

UChicago Oeconomica Econometrics Game

May 8, 2015

1 Introduction

Welcome to the UChicago Oeconomica Econometrics Game! The objective of this game is to test your knowledge of econometrics and offer you an opportunity to apply your skills in a policy relevant empirical setting. The question you are trying to answer in this game is: "What is the impact of obesity on health care costs?"

2 Obesity in the United States

Obesity is a condition characterized by abnormal or excessive fat accumulation that may impair health.¹ Notably, rates of obesity have increased dramatically over the past several decades. It is estimated that 34.9% of U.S. adults (roughly 78.6 million) are obese; one estimate of the annual medical cost of obesity in the U.S. was \$147 billion in 2008.²

Given the magnitude and stakes of the public health problem, many researchers have sought to quantify the exact causal relationship between obesity and medical expenditure. Despite there being consensus that obesity raises medical costs, there is still debate as to what extent and in what manners it does so. Recent research has been conducted by Finkelstein et al (2014)³, Grover et al (2014)⁴, Finkelstein et al (2005)⁵, and Finkelstein et al (2009)⁶.

3 Cawley and Meyerhoefer (2012)

One recent study aimed at more precisely estimating the cost of obesity deserves particular attention: Cawley and Meyerhoefer (2012)⁷.

"To estimate the impact of BMI⁸ and obesity on medical spending, Cawley and Meyerhoefer (2012) estimate a two-part model (2PM) of medical expenditure. The first part of the 2PM estimates

¹World Health Organization (2015)

²Centers for Disease Control and Prevention

³<http://pediatrics.aappublications.org/content/early/2014/04/02/peds.2014-0063.abstract>

⁴<http://www.sciencedirect.com/science/article/pii/S2213858714702293>

⁵<http://www.ncbi.nlm.nih.gov/pubmed/15760288>

⁶<http://content.healthaffairs.org/content/28/5/w822.long>

⁷<http://www.ncbi.nlm.nih.gov/pubmed/22094013>

⁸In the US, obesity is oftentimes defined as having a body mass index ("BMI") in excess of 30 kg/m^2 . A person's BMI is calculated by dividing a person's weight (expressed in kg) by the square of the person's height (expressed in meters).

the probability of positive medical expenditures, while the second part estimates the amount of medical expenditures conditional on having any. [They] specify the first part as a logit model and the second part as a gamma GLM with log link.” By employing this model, Cawley and Meyerhoefer (2012) are able to estimate both the extensive and intensive margins of medical expenditure. For our purposes, however, an estimate of the mean increase in per capita expenditure for the obese will suffice (*i.e.* two-stage least squares will suffice).

Because they are concerned about the potential for endogeneity of BMI/obesity, they use the BMI of the oldest child as an instrument. Their IV results indicate that obesity raises annual medical costs by \$2,741. This estimate is roughly twice the size of previous estimates, which if correct, suggests that obesity is a much larger contributor to total medical expenditure in the United States than previously thought.

4 Data

We rely on the Medical Expenditure Panel Survey (MEPS) for this analysis. The MEPS is a longitudinal survey that is conducted in several waves over a period of a couple years. Although the data are collected at the individual level, the survey is conducted at the household level. Cawley and Meyerhoefer (2012) take advantage of this feature of the MEPS in order to construct their instrument.

Cawley and Meyerhoefer (2012) limit their sample to “adults between the ages of 20 and 64 with biological children between the ages of 11 years and 20 years, and exclude pregnant women.” In order to better identify such adults, they rely on the restricted version of the MEPS. We only have access to the public use version of the MEPS; however, this should not be a severe limitation.

For convenience, the file `meps_2000_2012.dta` is provided to you. This file aggregates the MEPS for the years 2000 through 2012 and restricts the set of variables to those that are most relevant to the task at hand.

The raw MEPS data files for years 2000 through 2005 have been provided (h50, h60, h70, h79, h89, and h97, respectively). We have also included later years of data that Cawley and Meyerhoefer did not use in their analysis, specifically MEPS data for years 2006 through 2012 (h105, h113, h121, h129, h138, h147, and h155, respectively). Should you require any further data, for example CPI data or CPS data, feel free to download and use it.

5 Objectives

There are several main objectives:

1. Understand the econometric model of a peer-reviewed academic article.
2. Replicate the main findings (at least qualitatively) of the article.
3. Identify weaknesses in the econometric model (*e.g.* problems with identification, endogeneity, measurement error) and propose refinements and improvements.

4. Implement your proposed modifications and develop an econometric model that allows you to estimate the effect of obesity on medical expenditure.
5. State the necessary assumptions. Are these reasonable assumptions? Elaborate.
6. Estimate your model and compare your results to Cawley and Meyerhoefer (2012). As necessary, include other tables and graphs that augment the exposition.

6 Additional Comments

You will be judged on the creativity and quality of your work.