The impact of Sugar, Processed food and Fat consumption on Prevalence of Comorbidities in India: Testing Popkin's Nutrition in Transition hypothesis

Pratha Srivastava and Harshit Kumar

Prof Bharat Ramaswami

19th April, 2022

Abstract

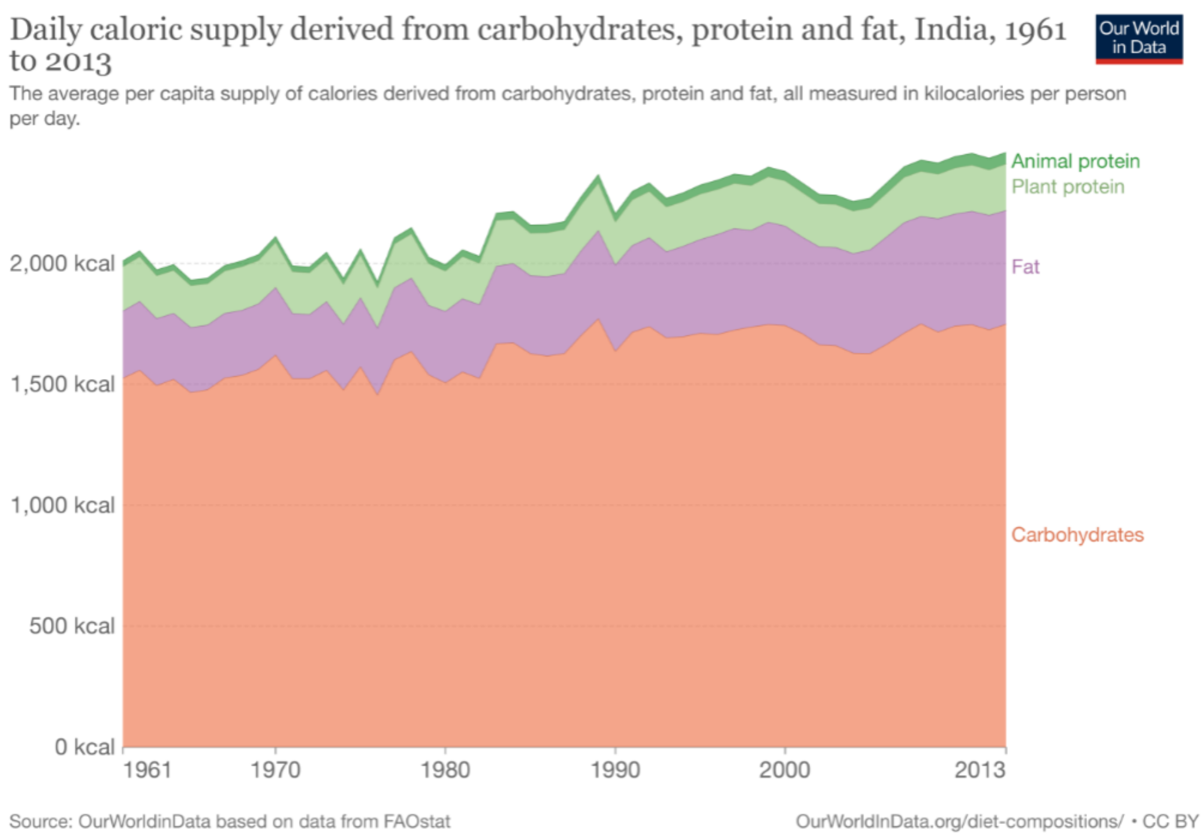
Barry Popkin's hypothesis states that as a country transitions from first to second stage, it is characterized by high consumption of animal source foods, processed foods, fats, sugar and less physical activities. Associated changes include increased incidences of comorbidities. We test the hypothesis for India using the IHDS-II dataset. We tested whether the households with higher consumption of processed foods, sugar and fats recorded in the 2011 survey caused higher probability of having co-morbidities. We rely on the 2SLS instrumental variable approach to test causality and use income as our instrument. We find that - based on our data, while increase in income led to an increase in meat consumption, sugar consumption and processed food consumption, our analysis does not confirm Barry Popkins hypothesis as we find most of the results inconclusive and statistically insignificant.

Introduction

In this paper we try to estimate the impact of sugar, oil and processed foods consumption on major morbidities in India. The Barry Popkins hypothesis states that a country goes through nutrition transition which consists of 3 stages . According to him, the first stage is characterized by poor countries, where people consume a diet mainly of plant based food that is inadequate in calories. Second stage occurs when countries get richer, people consume more calories overall, including more animal source foods, processed foods, fats, sugar and other sweeteners. People also perform less mandatory physical activity. Finally in the third stage people transition to eating less fat, sugar and processed food and more whole grains , fruits and vegetables. In addition people perform physical activity in their leisure time.

Based on the following graph, one can determine that between the period starting 2000 till 2013 India shows an improvement in average calorie supply, but this improvement is primarily in carbs and fats with a modest improvement in animal protein.We can say that India entered, what Popkins describes as the second stage. If India is in its second stage, according to Popkins, it is characterized by higher overall consumption of animal sourced food, processed food, sugar and other sweeteners and fewer people engaging in less mandatory physical activity for food production during this period.

A number of studies have concluded that associated changes of transition from first to second stage include -a general decline in morbidity but rise in chronic diseases like diabetes, heart disease, hypertension and cancer. We want to use IHDS -2 data to test if we can see households with more consumption of sugar, processed food and fats have higher rates of comorbidities. We want to test if the differing consumption of sugar, oil and processed food across households has a causal impact on prevalence of comorbidities. We understand that there may be a number of other factors that decide the existence of diseases, therefore in our estimation, we try to account for these by adding a bunch of covariates and fixed effects.Given that the Popkins hypothesis is based on the assumption that the transitions are a function of income, we use total income of the household as an instrument to establish causality. We explain more on the instrument used and covariates we have used in the empirical strategy and methodology section of this paper.



Literature review

A number of existing literature pieces study the prevalence of comorbidities in countries (developing) that are undergoing nutritional transition. Many of these studies have concluded that countries that see nutritional transitions, especially developing countries, have an increased

incidence of comorbidities. In developing countries, poverty, poor sanitation and hygiene are still common. L. Leocadio et al in their paper “The Transition from undernutrition to overnutrition under adverse environment and poverty:The risk of Chronic Diseases” conclude that populations living under such adverse environments and facing the nutritional transition have increased risks for chronic illnesses in later life, including diabetes, cardiovascular, and neurodegenerative diseases. The double burden of malnutrition is defined by the world health organization as “the coexistence of undernutrition along with overweight, obesity or diet-related non-communicable diseases (NCD), within individuals, households and populations, and across the life-course”[1] (L. Leocadio et al). At present, all high-income and many low- and middle-income countries are in a stage of the transition where nutrition-related noncommunicable diseases including obesity, type 2 diabetes, and hypertension are dominating adult morbidity and mortality and are very high or growing rapidly in prevalence[2].(Popkin)
It is important to study the effect of nutrition on co-morbidity from an economics lense because of the burden of these diseases. Cancer is the second most common disease for maximum deaths in India and the world(Jemal et al, 2007). An estimated 6,00,000 to 7,00,000 deaths were caused in India in 2012 alone, which is equivalent to the total number of deaths reported by the central government during the coronavirus pandemic.. There have been various studies that suggest that prevalence of cancer is partially affected by changing lifestyle choices, particular dietary intake i.e consumption choices(Sinha, Anderson 2003).

[1] L. Leocadio et al. The Transition from undernutrition to overnutrition under adverse environment and poverty:The risk of Chronic Diseases.
[2] Popkin, BM, Ng, SW. The nutrition transition to a stage of high obesity and noncommunicable disease prevalence dominated by ultra-processed foods is not inevitable.

The paper "Nutrition in transition:the changing global nutrition challenge" by Barry Popkins[3] suggests a basic shift in eating preferences, induced mainly by shifts in income, prices and food availability, but also by the modern food industry and the mass media. The remarkable shift in the occupation structure in lower-income countries from agricultural labor towards employment in manufacturing and services implies a reduction in energy expenditure. dietary and physical activity shifts are desirable in many ways. Yet, they carry with them many unwanted nutritional and health effects. This paradox and complexity makes it difficult to understand how to act in order to arrest the negative aspects of the nutrition transition. The changes are occurring most rapidly in lower-income countries, as is shown by the shifts in the distribution of the population, income and occupation patterns.

Data and Summary Statistics

We use the Indian Human Development Survey -II dataset for our estimation. IHDS is a nationally representative survey of 41,554 households conducted in 2004-2005. The initial IHDS sample consists of 26,734 rural and 14,820 urban households. The rural sample contains 13,900 rural households who were interviewed in 1993-94 in a previous survey by NCAER (Human Development Profile of India (HDPI)) and 27,654 new households. The urban sample was a stratified sample of towns and cities within states (or groups of states) selected by probability proportional to population (PPP). Of the 593 districts in India in 2001, 384 are included in IHDS. The sample is spread across 1,503 villages and 971 urban blocks. IHDS-II re-interviewed 83% of the original households as well as split households residing within the village and an additional sample of 2,134 households. The final sample size for IHDS-II is 42,152 households; 27,579

[3] Barry Popkin. Nutrition in transition:the changing global nutrition challenge. *Asia Pacific J Clin Nutr (2001) 10(Suppl.): S13–S18*

rural and 14,573 urban. These households are spread across 33 states and union territories, 384 districts, 1420 villages, and 1,042 urban blocksThe IHDS-II has data on the following areas which are crucial for our analysis- caste & community, consumption and standard of living, energy use ,income, agriculture, employment, government subsidies, education, social and cultural capital ,household & family structure, marriage, gender relations, fertility, health, village infrastructure We merge the data that was present separately for individual and household.

The list of variables used have been tabulated in Table I along with their summary statistics. Additional variables created for our estimation strategy have been discussed in the next section.

Empirical Strategy and Methodology

As the income of countries increases, the average calorie consumption increases and the nutritional composition of the diet also changes, tending towards more fats, sugar and processed foods. We use total income of a household as an instrument because income can affect the prevalence of comorbidities only through its effect on consumption patterns . We do acknowledge the fact that income may be impacting comorbidities through other channels, such as better access to healthcare, better sanitation etc. To account for this, we add other channels through which income may be impacting the existence of co-morbidities as covariates in our estimation equation. We also consider the fact that certain lifestyle choices such as alcohol and cigarette consumption may be significant determiners of prevalence of co-morbidities and therefore we use them as covariates. We use kilograms of sugar, process food and liters of oil consumed by each household instead of the amount spent on each of thes. We do this because the amount spent on these products would not help us identify the amount consumed. The amount spent would vary with prices and other factors. We use religion as a covariate because studies show that religion is correlated with comorbidities and income as well. Religion was found to have a significant cultural factor because Muslim women were found more likely to be overweight or obese than women from other religious groups -primarily Hindu (Griffiths and Bentely).[4] We also find that muslims perform worse off in almost all metric of employment and unemployment according to the NSSO data[5] , therefore income will also be correlated with religion We also include state fixed effects as there may be significant differences between states- such as some states may have better health awareness programs, perform better on development outcomes , difference in state institutions etc.

As stated earlier, we limit the scope of our comorbidities, based on our data set, to cataract, tuberculosis, High blood pressure, leprosy, cancer and asthma– with *comorb* taking the value 1 if the household member has any of the above stated comorbidities and 0 otherwise. In addition, we create new dummy variables to categorize the usage of substances such as alcohol and cigarettes using *alc_consumption* and *cig_consumption* with each taking the value 1 if the household member consumes them and 0 otherwise. We have assumed that consumption of any of these at any scale should impact the prevalence of comorbidities and thus have excluded the frequency of consumption. *insurance* serves as a dummy variable to indicate whether or not household members hold health insurance: govt or private. Multiple health insurances come with perks of free periodic tests and checkups throughout the course of the plan and we believed that having one would definitely help in increasing the chance of diagnosis of a comorbidity.

[4] Paula L. Griffiths, Margaret E. Bentley, The Nutrition Transition Is Underway in India
[5] Sunita Sanghi, A Srija. Employment trends among Religious Communities of India

Our analysis can be broken down into two stages, where we first attempt to understand the correlations between probability to have a comorbidity with sugar, oil and processed food consumption. In the first stage we run a multivariate OLS regression analysis using a linear probability model to understand if data validated our findings with the initial hypothesis.

$$comorb = \beta_0 + \beta_1 sugar\_consumption\_kg + \beta_2 processedfood\_consumptio +$$
$$\beta_3 oil\_vanaspati\_lt + \beta_4\, int\_coffee + \beta_{ij} X_i$$

where X_i includes covariates such as insurance, alcohol consumption, cigarette consumption, age, gender, region, insurance and so on.

After finding correlations for these particular variables, we include religion, access to public meetings, state fixed effects to further control our estimates and get robust results.

Our second part of analysis is our attempt to find a causal relationship between prevalence of comorbidities and consumption of sugar, processed food and oil . We use a two stage least squares regression using an instrumental variable to see the causal relationship. We argue that income directly impacts the consumption of sugar, processed foods, meat and so on which then affect the probability of having a comorbidity. Income also plays a role in the kind of lifestyle or elisure choices that an individual makes based on the resources available. This can include, and not limited to higher consumption of intoxicants, frequency of eating out, living in an urban or rural setting and so on. Furthermore, a higher income also leads to a better access to healthcare infrastructure and private facilities that may help in diagnosis and treatment of any of the comorbidities. With India in the second stage of transition, based on the Barry Popkins hypothesis, we expected income to have a causal relationship with higher comorbidity prevalence.

To use an IV estimation strategy, one must satisfy two conditions:

(1)instrumental relevance: corr$(Z_i, X_i) \neq 0$ and

(2)instrumental exogeneity: corr$(Z_i, u_i) = 0$

We check for the second condition first by running OLS regressions between comorb and total income. We do these regressions both with and without covariates to get unbiased estimates. As can be seen in Table II, We find the coefficient of total income to be very small: in the magnitudes of 10^{-8}. We therefore argue that income could serve as an exogenous instrument which was also affecting other covariates and so is relevant.

The equations for our 2sls regressions are:

$$sugar\ consumption = \beta_0 + \beta_1 tot\ income + \delta$$

$$comorb = \beta_0 + \beta_1 sugar\ consumption + \beta_{ij} X_i$$

Where δ and X_i are covariates

We do similar regressions for $processed\ foods\ consumption\ and\ vanaspati\ oil$, results for which have been discussed in detail in the following sections.

Results:

Table II compiles our results for the simple OLS regression as part of our first stage of analysis. As could be expected, age and gender has a direct effect on comorbidities with both being positively correlated and statistically significant at the 99% confidence interval. Furthermore, we find that people in our dataset living in urban conditions were much more vulnerable to comorbidities than people in living rural areas. This could be due to the prevalence of higher levels of pollutants, greenhouse emissions and other conditions. We find *urban* to be statistically significant at 99% confidence interval. As for consumption of intoxicants such as cigarette and alcohol, we find both of them positively correlated; however, cigarette consumption only is statistically significant. We also find religion to be a factor affecting the probability of comorbidities.

As for our primary variables of interest which are sugar consumption, meat consumption and other processed foods, most of our variables are in line with our initial hypothesis except meat consumption and oil/ vanaspati consumption which were negatively correlated. The mixed results lead us to our second part of analysis that uses instrumental variables for 2sls regression.

We perform three IV regressions, one each for sugar consumption, processed food consumption and meat consumption. Table IV, V and VI show our results in detail. As we can see in Table IV, sugar consumption is no longer a statistically significant variable and not in line with the simple OLS regression. Furthermore, insurance becomes a statistically significant variable implying that income does play a role in who can afford to get insurance and help in diagnosis of comorbidities which could further lead to an increase in probability of comorbidity. Table V shows that meat is now a statistically significant variable which could also be because increase in income would increase the affordability of meat consumption. Processed_foods in table VI also appears to no longer be statistically significant now and has similar implications as those put forth by Table IV.

Therefore, based on our data, while increase in income led to an increase in meat consumption, sugar consumption and processed food consumption, our analysis does not confirm Barry Popkins hypothesis as we find most of the results inconclusive and statistically insignificant. This could be due to multiple factors that could not be taken into consideration given the limited scope of the paper. As the dataset contains less than 5% of respondents who disclose that they have a comorbidity could also mean the results are weak.

Caveats:

Only a small size of the population in the dataset has comorbidities. Therefore the results we obtain are weak. It is also the reason for a small R square. As the diagnosis levels for comorbidities are still very low, there is a plausible deniability which limits our research and scope for research. Furthermore, as we use sugar consumption, meat consumption and processed foods in direct terms of kilogram consumption rather than normalized consumption for each individual in the household, we cannot be sure of how each individual consumes and how over consumption by a particular individual impacts consumption at the household level. A time series analysis with a bigger sample size could probably relay more robust results.

Discussion and conclusion-

Using the IHDS-II dataset we cannot conclusively state that increased consumption of sugar, processed food and oils has a causal impact on the presence of comorbidities in India. There are a number of reasons we see these results. Nevertheless the topic is of great significance and requires better datasets.

A number of studies conclude that calorie consumption dominated by high levels of processed foods, fats and sugar lead to negative public health outcomes. These detrimental public health outcomes can be reduced if countries successfully transition into the third stage. Depending on economic growth alone for this transition is not an ideal option. So far, departments under different ministries, directly or indirectly linked to salt, sugar, and fat markets in India, have facilitated a pro-investment environment without realizing their potential public health impacts. Policy initiatives can influence consumption of salt, sugar, and fat (visible and invisible forms).(Dasgupta, Rajib et al) [6]. Given that there is opportunity for businesses to make profits at the cost of public health, promotion of products for private benefit can be disguised as a public benefit of better food security. There is therefore an immediate need for better policy intervention for better public health outcome, better food security and the positive externality of economic growth that comes with it.

[6] Dasgupta, Rajib et al. "Sugar, salt, fat, and chronic disease epidemic in India: is there need for policy interventions?

References

Popkin, Barry . Nutrition in transition:the changing global nutrition challenge. *Asia Pacific J Clin Nutr (2001) 10(Suppl.): S13–S18*

Popkin, BM, Ng, SW. The nutrition transition to a stage of high obesity and noncommunicable disease prevalence dominated by ultra-processed foods is not inevitable. *Obesity Reviews*. 2022; 23(1):e13366. doi:10.1111/obr.13366

Dasgupta, Rajib et al. "Sugar, salt, fat, and chronic disease epidemic in India: is there need for policy interventions?." *Indian journal of community medicine : official publication of Indian Association of Preventive & Social Medicine* vol. 40,2 (2015): 71-4. doi:10.4103/0970-0218.153858 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4389505/

L. Leocadio et al. The Transition from undernutrition to overnutrition under adverse environment and poverty:The risk of Chronic Diseases. *Front. Nutr., 23 April 2021 | https://doi.org/10.3389/fnut.2021.676044*

Paula L. Griffiths, Margaret E. Bentley, The Nutrition Transition Is Underway in India, *The Journal of Nutrition*, Volume 131, Issue 10, October 2001, Pages 2692–2700, https://doi.org/10.1093/jn/131.10.2692

Popkin, Barry . Nutrition in transition:the changing global nutrition challenge. *Asia Pacific J Clin Nutr (2001) 10(Suppl.): S13–S18*

Popkin, BM, Ng, SW. The nutrition transition to a stage of high obesity and noncommunicable disease prevalence dominated by ultra-processed foods is not inevitable. *Obesity Reviews*. 2022; 23(1):e13366. doi:10.1111/obr.13366

Sinha R, Anderson D E, McDonald S S, Greenwald P. Cancer Risk and Diet in India . J Postgrad Med 2003;49:222-8

Sunita Sanghi, A Srija. Employment trends among Religious Communities of India https://ies.gov.in/pdfs/Employment_Trends_among_Religious_Communities_of_India.pdf

Appendix

TABLE I

Descriptive Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
comorb	204569	.065	.246	0	1
religion	204569	.298	.708	0	3
female	204569	.501	.5	0	1
cig consumption	204569	.865	.342	0	1
alc consumption	204569	.892	.31	0	1
insurance	204569	.108	.31	0	1
insurance	204569	.108	.31	0	1
urban	204569	.342	.474	0	1
EQ14 9.3 Major Morbidity Cataract	204568	.02	.193	0	2
EQ14 9.4 Major Morbidity Tuberculosis (	204568	.007	.11	0	2
EQ14 9.5 Major Morbidity High Blood Pre	204568	.062	.346	0	2

EQ14 9.6 Major Morbidity Heart disease	204568	.016	.177	0	2
EQ14 9.7 Major Morbidity Diabetes	204568	.034	.26	0	2
EQ14 9.8 Major Morbidity Leprosy	204568	.001	.046	0	2
EQ14 9.9 Major Morbidity Cancer	204568	.001	.05	0	2
EQ14 9.10 Major Morbidity Asthma	204568	.018	.19	0	2
HQ19 11.6 Highest adult education	204538	8.452	5.074	0	16
HQ24 14.16 Tea & coffee: Total value Rs	204463	102.994	106.254	0	2000
HQ24 14.17 Processed foods: Total value	204456	94.526	118.429	0	10200
HQ24 14.18 Paan/tobacco/intoxicants: Tot	204428	145.14	259.398	0	12000
HQ24 14.20 Restaurants/Eating out: Total	204452	93.371	286.651	0	20000
HQ23 14.9a Edible oil and vanaspati: Lit	148262	4.009	3.281	0	120

HQ23 14.7t Meat: Rs	82962	412.795	423.776	0	7500
HQ23 14.7a Meat: kg	117623	1.592	3.004	0	120
HQ23 14.3 Sugar: total Rs	204569	164.249	205.645	0	9600
HQ23 14.3a Sugar: kg	204410	5.172	6.04	0	300

TABLE II

Linear regression

comorb	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
tot_income	0	0	9.54	0	0	0	***
Constant	.061	.001	94.87	0	.06	.063	***

Mean dependent var	0.065	SD dependent var	0.246
R-squared	0.000	Number of obs	204569
F-test	90.994	Prob > F	0.000
Akaike crit. (AIC)	6783.390	Bayesian crit. (BIC)	6803.847

*** *p<.01, ** p<.05, * p<.1*

Linear regression

comorb	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
tot_income	0	0	1.34	.179	0	0	
age	.004	0	155.23	0	.004	.004	***
female	.009	.001	8.14	0	.007	.011	***
: base 0	0	.	.	.	.	.	
1	.009	.002	5.77	0	.006	.012	***
2	.004	.003	1.22	.221	-.002	.01	
3	.011	.003	4.07	0	.006	.016	***
urban	.018	.001	16.10	0	.016	.02	***
alcConstantumption	-.004	.002	-1.80	.071	-.009	0	*
cigConstantumption	.029	.002	12.59	0	.025	.034	***

eating_out	0	0	0.19	.845	0	0	
insurance	.012	.002	6.93	0	.008	.015	***
Constant	-.094	.002	-45.42	0	-.098	-.09	***

Mean dependent var	0.065	SD dependent var	0.246
R-squared	0.115	Number of obs	204448
F-test	2425.165	Prob > F	0.000
Akaike crit. (AIC)	-18187.844	Bayesian crit. (BIC)	-18065.108

**** p<.01, ** p<.05, * p<.1*

TABLE III

	(1)	(2)
	comorb	comorb
sugar_consumpti~g	0**	.001***

	(0)	(0)
MEAT_consumptio~g	-.001***	-.001***
	(0)	(0)
oil_vanaspati_lt	-.001***	-.002***
	(0)	(0)
processed_foods	0***	0***
	(0)	(0)
tea_coffee	0***	0***
	(0)	(0)
age		.004***
		(0)
female		.009***
		(.001)
0bn.religion		
1.religion		.015***
		(.002)
2.religion		-.032***

		(.005)
3.religion		.01***
		(.003)
urban		.015***
		(.001)
alc_consumption		-.002
		(.003)
cig_consumption		.02***
		(.003)
eating_out		0**
		(0)
insurance		.002
		(.002)
_cons	.057***	-.078***
	(.001)	(.003)
Observations	110672	110671
R-squared	.001	.108
other controls	no	Yes

state fixed effects	No	No
cluster	No	No

Standard errors are in parentheses

**** p<.01, ** p<.05, * p<.1*

TABLE IV

Instrumental variables 2SLS regression

comorb	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
sugarConstantumpti~g	0	0	1.35	.176	0	.001	
age	.004	0	155.24	0	.004	.004	***
female	.009	.001	8.14	0	.007	.011	***
: base 0	0	.	.	.	.	.	
1	.009	.002	5.68	0	.006	.012	***
2	.004	.003	1.36	.175	-.002	.011	
3	.01	.003	3.62	0	.005	.015	***
urban	.018	.001	16.47	0	.016	.02	***
alcConstantumption	-.004	.002	-1.80	.071	-.009	0	*

cigConstant umption	.029	.002	12.57	0	.025	.034	***
eating_out	0	0	0.33	.738	0	0	
insurance	.012	.002	7.07	0	.008	.015	***
Constant	-.096	.002	-38.86	0	-.101	-.091	***

Mean dependent var	0.065	SD dependent var	0.246
R-squared	0.115	Number of obs	204369
Chi-square	26648.774	Prob > chi2	0.000

****** p<.01, ** p<.05, * p<.1***

TABLE V

Instrumental variables 2SLS regression

comorb	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
MEATConsta ntumptio~g	.011	.005	2.26	.024	.001	.021	**
age	.004	0	98.97	0	.004	.004	***
female	.01	.001	7.17	0	.007	.013	***

: base 0	0	.	.	.	.	.	
1	-.019	.013	-1.40	.162	-.045	.007	
2	-.081	.02	-3.98	0	-.121	-.041	***
3	.02	.004	5.45	0	.013	.027	***
urban	.01	.002	4.72	0	.006	.015	***
alcConstantu mption	-.003	.003	-1.03	.305	-.01	.003	
cigConstantu mption	.02	.003	6.33	0	.014	.026	***
eating_out	0	0	-2.72	.007	0	0	***
insurance	.002	.002	1.01	.312	-.002	.007	
Constant	-.088	.004	-21.42	0	-.097	-.08	***

Mean dependent var	0.059	SD dependent var	0.237
R-squared	0.086	Number of obs	117601
Chi-square	13884.388	Prob > chi2	0.000

****** p<.01, ** p<.05, * p<.1***

TABLE VI

Instrumental variables 2SLS regression

comorb	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
processed_foods	0	0	1.34	.18	0	0	
age	.004	0	155.30	0	.004	.004	***
female	.009	.001	8.15	0	.007	.011	***
: base 0	0	.	.	.	.	.	
1	.009	.002	5.35	0	.005	.012	***
2	.003	.003	0.96	.338	-.003	.01	
3	.01	.003	3.68	0	.005	.015	***
urban	.017	.001	13.74	0	.015	.02	***
alcConstantumption	-.004	.002	-1.77	.076	-.009	0	*
cigConstantumption	.029	.002	12.48	0	.024	.034	***
eating_out	0	0	-0.65	.516	0	0	

insurance	.011	.002	6.75	0	.008	.015	***
Constant	-.096	.002	-40.07	0	-.1	-.091	***

Mean dependent var	0.065	SD dependent var	0.246
R-squared	0.115	Number of obs	204446
Chi-square	26680.034	Prob > chi2	0.000

****** p<.01, ** p<.05, * p<.1***