

Starting Guide to the Oregon Health Insurance Experiment (OHIE): Using the Data

Current as of: January 2, 2014

This document is intended to offer an overview of the Public Use Data files associated with the Oregon Health Insurance Experiment.

A User Guide and detailed codebooks are available via the Oregon Health Insurance Experiment website (<http://www.nber.org/oregon/data.html>). Users should consult these guides when using the data.

Project Description

In 2008, a group of uninsured low-income adults in Oregon was selected by lottery to be given the chance to apply for Medicaid. This lottery provides a unique opportunity to gauge the effects of expanding access to public health insurance on the health care use, financial strain, and health of low-income adults using a randomized controlled design. The Oregon Health Insurance Experiment followed and compared those selected in the lottery (treatments) with those not selected (controls).

Public Use Data Files Description

There are currently six separate datasets from the OHIE that have been made publicly available via the Inter-university Consortium for Political and Social Research (ICPSR). A list of available datasets is given below. In addition, we have made available code that replicates the portions of Finkelstein et al. (2012), Baicker et al. (2013), Taubman et al. (2014) that analyze these publicly available data. The unit of observation is the individual: records for each individual can be linked across datasets using the unique identifier **person_id** (found in every dataset).

1. Descriptive Variables

(File name: **oregonhie_descriptive_vars.dta**)

This dataset contains demographic characteristics that were recorded when individuals signed up for the lottery and lottery selection. Some of these variables are necessary to replicate Finkelstein et al 2012, Baicker et al (2013), and Taubman et al. (2014).

2. State Program Variables

(File name: **oregonhie_stateprograms_vars.dta**)

This dataset contains information from the state of Oregon on individuals' participation in the following state programs: Medicaid, the Supplemental Nutrition Assistance Program (SNAP), and Temporary Assistance to Needy Families (TANF). This dataset includes the insurance variables necessary to replicate Finkelstein et al 2012, Baicker et al (2013), and Taubman et al. (2014).

3. Initial Mail Survey

(File name: **oregonhie_survey0m_vars.dta**)

This dataset contains variables from a mail survey conducted from June-November 2008, referred to as the "Initial Mail Survey." The survey asked questions about health insurance as well as health care needs, experiences and costs. This dataset includes variables necessary to replicate parts of Finkelstein et al (2012).

4. Six Month Mail Survey

(File name: **oregonhie_survey6m_vars.dta**)

This dataset contains variables from a mail survey from January-May 2009, referred to as the “Six Month Survey.” The survey asked questions about health insurance as well as health care needs, experiences and costs. This dataset includes variables necessary to replicate parts of Finkelstein et al (2012).

5. Twelve Month Mail Survey

(File name: **oregonhie_survey12m_vars.dta**)

This dataset contains variables from a mail survey which began in July 2009 with intensive follow up continuing until March 2010. This survey is referred to as the “Twelve Month Survey.” The survey contained questions about health insurance as well as health care needs, experiences and costs. This dataset includes variables necessary to replicate parts of Finkelstein et al (2012).

6. In-person Survey

(File name: **oregonhie_inperson_vars.dta**)

This dataset contains variables from a series of in-person interviews conducted from August 2009 to October 2011. It is referred to as the “In-person survey.” The survey contained questions about health insurance as well as health care needs, experiences and costs and recorded a series of physical measurements of the survey participant. The dataset contains only the variables necessary to replicate the analysis in Baicker et al (2013).

7. Emergency Department Data

(File name: **oregonhie_ed_vars.dta**)

This dataset contains variables derived from administrative data of all visits to twelve hospital emergency departments in the area of Portland, OR. The variables capture emergency department use between January 1, 2007 and September 30, 2009. This dataset includes variables necessary to replicate the parts of Taubman et al (2014).

Linking Individuals Across Datasets

Records for each individual can be linked across datasets using the unique identifier **person_id** (found in every dataset). No other variables are repeated across datasets. For example, the variable for lottery selection (**treatment**) is found in Descriptive Statistics and no other dataset.

Treatment

The treatment of the OHIE was selection by the lottery. Lottery selection gave an individual the ability to apply for Medicaid. The treatment variable is called **treatment** and can be found in the Descriptive Statistics dataset. This variable can be used to estimate intent to treat estimates of the effect of lottery selection.

In addition to estimates of the effect of lottery selection, estimates of the effect of Medicaid coverage may also be of interest. Variables summarizing Medicaid enrollment information can be found in the State Programs Variables dataset. These variables can be used to estimate local average treatment effect (two-stage least-squares instrumental variable) estimates of the effect of Medicaid coverage.

Selection by the lottery was random conditional on the number of household members on the lottery list (**numhh_list**). It is necessary to control for this in all experimental analyses using indicator variables or another approach. Because treatment occurs at the household level, all analyses from the Oregon Health Insurance Experiment cluster standard errors at the household (identified by **household_id**). Both **numhh_list** and **household_id** are available in the Descriptive Statistics dataset.

Weights

Throughout much of the OHIE analysis, weights were used. Weights that account for the probability of being sampled for the six-month mail survey are given by the variable **weight_6m** in the Six Month Mail Survey dataset. Weights that account for the probability of being sampled and survey procedures are given by the variable **weight_12m** in the Twelve Month Mail Survey dataset. Weights that account for the probability of being sampled and fielding procedures are given by the variable **weight_total_inp** in the In-Person Survey dataset.

References and Replication Code

Finkelstein, Amy et al “The Oregon Health Insurance Experiment: Evidence from the First Year”. *The Quarterly Journal of Economics*. August 2012. Vol 127(3). (Online Appendix also available)

Code has been made available that replicates the analyses in Finkelstein et al (2012) which are performed on the publicly available data. The code includes a Stata .do file (oregon_hie_qje_replication.do) which calls several subprograms. A list of the replicated analyses and instructions for running the .do file can be found at the top of this master .do file.

Baicker, Katherine et al “The Oregon Experiment – Effects of Medicaid on Clinical Outcomes”. *New England Journal of Medicine*. 2 May 2013. Vol 368(18). (Online Appendix also available)

Code has been made available that replicates the analyses in Baicker et al (2013) which are performed on the publicly available data. The code includes a Stata .do file (oregon_hie_nejm_replication.do) which calls several subprograms. A list of the replicated analyses along with instructions for running the .do file can be found at the top of this master .do file.

Taubman, Sarah et al “Medicaid Increases Emergency Department Use: Evidence from Oregon’s Health Insurance Experiment”. *Science* 2 Jan 2014. (Online Appendix also available)

Code has been made available that replicates the analyses in Taubman et al (2014) which are performed on the publicly available data. The code includes a Stata .do file (oregon_hie_science_replication.do) which calls several subprograms. A list of the replicated analyses along with instructions for running the .do file can be found at the top of this master .do file.

Citation for Public Use Data

If obtained through Oregon Health Insurance website: Finkelstein, Amy. Oregon Health Insurance Experiment Public Use Data, 2013. Available at <http://www.nber.org/oregon/data.html>.

If obtained through ICPSR: Finkelstein, Amy. Oregon Health Insurance Experiment Public Use Data. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2013.