

Inequality Pset 4

Julia Du

2021-05-27

Contents

Load necessary libraries	1
Question 1	1
Question 2	1
Question 3	3
Question 4	4
Question 5: Data exercise	5
Question 5a & 5b	5
Question 5c(i): 1st stage eqn by OLS	6
Question 5c(ii): 1st stage OLS results	7
Question 5d(i): Structural eqn by 2SLS	7
Question 5d(ii): Intent-to-treat eqn by OLS	9
Question 5d(iii): $\hat{\pi}$ Coefficients	10
Question 5d(iii): $\hat{\beta}$ coefficients	11
Question 5d(iv)	11

Load necessary libraries

```
library(tidyverse)
library(lubridate)
library(tinytex)
library(stargazer)
library(lfe)

theme_set(theme_minimal())
```

Question 1

The research question is the effect of expanding access to public health insurance (Medicaid) on the health care use, financial strain, and health of low-income adults.

It's interesting because the estimates of public health insurance apply to able-bodied uninsured adults below 100% of the poverty line, which is a population of significant policy interest. It's also interesting there hasn't been a lot of research on benefits of health insurance outside health care utilization. This question lets us look at the other, "risk-spreading" benefits of health insurance, like less financial strain or less emotional stress/worry. Finally, it's interesting because the Oregon experiment is a randomized control trial (RCT), making it the gold standard for internal validity. If we find in our data something contrary to our expectations, RCTs make us confident enough in our experimental design that we don't chalk up surprises to wrong empirical specifications.

Question 2

(2a)

The identifying assumption is that, in the absence of treatment, the treatment group would've had the same distribution of outcomes as the control group.

There are 3 possible violations:

- nonrandom assignment (i.e. random assignment is done incorrectly)
 - i.e. that the assignment of the ability to apply for Medicaid (i.e. lottery results) were randomized & that the treatment/control individuals in the subsamples were not differentially selected from the full sample
 - test: covariate balance test
- differential participation in experiment across treatment groups
 - i.e. attrition rates or people's decision to drop out of experiment (i.e. not respond to survey) is a function of their assigned treatment group. in other words, the decision to drop out is NOT random
 - test: covariate balance test among those who participated, looking @ participation rates across treatment groups
- differential reporting of outcomes across groups
 - e.g. people randomized into opportunity to apply for Medicaid (win lottery, this is the treatment group) may tend to report their health care usage more than those who lose the lottery

In this context, the line between differential participating and differential reporting is a blurry (i.e. pretty much the same thing), per office hours with Prof. Deshpande.

(2b)

To ensure that the lottery indeed randomly selected people, Finkelstein et al. verified the selection process with independent computer simulations within sampling error (pg. 1074). They also demonstrated that this selection procedure created survey samples with a balance of treatment and control traits (pg. 1074) – essentially a covariate balance test.

They also looked to address potential differential participation in experiment, as reflected by response rates. They used a covariate balance test for those who participated, looking at the treatment-control balance for three different samples (the hospital/admin data sample, credit report sample, survey data sample) in Table II. They found that the estimated difference in response rates is quite small. Also, in each of the three samples, they couldn't reject the null of treatment-control balance on the lottery list variables,

the pre randomization measures, or combined set of variables (pg. 1076). This indicates that differential participation in experiment doesn't seem to be an issue here.

(2c)

Differential reporting, when non-random, could lead to biased estimates of the effect of Medicaid on outcomes (healthcare use, financial strain, health), thereby presenting a threat to validity. For example, people randomized into treatment (i.e. those who are selected to have the chance to apply for Medicaid by winning the lottery) may tend to fill out the health surveys more than those who lose the lottery (recall that in this experiment, there isn't really a distinction between differential participation and reporting). We'll see in Question (5b) that lottery winners (the treated) actually do respond to surveys a bit less than lottery losers (control), but we aren't that concerned given how small the difference is (more discussion in 5b).

Also, in this context, we're using administrative data and credit/survey data. There shouldn't be an issue with differential reporting in the admin data as we can see everybody's outcomes. There could be a bigger problem with differential reporting (when it's non-random, e.g. lottery winners report less than lottery losers) if we were solely reliant on credit/survey data, where we don't see everybody's outcomes – but luckily, we have admin data as well.

(2d)

In this case, imperfect take-up is not a threat to validity (i.e. does not lead to biased estimates) because we can observe outcomes with our administrative data. Unlike the RAND experiment, where researchers had no way to know the outcomes of people who dropped out (and thus imperfect take-up did pose a threat to validity and produce downward-biased estimates), we can collect admin data on everyone's outcomes in the OHIE context. Our estimates of the effect of Medicaid on outcomes will just be a local average treatment effect (LATE), estimated off a particular group of people who comply. The first stage estimate is just smaller.

(2e)

Other violations of the exclusion restriction considered is the differential treatment of participants following random assignment. We have to make sure that the control and treatment groups aren't differentially affected by something else after random assignment, so that the only way winning the lottery can affect the outcome variables is through expanded Medicaid access. While it's hard to test for the direct effect of winning the lottery, the researchers do look at enrollment in other public programs. The idea is that if treated participants enroll in TANF or SNAP more than the control after winning the lottery, then the observed effect on outcome cannot be wholly chalked up to Medicaid access (but to Medicaid **and** these other programs). Thus, we want to make sure that winning the lottery doesn't lead to differential enrollment in public programs. We find winning the lottery has no effect on TANF enrollment, but SNAP enrollment increasing by a significant 2pp among winners. However, this is small enough that we can probably ignore it and conclude the exclusion restriction isn't being violated.

Question 3

(3a)

- structural:

$$y_{ihj} = \pi_0 + \pi_1 Insurance_{ih} + X_{ih}\pi_2 + V_{ih}\pi_3 + v_{ihj}$$

- first-stage:

$$Insurance_{ih} = \delta_0 + \delta_1 Lottery_h + X_{ih}\delta_2 + V_{ih}\delta_3 + \mu_{ihj}$$

- reduced form:

$$y_{ihj} = \beta_0 + \beta_1(\text{Lottery}_h) + X_{ih}\beta_2 + V_{ih}\beta_3 + \varepsilon_{ihj}$$

There is ambiguity on the correct endogenous variable as we don't know through which channels Medicaid affects the outcome. If being on Medicaid has a cumulative effect, we'd want to use months on Medicaid as the endogenous variable. If the only thing that matters for outcomes now is if I'm on Medicaid now, maybe we use "on Medicaid at end of study" as the endogenous variable. If I think ever being on Medicaid will affect me permanently, then I use that as the endogenous variable.

We should note also that the differing definitions of endogenous variables can give bounds for the LATE estimate of Medicaid. Using "on Medicaid at end of study" gives a lower bound on the 1st stage estimate – and thus an upper bound on the LATE estimate – while "ever being on Medicaid during study" gives a lower bound on the LATE estimate.

(3b)

X_{ih} are the set of covariates that are correlated with treatment probability (and potentially with the outcome). They include household size dummies, survey wave dummies, and the interactions between survey wave dummies & household size dummies. **They must be controlled for to give an unbiased estimate of β_1** , which is the main coefficient of interest and is ultimately used to describe the relationship between winning the lottery and the outcome. Random assignment is conditional on the number of family members on the lottery list (since individuals who win are disproportionately drawn from larger households). Also, the fraction of treatment individuals varies across the 7 survey waves, so we must include survey wave as a control.

We also include **other covariates V_{ih} to improve power** by accounting for chance differences in variables between treatment and control groups. These aren't needed for an unbiased β_1 estimate, but they help with precision.

Note: β_1 itself gives the average difference in (adjusted) means between the treatment group (the lottery winners) and the control group (those not selected by the lottery); it is the impact of being able to apply for Medicaid through the lottery.

(3c)

The first-stage estimate for "ever on Medicaid" is 0.256, indicating that winning the lottery is associated with a 25.6% increase in the probability of having insurance (specifically Medicaid) during the study period. It is much less than 1 because not all lottery winners became Medicaid participants (only 60% of winners sent in the paperwork to enroll in program, and only 50% of that group actually qualified – resulting in 30% of lottery winners actually receiving Medicaid).

(3d)

OHIE's population of interest is able-bodied uninsured adults below 100% of the federal poverty line who express interest in insurance coverage. In this study, the overall lottery population is older, whiter, and in worse-health than the general low-income, uninsured adult US population. Relative to the overall lottery population, the study's compliers are even older, "more likely white, in worse health, and in lower socioeconomic status" (p. 1078). Thus, this makes us quite skeptical about extrapolating results to populations like children and pregnant women, who are quite different from OHIE's complier population.

We are less skeptical about extrapolating OHIE's results to the ACA because the ACA's target is low-income adults. OHIE gives a pretty good sense of what the ACA expansions are going to do for low-income adults, though we should be careful since OHIE's lottery **and** complier population is whiter, older, and in worse health than the overall low-income adult population targeted by the ACA.

Question 4

(4a)

part i.

Ever being on Medicaid is associated with a 2.1 percentage point increase in the probability of having a hospital admission (making the overall probability 30%), an 8.8 percentage point (overall 15%) increase in the probability of taking any prescription drugs, and a 21 percentage point (overall 35%) increase in the probability of having an outpatient visit. They couldn't reject the null of no change in emergency room utilization, but the confidence intervals allow for substantial effects in either direction (i.e. ER visits may decrease or increase due to being on Medicaid). In addition, insurance is associated with 0.3 standard deviation increase in reported compliance with recommended preventive care such as mammograms and cholesterol monitoring.

part ii.

Ever being on Medicaid results in decreased exposure to medical liabilities and out-of-pocket medical expenses, including a 6.4 percentage point decline in the probability of having an unpaid medical bill sent to a collections agency (making the overall probability 25%) and a 20 percentage point (overall 35%) decline in having any out-of-pocket medical expenditures. Because much medical debt is never paid, expanded coverage appears to affect (specifically, reduce) the financial strain on the newly insured and their medical providers (or whomever they pass the costs on to).

part iii.

Ever being on Medicaid is associated with improvements across the board in measures of self-reported physical and mental health, averaging a 0.2 standard deviation improvement. This improvement seems to partly reflect a general sense of improved well-being. A year after random assignment, Medicaid is associated with a 32% increase in self-reported overall happiness.

(4b)

I think which article headline you find compelling depends on which results you're looking at. The NPR article discussing Medicaid's big effects on people's lives seems more correct when focusing on the QJE 2012 paper (the one we read for this pset). In (4a), we see that ever being on Medicaid in OHIE decreases financial strain and increases self-reported physical and mental health alongside healthcare utilization. This jives with what we heard in class, though other studies show that Medicaid had no statistically significant effect on physical health. Additionally, Finkelstein et al. 2015 & Finkelstein et al. 2017 show that Medicaid value to recipients is substantially lower than Medicaid spending due to existing charity care and bad debt. The low-income insured have substantial implicit insurance even without Medicaid. This is probably where the Washington Post article headline came from.

Overall, I'd still say that the NPR headline seems more accurate, just because we spent the most time with the QJE 2012 paper and I understand the changes in people's lives that Medicaid can bring about – but the Washington Post headline is still valid.

Question 5: Data exercise

Just so you know, the dummy syntax for the `felm` command is as follows:

```
felm(causal relation of interest | fixed effects | IVs | clusters, data = your_data)
```

Question 5a & 5b

```
q4 <- read_csv("./dataexercise_pset4/pset_experiment_data.csv")
```

```
# 5b
```

```
q4 <- q4 %>%
  drop_na(treatment) %>%
  mutate(
    treatment = if_else(treatment == "Selected", 1, 0),
    returned_12m = if_else(returned_12m == "Yes", 1, 0))
```

```
# calculate avg survey response rate
```

```
q4 %>%
  drop_na(returned_12m) %>% # CAN I DROP NA
  group_by(treatment) %>%
  summarize(response_rate = mean(returned_12m))
```

```
## # A tibble: 2 x 2
##   treatment response_rate
##   <dbl>         <dbl>
## 1         0         0.415
## 2         1         0.399
```

```
# t-test to see if difference in rate is sig
```

```
q4_treatonly <- q4 %>%
  filter(treatment == 1)
q4_controlonly <- q4 %>%
  filter(treatment == 0)

t.test(q4_treatonly$returned_12m, q4_controlonly$returned_12m)
```

```
##
## Welch Two Sample t-test
##
## data: q4_treatonly$returned_12m and q4_controlonly$returned_12m
## t = -3.9564, df = 58341, p-value = 7.618e-05
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -0.024056231 -0.008117376
## sample estimates:
## mean of x mean of y
## 0.3991686 0.4152554
```

(5b)

The response rates are pretty similar across the control and treatment groups (41.5% and 39.9%, respectively). Given this closeness, I'd say the difference in survey response rates isn't concerning.

This is supported by the t-test. We see that although we reject the null (i.e. there is a significant difference between the means), the 95% confidence interval contains values quite close to 0. Thus, though there is evidence of differential response rates, the response rates are close enough (much smaller than in the RAND experiment) that it doesn't seem like an existential threat to our experiment. Discussion of Table II in the paper on pg. 1075-76 lends further credence to the idea that this small difference in response rates doesn't disqualify the experiment's results.

Question 5c(i): 1st stage eqn by OLS

```
q4 <- q4 %>%
  mutate(
    ohp_all_mo_survey = parse_number(ohp_all_mo_survey),
    ohp_all_ever_survey = if_else(ohp_all_ever_survey == "Enrolled", 1, 0))

dummies_q4 <- q4 %>%
  select(starts_with ("ddd")) %>%
  colnames()

# drop NAs for weights to
q4_test <- q4 %>%
  drop_na(weight_12m) %>%
  filter(sample_12m_resp == "12m mail survey responder")

q4_1ststage_ols_ever <- felm(as.formula(
  paste("ohp_all_ever_survey ~ treatment", "+",
    paste(dummies_q4, collapse = " + "),
    "| 0 | 0 | household_id", sep = "")),
  data = q4_test, weights = q4_test$weight_12m)

q4_1ststage_ols_mo <- felm(as.formula(
  paste("ohp_all_mo_survey ~ treatment", "+",
    paste(dummies_q4, collapse = " + "),
    "| 0 | 0 | household_id", sep = "")),
  data = q4_test, weights = q4_test$weight_12m)

q4_1ststage_ols_end <- felm(as.formula(
  paste("ohp_all_end_survey ~ treatment", "+",
    paste(dummies_q4, collapse = " + "),
    "| 0 | 0 | household_id", sep = "")),
  data = q4_test, weights = q4_test$weight_12m)

#q4_1ststage_ols_end %>%
# summary("robust")
```

Question 5c(ii): 1st stage OLS results

```
stargazer(q4_1ststage_ols_ever, q4_1ststage_ols_mo, q4_1ststage_ols_end,
  type = "latex",
  keep = "treatment",
  title = "Question 5c(ii)",
  dep.var.labels = c("Ever on Medicaid", "Number of months on Medicaid",
    "On Medicaid at end"),
  omit.stat = c("f", "rsq", "adj.rsq", "ser"),
  dep.var.caption = "OLS 1st stage",
  digits = 4)
```

% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu
% Date and time: Thu, May 27, 2021 - 12:53:14 AM

Table 1: Question 5c(ii)

	OLS 1st stage		
	Ever on Medicaid (1)	Number of months on Medicaid (2)	On Medicaid at end (3)
treatment	0.2902*** (0.0066)	3.9427*** (0.0896)	0.1890*** (0.0061)
Observations	23,741	23,741	23,741

Note:

*p<0.1; **p<0.05; ***p<0.01

Question 5d(i): Structural eqn by 2SLS

*#healthnotpoor has 1, 0, & NA. notbaddays is dbl w/ NAs - don't need if_else for them
note: if_else command preserves NAs*

```
q4 <- q4 %>%
  mutate(
    rx_any_12m = if_else(rx_any_12m == "Yes", 1, 0),
    doc_any_12m = if_else(doc_any_12m == "Yes", 1, 0),
    er_any_12m = if_else(er_any_12m == "Yes", 1, 0),
    hosp_any_12m = if_else(hosp_any_12m == "Yes", 1, 0),
    cost_any_oop_12m = if_else(cost_any_oop_12m == "Yes", 1, 0),
    cost_any_owe_12m = if_else(cost_any_owe_12m == "Yes", 1, 0),
  )

# drop NAs for weights too & select only survey data
# (have to rerun same command as before since we changed q4 above)
q4_test_1 <- q4 %>%
  #drop_na(weight_12m) %>%
  filter(sample_12m_resp == "12m mail survey responder")

q4_struct_ols_a <- felm(as.formula(
  paste("rx_any_12m ~ ",
        paste(dummies_q4, collapse = " + "),
        "| 0 | (ohp_all_ever_survey ~ treatment) | household_id", sep = "")),
  data = q4_test_1, weights = q4_test_1$weight_12m)

q4_struct_ols_a <- felm(as.formula(
  paste("rx_any_12m ~ ",
        paste(dummies_q4, collapse = " + "),
        "| 0 | (ohp_all_ever_survey ~ treatment) | household_id", sep = "")),
  data = q4_test_1, weights = q4_test_1$weight_12m)

q4_struct_ols_b <- felm(as.formula(
  paste("doc_any_12m ~ ",
        paste(dummies_q4, collapse = " + "),
        "| 0 | (ohp_all_ever_survey ~ treatment) | household_id", sep = "")),
  data = q4_test_1, weights = q4_test_1$weight_12m)
```



```

q4_struct_ols_c <- felm(as.formula(
  paste("er_any_12m ~ ",
        paste(dummies_q4, collapse = " + "),
        "| 0 | (ohp_all_ever_survey ~ treatment) | household_id", sep = "")),
  data = q4_test_1, weights = q4_test_1$weight_12m)

q4_struct_ols_d <- felm(as.formula(
  paste("hosp_any_12m ~ ",
        paste(dummies_q4, collapse = " + "),
        "| 0 | (ohp_all_ever_survey ~ treatment) | household_id", sep = "")),
  data = q4_test_1, weights = q4_test_1$weight_12m)

q4_struct_ols_e <- felm(as.formula(
  paste("cost_any_oop_12m ~ ",
        paste(dummies_q4, collapse = " + "),
        "| 0 | (ohp_all_ever_survey ~ treatment) | household_id", sep = "")),
  data = q4_test_1, weights = q4_test_1$weight_12m)

q4_struct_ols_f <- felm(as.formula(
  paste("cost_any_owe_12m ~ ",
        paste(dummies_q4, collapse = " + "),
        "| 0 | (ohp_all_ever_survey ~ treatment) | household_id", sep = "")),
  data = q4_test_1, weights = q4_test_1$weight_12m)

q4_struct_ols_g <- felm(as.formula(
  paste("health_notpoor_12m ~ ",
        paste(dummies_q4, collapse = " + "),
        "| 0 | (ohp_all_ever_survey ~ treatment) | household_id", sep = "")),
  data = q4_test_1, weights = q4_test_1$weight_12m)

q4_struct_ols_h <- felm(as.formula(
  paste("notbaddays_tot_12m ~ ",
        paste(dummies_q4, collapse = " + "),
        "| 0 | (ohp_all_ever_survey ~ treatment) | household_id", sep = "")),
  data = q4_test_1, weights = q4_test_1$weight_12m)

#q4_struct_ols_f %>%
# summary("robust")

```

Question 5d(ii): Intent-to-treat eqn by OLS

```

q4_itt_ols_a <- felm(as.formula(
  paste("rx_any_12m ~ treatment + ",
        paste(dummies_q4, collapse = " + "),
        "| 0 | 0 | household_id", sep = "")),
  data = q4_test_1, weights = q4_test_1$weight_12m)

q4_itt_ols_b <- felm(as.formula(
  paste("doc_any_12m ~ treatment + ",
        paste(dummies_q4, collapse = " + "),
        "| 0 | 0 | household_id", sep = "")),
  data = q4_test_1, weights = q4_test_1$weight_12m)

```

```

data = q4_test_1, weights = q4_test_1$weight_12m)

q4_itt_ols_c <- felm(as.formula(
  paste("er_any_12m ~ treatment + ",
        paste(dummies_q4, collapse = " + "),
        "| 0 | 0 | household_id", sep = "")),
  data = q4_test_1, weights = q4_test_1$weight_12m)

q4_itt_ols_d <- felm(as.formula(
  paste("hosp_any_12m ~ treatment + ",
        paste(dummies_q4, collapse = " + "),
        "| 0 | 0 | household_id", sep = "")),
  data = q4_test_1, weights = q4_test_1$weight_12m)

q4_itt_ols_e <- felm(as.formula(
  paste("cost_any_oop_12m ~ treatment + ",
        paste(dummies_q4, collapse = " + "),
        "| 0 | 0 | household_id", sep = "")),
  data = q4_test_1, weights = q4_test_1$weight_12m)

q4_itt_ols_f <- felm(as.formula(
  paste("cost_any_owe_12m ~ treatment + ",
        paste(dummies_q4, collapse = " + "),
        "| 0 | 0 | household_id", sep = "")),
  data = q4_test_1, weights = q4_test_1$weight_12m)

q4_itt_ols_g <- felm(as.formula(
  paste("health_notpoor_12m ~ treatment + ",
        paste(dummies_q4, collapse = " + "),
        "| 0 | 0 | household_id", sep = "")),
  data = q4_test_1, weights = q4_test_1$weight_12m)

q4_itt_ols_h <- felm(as.formula(
  paste("notbaddays_tot_12m ~ treatment + ",
        paste(dummies_q4, collapse = " + "),
        "| 0 | 0 | household_id", sep = "")),
  data = q4_test_1, weights = q4_test_1$weight_12m)

#q4_itt_ols_a %>%
# summary("robust")

```

For Question 5d(iii), I created 2 separate tables (one of $\hat{\pi}$ estimates for the structural equation, one of $\hat{\beta}$ estimates for the ITT equation) since combining all 16 regressions into 1 table produces something illegible.

Question 5d(iii): $\hat{\pi}$ Coefficients

```

stargazer(q4_struct_ols_a, q4_struct_ols_b, q4_struct_ols_c, q4_struct_ols_d,
  q4_struct_ols_e, q4_struct_ols_f, q4_struct_ols_g, q4_struct_ols_h,
  type = "latex",
  omit = c(dummies_q4, "Constant"),
  title = "Question 5d(iii): Pi Estimates (Structural Eqn)",

```

```

covariate.labels = "Ever on Medicaid",
dep.var.labels = c("rx", "doc", "er", "hosp", "cost oop", "cost owe",
                  "health", "notbad"),
omit.stat = c("f", "rsq", "adj.rsq", "ser"),
dep.var.caption = "",
digits = 4,
font.size = "small",
column.sep.width = "0pt")

```

% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu
 % Date and time: Thu, May 27, 2021 - 12:53:19 AM

Table 2: Question 5d(iii): Pi Estimates (Structural Eqn)

	rx	doc	er	hosp	cost oop	cost owe	health	notbad
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ever on Medicaid	0.0878*** (0.0288)	0.2124*** (0.0252)	0.0223 (0.0231)	0.0077 (0.0136)	-0.1995*** (0.0262)	-0.1798*** (0.0265)	0.0990*** (0.0176)	1.3171** (0.5629)
Observations	18,308	23,492	23,514	23,573	23,426	23,451	23,361	21,881

Note:

*p<0.1; **p<0.05; ***p<0.01

Question 5d(iii): $\hat{\beta}$ coefficients

```

stargazer(q4_itt_ols_a, q4_itt_ols_b, q4_itt_ols_c, q4_itt_ols_d, q4_itt_ols_e,
          q4_itt_ols_f, q4_itt_ols_g, q4_itt_ols_h,
          type = "latex",
          omit = c(dummies_q4, "Constant"),
          title = "Question 5d(iii): Beta Estimates (ITT Equation)",
          covariate.labels = "Ever on Medicaid",
          dep.var.labels = c("rx", "doc", "er", "hosp", "cost oop", "cost owe",
                            "health", "notbad"),
          omit.stat = c("f", "rsq", "adj.rsq", "ser"),
          dep.var.caption = "",
          digits = 4,
          font.size = "small",
          column.sep.width = "0pt")

```

% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu
 % Date and time: Thu, May 27, 2021 - 12:53:19 AM

Question 5d(iv)

For equation (3), π_1 **represents the LATE of insurance** – it is the causal impact of insurance on the outcome variable among the group of individuals who get on Medicaid through winning the lottery and who wouldn't get Medicaid without winning the lottery (i.e. the complier group).

Ex: For the compliers, insurance is associated with an 8.78% increase in the probability of currently taking any prescription meds.

Table 3: Question 5d(iii): Beta Estimates (ITT Equation)

	rx	doc	er	hosp	cost oop	cost owe	health	notbad
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ever on Medicaid	0.0252*** (0.0083)	0.0617*** (0.0074)	0.0065 (0.0067)	0.0022 (0.0040)	-0.0580*** (0.0077)	-0.0523*** (0.0076)	0.0288*** (0.0051)	0.3810** (0.1618)
Observations	18,308	23,492	23,514	23,573	23,426	23,451	23,361	21,881

Note:

*p<0.1; **p<0.05; ***p<0.01

For equation (1), β_1 represents the causal impact of being able to apply for Medicaid through the lottery (i.e. winning the lottery) on the outcome variable.

Ex: Winning the lottery is associated with an 2.52% increase in the probability of currently taking any prescription meds.