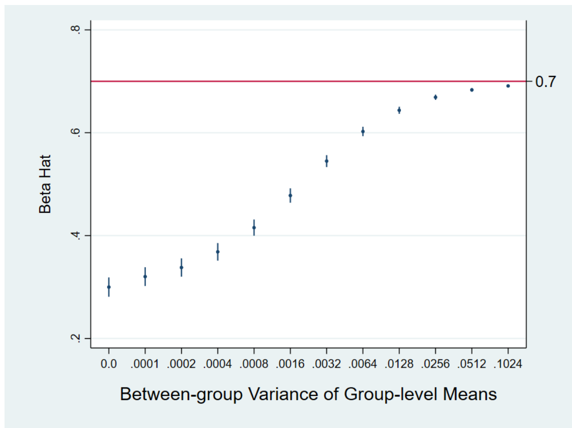

Figure 5: Between-group Variance of Group-level Means and The Beta Hat. The horizontal axis represents the between-group variance of group-level means of log household income. This figure shows that the beta hat approaches 0.7 (the income inelasticity of consumption) as this variance goes to infinity.

Figure 6 shows that the between-group variation in group-level means "stretches" the scatter plot along with the consumption function and makes the axis of symmetry of the graph parallel to the consumption function. In this example simulation, the between-group variance of group-level means is set to one ($m^2 = 1$), which is much greater than the maximum value in the previous series of simulations. The scatter plots of the baseline simulation and the example simulation for consumption are shown in the left and the right panels, respectively, in Figure 6. The teal lines in the two panels are the fitted lines of each baseline regression. The plot of the baseline simulation here shows a different shape compared with its plot in the right panel of Figure 1. That is because the ranges of y-axes are different in the two panels. The range in Figure 6 is determined by the example simulation, in which the 80th percentiles of cell-level income distributions have a more significant variance than the baseline simulation.

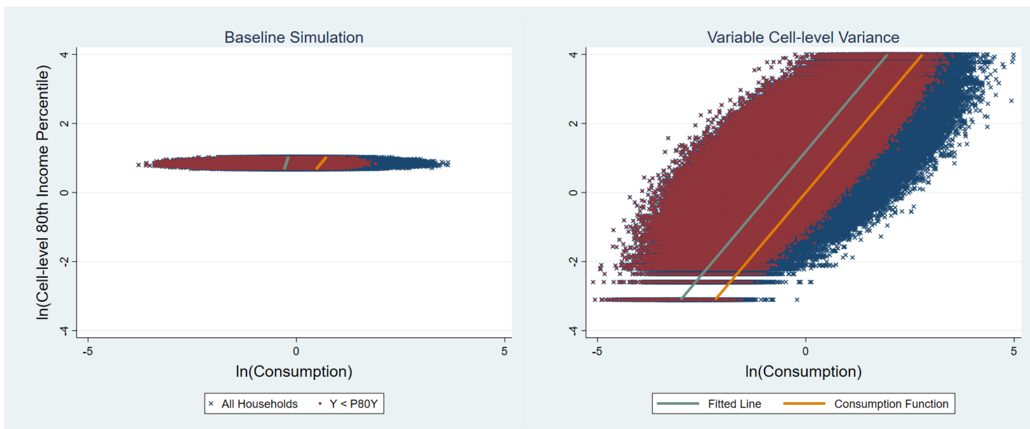

Figure 6: Between-group Variance of Group-level Means and Beta Hat – An Example. This figure illustrates that the between-group variation in the group-level means of log household income raises the beta hat. The left panel shows the case in which there is no such between-group variation, while the right panel shows the case in which there is such variation.

The graph of the example simulation looks like a transformation of the baseline simulation. First, the means of the whole sample, zero, are equal between the two simulations. Second, it looks like the range and the variance of each cell in the example simulation are still equal to those of the corresponding cell in the baseline simulation. Third, cell-level means of the example simulation, and thus the cell-level

80th percentiles, vary across cells to a larger extent than the baseline simulation. Therefore, holding the cell-level variance constant, the variation in the first term of equation (1) dominates the variation in the second term, which explains why the mechanical correlation becomes stronger for the observations in the whole sample, represented by the blue dots. As to the estimation sample of the baseline regression, we can see from the right panel that the log 80^{th} percentile, compared with the left panel, is also more correlated with the cell-level means of the maroon dots, which explains why the confidence interval in Figure 5 becomes smaller as the between-group variance of group-level means increases.

To understand the slope of the fitted line, note that the two boundaries of blue area in the left panel are more vertical than the boundaries in the right panel. Therefore, the fitted line of the blue area in the left panel is more likely to be vertical than that in the right panel. Also, the axis of symmetry of the blue area in the right panel (the line passing through all the group means, which is not plotted in the figure) is parallel to the consumption function, so the fitted line of the maroon area in the right panel will also be parallel to the consumption function. This perspective explains why the beta hat approaches the income elasticity of consumption (0.7).

In conclusion, the level of the beta hat will approach the income elasticity of consumption, and the confidence interval of the beta hat in the baseline regression will become smaller, as the between-group variance of group-level means in individual log income goes to infinity. Due to its effect, this variance will be an important factor to consider when constructing the simulations in chapters 4 and 5.

The fourth experiment shows that the beta hat approaches a negative value (or a positive value) as the between-group variance of group-level variances of log household income goes to zero (or infinity). In this experiment, I construct eleven simulations using multiple between-group variances of the group-level variances. The data generating process is:

$$lnY_{istj} \sim N(0, \sigma^2_{stj}), \quad \sigma^2_{stj} \sim \chi^2(k_j) \; and \; 2k_j = 2 \times 3^0, \; \dots, \; 2 \times 3^{10}$$

In each simulation j, the mean of log income is held constant at zero for all the cells, while the group-level variance σ^2 can change across location-period cells (s, j). The variance of a cell is drawn from a chi-squared distribution, whose degree of freedom k_j can take different values in different simulations. In other words, different simulations differ in the between-group variance of group-level means k and the between-group variance of group-level variances $2k$. Even if the variance $2k$ is a one-to-one function of the mean k, as we will see below, it is the variance $2k$ that plays a major role in this experiment instead of the mean k. Eleven simulations are constructed, and the eleven between-group variances of group-level variances form a geometric series with a coefficient of 2 and a common ratio of 3. The specification of the regressions in this experiment is the same as the baseline regression.

The between-cell variance of cell-level variances of log household income $2k$, or $var(\sigma^2_{st})$, is positively associated with the variance of the second term in equation (1) $var(c_p\sigma_t)$. Therefore, holding the variance of the first term in equation (1) constant, an increase in the variation in the second term will weaken the correlation between the group-level mean and the group-level quantile, and thus weaken the correlation between the individual-level variable lnY_{ist} (and thus individual log consumption lnC_{ist}) and the group-level quantile for the whole sample. As shown below, this rationale can also be applied to a sub-sample that is defined by a quantile.

The estimation results of the entire series are plotted in Figure 7. The confidence interval becomes smaller as the between-cell variance of the cell-level variances of log household income decreases. More importantly, the beta hat

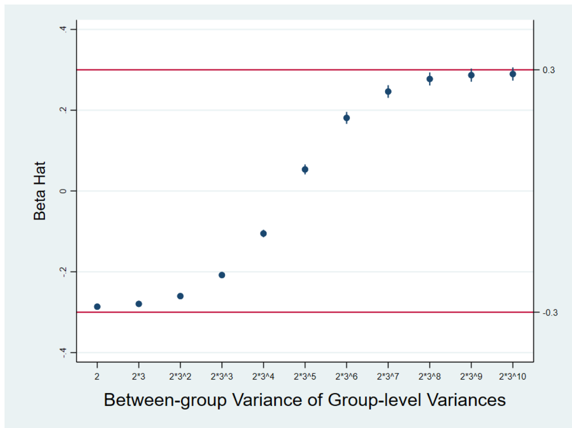

Figure 7: Between-group Variance of Group-level Variances and The Beta Hat. The horizontal axis represents the between-group variance of group-level variances of log household income. This figure shows the beta hat approaches a negative value as this variance goes to zero

approaches a negative number (or a positive number) as the between-group variance of the group-level variance goes to zero (or infinity). Note that the limiting case at the left endpoint here is not the baseline simulation where the variance of log income is one. In this limiting case, both the between-group mean k and the between-group variance $2k$ decline to zero. Since any group-level variance must not be negative, a zero-mean implies that group-level variances should be zero in all the groups. The limiting case at the right endpoint is not identical to the baseline simulation ex-ante. The baseline simulation has no ex-ante between-group variation in group-level variances, while the between-group variances of group-level variances $2k$ in the simulations of this experiment are strictly greater than zero. However, this limiting case, as shown below, will behave similarly to the baseline simulation in some respects ex-post. Equation (1) does not explain why the beta hat in the baseline regression approaches a negative number (or a positive number) when the between-group variance of the group-level variance goes to zero (or infinity). Next, I will use an example simulation to illustrate the underlying reason.

Figure 8 shows that increases in the between-group variation in group-level variances $2k$ let the scatter plot change from a triangle to an olive. This change, along with the asymmetric range of the regression sample, causes the change in the beta hat

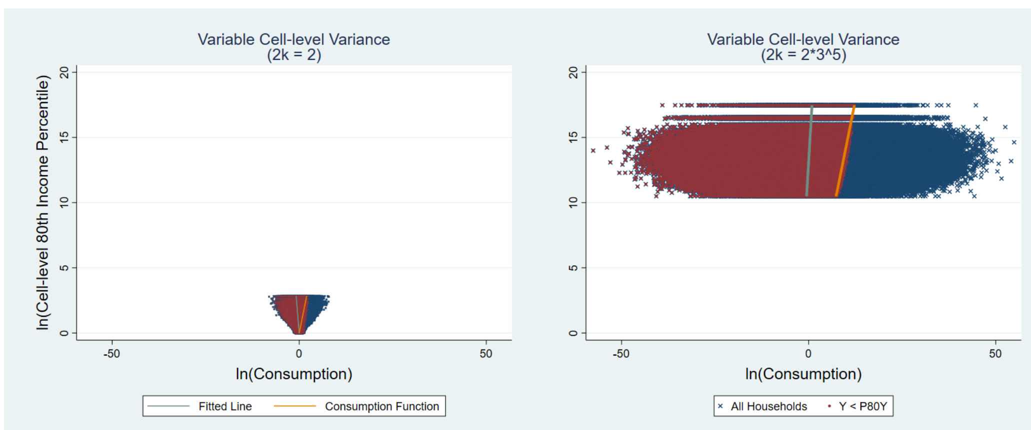

Figure 8: Between-group Variance of Group-level Variances and The Beta Hat – An Example. This figure illustrates that the between-group variation in group-level variances of log household income reduces the beta hat. The left panel shows the case in which there is little such between-group variation, while the right panel shows the case in which there is plenty.

as the variance $2k$ changes. Figure 8 represents the scatter plots of two example simulations. The two between-group variances are equal to 2 and 2×3^5, respectively.

As we can see, a smaller between-group variance of group-level variances $2k$ makes the blue area a triangle in the left panel, and a larger variance $2k$ makes an olive in the right panel. The mechanical correlation in the maroon area is stronger in the left panel than the right panel because the left boundary of the maroon area is a clear-cut straight line in the left panel. On the contrary, the data points on the left boundary in the right panel are noisier due to a bigger between-group variance of group-level variances $2k$. Also, the fitted line of these points would be curved. As to the right boundaries, they are clear-cut straight lines in both panels. These two features of the left boundary explain why the confidence interval of the beta hat increases with the variance $2k$.

If I were to fit a linear line for the data points on the left boundaries in both panels, one could see that the fitted line in the left panel is flatter than the right boundary and that the fitted line in the right panel is nearly vertical. The right boundary in the right panel is closer to the consumption function than the one in the left panel, but the slopes in both panels are positive and do not differ so much as the left boundaries. Therefore, the difference in the slope of the left boundary explains why the beta hats are negative in some simulations while positive in others. Suppose I were to continue increasing the variance $2k$. In that case, the graph of the whole sample should become longitudinal symmetric with the axis of symmetry being the between-group median of the log group-level 80th percentiles of the income distribution, just like the olive shapes in Figure 1 (This presumably results from the fact that the limit of a chi-squared distribution is approximately a normal distribution). As a result, the slope of the fitted line of the regression sample would approach a value similar to the beta hat in the baseline simulation as the variance $2k$ goes to infinity.

The change in the shape of the whole graph, along with the asymmetric range of the regression sample, explains the change in the beta hat as the variance $2k$ increases.

In conclusion, the beta hat in the baseline regression approaches a negative number (or a positive number), and the confidence interval becomes smaller (or bigger), as the between-group variance of group-level variances in log household income goes to zero (or infinity). Due to its impact, this variance and the range of household income that defines the estimation sample will be important factors to consider when constructing the simulations and analyzing techniques embedded in the three quantile-based consumption spillover tests in chapters 4 and 5.

In this chapter, I have shown that the size of the mechanical correlation bias in a quantile-based regression is affected by these four factors: the range of household income that defines a regression sample, the weight of measurement error in log household income, the between-group variance of group-level means in log household income, and the between-group variance of group-level variances in log household income. I will hold these factors constant when comparing the effects of different techniques on the beta hat in the chapters below.

Chapter 3

SUMMARY OF THE QUANTILE-BASED CONSUMPTION SPILLOVER TESTS IN THE LITERATURE

So far, I have shown that the quantile-based consumption spillover tests can be biased by the mechanical correlation between household incomes and aggregate quantiles of the income distribution of the group to which the observation belongs. I have also found out the determinants of this mechanical correlation bias. In this chapter, I will break down the full specifications of three quantile-based consumption spillover tests in the current literature into different transformations of the quantile and other variants to the baseline regression, i.e., the embedded techniques, and then analyze the efficacy of these techniques in the chapters below.

Three studies use quantiles or their transformations to test the consumption spillover hypotheses: Drechsel-Grau and Schmid (2014), Quintana-Domeque and Wohlfart (2016), and Bertrand and Morse (2016). I will call these tests DGS, QDW, and BM, respectively. Their methods differ in many respects: the form of the dependent variable, how to define reference groups, how to evaluate the reference group, whether use instrumental variables, what other controls to include, and how to define the regression sample. Table 3 presents a summary of these tests.

In DGS and QDW, the dependent variable is the change in the logarithm of household consumption, while it is the level of the log household consumption in BM. The ways that these authors define reference groups differ in the following dimensions: whether observations are ranked by income or consumption, whether the

Table 2: Summary of the Three Quantile-based Consumption Spillover Tests

	Drechsel-Grau and Schmid (2014)	Quintana-Domeque and Wohlfart (2016)	Bertrand and Morse (2016)
Dependent variable	*growth* in household consumption	*growth in household consumption*	log household consumption
Reference group	the group of observations in all the *higher consumption* classes of the *national* population in the given period	the group of observations in the *top 20%* of local income distribution in the given period	the group of observations used to calculate the 80th (and the 20th, 50th) percentiles of the local income distribution in the recent three periods from *another sample*
Reference variable	consumption growth of the reference group	*the predicted consumption* growth of the reference group	log income of the reference group
Instrumental variable	N/A	*the lagged* average income of the reference group, *the lagged* growth in the reference group's threshold	N/A
Other Controls	HOP and its interaction with the reference variable, household income growth, location and time fixed effects, and demographics	location and time fixed effects, and demographics	income bucket fixed effects, local unemployment rate, location-specific time trends, cluster at the local level, location and time fixed effects, and demographics
Sample	less than the 95th percentile of the national consumption distribution in the given period	less than the 80th percentile of the local income distribution in the given period	less than the 80th percentile of the local income distribution in the given period

reference group is an interval or a point of distribution, whether the reference group is relative or absolute, whether the reference group is defined locally or nationally, whether the reference group is generated within the sample or out of the sample, and whether the reference group includes previous periods. DGS uses consumption to rank observations, while QDW and BM use income. In DGS and QDW, a reference group is an interval, i.e., a group of observations, while it is a point in the distribution, i.e., a percentile, in BM. The reference groups are relative in DGS but absolute in QDW and BM. Specifically, the rank of a household's relative reference group changes as the rank of the household in question changes, while the rank of an absolute reference group is held constant. DGS ranks the observations into eleven classes using quantiles. Observations below the ninth decile are divided into nine classes according to the nine deciles, and the observations above are divided into two classes using the 95th percentile. Each household's reference group consists of all the households that are in the higher classes. The only exception is that the top 5% do not have a reference group. On the contrary, QDW and BM use a percentile to divide the population into two groups, with one group being the sample of the regression and the other group or its threshold being the reference group.

The reference groups in these three papers also differ in other aspects. DGS ranks observations in the entire country (Germany) all together while QDW and BM rank the observations in each local geographic area (counties in the United Kingdom and states in the United States) separately. DGS and QDW define the reference groups by using the observations that consist of the regression sample, while BM takes the reference variable from a second sample. Specifically, the regression sample of BM is from the Consumer Expenditure Survey (CEX), while the percentiles of income distributions of BM are taken from the Current Population Survey (CPS). In DGS and QDW, the reference group only involves the current period, but the reference group in BM comprises the 80th percentiles in the recent three periods. Finally, the reference groups in these papers are, respectively, the set of observations in all the higher consumption classes of the national population in the given period (DGS), the set of observations in the top 20% of the local income distribution in the given period (QDW), and the set of observations from another sample used to calculate the 80th percentile of the local income distribution in the recent three periods (BM)

The variable that these researchers use to evaluate their reference groups may differ from the variable to rank the observations. These two types of variables are the same in DGS (consumption) and BM (income), while QDW uses incomes to rank the observations but consumption to evaluate the reference groups. In all three papers, the form of the reference variables is the same as their dependent variables. DGS and QDW use growth, while BM uses the level of the logarithm. Evaluating the reference group of DGS and BM generates the reference variable, i.e., the independent variables of interest, of the two tests. The reference variables in the two tests are the change in the logarithm of the average consumption of all the higher classes than the one to which the observation belongs in the national consumption distribution of a given period (DGS), and the logarithm of the average 80th percentiles of the local income distribution in periods t, t-1, and t-2 obtained from the second sub-sample (BM), respectively.

The regression sample in DGS is different from those in QDW and BM. In DGS, the regression sample comprises observations whose consumption is less than the 95th percentile of the national consumption distribution in the given period. In QDW and BM, the regression sample is a set of observations whose incomes are less than the 80th percentile of the income distribution of the location-period cell to which the observation belongs. The percentile in BM is calculated from a different sample.

As to controls, all the tests control for demographics, location fixed effects, and time fixed effects. Other controls in DGS include HOP, its interaction with the reference variable, and household income growth. HOP is an indicator variable that is equal to one when a household changes its class from the previous period and is zero if not. BM controls for reference variables at the other two quantiles: the 20th and 50th percentiles. BM also controls for a series of location-specific time trends, a series of income bucket dummies, and local unemployment rates. The standard errors are clustered at the local level in BM.

In this chapter, I list the differences between the specifications of the three consumption spillover tests that use a transformation of quantiles as the variable of interest. Next, I will show the effectiveness of these differences (or embedded techniques) in lowering the beta hat using an extreme simulation in chapter 4 and realistic simulations in chapter 5 and find out which technique I can use.

Chapter 4

BREAKDOWN OF THE FOUR TESTS IN AN EXTREME SIMULATION

So far, I have shown the mechanical correlation can bias the estimate of consumption spillover effects using a quantile-based regression. I have pointed out which factor can affect this bias. Also, I have listed all the techniques embedded in the current quantile-based consumption spillover tests. In this chapter, I will show the efficacy of these techniques in an extreme simulation. The simulation is the extreme one because both the between-group variance of group-level means and the between-group variance of group-level variances of the log household income are at their extreme levels. I will say that "a technique can lower the beta hat to a level of moderate economic significance" if the order of the magnitude of the absolute value of the beta hat in the regression with this technique is negative two. I will say that "a technique can lower the beta hat to a level of hardly any economic significance" if the order is less than negative two.

4.1 Setup of the extreme simulation and preview of the results

To determine which technique above can eliminate the mechanical correlation bias, I construct a new simulation on top of the baseline simulation:

$$lnC_{ist} = 0.7lnY_{ist} + 0.3\epsilon_{ist},\ lnY_{it} \sim N(\mu_{st},\ \sigma_{st}^2\ )\ \ \mu_{st} \sim N\ (0,\ 1)\ \ \sigma_{st}^2 \sim \chi^2\ (1)\ \ \epsilon_{ist} \sim N(0,\ 1) \quad (7)$$

$$lnY_{ist}^* = 0.8lnY_{ist} + 0.6\nu_{ist},\quad \nu_{ist} \sim N\ (\mu_{st},\ \sigma_{st}^2) \quad (8)$$

This simulation adds three additional features to the baseline simulation for log income: measurement error, between-group variation in group-level means, and between-group variation in group variances. In this new simulation, the number of location-period cells and the number of households in each location-period cell are the same as those in the baseline simulation described by equation (2). The mean and variance of the log household incomes are cell-specific. Cell-level means μ_{st} are drawn from the standard normal distribution, which leads to a sufficiently high between-cell variance of the cell-level means and guarantees a positive, mechanical correlation if everything else is controlled for. Cell-level variances are drawn from the chi-square distribution of degree one so that the between-cell variance of the cell-level variances is sufficiently low, which will generate a negative mechanical correlation if everything else is controlled for. All the parameters in the consumption function are the same as those in the baseline simulation. The measurement error v_{ist} is added to the income in the same way as before: the sum of squares of the two weights is equal to one. The weight of the error term b is 0.6. The error term in the given cell follows the same distribution as the log income in the given cell. I will use the observed income Y^* for all the regressions in this chapter.

The summary statistics are presented in Table 3. As it is shown, the error term increases the between-cell variance of the cell-level means of log income from 0.98 to 1.92, whose effects on the estimate are limited since 0.98 is already sufficiently high. The between-cell variance of the cell-level variances of the observed log income remains the same as the true log income.

Before going to the details of all the single techniques, let us look at how the full specifications of the three tests work with the simulated data. It turns out none of them produce a beta hat that is statistically insignificantly different from zero. The estimation results are presented in Table 4. The beta hat of DGS is 1.070, which is the

Table 3: Summary Statistics of the Extreme Simulation

	Mean	S.d.	Var
		Individual-level	
ln(True Income)	-0.00924	1.41	1.98
ln(Observed Income)	-0.0132	1.71	2.92
Observations	2502500		
		Cell-Level	
Between-cell mean of ln(True Income)	-0.00924	0.99	0.98
Between-cell s.d. of ln(True Income)	0.803	0.60	0.36
Between-cell variance of ln(True Income)	1.005	1.40	1.96
Between-cell mean of ln(Observed Income)	-0.0132	1.39	1.92
Between-cell s.d. of ln(Observed Income)	0.803	0.60	0.36
Between-cell variance of ln(Observed Income)	1.005	1.40	1.95
Observations	2500		

Notes: The data is generated by an extreme simulation. The data generating process is described by equations (7) and (8).

highest one among all the regressions. The beta hat of QDW, 0.562, is in the middle. In BM, the beta hat of the 80^{th} percentile is -0.0397, which has the lowest absolute value among all three regressions. The absolute values of the coefficients on the other two percentiles in BM are at the same order of magnitude as its 80^{th} percentile. All three of the estimated coefficients in BM are statistically significantly different from zero, even though it is only moderately significant in the economic sense. None of the three tests discussed here can reduce the mechanical correlation bias to an economically insignificant level in this extreme simulation.

Next, I am going to break down these full specifications to multiple variants to their corresponding baseline regressions and find out which embedded technique reduces, maintains, or even increases the level of the beta hat compared to the level in their baseline regressions. The embedded techniques of the three quantile-based consumption spillover tests will be examined in the order of DGS, QDW, and BM. For each test, I will run a baseline regression first. Then, I will apply one technique at a time in the baseline regression to check how the new beta hat compares to the

Table 4: The Three Tests in the Extreme Simulation

	(1)	(2)	(3)
Dependent variable	*Dln(Consumption)*	*Dln(Consumption)*	*ln(Consumption)*
Sample:	C<P95C	Y<P80Y	Y1<P80Y2
Unit of group	*Period*	*Location-period Cell*	*Location-period Cell*
	DGS	QDW	BM
DlnC(All the higher consumption classes of the entire population)	1.070***		
	(0.00396)		
HatDlnC(Top 20% Incomes)		0.562***	
		(0.000966)	
lnY(The 80th income percentiles in the recent three periods of sample 2)			-0.0397***
			(0.00893)
lnY(The 20th income percentiles in the recent three periods of sample 2)			-0.0338***
			(0.0105)
lnY(The 50th income percentiles in the recent three periods of sample 2)			0.0662***
			(0.0186)
Half of the whole sample	No	No	Yes
Location and period FEs	Yes	Yes	Yes
State-specific trends	No	No	Yes
lnY, DlnY, or income bucket FEs	Yes	No	Yes
Reference variable*HOP	Yes	No	No
Instrumental variable	No	Yes	No
Clustered s.e.	No	No	Yes
R^2	0.869	0.211	0.698
Observations	2213267	1922131	961386

Notes: The data is generated by an extreme simulation. The data generating process is described by equations (7) and (8). The dependent variable in columns 1 – 2 (or column 3) is the change in (or the level of) the logarithm of household consumption in a period. The regression sample in column 1 (or columns 2 - 3) is the set of observations whose consumption (or income) less than the 95th (or 80th) percentile of the consumption (or income) distribution of the time group (location-period cell) to which the observation belongs. A group within which the observations are ranked to get the distribution in column 1 (or columns 2 – 3) is the set of observations in the given period in the whole country (or the local area in which the observation resides). For each reference variable, the reference group is in the parenthesis. A "percentile" in the reference group indicates the observations used to calculate the percentile. The variable used to evaluate the reference group and the operators of the logarithm ("ln"), the first difference ("D"), and the predicted value ("Hat") are outside the parenthesis. The reference variables in the three columns are the change in the logarithm of the average consumption of all the higher classes than the one to which the observation belongs in the national consumption distribution of a given period (DGS), the predicted value of the change in the logarithm of the average consumption of the observations whose incomes are greater than the 80th percentile of the income distribution of the location-period cell to which the observation belongs (QDW), and the logarithm of the average 80th (and 20th, 50th) percentiles of the local income distribution in periods t, t-1, and t-2 obtained from the second sub-sample (BM), respectively. The instruments in QDW are the lagged average income of the reference group and the lagged growth in the threshold of the reference group. Standard errors are in parenthesis. ***, **, * denote statistical significance at the 1, 5, and 10% level, respectively.

original one. The baseline regressions of all the tests are the regression of log household consumption on the 80th percentile of the cell-level income distribution, but there are three nuances. First, in DGS, the 80th percentile is defined within the *national* income distribution, but *local* distributions are used in the other baseline regressions. Second, the sample for the baseline regression of DGS will be the set of the observations whose incomes are less than the 95th percentile of the national income distribution, but the 80th percentile of the local distribution will be used for baseline regressions of the other two tests. Third, the standard errors in the baseline regression of BM are clustered at the local level but not in the other two tests.

4.2 Break Down Drechsel-Grau and Schmid (2014)

Though none of the techniques embedded in DGS can make the beta hat statistically significantly different from zero, three techniques can lower the beta hat: evaluating the reference group by consumption, ranking observations according to their consumption, and controlling for income.

The estimation results are presented in Table 5. In the baseline regression, the dependent variable is the logarithm of household consumption in a period. The reference variable is the logarithm of the 80th percentile of the national income distribution in the given period. The regression sample is the set of observations whose current incomes are less than the current 95th percentile of the national income distribution. The beta hat of the baseline regression (column 1) is 0.298.

When the technique of evaluating the reference group by consumption is examined (column 3), all the other ingredients of the new regression remain the same as the baseline regression. Particularly, the reference group itself still consists of the observations whose incomes are used to calculate the 80th percentile of the national income distribution, but consumption is used to evaluate these observations. Therefore, the reference variable becomes the logarithm of the average consumption of the observations used to calculate the 80th percentile of the national income.

Table 5: Break Down Drechsel-Grau and Schmid (2014)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent variable	*lnC*	*DlnC*	*lnC*	*lnC*	*lnC*	*lnC*	*lnC*	*lnC*	*DlnC*
Sample:	Y<P95Y	Y<P95Y	Y<P95Y	C<P95C	Y<P95Y	Y<P95Y	Y<P95Y	Y<P95Y	C<P95C
Unit of group					*Period*				
	Baseline	First Difference	Evaluated by C	Ranked by C	Interval Reference	Relative Reference	Control for lnY	Fixed Effects	Complete Specification
lnY(The 80th income percentile)	0.298***						-0.0126***	0.299***	
	(0.00241)						(0.00132)	(0.00240)	
DlnY(The 80th income percentile)		0.285***							
		(0.00242)							
lnC(The 80th income percentile)			0.00963***						
			(0.00112)						
lnY(The 80th consumption percentile)				0.0456***					
				(0.000943)					
lnY(Incomes between the 75th and 85th percentiles)					0.298***				
					(0.00242)				
lnY(The next-higher income class threshold)						0.531***			
						(0.00137)			
DlnC(All the higher consumption classes)									1.070***
									(0.00396)
Reference variable*HOP	No	No	No	No	No	Yes	No	No	Yes
lnY or DlnY	No	No	No	No	No	No	Yes	No	Yes
Location and period FEs	No	No	No	No	No	No	No	Yes	Yes
Observations	2377350	2329803	2377350	2377350	2377350	2377350	2377350	2377350	2213267
R-squared	0.00639	0.00592	0.0000310	0.000983	0.00636	0.661	0.706	0.0117	0.869

Notes: The data is generated by an extreme simulation. The data generating process is described by equations (7) and (8). The dependent variable in columns 2 and 9 is the change in the logarithm of the household consumption in a period, while it is the level in other columns. The sample in all the regressions is the set of observations that are lower than the 95th percentile of the national distribution in the given period. The sample is defined by consumption in columns 4 and 9, while by income in other columns. A group within which the observations are ranked to get the distribution in all the columns is the set of observations in the given period in the whole country. For each reference variable, the reference group is in the parenthesis. A "percentile" in the reference group indicates the observations used to calculate the percentile. The variable used to evaluate the reference group and the operators of the logarithm ("ln") and the first difference ("D") are outside the parenthesis. Standard errors are in parenthesis. ***, **, * denote statistical significance at the 1, 5, and 10% level, respectively.

distribution. The beta hat of this new regression decreases from 0.298 in the baseline regression to 0.0904

When the technique of ranking observations by consumption is examined (column 4), the reference group becomes the set of observations used to calculate the 80th percentile of the national consumption distribution in the given period, but it is still evaluated by income. Therefore, the reference variable becomes the log average income of the observations used to calculate the 80th percentile of the national consumption distribution. As for the regression sample, I also replace income with consumption to define it, which becomes the set of the observations whose current consumption is less than the current 95th percentile of the national consumption distribution. The beta hat falls from 0.298 in the baseline regression to 0.0456.

When the technique of controlling for income is examined (column 7), nothing else in the regression changes, except that the logarithm of household income is controlled for. The beta hat drops from 0.298 in the baseline regression to -0.0126.

The techniques of the first difference, interval reference group, fixed effects, and relative reference groups turn out to be ineffective in lowering the beta hat in this extreme simulation. When the technique of the first difference is examined (column 2), both the dependent variable and the reference variable become their first differences. The beta hat slightly decreases from 0.298 in the baseline regression to 0.285. When the technique of interval reference group is applied (column 5), the reference group becomes the set of the observations whose incomes are between the 75th and 85th percentiles of the national income distribution. The reference variable becomes the logarithm of the average income of the reference group. The beta hat in this regression remains the same as in the baseline regression. When the technique of fixed effects is examined (column 8), I control for the location fixed effects and the time fixed effects. The beta hat in this regression is nearly identical to that in the baseline regression.

Out of all the DGS techniques, only one increases the beta hat, which is the technique of relative reference group (column 6). In this specification, the reference group of an observation becomes the next- higher class of the observation. For example, the 20th percentile is the next-higher class for the observations whose incomes are between the 10th and the 20th percentiles. If this household moves into the interval between the 20th and the 30th percentiles in the next period, its reference group will change to the 30th percentile correspondingly. As to the observations between the 90th and 95th percentile, their reference group is the 95th percentile. In the spirit of DGS, I also control for the interaction between the reference variable and the indicator variable HOP. The beta hat in this regression rises from 0.298 in the baseline regression to 0.531.

When all DGS techniques are applied (column 9), the specification reads:

$$\Delta lnC_{ist} = \alpha + \beta \Delta ln\bar{C}\ (HigherConsumptionClasses)_t$$
$$+ \phi \Delta ln\bar{C}\ (HigherConsumptionClass)_t \times HOP_{it} + \psi \Delta lnY^{*}_{ist} + \gamma_s + \delta_t + \epsilon_{ist},$$
$$if\ C_{ist} < 95thConsumptionPercentile_t \qquad (9)$$
$$HOP_{it} = \begin{cases} 1, & if\ \Delta ConsumptionClass_{ist} \neq 0 \\ 0, & if\ \Delta ConsumptionClass_{ist} = 0 \end{cases}$$

The dependent variable becomes the first difference of log household consumption in the current period. The reference variable becomes the change in the logarithm of the average consumption in all the classes of the national consumption distribution that are higher than the one to which the observation belongs in the current period $\Delta ln\bar{C}\ (HigherConsumptionClasses)_t$. Consumption distribution is used to define the regression sample, becoming the set of observations whose current consumption is less than the current 95th percentile of the national consumption distribution. Other controls include an interaction between the reference variable and the indicator variable HOP, household income growth, time dummies, and location dummies. In this regression, the beta hat increases from 0.298 in the baseline regression to 1.070.

This breakdown of DGS indicates that the techniques of evaluating the reference group by consumption, ranking observations according to their consumption, and controlling for household income have the potential to eliminate the mechanical correlation bias. However, the mitigating effects of these techniques may be offset by other techniques, e.g., relative reference groups, which are shown to have an inflating effect.

4.3 Break Down Quintana-Domeque and Wohlfart (2016)

A breakdown of all the techniques embedded in QDW shows that only the techniques of lagged reference variable and interval reference group can lower the beta hat, even if the estimated coefficients are still statistically significantly different from zero.

The estimation results are presented in Table 6. In the baseline regression (column 1), the dependent variable is the logarithm of household consumption in a period, the reference variable is the logarithm of the 80th percentile of the local income distribution in the given period, and the regression sample is the observations whose incomes are less than their reference variables. The beta hat of the baseline regression is 0.419[1].

When lagged reference variable technique is examined (column 6), the reference variable is lagged by one period, becoming the logarithm of the 80th percentile of the income distribution of the location-period cell to which the household belongs in the previous period. All the other ingredients of this specification remain the same as the baseline regression. Particularly, the regression sample is still defined using current values. The beta hat in this regression decreases from 0.419 in the baseline regression to 0.00765. When the interval reference group technique is

Table 6: Break Down Quintana-Domeque and Wohlfart (2016)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent variable	*lnC*	*DlnC*	*lnC*	*lnC*	*lnC*	*lnC*	*lnC*	*lnC*	*DlnC*
Sample:	Y<P80Y	Y<P80Y	Y<P80Y	Y<P80Y	Y<P80Y	Y<P80Y	Y<P80Y	Y<P80Y	Y<P80Y
Unit of group	*Location-period Cell*								
	Baseline	First Different	Level and Difference	Interval Reference	Evaluated by C	Lagged Reference Variable	Instrumental Variables	Fixed Effects	Complete Specification
lnY(The 80th income percentile)	0.419***		0.425***					0.421***	
	(0.000367)		(0.000521)					(0.000372)	
DlnY(The 80th income percentile)		0.429***	-0.00668***						
		(0.000381)	(0.000369)						
lnY(Top 20% incomes)				0.278***					
				(0.000349)					
lnC(The 80th income percentile)					0.539***				
					(0.000664)				
LaglnY(The 80th income percentile)						0.00765***			
						(0.000475)			
HatlnC(The 80th consumption percentile)							0.820***		
							(0.000799)		
HatDlnC(Top 20% Incomes)									0.578***
									(0.00112)
Location and period FEs	No	No	No	No	No	No	No	Yes	Yes
Observations	2000292	1960285	1960285	2000292	2000292	1960285	2000292	2000292	1920280
R-squared	0.395	0.393	0.396	0.241	0.248	0.000133	0.363	0.400	0.209

Notes: The data is generated by an extreme simulation. The data generating process is described by equations (7) and (8). The dependent variable in columns 2 and 9 is the change in the logarithm of the household consumption in a period, while it is the level in other columns. The sample in all the regressions is the set of observations that are lower than the 80th percentile of the income distribution of the location-period cell to which the observation belongs. A group within which the observations are ranked to get the distribution in all the columns is the set of observations in the given period in the local area in which the observation resides. For each reference variable, the reference group is in the parenthesis. A "percentile" in the reference group indicates the observations used to calculate the percentile. The variable used to evaluate the reference group and the operator of the logarithm ("ln"), the first difference ("D"), the lagged value ("Lag"), and the predicted value ("Hat") are outside the parenthesis. The instruments in column 7 are the lagged average income of the reference group and the lagged growth in the threshold of the reference group. Standard errors are in parenthesis. ***, **, * denote statistical significance at the 1, 5, and 10% level, respectively.

examined (column 4), the reference group, following the spirit of QDW, becomes the set of the observations whose incomes are greater than the 80th percentile of the cell-level income distribution. Accordingly, the reference variable becomes the log average income of the new reference group. The beta hat in this regression falls from 0.419 in the baseline regression to 0.278.

The techniques of fixed effects, first difference, using the first difference and level simultaneously, evaluating the reference group by consumption, and instrument variable are not effective in lowering the beta hat in this simulation. The fixed effects technique is applied here (column 8) in the same way as in the breakdown of DGS, which barely changes the beta hat here. In fact, these two techniques slightly increase the beta hat: the first difference and using the level and first difference simultaneously. When both the dependent and independent variables are transformed into their first differences (column 2), the beta hat increases from 0.419 in the baseline regression to 0.429. When "the level reference variable" increases to 0.425 while the estimated coefficient of "the difference reference variable" is -0.00668, both significantly different from zero. The techniques of evaluating the reference group by consumption and instrument variables raise the beta hat to a larger degree. When the reference group is evaluated by consumption (column 5), the beta hat increases from 0.419 in the baseline regression to 0.539. When the technique of instrument variable is examined (column 7), the predicted value of the 80th percentile of the cell-level consumption distribution becomes the new reference variable, which is instrumented by the 80th income percentile. The beta hat in this regression increases from 0.419 in the baseline regression to 0.820.

When all the techniques in QDW are applied (column 9), the specification reads:

$$\Delta lnC_{ist} = \alpha + \beta \Delta ln\widehat{\bar{C}}\,(Top20\%Income^{*})_{st} + \gamma_s + \delta_t + \epsilon_{ist}, \quad if\ Y_{ist} < 80thIncomePercentile^{*}_{st} \qquad (10)$$

$GMM\ first\ step$:

$$\Delta ln\bar{C}\ (Top20\%Income^{*})_{st} = \theta + \phi\ ln\bar{Y}(Top20\%Income\ )^{*}{}_{s,t-1} + \psi\Delta ln\bar{Y}(80thIncomePercentile\ )^{*}{}_{s,t-1} + \nu_{ist}$$

The dependent variable becomes the first difference of the log household consumption in a period. The reference variable becomes the change in the logarithm of the average consumption of the observations whose incomes are greater than the 80th percentile of the income distribution of the location-period cell to which the observation belongs $\Delta ln\bar{C}\ (Top20\%Income^{*})_{st}$, and this reference variable is instrumented by the lagged logarithm of the average income of the top 20 percent $lnY(Top20\%Income\ )^{*}{}_{s,t-1}$ and the lagged growth of the 80th income percentile $lnY(Top20\%Income\ )^{*}{}_{s,t-1}$. Other controls include time dummies, and location dummies. The regression sample still consists of the observations whose incomes are less than the 80th percentiles of the income distribution of the location-period cells to which the observations belong. In this regression, the beta hat increases from 0.419 in the baseline regression to 0.562.

This breakdown of QDW shows that the technique of lagged reference variables has the potential to reduce the mechanical correlation bias. However, the techniques of evaluating the reference group by consumption and instrument variable may enhance the mechanical correlation bias.[2]

4.4 Break Down Bertrand and Morse (2016)

Even though they cannot make the coefficient statistically insignificantly different from zero, a breakdown of all the techniques embedded in BM (Table 7) shows that these three techniques can lower the beta hat: income bucket fixed effects, multiple reference variables, and moving average. In this exercise, the baseline

Table 7: Break Down Bertrand and Morse (2016)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variable	*lnC*	*lnC*	*lnC*	*lnC*	*lnC*	*lnC*	*lnC*
Sample:	Y<P80Y	Y1<P80Y2	Y<P80Y	Y<P80Y	Y<P80Y	Y<P80Y	Y1<P80Y2
Unit of group	*Location-period Cell*						
	Baseline	Exogenous Reference	Income Bucket	3-period Moving Average	Fixed Effects and Trends	Multiple References	Complete Specification
lnY(The 80th income percentile)	0.419***		-0.0365***		0.420***	0.0335***	
	(0.00478)		(0.00237)		(0.00512)	(0.00792)	
lnY(The 80th income percentile of sample 2)		0.418***					
		(0.00468)					
lnY(The 80th income percentiles in the recent three periods)				0.310***			
				(0.0111)			
lnY(The 50th income percentile)						0.200***	
						(0.0162)	
lnY(The 20th income percentile)						0.266***	
						(0.00828)	
lnY(The 80th income percentiles in the recent three periods of sample 2)							-0.0397***
							(0.00893)
lnY(The 20th income percentiles in the recent three periods of sample 2)							-0.0338***
							(0.0105)
lnY(The 50th income percentiles in the recent three periods of sample 2)							0.0662***
							(0.0186)
Cluster s.e.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location and period FEs	No	No	No	No	Yes	No	Yes
Location-specific time trends	No	No	No	No	Yes	No	Yes
Income bucket FEs	No	No	Yes	No	No	No	Yes
Half of the whole sample	No	Yes	No	No	No	No	Yes
R-squared	0.395	0.396	0.699	0.0979	0.402	0.525	0.698
Observations	2000292	1001199	2000292	1920280	2000292	2000292	961386

Notes: The data is generated by an extreme simulation. The data generating process is described by equations (7) and (8). The dependent variable in all the columns is the logarithm of the household consumption in a period. The sample in all the regressions is the set of observations whose incomes are less than the 80th percentile of the income distribution of the location-period cell to which the observation belongs. A group within which the observations are ranked to get the distribution in all the columns is the set of observations in the given period in the local area in which the observation resides. For each reference variable, the reference group is in the parenthesis. A "percentile" in the reference group indicates the observations used to calculate the percentile. The variable used to evaluate the reference group and the logarithm operator ("ln") are outside the parenthesis. All the standard errors are clustered at the local level. Standard errors are in parenthesis. ***, **, * denote statistical significance at the 1, 5, and 10% level, respectively.

regression is the same as the baseline regression of QDW, except that standard errors are cluster at the local level in BM.

To examine the effects of income bucket fixed effects (column 3), I generate 400, which is about the same number as in BM, income buckets for the observations whose incomes are between the 5th and 95th percentiles of the entire income distribution. Indicator variables for these buckets, plus the other two buckets for the top and bottom five percent, are added onto the baseline. The beta hat in this regression decreases from 0.419 in the baseline regression to -0.0365. In the regression with multiple reference variables (column 6), I add the 20th and 50th percentiles of the cell-level income distribution to the baseline regression. The estimated coefficient of the 80th percentile of the cell-level income distribution decreases from 0.419 in the baseline regression to 0.0335. The coefficients on the 20th and 80th percentiles are 0.200 and 0.266, respectively, both of which are statistically significantly different from zero. In the regression of 3-period moving average (column 4), the reference variable becomes the logarithm of the average 80th percentile of the cell-level income distribution in the current period, period t -1, and period t - 2. The beta hat decreases in this regression from 0.419 in the baseline regression to 0.310.

The rest of the techniques in BM do not change the beta hat. When the exogenous reference technique is examined (column 2), I split the whole sample into two sub-samples randomly. One half is used as the regression sample, and the other half is used to calculate the exogenous cell-level 80th percentiles. The regression sample becomes the first sub-sample observations whose incomes are less than the exogenous cell-level 80th percentiles. The beta hat in this regression (0.418) is nearly identical to the beta hat (0.419) in the baseline regression. In the regression of fixed effects and trends (column 5), I control for the location fixed effects, the time fixed effects, and a series of location-specific time trends, which are the interactions between location indicator variables and a time trend. The estimated coefficient of this regression is 0.420, which is practically identical to that in the baseline regression.

When all the techniques are applied simultaneously (column 7), the specification reads:

$$lnC_{ist} = \alpha + \beta_1 l\,n\,\bar{Y}(80thIncomePercentileofSample2\,)^*{}_{t,t-2|s} + \beta_2 l\,n\,\bar{Y}(20thIncomePercentileofSample2\,)^*{}_{t,t-2|s} + \beta_3 l\,n\,\bar{Y}(50thIncomePercentileofSample2\,)^*{}_{t,t-2|s} + \eta^Y_{ist} + t * \gamma_s + \gamma_s + \delta_t + \epsilon_{ist},$$
$$if\ Y^*_{ist} < 80thIncomePercentileofSample2^*_{st} \quad (11)$$

The dependent variable is still the log household consumption in a period. The reference variable becomes the logarithm of the average income of the observations from the second sub-sample that are used to calculate the 80^{th} (and 20^{th}, 50^{th}) percentiles of the local income distribution in period t, t-1, and t-2. Other controls include time dummies, location dummies, and interactions between a time trend and location dummies. The income bucket fixed effects are also controlled for. The regression sample becomes the set of observations in the first-sub sample whose incomes are less than the exogenous 80^{th} percentile of the income distribution of the location-period cell to which the observation belongs. The beta hat of the 80^{th} percentile in this regression falls from 0.419 in the baseline regression to -0.0397. The coefficients of the 20^{th} and 50^{th} percentiles are -0.0338 and 0.0662, respectively. All the coefficients are significantly different from zero.

This breakdown of BM shows that the techniques of income buckets, 3-period moving average, and multiple references have the potential to decrease the mechanical correlation bias for the 80^{th} percentile. No techniques in BM are shown to increase this bias.

The analysis of all the techniques embedded in the three quantile-based consumption spillover tests above shows some interesting results. First, none of the full tests can make the beta hat statistically insignificantly different from zero, but BM can reduce the beta hat to a level that only has moderate economic significance.

Second, none of the single techniques can make the beta hat statistically insignificantly different from zero either, but some of them can at least reduce the beta hat to the level of moderate economic significance. The technique of lagged reference variable in QDW is the most effective in lowering the beta hat. The techniques of controlling for income or income bucket fixed effects, ranking observations according to their consumption, and multiple references are the second effective. Some techniques, including evaluating reference groups by consumption and interval reference group, are effective in one test but ineffective in another one due to some unknown reasons. The technique of a 3-period moving average is only slightly effective. All the other techniques either do not change the level of the beta hat or even increase it.

4.5 Break Down All My Techniques

I propose the following adjustments to the baseline regression: symmetric range, controlling for income, the first difference, lagged reference groups, and ranking observations by consumption. I choose these techniques not according to their performance in this extreme simulation but the more realistic simulations to be shown chapter 5. However, it is interesting to show how these techniques perform in this extreme simulation.

The estimation results are presented in Table 8. The specification and results of the baseline regression (column 1) are the same as that in QDW. In the regression of symmetric range (column 2), the regression sample becomes set of the observations whose incomes are between the 20th and 80th percentiles of income distribution of the location-period cells to which they belong. I choose this technique because an asymmetric range will bias the estimate of the consumption spillover effects, as shown

Table 8: Break Down All My Test

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dependent variable	*lnC*	*lnC*	*lnC*	*DlnC*	*lnC*	*lnC*	*DlnC*	*DlnC*	*DlnC*	*DlnC*
Sample:	Y<P80Y	P20Y< Y<P80Y	Y<P80Y	Y<P80Y	LagY< LagP80Y	C<P80C	LagP20C <LagC< LagP80C	LagP20C <LagC< LagP80C	LagP20C <LagC< LagP80C	LagP20C <LagC< LagP80C
Unit of group	*Location-period Cell*									
	Baseline	Symmetric Range	lnY	First Difference	Lagged Reference Groups	Ranked by C	Full Specification	Cluster	FEs and Trends	Multiple References
lnY(The 80th income percentile)	0.419***	0.445***	-0.0371***							
	(0.000367)	(0.000350)	(0.000393)							
lnY(The 80th income percentile)				0.429***						
				(0.000381)						
lnY(The previous 80th income percentile)					0.331***					
					(0.000366)					
lnY(The 80th consumption percentile)						0.358***				
						(0.000323)				
DlnY(The previous 80th consumption percentile)							0.0256***	0.0256***	0.0264***	0.0153***
							(0.000322)	(0.000783)	(0.000328)	(0.000359)
DlnY(The previous 20th consumption percentile)										0.0131***
										(0.000351)
DlnY(The previous 50th consumption percentile)										0.0151***
										(0.000390)
lnY or DlnY	No	No	Yes	No	No	No	Yes	Yes	Yes	Yes
Cluster s.e.	No	No	No	No	No	No	No	Yes	No	No
Location and time FEs	No	No	No	No	No	No	No	No	Yes	No
Location-specific time trends	No	No	No	No	No	No	No	No	Yes	No
R-squared	0.395	0.519	0.717	0.393	0.295	0.381	0.752	0.752	0.752	0.753
Observations	2000292	1498097	2000292	1960285	1960284	2000505	1468505	1468505	1468505	1468505

Notes: The data is generated by an extreme simulation. The data generating process is described by equations (7) and (8). The dependent variable in columns 4 and 7 - 10 is the change in the logarithm of the household consumption in a period, while it is the level in other columns. The sample in columns 1 – 5 is defined by income, while by consumption in other columns. The sample in columns 1 – 4 and 6 is defined by current values, while previous values in the other columns. The sample in column 2 is defined by the 20th and 80th percentiles, while only by the 80th in the other columns. In all the columns, a group within which the observations are ranked to get the distribution is the set of observations in the given period in the local area in which the observation resides. For each reference variable, the reference group is in the parenthesis. A “percentile” in the reference group indicates the observations used to calculate the percentile. The variable used to evaluate the reference group and the operators of the logarithm (“ln”) and the first difference (“D”) are outside the parenthesis. Standard errors in column 8 are clustered at the local level. Standard errors are in parenthesis. ***, **, * denote statistical significance at the 1, 5, and 10% level, respectively.

in Chapter 2. The beta hat increases from 0.419 in the baseline to 0.445. In the regression of log household income being controlled for (column 3), the beta hat decreases from 0.419 in the baseline to - 0.0375. In the regression of the first difference (column 4), both the dependent and independent variables become their first differences and the estimated coefficient increases from 0.419 in the baseline regression to 0.429. As I will show chapter 5, this technique will decrease the beta hat if the simulation is calibrated to match the actual data. In the regression of lagged reference groups (column 5), the reference group becomes the set of the households used to calculate the 80th percentile in the previous period, but they are still evaluated by the current income. Therefore, the reference variable is the logarithm of the current income of the 80th percentile of the local income distribution in the previous period. The regression sample correspondingly becomes the set of the observations whose lagged incomes are less than the 80th percentile of the local income distribution in the previous period. The beta hat decreases from 0.419 in the baseline regression to 0.381. In the regression of ranking observations by consumption (column 6), the reference group becomes the 80th percentile of the cell-level consumption distribution, but they are still evaluated by income. Therefore, the reference variable becomes the logarithm of the average income of the observations used to calculate the 80th percentile of the cell-level consumption distribution. The regression sample correspondingly becomes the set of the observations whose consumption is less than the 80th percentile of the cell-level consumption distribution. In this regression, the beta hat decreases from 0.419 in the baseline regression to 0.375. When all my techniques are applied simultaneously (column 7), the specification reads:

$$\Delta lnC_{ist} = \beta_0 + \beta_1 \Delta ln\overline{Y^*}(Lag80thConsumptionPercentile\,)_{st} + \beta_2 \Delta lnY^*_{ist} + \epsilon_{ist},$$

$$if\ 20thConsumptionPercentile_{s,t-1} < \ C_{is,t-1} < 80thConsumptionPercentile_{s,t-1} \quad (12)$$

The dependent variable becomes the first difference of log household consumption in a period. The reference variable becomes the first difference of the logarithm of the income of the 80th percentile of the cell-level consumption distribution defined in the previous period. The regression sample becomes the observations whose lagged consumption is between the lagged 20th and the lagged 80th percentiles of the cell-level consumption distributions. In this regression, the estimated coefficient decreases from 0.419 in the baseline regression to 0.0279. The former value has one less order of magnitude than the latter value.

I also examine if some of the other techniques can lower the beta hat further on top of the full specification that I propose. The techniques of clustering the standard errors (column 8) and fixed effects and trends (column 9) do not decrease the beta hat on top of the full specification. The multiple reference variables technique does lower the estimated coefficient of the 80th percentile from 0.0297 in my full specification to 0.0139, but the new beta hat is at the same order of magnitude as the original one.

In this chapter, I have shown that none of the single techniques embedded in the current literature, their full specifications, nor the specification that I propose can lower the beta hat to a level that has barely any statistical significance given a certain level of measurement error, a sufficiently large level of variation in group means, and a sufficiently low level of of variation in group variances. However, some techniques, including the full specification that I propose, can diminish the economic significance of the beta hat to a moderate or even negligible level. It is interesting to see how these techniques perform in a more realistic and more complex simulation.

Chapter 5

BREAKDOWN OF THE FOUR TESTS IN A REALISTIC SIMULATION

So far, I have shown that the mechanical correlation bias exists, pointed out its determinants, listed the techniques embedded in the current quantile-based consumption spillover test, and examined the performance of the four quantile-based consumption spillover tests using an extreme but simple simulation. Given that the estimate may be affected by group variation and measurement error of the data, I will now re-examine the performance using a realistic simulation. The new simulation is more realistic than the extreme simulation because its between-group variances and consumption function will be calibrated to match the actual data. Also, since the true weight of the measurement error is unknown, I will use multiple weights to construct a series of simulations. These simulations are more complex than the extreme simulation because now they feature business cycles, economic growth, and geographical heterogeneity. Since these simulations are calibrated to match the actual data, I will call them realistic simulations. As in the extreme simulation, I will examine the efficacy of the four quantile-based consumption spillover tests by looking at their full specifications and each embedded technique, respectively.

5.1 The CEX Data

I use the data from the Consumer Expenditure Survey (CEX) of the Bureau of Labor Statistics (BLS) from 1996 to 2007. The dataset is a rotating panel, and there are four panels at any given time. Each panel has around 1,500 consumer units. A consumer unit in a panel was interviewed for five consecutive quarters, with the first one being a bounding interview. Every subsequent quarter, one of the four panels will finish their interviews, and a new panel will be rotated in. I use the data of the second and fifth interviews.

I made several modifications for the data to be applicable to the model I propose. I drop the households that have nonpositive income. I use the variable FINCBTAX as household income. FINCBTAX is nominal income before tax in the past 12 months. The CEX only records FINCBTAX in the second and the fifth interviews. I deflate all the monetary variables to the 2017 level using the quarterly division-level Consumer Price Index from the BLS.

I drop the households that do not have location information. The smallest geographically identifiable location information in the CEX available to the public is the state in which the household is located. The location variable that I use is census divisions, partly because "[T]he CE sample was not designed to produce precise estimates for individual states" (U.S. Department of Labor 2008, page 294). I also drop the households whose expenditure is equal to or less than 0. I use the variable ETOTAL, which is the total outlay in the current quarter, as the measure of expenditure. I multiply the quarterly expenditure by four to compare its mean with the annual income. The expenditure not only includes the outlays on goods and services that are not financed but also the financial payments if the outlays are financed, including down payments, reduction in principle, interest payments, and fees. The expenditure does not include residential investment, i.e., the purchasing price or the down payment for housing.

Following Bertrand and Morse (2016), I also drop the households whose share of a type of expenditure, except for food and shelter, exceeds half of the total expenditure in a quarter; the households whose expenditure on shelter exceeds the total expenditure; the households that have 0 food expenditure; and the households that did not participate in the second or the fifth interview.

To avoid the influence of outliers, I drop the top 1% and the bottom 1% in the distribution of expenditure growth of the whole sample. Then, I drop the top 1% and the bottom 1% in the distribution of income growth of the whole sample.

Out of all households that finished both interviews in the sample period, about 32% are dropped. I end up with 37,365 observations. On average, there are 95.86 observations left in each division-quarter cell. Table 9 presents the summary statistics of the whole sample.

Table 9: Summary Statistics of CEX Data and the Realistic Simulation Data

	CEX		Simulation	
	Mean	Sd	Mean	Sd
	Between-individual			
ln(Income) in wave 2	10.76	1.08	10.78	1.09
ln(Expenditure) in wave 2	10.70	0.72	10.71	0.72
ln(Income) in wave 5	10.80	1.05	10.81	1.05
ln(Expenditure) in wave 5	10.73	0.71	10.73	0.72
Observations	37365		747300	
	Between-cell			
Cell-level mean of ln(Income) in wave 2	10.76	0.19	10.78	0.19
Cell-level s.d. of ln(Income) in wave 2	1.052	0.15	1.060	0.15
Cell-level mean of ln(Expenditure) in wave 2	10.70	0.13	10.71	0.092
Cell-level s.d. of ln(Expenditure) in wave 2	0.700	0.056	0.710	0.052
Cell-level mean of ln(Income) in wave 5	10.79	0.18	10.81	0.18
Cell-level s.d. of ln(Income) in wave 5	1.020	0.14	1.027	0.14
Cell-level mean of ln(Expenditure) in wave 5	10.73	0.13	10.73	0.092
Cell-level s.d. of ln(Expenditure) in wave 5	0.699	0.057	0.707	0.052
Observations	402			

Notes: The CEX data is from the Consumer Expenditure Survey of the Bureau of Labor Statistics (BLS) from 1996 to 2007. The simulated data is from the realistic simulation, described by equations (13) and (14).

5.2 Setup of the Simulation and the Summary Statistics

The procedure to generate the simulated data is meant to make sure the between-group variance of group-level means of log household income, the between-group variance of group-level variances of log household income, and the consumption function of the simulated data match the actual data. Because the true weight of measurement error is unknown, I will construct a series of simulations with different measurement error weights.

I copy the identifier, interview time, wave, and geographic location information of each observation directly from the CEX data set and paste them in a blank one as the start of the simulation. Then, I replace each observation with 20 copies of the observation and assign new IDs to all the observations. To construct the location-period cell, I use census divisions as the geographic unit and quarters as the time unit. There are nine census divisions in the United States, 48 quarters in the sample period, and thus 432 division-quarter cells theoretically in the sample. However, the 48 quarters only include 45 complete panels for each wave due to the incomplete panels at the beginning and the end of the sample period. Also, some cells are omitted because there are no observations in those cells. Eventually, I end up with 402 cells for each wave. The next step is to generate income and consumption for each observation.

The logarithm of household income is simulated to generate the same level of between-group variations in log household income as the CEX data, including the between-group variance of group-level means and between-group variance of group-level variances. The data generating process of log household income is:

$$lnY_{igw} \sim N\left(\mu_{gw}^{CEX},\ \left(\sigma_{gw}^{CEX}\right)^2\right) \quad (13)$$

The log income of a household i in a given division-quarter cell g and a given wave w is drawn from a normal distribution, whose mean and variance are equal to those of the corresponding cell and wave in the CEX data.

The logarithm of consumption is simulated to follow the consumption function of the CEX data. The simulation includes two steps:

$$lnC_{igw}^{CEX} = \alpha_{0w} + \alpha_{1w} lnY_{igw}^{CEX} + \epsilon_{igw}$$

$$lnC_{igw} = \hat{\alpha}_{0w} + \hat{\alpha}_{1w} lnY_{igw} + \nu_{igw},\ \nu_{igw} \sim N\left(0, \left(\sigma_{gw}^{e}\right)^{2}\right) \quad (14)$$

First, I estimate the consumption functions using the CEX data for each wave, respectively, and obtain the estimated intercept $\hat{\alpha}_0$, estimated income elasticity of consumption $\hat{\alpha}_1$, and variance of the OLS residuals $(\sigma^e)^2$ for each wave. Second, I use the estimation results calculated in the first step and the simulated household income to construct household consumption. The log consumption of a given household i in a given division-quarter cell g and a given wave w is equal to the sum of three terms: the estimated intercept of this wave $\hat{\alpha}_{0w}$, the product of the estimated income elasticity of consumption in this wave $\hat{\alpha}_{1w}$ and the simulated log income of the observation lnY_{igw}, and a noise term ν_{igw}, which is drawn from a normal distribution of mean zero and variance equal to the variance of the OLS residuals in the first step for this cell and this wave $\left(\sigma_{gw}^{e}\right)^{2}$.

Table 9 presents the summary statistics of the CEX data and the simulated data at both the individual and cell levels for the two waves. At the individual level, the simulated data has the same mean and standard deviation for both the log income and the log consumption as the CEX data in both waves. As to the between-cell variation of log household income, both the between-cell standard deviations of cell-level means and the between-cell standard deviation of cell-level standard deviations are equal between the simulated data and the CEX data in both waves.

Table 10 compares the estimates of the consumption functions between the CEX data and the simulated data in both waves, in which columns 1 and 3 compare the functions for the second wave and columns 4 and 6 for the fifth wave. In both waves, the estimated coefficients are identical across the data sets. However, the standard error of the beta hat of the simulated data is much smaller than that of theCEX data in either wave. That is simply because the simulated data has more observations than the CEX data. In columns 2 and 5, I report estimation results of the

Table 10: Consumption Functions of the Realistic Simulation

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable	*ln(Consumption)*					
Unit of group	*Location-period cell*					
	CEX Wave 2	CEX Wave 2 (20N)	Simulation Wave 2	CEX Wave 5	CEX Wave 5 (20N)	Simulation Wave 5
ln(Income)	0.483***	0.483***	0.483***	0.503***	0.503***	0.503***
	(0.00234)	(0.000524)	(0.000522)	(0.00237)	(0.000531)	(0.000530)
Constant	5.500***	5.500***	5.505***	5.295***	5.295***	5.299***
	(0.0253)	(0.00566)	(0.00566)	(0.0258)	(0.00576)	(0.00575)
R-squared	0.533	0.533	0.534	0.546	0.546	0.547
Observations	37365	747300	747300	37365	747300	747300

Notes: The dependent variable is the logarithm of consumption in a period in all the columns. The sample in all the columns is the set of observations whose incomes are less than the 80th percentile of the income distribution of the location-period cell to which the observation belongs. The data in columns 1 and 4 are the CEX data for waves 2 and 5, respectively. In the data of columns 2 and 5, each observation in columns 1 and 4 is replaced by 20 copies of the observation, respectively. The data in columns 3 and 6 are from the extreme simulation for waves 2 and 5, respectively. Standard errors are in parenthesis. ***, **, * denote statistical significance at the 1, 5, and 10% level, respectively.

expanded CEX data, in which each observation, including its income and consumption, in the original CEX dataset is replaced by 20 copies of the observation.

The beta hats in these two columns are equal to their benchmarks in columns 1 and 4. However, the standard errors of the expanded CEX data, which are the benchmarks for the simulated data, are equal to their counterparties in columns 3 and 6, respectively. Moreover, the estimated intercepts and the R-squared are approximately equal between the CEX and the simulated data. The standard errors of the intercepts are approximately equal between the expanded CEX and the simulated data.

Figure 9 presents scatter plots of consumption for both the CEX data and the simulated data. To have a better visual comparison, in the scatter plot of the simulated data, I only include 5% of all the data points, randomly drawn from the whole sample following a uniform distribution.

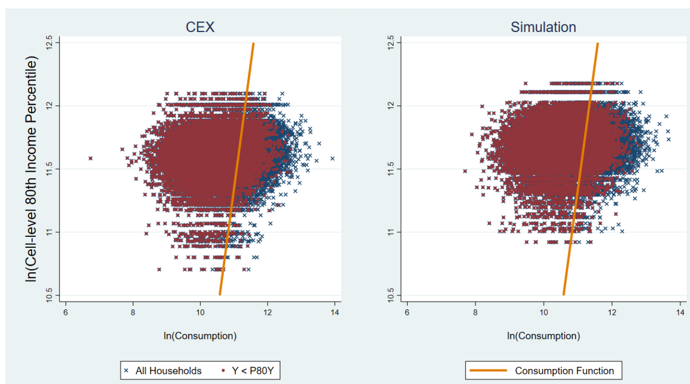

Figure 9: CEX Data vs. Simulated Data. This figure compares the scatter plots of the CEX data (left panel) with the simulated data (right panel). The CEX data is from the Consumer Expenditure Survey of the Bureau of Labor Statistics (BLS) from 1996 to 2007. The simulated data is from the realistic simulation, described by equations (13) and (14). Only the randomly chosen five percent of all the data points from the simulated data are displayed to make the visual comparison.

Figure 9 also presents the consumption functions (the dark orange lines) estimated using the respective data sets.

Next, I add measurement error to the simulated income as follows:

$$lnY^{*}_{igw} = a \cdot lnY_{igw} + b \cdot u_{igw}, u_{igw} \sim N\left(\mu^{lnY}_{gw}, \left(\sigma^{lnY}_{gw}\right)^2\right) \text{ and } a = \sqrt{1 - b^2} \quad (15)$$

The observed log income lnY^{*} is constructed in the same way as in chapter 4, ensuring that its between-group variance of group-level variances remains the same as true log income lnY and the pure effects of measurement error can be isolated. The measurement error in a given cell and a given wave u_{igw} follows a normal distribution whose mean and variance are equal to those of the CEX data in the same cell and wave $\left(\mu^{lnY}_{gw}, \left(\sigma^{lnY}_{gw}\right)^2\right)$. The sum of squares of the coefficient on the true log income a and the coefficient on the error term b is equal to one. As a result, the variance of log income in each cell of the simulation is the same as the corresponding cell in the CEX data. More important, the between-cell variance of the cell-level variances of the log income remains unchanged after this transformation. One caveat of this approach is that the mean of the log income in every cell, and thus the between-cell variance of the cell-level means, is different between the simulated data and the CEX data. Therefore, an alternative approach, $a = 1 - b$, which fixes this issue on cell-level means but gives up the cell-level variances, is applied to construct the simulation in the appendix. The results of the two approaches turn out to be equivalent. Lastly, since the true weight of the measurement error is unknown, I try a series of values for the coefficient on the error term b, ranging from zero to 0.9 with an incremental of 0.1.

Figure 10 shows the evolution of the two between-cell variations for each wave. The between-group variance of group-level variances remains constant as the

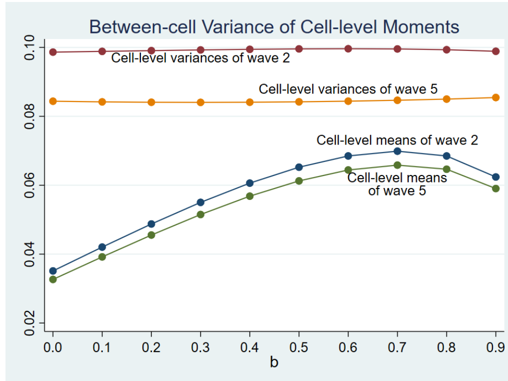

Figure 10: Between-cell Variance of Cell-level Moments of the Observed Log Income. The data is from the realistic simulation, described by equations (13) - (15). The coefficient b on the horizontal axis represents the weight of measurement error, as defined in equation (15). The figure presents the between-cell variances of cell-level means and cell-level variances of the observed log household income for both waves.

weight of measurement error increases in each wave. However, the curves of the between-cell variance of the cell-level means in both waves are invertedly U-shaped.

In conclusion, the setup of this realistic simulation makes sure the between-cell variance of the cell-level means of the true log household income, the between-group variance of group-level variances of the true log household income, and the consumption function match the actual data. Adding measurement error to log household income changes either its between-cell variance of the cell-level means or its between-group variance of group-level variances even though the estimation results of the two approaches are similar. I choose to maintain the between-cell variance of the cell-level means in the main text and leave the results of maintaining the between-group variance of the group-level variance in the appendix.

5.3 Preview of the Four Tests

Fifty seeds are used to construct 50 simulations, to make sure estimation results below are robust to different seeds chosen for the random number generator. These seeds are randomly drawn from a uniform distribution with a range from -2^{31} to 2^{31}, which is the largest range allowed in Stata. I estimate the beta hat 50 times using these simulated data sets and compare the group mean of the beta hats with zero using a one-sample t-test for each quantile-based consumption spillover test, respectively.

$$H_0: \hat{\beta} = 0$$

The means and their 95% confidence intervals of these t-tests are presented in Figure 11.

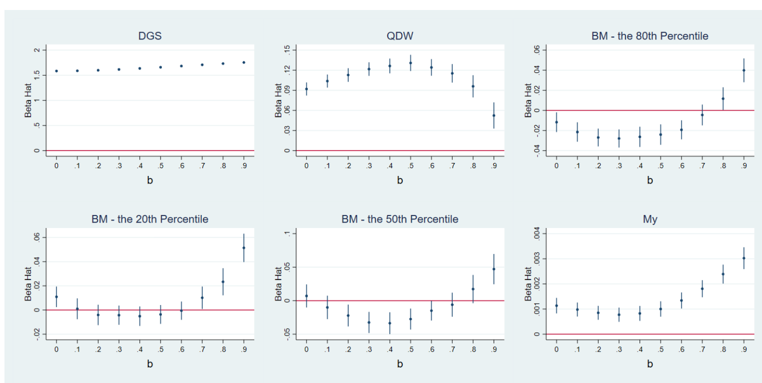

Figure 11: Full Specifications of All the Four Tests. This figure compares the beta hats estimated by the full specifications of all four tests at different weights of measurement error. The data is from the realistic simulation, described by equations (13) - (15). 50 such simulations are constructed with 50 different seeds. In each panel, the horizontal axis labels the mean and the 95% confidence interval of the 50 beta hats from a one-sample t-test. The coefficient b on the horizontal axis in all the panels represents the weight of measurement error, as defined in equation (15). Details of the specifications of DGS, QDW, and BM can be found in Table 4. Details of my specification can be found in column 7 of Table 8.

The estimation results show that my test is the best one in terms of reducing the beta hat. As shown in the figure, DGS generates the highest level of the beta hat, which starts from 1.59 when there is no measurement error and increases to above 1.76 when the weight of the measurement error is 0.9. The graphs of the other tests are either U-shaped or invertedly U-shaped. QDW produces the second-highest beta hat, ranging between 0.04 and 0.14. BM is the second-best test. The beta hats on the 80th, 20th, and 50th percentiles fluctuate between -0.05 and 0.05. Among all three percentiles, the beta hats on the 80th and 50th percentiles are more likely to be negative, while the beta hat on the 20th percentile is more likely to be positive. The absolute values of these beta hats are smaller than those in DGS and QDW, but these beta hats still have moderate economic significance. My test performs the best in lowering the beta hat, which ranges from 0.0008 to 0.003 and has barely any economic significance. Overall, the beta hat of DGS has the highest order of magnitude, QDW the second-highest, BM the third-highest, and my test the lowest.

5.4 Break Down the Four Tests

So far, I have constructed a realistic simulation that matches the CEX data. I have also compared the full specifications of the four quantile-based consumption spillover tests using the simulated data. In this chapter, I will break down these full specifications to multiple variants to their corresponding baseline regressions and examine these embedded techniques in the same way as in chapter 4. This analysis will help me choose techniques to eliminate the mechanical correlation bias from estimating the consumption spillover effects using the CEX data.

A breakdown of DGS with these realistic simulations (Figure 12) shows similar results as the extreme simulation, but there are some interesting differences.

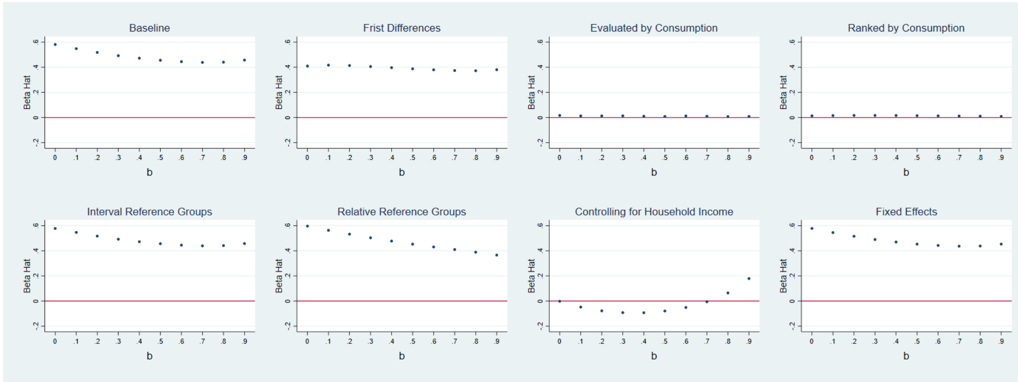

Figure 12: Break Down Drechsel-Grau and Schmid (2014). This figure compares the beta hats estimated by the different techniques embedded in the full specification of Drechsel-Grau and Schmid (2014) with its baseline at different weights of measurement error. The data is from the realistic simulation, described by equations (13) - (15). 50 such simulations are constructed with 50 different seeds. In each panel, the horizontal axis labels the mean and the 95% confidence interval of the 50 beta hats from a one-sample t-test. The coefficient b on the horizontal axis in all the panels represents the weight of measurement error, as defined in equation (15). Details of these specifications can be found in Table 5

The same three techniques, including evaluating the reference group by consumption, ranking observations by consumption, and controlling for income, can lower the beta hat to a very low level as in the extreme simulation. However, the efficacy of controlling for income depends on the weight of measurement error, while the first two techniques do not. Recall that in the extreme simulation, where there is only one level for the weight, the beta hat is a negative value when household income is controlled for. With the multiple levels for the weight, we can see that the beta hat is not always negative when income is controlled for. The beta hat is very close to zero when there is no measurement error in the data. As the weight of the measurement error increases, the beta hat first decreases to a negative value and eventually increases to a positive value. Besides the technique of controlling for income, the realistic simulations show that the technique of the first difference also has different effects at different weights of measurement error. In the extreme simulation, the technique of the first difference slightly decreases the beta hat, while the realistic simulations show that the effect is bigger when the weight of the measurement error is smaller. The technique of relative reference also shows different effects compared to the extreme simulation, in which it raises the beta hat. However, this technique does not change the beta hat until the weight of measurement error increases to 0.7 in the realistic simulations. The rest of the DGS techniques, including interval reference groups, and fixed effects, turn out to be as ineffective as in the extreme simulation in regard to lowering the beta hat. Even though the three techniques shown effectively lowering the beta hat in the extreme simulation are still shown effective in the realistic simulation, it is still unclear if they will remain effective when the upper bound of the regression sample becomes the percentile of the reference group, and the unit of a group becomes the location-period cell.

A breakdown of QDW with the data from the realistic simulations (Figure 13) also shows some interesting differences from the extreme one. The technique of evaluating the reference group by consumption, which raises the beta hat in the extreme simulation, now lowers it to a level with barely any economic significance for all the weights of measurement error. In contrast, the technique of lagged reference variable, which lowers the beta to an economically insignificant level in the extreme simulation, only slightly lowers it now. The technique of the first difference here turns to lower the beta hat slightly from raising it in the extreme simulation. Also, the technique of fixed effects, instead of ineffective in the extreme simulation, lowers the beta hat now, and its effects are greater when the weight of measurement error is relatively small. The interval reference group technique still lowers the beta hat moderately at all the weights of measurement error, but the efficacy decreases as the weight of measurement error increases. All the other techniques, including using the

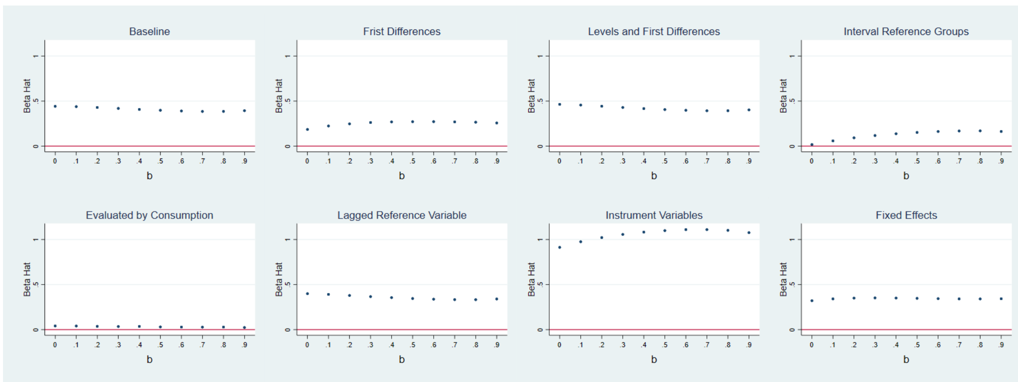

Figure 13: Break Down Quintana-Domeque and Wohlfart (2016). This figure compares the beta hats estimated by the different techniques embedded in the full specification of Quintana-Domeque and Wohlfart (2016) with its baseline at different weights of measurement error. The data is from the realistic simulation, described by equations (13) - (15). 50 such simulations are constructed with 50 different seeds. In each panel, the horizontal axis labels the mean and the 95% confidence interval of the 50 beta hats from a one-sample t-test. The coefficient b on the horizontal axis in all the panels represents the weight of measurement error, as defined in equation (15). Details of these specifications can be found in Table 6.

first difference and level simultaneously, and instrument variable, as in the extreme simulation, do not lower the beta hat in the realistic simulations. Additionally, when the technique of instrument variable is applied, the beta hat increases as the weight of the measurement error increases. In the end, the techniques of the first difference and evaluating reference group by consumption are shown to be able to lower the beta hat by a meaningful amount in the realistic simulation.

A breakdown of BM (Figure 14) using the data from these realistic simulations also shows interesting results compared to the extreme simulation. The specifications here are the same as those used for the extreme simulation except that the group-level unemployment rate is control for in all the regressions. The technique of income bucket fixed effects still lowers the beta hat to a very large extent as in the extreme

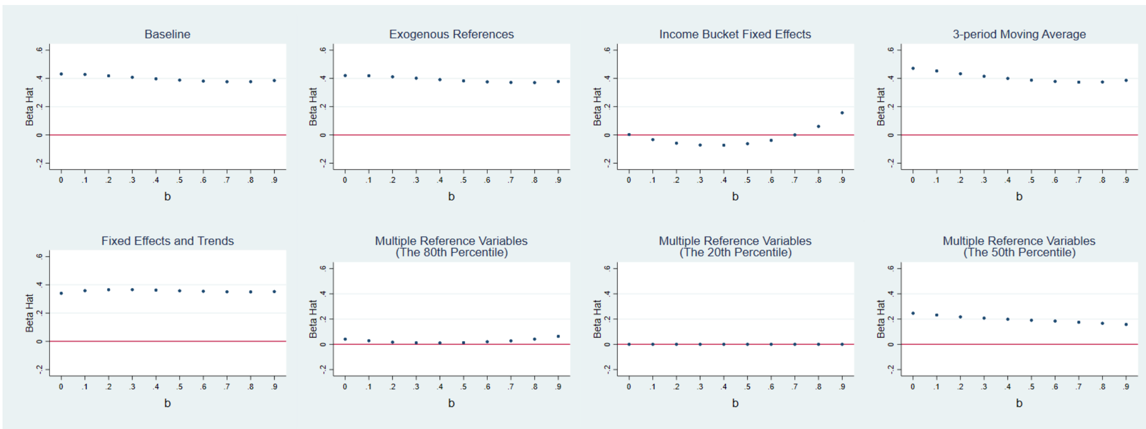

Figure 14: Break Down Bertrand and Morse (2016). This figure compares the beta hats estimated by the different techniques embedded in the full specification of Bertrand and Morse (2016) with its baseline at different weights of measurement error. The data is from the realistic simulation, described by equations (13) - (15). 50 such simulations are constructed with 50 different seeds. In each panel, the horizontal axis labels the mean and the 95% confidence interval of the 50 beta hats from a one-sample t-test. The coefficient b on the horizontal axis in all the panels represents the weight of measurement error, as defined in equation (15). Details of these specifications can be found in Table 7.

simulation. The weight of measurement error affects the efficacy of the technique of income bucket fixed effects the same way as it affects the technique of controlling for income. The technique of multiple reference variables lowers the beta hat on the 80th

percentile as much as it does in the extreme simulation. However, the coefficients on the 20th and 50th percentiles are still too high to be ignored. Instead of slightly lowering the beta hat in the extreme simulation, the technique of moving average now increases the beta hat when the weight of measurement error is relatively low. The technique of using fixed effects and location-specific time trends simultaneously shows the same behavior as the technique of using fixed effects only in QDW. Besides, the technique of exogenous reference groups remains ineffective as in the extreme simulation. In the end, only the techniques of income buckets and multiple reference groups can lower the beta hat of the 80th percentile by a meaningful amount in the realistic simulation[3].

The comparison of the breakdowns between chapters 4 and 5 shed some light on the efficacy of the techniques that the three quantile-based consumption spillover tests use to meet the two conditions that Angrist (2014) proposes. QDW uses lagged reference variable to ensure the exogeneity of the reference group (i.e., the peers). This technique works well in the extreme simulation when there is no serial correlation. However, the technique fails in the realistic simulations in which the features of the business cycle and economic growth are built. This suggests that the technique of lagged reference variable may need the help with time fixed effects or the first difference to lower the beta hat in the realistic simulation. However, their deflating effects are overwhelmed by other inflating techniques, such as instrument variables. BM uses an out-of-sample reference group to ensure the exogeneity of the reference group. As shown in both simulations, this technique does not affect the beta hat, probably because the distributions that the log household income follows are the same across the corresponding cells in the two sub-samples. To satisfy the separation condition, both QDW and BM only include the subjects in the regression sample. In other words, no observation in the regression sample is a subject and a peer at the same time. However, DGS uses relative reference groups, so a class could compare themselves with higher classes or be compared by lowers classes at the same time, which inflates the mechanical correlation bias as shown in columns 1 and 6 of Table 5.

In conclusion, by breaking down the techniques in the three quantile-based consumption spillover tests above, I find these techniques can reduce the mechanical correlation bias: controlling for income or income buckets, evaluating the reference group by consumption, and lagged reference variable combined with the first difference or fixed effects.

Given all the information above and subsequent trials, I eventually choose the techniques of symmetric range, controlling for household income, the first difference, holding the observations' ranks constant at the level in the previous period, and ranking the observations according to their consumption. A breakdown of my proposed techniques with these realistic simulations (Figure 15) also shows some interesting differences from the extreme simulation.

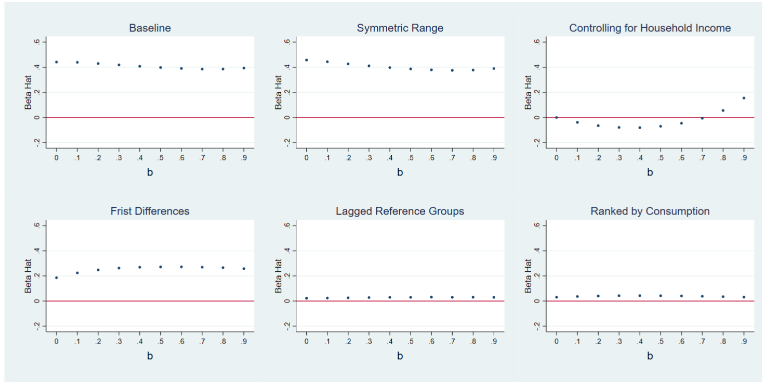

Figure 15: Break Down My Test. This figure compares the beta hats estimated by the different techniques embedded in the full specification that I propose with its baseline at different weights of measurement error. The data is from the realistic simulation, described by equations (13) - (15). 50 such simulations are constructed with 50 different seeds. In each panel, the horizontal axis labels the mean and the 95% confidence interval of the 50 beta hats from a one-sample t-test. The coefficient b on the horizontal axis in all the panels represents the weight of measurement error, as defined in equation (15). Details of these specifications can be found in Table 8

The technique of lagged rank and the technique of ranking observations by consumption, which only lowers the beta hat slightly in the extreme simulation, now lower it to a level with barely any economic significance. The technique of first difference and the technique of controlling for household income behave the same way as in the breakdown of DGS and QDW using the data from the realistic simulations.

My full test may be able to detect consumption spillovers. A technique that eliminates the effects of the mechanical correlation from estimating the unbiased consumption spillover effects may eliminate the spillover effects, which is not the case for the full specification that I propose. To show that my test can preserve consumption spillover, I add a type of spillover to the simulations above and re-run the baseline specification (equation 4) and the full specification that I propose (equation 12). Spillovers are added in the following way:

$$lnC^*_{ist} = lnC_{ist} + 0.1 \cdot lnY(80thIncomePercentile)_{st},$$

$$if\ Y_{ist} < 80thIncomePercentile_{st} \qquad (16)$$

where ten percent of the 80th percentile of the local income distribution in a given period is added to the original log consumption of the observations whose incomes are less than the 80th percentile of the income distribution in that location-period cell. The new consumption C^*_{ist} is used to run the regressions. There is only one original extreme simulation, but there are five hundred possible original realistic simulations to choose. I only pick one out of the five hundred to show my test can preserve the spillovers described by equation (16). The seed is the first one in the original list of seeds. The weight of measurement error is set to be 0.5. Table 11 presents the estimation results. The spillovers increase the beta hats of the baseline regressions in both the extreme simulation (column 1 vs. 5) and the realistic simulation (column 3 vs. 7) by about 0.1. As to the beta hats in my tests, the spillovers raise the beta hat in the extreme simulation by about 0.04 (column 2 vs. 6) and the beta hat in the realistic simulation by about 0.01, or by one order of magnitude (column 4 vs. 8). More important, the spillovers not only increase the economic significance of the beta hat in my test with the data of the realistic simulation but also increases its statistical significance to the level of one percent from the level below ten percent.

The estimation results of the full specifications of the four quantile-based consumption spillover tests using the CEX data (Table 12) echo the results of the realistic simulations. The specifications here are the same as those used for the realistic simulations, except that I control for the cell-level unemployment rate in the

BM as it does in its paper (Bertrand and Morse 2016), which does not change the main results. As shown in Table 12, the DGS produces a beta hat of the highest order of magnitude among the four tests, QDW the second, BM the third, and my test the

Table 11: My Test with the Presence of Spillovers

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent variable	*lnC*	*DlnC*	*lnC*	*DlnC*	*lnC*	*DlnC*	*lnC*	*DlnC*
Sample	Y<P80Y	LagP20C< LagC< LagP80C	Y<P80Y	LagP20C< LagC< LagP80C	Y<P80Y	LagP20C< LagC< LagP80C	Y<P80Y	LagP20C< LagC< LagP80C
Unit of group	Location-period	Location-period	Division-quarter	Division-quarter	Location-period	Location-period	Division-quarter	Division-quarter
	Without Spillovers				With Spillovers			
	Extreme Simulation		Realistic Simulation		Extreme Simulation		Realistic Simulation	
	Baseline	My	Baseline	My	Baseline	My	Baseline	My
lnY(The 80th income percentile)	0.419***				0.519***			
	(0.000367)				(0.000367)			
DlnY(The previous 80th consumption percentile)		0.0256***				0.0633***		
		(0.000322)				(0.000329)		
lnY(The 80th income percentile)			0.404***				0.497***	
			(0.00408)				(0.00406)	
DlnY(The previous 80th consumption percentile)				0.00124				0.0125***
				(0.000914)				(0.00201)
lnY or DlnY	No	Yes	No	Yes	No	Yes	No	Yes
R-squared	0.395	0.752	0.0161	0.286	0.501	0.765	0.0244	0.000300
Observations	2000292	1468505	597840	448380	2000292	1443379	597840	134755

Notes: The data in columns 1, 2, 5, and 6 is generated by an extreme simulation with consumption spillovers. The data generating process is described by equations (7), (8), and (16). The data in columns 3, 4, 7, and 8 is generated by a realistic simulation. The data generating process is described by equations (13) – (16). The dependent variable lnC denotes the logarithm of the household consumption in a period, while DlnC denotes the first difference. In columns 1, 2, 5, and 6, the sample is the set of observations whose incomes are less than the 80th percentile of the income distribution. A group within which the observations are ranked to get the distribution in those columns is the set of observations in the given period and the local area in which the observation resides. In columns 3, 4, 7, and 8, the sample is the set of the observations whose consumption in the previous period is less than the 80th percentile of consumption distribution in the previous period. A group within which the observations are ranked to get the distribution in these columns is the set of observations in the given quarter and the census division in which the observation resides. For each reference variable, the reference group is in the parenthesis. A "percentile" in the reference group refers to the observations used to calculate the percentile. The variable used to evaluate the reference group and the operators of the logarithm ("ln") and the first difference ("D") are outside the parenthesis. Standard errors are in parenthesis. ***, **, * denote statistical significance at the 1, 5, and 10% level, respectively.

lowest. Like the beta hats in the simulations, the beta hat of the 80th percentile in BM is negative, the 20th positive, and 50th negative. These similarities suggest that these estimation results using the CEX data may be generated by the same mechanism as the simulation data, which does not involve consumption spillover. Therefore, the preference for "keeping up with the Joneses" is not the only alternative hypothesis for any statistically significant estimated coefficient using the actual data. Among the four estimates using the CEX data, the beta hats of the DGS and QDW are statistically significantly different from zero at one percent and five percent levels, respectively. The mechanical correlation could also explain these results. To get the unbiased spillover effects, econometricians must eliminate this mechanical correlation bias. The previous simulation shows that the BM and my test can reduce the mechanical correlation bias to a moderate and a negligible level, respectively. Using the CEX data, the beta hats of these two methods are at the same order of magnitude as they are in the simulation, respectively, and statistically insignificantly different from zero. In other words, after eliminating the second alternative hypothesis, there is no evidence to reject the null hypothesis that the beta hat is equal to zero, and there is no reason to accept the first alternative hypothesis that people have the preference of comparing themselves with the reference groups defined in those tests. The estimation results of BM here contradict the findings in their original paper (Bertrand and Morse 2016). Given that my method performs better than BM in the realistic simulations, I would rather believe that there is no evidence of consumption spillovers in the CEX data at the division-quarter level.

Table 12: The Four Tests with the CEX Data

	(1)	(2)	(3)	(4)
Dependent variable	*Dln(Consumption)*	*Dln(Consumption)*	*ln(Consumption)*	*Dln(Consumption)*
Sample:	C<P95C	Y<P80Y	Y1<P80Y2	LagC<LagP80C
Unit of group	*Quarter*	*Division-quarter Cell*	*Division-quarter Cell*	*Division-quarter Cell*
	DGS	QDW	BM	My
DlnC(Higher consumption classes)	1.384***			
	(0.0385)			
HatDlnC(Top 20% income)		0.468**		
		(0.183)		
lnY(The 80th income percentiles in the recent three periods of sample 2)			-0.0497	
			(0.0349)	
lnY(The 20th income percentiles in the recent three periods of sample 2)			0.0219	
			(0.0320)	
lnY(The 50th income percentiles in the recent three periods of sample 2)			-0.00122	
			(0.0423)	
DlnY(Previous 80th consumption percentile)				0.00315
				(0.00359)
Half of the whole sample	No	No	Yes	No
Location and period FEs	Yes	Yes	Yes	Yes
Division-specific trends	No	No	Yes	No
lnY, DlnY, or income bucket FEs	Yes	No	Yes	Yes
Reference variable*HOP	Yes	No	No	No
Instrumental variable	No	Yes	No	No
Clustered s.e.	No	No	Yes	No
Demographic characteristics	Yes	Yes	Yes	Yes
Unemployment rate	No	No	Yes	No
Observations	34570	27667	14429	22102
R-squared	0.721	0.00687	0.588	0.0472

Notes: The dependent variable in columns 1 – 2 and 4 (or column 3) is the change in (or the level of) the logarithm of household consumption in a period. The regression sample in column 1 (or columns 2 - 4) is the set of observations whose consumption (or income) less than the 95th (or 80th) percentile of the consumption (or income) distribution of the time group (location-period cell) to which the observation belongs. A group within which the observations are ranked to get the distribution in column 1 (or columns 2 – 4) is the set of observations in the given quarter in the whole country (or the census division in which the observation resides). For each reference variable, the reference group is in the parenthesis. A "percentile" in the reference group indicates the observations used to calculate the percentile. The variable used to evaluate the reference group and the operators of the logarithm ("ln"), the first difference ("D"), and the predicted value ("Hat") are outside the parenthesis. The reference variables in the four columns are the change in the logarithm of the average consumption of all the higher classes than the one to which the observation belongs in the national consumption distribution of a given period (DGS), the predicted value of the change in the logarithm of the average consumption of the observations whose incomes are greater than the 80th percentile of the income distribution of the location-period cell to which the observation belongs (QDW), the logarithm of the average 80th (and 20th, 50th) percentiles of the local income distribution in periods t, t-1, and t-2 obtained from the second sub-sample (BM), and the change in the logarithm of the average income of the 80th percentile of the consumption distribution of the location-period cell to which the observation belongs in the previous period (My), respectively. The instruments in QDW are the lagged average income of the reference group and the lagged growth in the threshold of the reference group. Standard errors are in parenthesis. ***, **, * denote statistical significance at the 1, 5, and 10% level, respectively.

Chapter 6

CONCLUSION

Many tests in the literature argue that consumption spillovers cause positive correlations between individual consumption levels and aggregate income quantiles (or their transformations). However, I show that this correlation might be subject to a mechanical correlation bias. I find that the size of this mechanical correlation bias is determined by the range of household income that defines the regression sample, the weight of measurement error in the household income, the between-group variance of group-level means of log household income, and the between-group variance of group-level variances of log household income. Even though some quantile-based consumption spillover tests in the current literature (Drechsel-Grau and Schmid 2014; Quintana-Domeque and Wohlfart 2016; Bertrand and Morse 2016) follow the Angrist (2014) and take some measures to eliminate this mechanical correlation bias, I show that these measures fail in some cases using an extreme simulation and a series of realistic simulations.

I break down the full specifications of the current quantile-based consumption spillover tests into different single variants to the simplest univariate regression. I construct an extreme simulation in which the variance of the group means is sufficiently high, the between-group variance of group-level variances is sufficiently low, and the weight of measurement error is at a certain level. The extreme simulation shows that none of the three full specifications can reduce the mechanical correlation bias to an economically insignificant level. The simulation also shows that the technique of lagged reference variables, controlling for income or income bucket fixed effects, ranking observations according to their consumption, and multiple references can reduce the bias. Then I calibrate the data to match the between-group variation in the CEX data and examine the efficacy of different techniques embedded in the three quantile-based consumption spillover tests in lowering the mechanical correlation bias. I propose that econometricians should simultaneously use the techniques of symmetric ranges, first difference, controlling for income growth, lagged rank, and ranking observations by consumption to eliminate the mechanical correlation from estimating the consumption spillover effects. Using these new techniques, I find no evidence for consumption spillover at the level of census division in the quarterly CEX data.

The estimation results of my test only show that there is no evidence for upward-looking comparisons in a particular dataset, while it does not exclude possibilities of other types of consumption spillovers in other datasets. A different sample period, a different definition of consumption, or a different definition of the subjects or peers may change the estimation results. Consumption spillovers may only exist in a special period. Stefani (2020) replicates Bertrand and Morse (2016) using the data from 1996 to 2007 obtained from their online appendix (2016) and finds that the estimated coefficient is insignificantly different from zero. This result indicates that the upward-looking comparisons may only exist in the rest of the sample period (1980 – 1995) in Bertrand and Morse (2016). The consumption spillovers may only exist for a special group of subjects. Stefani (2020) shows that the income of rich households is positively correlated with the expenditure of the homeowners while uncorrelated with renters. Consumption spillovers may only happen to visible consumption. For example, Charles, Hurst, and Roussanov (2009) find evidence for spillovers of visible goods, Kuhn et al. (2011) find evidence for cars, and Quintana-Domeque and Wohlfart (2016) find evidence for food away from home. The consumption spillover may only exist in one's social network instead of a large geographic area such as the census divisions. For example, Kuhn et al. (2011) and Agarwal, Mikhed, and Scholnick (2020) find evidence for consumption spillovers in one's neighborhood, and De Giorgi, Frederiksen, and Pistaferri (2020) in one's workplace. If the strength of consumption spillovers declines as the size of the reference groups increases, what would be the largest size of a reference group? If there were no consumption spillovers at the national level, what would cause the positive correlations between the national income inequality and the household debt-to-income ratio?

On the one hand, the cause of the positive correlation could be the rise in within-region inequality. Changes in inequality can be decomposed into changes in between-group inequality and changes in within-group inequality. If the majority of the changes in the national income inequality since 1980 were accounted for by the changes in the within-region inequality, consumption spillovers could still explain the correlation between the national inequality and household indebtedness. On the other hand, the correlation could be explained by a different, supply-side story. Gaubert et al. (2021) show that both the between-state and the between-country income inequalities have been rising since 1980. Suppose the rise in the between-region inequality was the main reason for the rise in household-level inequality across the United States. In that case, rich regions could supply loanable funds to poor regions, increasing the household indebtedness in the poor regions. This supply-side story is consistent with the evidence found in Coibion et al. (2020). Of course, the supply-side story could be combined with the demand-side story, generating a third story. The rise in national inequality might be caused by the rise in incomes of the rich people in a few rich regions, which could increase both the between-region inequality and the within-region inequalities in these rich regions. The rich people in rich regions could supply funds to poor people across the country, and poor people in rich regions might demand funds due to the consumption spillovers from the rich people in the rich regions. The validity of all these theories depends on the facts about the maximum size of a reference group and the decomposition of the rise in national inequality, which will be left to future studies.

ENDNOTES

[1] This beta hat is greater than the beta hat in the baseline of DGS. This could be caused by either of the two differences between the two regression. The first difference is that the unit of a group within which the observations are ranked is the period in DGS, while the unit is the location-period cell in QDW. The second difference involves the relationship between the percentile of the reference group and the upper bound of the regression sample. The upper bound of the regression sample in the DGS is the 95th percentile, which is greater than the percentile of the reference group (80th) in the regression; the upper bound in the QDW is the 80th percentile, which is equal to the percentile of the reference group in the regression. It is unclear which factor causes the difference in the beta hats.

[2] The techniques of evaluating the reference group by consumption and the interval reference group here show opposite effects compared to the breakdown of DGS. The difference in the technique of evaluating the reference group by consumption may be caused by the two differences between the corresponding regressions in the two breakdowns: 1) the unit of a group, and 2) the relationship between the upper bound of the regression sample and the percentile of the reference group. Apart from these two differences, the difference in the effects of interval reference groups between the two breakdowns has a third, or even fourth, factor to consider: the mid-point of the interval in QDW, the 90th percentile, is higher than the 80th percentile, which is the percentile of the reference group and the upper bound of the regression sample. Therefore, it is unclear which factor reduces the beta hat in QDW, and the technique of interval reference group still may not qualify as a candidate to solve the mechanical correlation problem.

[3] The technique of interval reference group still may not qualify as a candidate to lower the beta hat due to the same reason stated in the previous chapter.

REFERENCES

Agarwal, Sumit, Vyacheslav Mikhed, and Barry Scholnick. 2020. "Peers' Income and Financial Distress: Evidence from Lottery Winners and Neighboring Bankruptcies." *The Review of Financial Studies* 33 (1): 433–72. https://doi.org/10.1093/rfs/hhz047.

Alvarez-Cuadrado, Francisco, and Irakli Japaridze. 2017. "Trickle-down Consumption, Financial Deregulation, Inequality, and Indebtedness." *Journal of Economic Behavior and Organization* 134 (40): 1–26. https://doi.org/10.1016/j.jebo.2016.12.007.

Angrist, Joshua D. 2014. "The Perils of Peer Effects." *Labour Economics* 30: 98–108. https://doi.org/10.1016/j.labeco.2014.05.008.

Belabed, Christian A., Thomas Theobald, and Till van Treeck. 2018. "Income Distribution and Current Account Imbalances." *Cambridge Journal of Economics* 42 (1): 47–94. https://doi.org/10.1093/cje/bew052.

Bertrand, Marianne, and Adair Morse. 2016. "Trickle-down Consumption." *Review of Economics and Statistics* 98 (5): 863–79. https://doi.org/10.1162/REST_a_00613.

Bordo, Michael D., and Christopher M. Meissner. 2012. "Does Inequality Lead to a Financial Crisis?" *Journal of International Money and Finance* 31 (8): 2147–61. https://doi.org/10.1016/j.jimonfin.2012.05.006.

Büyükkarabacak, Berrak, and Neven T. Valev. 2010. "The Role of Household and Business Credit in Banking Crises." *Journal of Banking and Finance* 34 (6): 1247–56. https://doi.org/10.1016/j.jbankfin.2009.11.022.

Cardaci, Alberto. 2018. "Inequality, Household Debt and Financial Instability: An Agent-Based Perspective." *Journal of Economic Behavior and Organization* 149 (January): 434–58. https://doi.org/10.1016/j.jebo.2018.01.010.

Charles, Kerwin Kofi, Erik Hurst, and Nikolai Roussanov. 2009. "Conspicuous Consumption and Race." *Quarterly Journal of Economics* 124 (2): 425–67. https://doi.org/10.1162/qjec.2009.124.2.425.

Coibion, Olivier, Yuriy Gorodnichenko, Marianna Kudlyak, and John Mondragon.

2020. “Greater Inequality and Household Borrowing: New Evidence from Household Data.” *Journal of the European Economic Association* 18 (6): 2922–71. https://doi.org/10.1093/jeea/jvz076.

Daly, Mary C., Daniel J. Wilson, and Norman J. Johnson. 2013. “Relative Status and Well-Being: Evidence from U.S. Suicide Deaths.” *Review of Economics and Statistics* 95 (5): 1480–1500. https://doi.org/10.1162/REST_a_00355.

Drechsel-Grau, Moritz, and Kai D. Schmid. 2014. “Consumption-Savings Decisions under Upward-Looking Comparisons.” *Journal of Economic Behavior and Organization* 106: 254–68. https://doi.org/10.1016/j.jebo.2014.07.006.

Frank, Robert H. 2014. “Expenditure Cascades.” *Review of Behavioral Economics* 1 (1–2): 55–73. https://doi.org/10.1561/105.00000003.

Gaubert, Cecile, Patrick M Kline, Damián Vergara, and Danny Yagan. 2021. “Trends in U.S. Spatial Inequality: Concentrating Affluence and a Democratization of Poverty.” Working Paper Series. https://doi.org/10.3386/w28385.

Georgarakos, Dimitris, Michael Haliassos, and Giacomo Pasini. 2014. “Household Debt and Social Interactions.” *Review of Financial Studies* 27 (5): 1404–33. https://doi.org/10.1093/rfs/hhu014.

Giorgi, Giacomo De, Anders Frederiksen, and Luigi Pistaferri. 2020. “Consumption Network Effects.” *The Review of Economic Studies* 87 (1): 130–63. https://doi.org/10.1093/restud/rdz026.

Guven, Cahit, and Bent E. Sørensen. 2012. *Subjective Well-Being: Keeping Up with the Perception of the Joneses*. *Social Indicators Research*. Vol. 109. https://doi.org/10.1007/s11205-011-9910-x.

Jordà, Òscar, Moritz Schularick, and Alan M. Taylor. 2016. “Sovereigns versus Banks: Credit, Crises, and Consequences.” *Journal of the European Economic Association* 14 (1): 45–79. https://doi.org/10.1111/jeea.12144.

Kuhn, Peter, Peter Kooreman, Adriaan Soetevent, and Arie Kapteyn. 2011. “The Effects of Lottery Prizes on Winners and Their Neighbors: Evidence from the Dutch Postcode Lottery.” *American Economic Review* 101 (5): 2226–47. https://doi.org/10.1257/aer.101.5.2226.

Luttmer, Erzo F.P. 2005. “Neighbors as Negatives: Relative Earnings and Well-Being.” *Quarterly Journal of Economics* 120 (3): 963–1002. https://doi.org/10.1162/003355305774268255.

Manski, Charles F. 1993. "Identification of Endogenous Social Effects the Reflection Problem." *Review of Economic Studies* 60 (3): 531–42. https://doi.org/10.2307/2298123.

Mian, Atif, and Amir Sufi. 2010. "Household Leverage and the Recession of 2007-09." *IMF Economic Review* 58 (1): 74–117. https://doi.org/10.1057/imfer.2010.2.

Mian, Atif, Amir Sufi, and Emil Verner. 2017. "Household Debt and Business Cycles Worldwide." *Quarterly Journal of Economics* 132 (4): 1755–1817. https://doi.org/10.1093/qje/qjx017.

Perugini, Cristiano, Jens Hölscher, and Simon Collie. 2015. "Inequality, Credit and Financial Crises." *Cambridge Journal of Economics* 40 (1): 227–57. https://doi.org/10.1093/cje/beu075.

Piketty, Thomas, and Emmanuel Saez. 2003. "Income Inequality in the United States, 1913-1998." *Quarterly Journal of Economics* 118 (1): 1–39. https://doi.org/10.1162/00335530360535135.

Quintana-Domeque, Climent, and Johannes Wohlfart. 2016. "'Relative Concerns for Consumption at the Top': An Intertemporal Analysis for the UK." *Journal of Economic Behavior and Organization* 129: 172–94. https://doi.org/10.1016/j.jebo.2016.06.005.

Schularick, Moritz, and Alan M. Taylor. 2012. "Credit Booms Gone Bust: Monetary Policy, Leverage Cycles, and Financial Crises, 1870-2008." *American Economic Review* 102 (2): 1029–61. https://doi.org/10.1257/aer.102.2.1029.

Stefani, Alessia De. 2020. "Debt, Inequality and House Prices: Explaining the Dynamics of Household Borrowing Prior to the Great Recession." *Journal of Housing Economics*. Vol. 47. https://doi.org/10.1016/j.jhe.2018.09.001.

Stockhammer, Engelbert, and Rafael Wildauer. 2018. "Expenditure Cascades, Low Interest Rates or Property Booms? Determinants of Household Debt in OECD Countries." *Review of Behavioral Economics* 5 (2): 85–121. https://doi.org/10.1561/105.00000083.

Thompson, Jeffrey. 2018. "Rising Top Incomes and Increased Borrowing in the Rest of the Distribution." *Economic Inquiry* 56 (2): 686–708. https://doi.org/10.1111/ecin.12520.

U.S. Department of Labor, Bureau of Labor Statistics. 2008. "Consumer Expenditure Survey, Interview Survey, 2006."

Wildauer, Rafael. 2016. "Determinants of US Household Debt: New Evidence from the SCF." Vol. 6. https://ideas.repec.org/p/ris/kngedp/2016_006.html.

APPENDIX

THE SECOND APPROACH TO ADD NOISE TO LOG INCOME

In the main text, I add measurement error to the logged household income as a way to maintain the between-cell variance of cell-level variances of the logged income at its level before adding the error regardless of the weight of measurement error, while its between-cell variance of the cell-level means can change as the weight of the measurement error changes. In this appendix, I will take a second approach to add measurement error. In this approach, I can fix the between-cell variance of the first-moment of the observed logged income to the level of the true logged income at all the weights of measurement error, but I have to allow the between-cell variance of the second moment of the observed logged income to vary as the weight changes. The second approach to adding measurement error is:

$$lnY^{*}_{igw} = a \cdot lnY_{igw} + b \cdot u_{igw}, u_{igw} \sim N\left(\mu^{lnY}_{gw}, \left(\sigma^{lnY}_{gw}\right)^{2}\right) \; and \; a = 1 - b \quad (17)$$

The only difference between the two approaches is the relationship between the coefficient on the true logged income a and the coefficient on the error term b. In the previous approach, as described by equation (16), the sum of squares of the two coefficients is equal to one, but their sum is equal to one in the new approach. Since the error term u follows the same distribution as the true logged income lnY in each cell and each wave, this setup will preserve the mean of true logged income in each cell and wave, and thus the between-cell variance of the cell-level mean of true logged income will be maintained in both waves.

As in the main text, the individual level variance and the between-cell variance of the first and the second cell-level moments of the observed logged income at different weights of measurement error in the second approach are plotted in Figure A1. Instead of being invertedly U-shaped in Figure 10, the individual-level variance curves in Figure A1 become u-shaped for both waves. The curve of the between-cell variance of cell-level means in Figure A1, compared to the u-shaped curve in Figure 10, becomes horizontal for each wave. The curve of between-cell variance of cell-level variances, which used to be horizontal in Figure 10, now becomes u-shaped in Figure A1 in each wave.

As in the main text, the means and the 95% confidence intervals from the t-tests for the beta hats in the full specifications from all the four consumption spillover tests in the second approach are displayed in Figure A2. The order of magnitudes of the absolute values of the beta hats in the second approach are the same as those in the first approach: one in DGS, -1 in QDW, -2 in BM, and -3 in my test.

Figures A2-A5, as Figures 11-14 in the main text, show the efficacy of the different techniques embedded in the full specification of DGS, QDW, BM, and my test, in the second approach. These estimation results in the second approach only have one difference from those in the second approach. For all four tests, the curves in the baseline regressions become invertedly u-shaped here in the second approach instead of being u-shaped in the first approach. All the techniques embedded in the full specifications of the four consumption spillover tests in the second approach perform the same way as in the first approach.

Table A1, as Table 1 in the main text, check if my test preserves the consumption spillovers. The results here is the same as the main text.

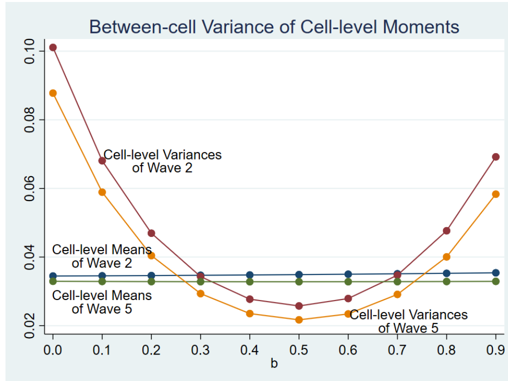

Figure A1: The Between-cell Variance of Cell-level Moments of the Observed Log Income – Approach 2. The data is from the realistic simulation, described by equations (13), (14), and (17). The coefficient b on the horizontal axis in both panels represents the weight of measurement error, as defined in equation (17). The left panel presents the variance of the observed log household income at different weights for both waves. The right panel presents the between-cell variances of cell-level means and cell-level variances of the observed log household income for both waves.

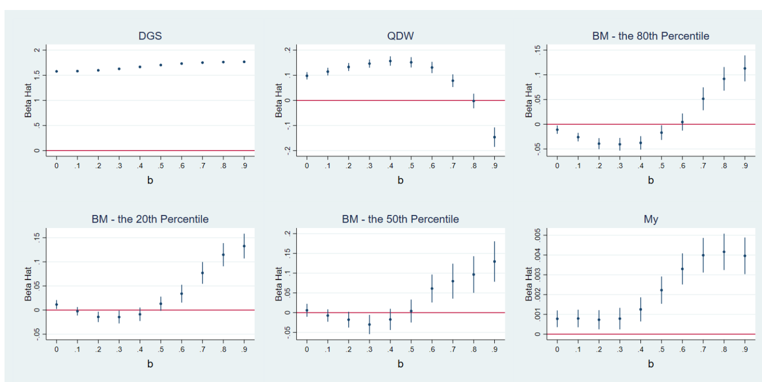

Figure A2: Full Specifications of All the Four Tests – Approach 2. This figure compares the beta hats estimated by the full specifications of all four tests at different weights of measurement error. The data is from the realistic simulation, described by equations (13), (14), and (17). Fifty such simulations are constructed with 50 different seeds. In each panel, the horizontal axis labels the mean and the 95% confidence interval of the 50 beta hats from a one-sample t-test. The coefficient b on the horizontal axis in all the panels represents the weight of measurement error, as defined in equation (17). Details of these specifications can be found in Table 4.

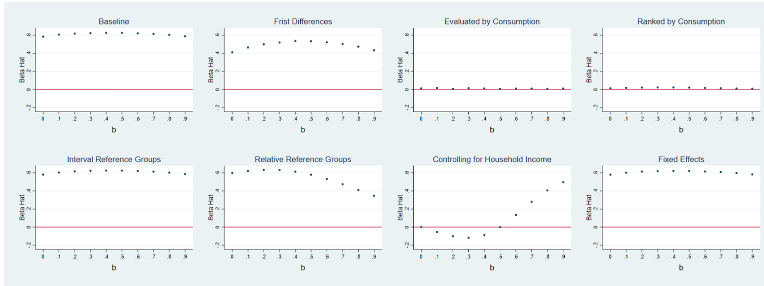

Figure A3: Break Down Drechsel-Grau and Schmid (2014) - Approach 2. This figure compares the beta hats estimated by the different techniques embedded in the full specification of Drechsel-Grau and Schmid (2014) with its baseline at different weights of measurement error. The data is from the realistic simulation, described by equations (13), (14), and (17). Fifty such simulations are constructed with 50 different seeds. In each panel, the horizontal axis labels the mean and the 95% confidence interval of the 50 beta hats from a one-sample t-test. The coefficient b on the horizontal axis in all the panels represents the weight of measurement error, as defined in equation (16). Details of these specifications can be found in Table 5.

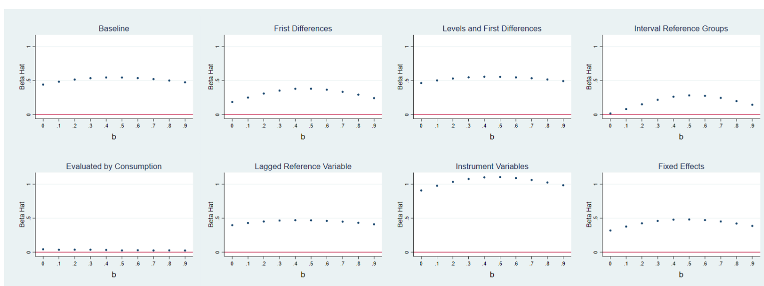

Figure A4: Break Down Quintana-Domeque and Wohlfart (2016) – Approach 2. This figure compares the beta hats estimated by the different techniques embedded in the full specification of Quintana-Domeque and Wohlfart (2016) with its baseline at different weights of measurement error. The data is from the realistic simulation, described by equations (13), (14), and (17). Fifty such simulations are constructed with 50 different seeds. In each panel, the horizontal axis labels the mean and the 95% confidence interval of the 50 beta hats from a one-sample t-test. The coefficient b on the horizontal axis in all the panels represents the weight of measurement error, as defined in equation (17). Details of these specifications can be found in Table 6.

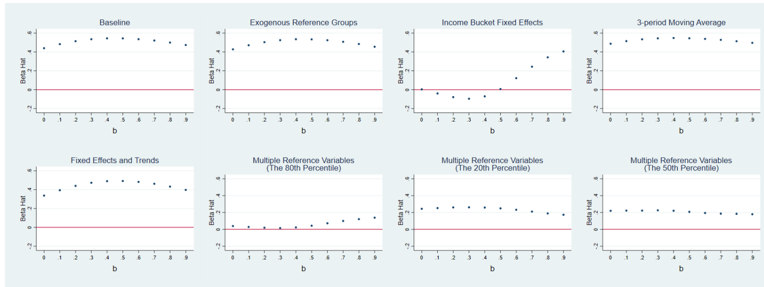

Figure A5: Break Down Bertrand and Morse (2016) – Approach 2. This figure compares the beta hats estimated by the different techniques embedded in the full specification of Bertrand and Morse (2016) with its baseline at different weights of measurement error. The data is from the realistic simulation, described by equations (13), (14), and (17). Fifty such simulations are constructed with 50 different seeds. In each panel, the horizontal axis labels the mean and the 95% confidence interval of the 50 beta hats from a one-sample t-test. The coefficient b on the horizontal axis in all the panels represents the weight of measurement error, as defined in equation (17). Details of these specifications can be found in Table 7.

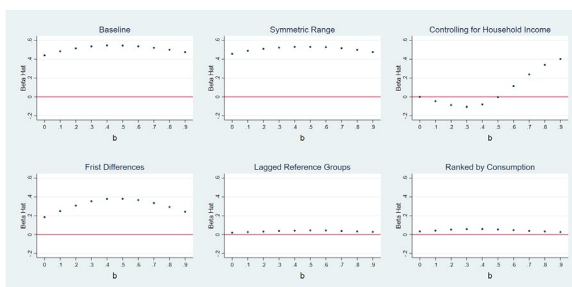

Figure A6: Break Down My Test – Approach 2. This figure compares the beta hats estimated by the different techniques embedded in the full specification that I propose with its baseline at different weights of measurement error. The data is from the realistic simulation, described by equations (13), (14), and (17). Fifty such simulations are constructed with 50 different seeds. In each panel, the horizontal axis labels the mean and the 95% confidence interval of the 50 beta hats from a one-sample t-test. The coefficient b on the horizontal axis in all the panels represents the weight of measurement error, as defined in equation (17). Details of these specifications can be found in Table 8.

Table A1: My Test with the Presence of Spillovers – A Second Approach

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent variable	*lnC*	*DlnC*	*lnC*	*DlnC*	*lnC*	*DlnC*	*lnC*	*DlnC*
Sample	Y<P80Y	LagP20C< LagC< LagP80C	Y<P80Y	LagP20C< LagC< LagP80C	Y<P80Y	LagP20C< LagC< LagP80C	Y<P80Y	LagP20C< LagC< LagP80C
Unit of group	Location-period	Location-period	Division-quarter	Division-quarter	Location-period	Location-period	Division-quarter	Division-quarter
	Without Spillovers				With Spillovers			
	Extreme Simulation		Realistic Simulation		Extreme Simulation		Realistic Simulation	
	Baseline	My	Baseline	My	Baseline	My	Baseline	My
lnY(The 80th income percentile)	0.419***				0.519***			
	(0.000367)				(0.000367)			
DlnY(The previous 80th consumption percentile)		0.0256***				0.0633***		
		(0.000322)				(0.000329)		
lnY(The 80th income percentile)			0.550***				0.639***	
			(0.00575)				(0.00564)	
DlnY(The previous 80th consumption percentile)				0.00131				0.0131***
				(0.00146)				(0.00220)
lnY or DlnY	No	Yes	No	Yes	No	Yes	No	Yes
R-squared	0.395	0.752	0.0151	0.182	0.501	0.765	0.0210	0.00290
Observations	2000292	1468505	597839	448379	2000292	1443379	597840	230549

Notes: The data in columns 1, 2, 5, and 6 is generated by an extreme simulation with consumption spillovers. The data generating process is described by equations (7) , (16), and (17). The data in columns 3, 4, 7, and 8 is generated by a realistic simulation. The data generating process is described by equations (13), (14), (16), and (17). The dependent variable lnC denotes the logarithm of the household consumption in a period, while DlnC denotes the first difference. In columns 1, 2, 5, and 6, the sample is the set of observations whose incomes are less than the 80th percentile of the income distribution. A group within which the observations are ranked to get the distribution in those columns is the set of observations in the given period and the local area in which the observation resides. In columns 3, 4, 7, and 8, the sample is the set of the observations whose consumption in the previous period is less than the 80th percentile of consumption distribution in the previous period. A group within which the observations are ranked to get the distribution in these columns is the set of observations in the given quarter and the census division in which the observation resides. For each reference variable, the reference group is in the parenthesis. A "percentile" in the reference group refers to the observations used to calculate the percentile. The variable used to evaluate the reference group and the operators of the logarithm ("ln") and the first difference ("D") are outside the parenthesis. Standard errors are in parenthesis. ***, **, * denote statistical significance at the 1, 5, and 10% level, respectively.