

Field Experiments: Design, Analysis and Interpretation

Solutions for Chapter 11 Exercises

Alan S. Gerber and Donald P. Green*

Question 1

Important concepts:

- a) Explain the distinction between a sample average treatment effect and a population average treatment effect. Why might a researcher be primarily interested in one rather than the other?

Answer:

Define a population as a set of subjects from which an experimental sample is drawn. Depending on how a sample is drawn, the ATE for the sample may be similar or different from the ATE for the broader population; large, random samples tend to have similar ATEs to their parent populations. Researchers may be interested in the ATE for the sample because their primary goal is to figure out how the subjects in the experiment respond to the treatment. Or researchers may be interested in the ATE for the population because they seek to draw generalizations about how the intervention would work were it applied to others in the population.

- b) What is a meta-analysis? Why is meta-analysis a better way to summarize research findings than comparing the number of studies that show significant estimated treatment effects to the number of studies that show insignificant estimated treatment effects?

Answer:

Meta-analysis refers to statistical procedures designed to summarize the results of research literatures. Meta-analysis is sometimes described as a “systematic” method for constructing a literature review because it summarizes research findings based on a replicable formula. Specifically, when meta-analysis is used to pool several studies, each study’s experimental result is weighted according to a formula that follows from an underlying statistical model. In this chapter, the model involves random sampling from a large population, and the formula (fixed effects meta analysis) weights each study to the inverse of its precision, or squared standard error. This procedure is superior to a count of studies that show significant or insignificant results because the latter potentially accords too much weight to small studies that produce statistically insignificant results and too little weight to large studies that convincingly demonstrate an effect when other, smaller studies fail to do so. Another advantage of meta-analysis over this head-count method is that meta-analysis generates a point estimate and confidence interval, which is more informative than a summary statement about statistical significance.

- c) Using equations (11.2), (11.3), and (11.4), provide a hypothetical example to illustrate how uncertainty about the possibility of bias affects the way in which prior beliefs are updated in light of new evidence.

Answer:

*Solutions prepared by Peter M. Aronow and revised by Alexander Coppock

Suppose that a researcher were to conduct a study on the effects of SAT prep classes on SAT scores using an observational design that compares a national random sample of high school seniors who take the class to those who do not. The researcher's normal prior about the ATE is centered at 30 points with a standard deviation of 15 points. The researcher's normal prior about the bias of the design is 15 points with a standard deviation of 10 points. The study's results suggest that the course increases performance by 65 points with a standard deviation of 5 points. In other words, $g = 30$, $\sigma_g^2 = 225$, $\beta = 15$, $\sigma_\beta^2 = 100$, $x_e = 65$, and $\sigma_{x_e}^2 = 25$. Plugging these numbers into equation (11.3) gives:

$$\sigma_{\bar{\tau}|x_e}^2 = \frac{1}{\frac{1}{\sigma_g^2} + \frac{1}{\sigma_\beta^2 + \sigma_{x_e}^2}} = \frac{1}{\frac{1}{225} + \frac{1}{100+25}} = 80.36$$

Plugging these numbers into equation (11.4) gives:

$$p_1 = \frac{\sigma_{\bar{\tau}|x_e}^2}{\sigma_g^2} = \frac{\sigma_\beta^2 + \sigma_{x_e}^2}{\sigma_g^2 + \sigma_\beta^2 + \sigma_{x_e}^2} = \frac{100 + 25}{225 + 100 + 25} = 0.357p_2 = \frac{\sigma_{\bar{\tau}|x_e}^2}{\sigma_\beta^2 + \sigma_{x_e}^2} = \frac{\sigma_g^2}{\sigma_g^2 + \sigma_\beta^2 + \sigma_{x_e}^2} = 1 - p_1 = 0.643$$

Finally, plugging these numbers into equation (11.2) gives the posterior estimate:

$$\begin{aligned} E[\bar{\tau}|X_e = x_e] &= p_1 * g + p_2(x_e - \beta) \\ &= 0.357 * 30 + 0.643 * (65 - 15) = 42.87 \end{aligned}$$

In the absence of uncertainty about bias (i.e., if $\sigma_\beta^2 = 0$), the weight given to the new evidence (p_2) would have been much greater: $\frac{\sigma_g^2}{\sigma_g^2 + \sigma_\beta^2 + \sigma_{x_e}^2} = \frac{225}{225+0+25} = 0.9$. The posterior would have more strongly shaped by the observational results:

$$\begin{aligned} E[\bar{\tau}|X_e = x_e] &= p_1 * g + p_2(x_e - \beta) \\ &= 0.1 * 30 + 0.9 * (65 - 15) = 48 \end{aligned}$$

d) What does it mean to conduct a hypothesis test that compares two “nested” models?

Answer:

Models are said to be “nested” when one model can be written as a special case of another model. For example, if one conducts an experiment with three groups, a control group and two treatments, one could estimate the ATE of each treatment, or one could estimate a nested model in which both treatments are assumed to have the same ATE. When expressed in regression form (with indicator variables for each treatment), the first model is

$$Y_i = \beta_0 + \beta_1 D_{1i} + \beta_2 D_{2i} + \epsilon_i$$

and the second model is:

$$Y_i = \beta_0 + \beta_1 (D_{1i} + D_{2i}) + \epsilon_i$$

Question 2

Identify a research literature where an experiment was replicated at least once. Carefully consider the manner in which subjects and contexts were selected, as well as the methods used to administer treatments and measure outcomes. If publication bias is a concern, note that, too. Based on your close reading of these studies, assess whether the estimated ATE from each study can be said to constitute an independent random sample from some larger population.

Answer:

Answers will vary widely.

Question 3

Suppose one were to sample N subjects at random from a population of N^* people. An experiment is performed whereby m of the N subjects are assigned to receive a treatment, and the remaining $N-m$ are assigned to the control group. Suppose that sometime after the treatment is administered, outcomes are measured for all N^* people.

- a) Suppose one estimates the population ATE by comparing the mean outcome among the m subjects in the treatment group to the mean outcome among the $N^* - m$ subjects who were not assigned to the treatment. Is this estimator unbiased?

Answer:

Yes. The subjects assigned to the treatment and control groups are each random samples from the pool of N^* subjects in the population. Therefore, they have the same expected potential outcomes.

- b) Would the appropriate standard error of this difference-in-means estimator be equation (11.1), equation (3.4), or neither?

Answer:

The correct formula is a modified version of equation (3.4) in which N^* replaces N .

Question 4

Suppose that $N = 2m$. Using equations (3.4) and (11.1), show that $SE(\widehat{SATE}) \approx SE(PATE)$ when the treatment effect is constant across subjects. Hint: In this case, $Var(Y_i(0)) = Var(Y_i(1)) = Cov(Y_i(0), Y_i(1))$.

Answer:

Substituting for $Var(Y_i(1))$ and for $Cov(Y_i(0), Y_i(1))$ and imposing the constraint that $N = 2m$ allows us to re-write equation (3.4) as:

$$\begin{aligned} SE(\widehat{SATE}) &= \sqrt{\frac{1}{N-1} (Var(Y_i(0)) + Var(Y_i(1)) + 2 * Cov(Y_i(0), Y_i(1)))} \\ &= 2\sqrt{\frac{Var(Y_i(0))}{N-1}} \end{aligned}$$

Similarly, equation (11.1) may be re-written:

$$\begin{aligned} SE(\widehat{PATE}) &= \sqrt{\frac{Var(Y_i(0))}{m} + \frac{Var(Y_i(0))}{m}} \\ &= \sqrt{\frac{2 * Var(Y_i(0))}{N} + \frac{2 * Var(Y_i(0))}{N}} \\ &= 2\sqrt{\frac{Var(Y_i(0))}{N}} \end{aligned}$$

The two formulas converge as N increases.

Question 5

Using the Bayesian updating equations, show algebraically how the priors represented in Figure 11.1 combine with the experimental results depicted in order to form a posterior distribution with a mean of 8 and a standard deviation of 0.89.

Answer:

In this example, $g = 0$, $\sigma_g^2 = 4$, $\beta = 0$, $\sigma_\beta^2 = 0$, $x_e = 10$, and $\sigma_{x_e}^2 = 1$. Plugging these numbers into equation (11.3) gives:

$$\sigma_{\bar{\tau}|x_e}^2 = \frac{1}{\frac{1}{\sigma_g^2} + \frac{1}{\sigma_\beta^2 + \sigma_{x_e}^2}} = \frac{1}{\frac{1}{4} + \frac{1}{0+1}} = 0.8$$

Plugging these numbers into equation (11.4) gives:

$$\begin{aligned} p_1 &= \frac{\sigma_{\bar{\tau}|x_e}^2}{\sigma_g^2} = \frac{\sigma_\beta^2 + \sigma_{x_e}^2}{\sigma_g^2 + \sigma_\beta^2 + \sigma_{x_e}^2} = \frac{0 + 1}{4 + 0 + 1} = 0.2 \\ p_2 &= \frac{\sigma_{\bar{\tau}|x_e}^2}{\sigma_\beta^2 + \sigma_{x_e}^2} = \frac{\sigma_g^2}{\sigma_g^2 + \sigma_\beta^2 + \sigma_{x_e}^2} = 1 - p_1 = 0.8 \end{aligned}$$

Finally, plugging these numbers into equation (11.2) gives the posterior estimate:

$$\begin{aligned} E[\bar{\tau}|X_e = x_e] &= p_1 * g + p_2(x_e - \beta) \\ &= 0.2 * 0 + 0.8 * (10) = 8 \end{aligned}$$

Thus, Figure 11.1 depicts the posterior as centered at 8 with a standard error of $\sqrt{0.8} = 0.89$.

Question 6

Prior to a 2006 primary election, Gerber, Green, and Larimer sent a sample of registered voters in Michigan an encouragement to vote that disclosed whether each person registered to vote at that address had voted in previous elections.¹ (The self mailing is described in Chapter 10.) A similar mailing was sent to Michigan voters in 2007, prior to municipal elections in small cities and towns,² and to Illinois voters in 2009, prior to a special congressional election.³ For comparability, we restrict each of the samples to the set of households containing just one registered voter.

¹Gerber, Green, and Larimer 2008.

²Gerber, Green, and Larimer 2010.

³Sinclair, McConnell, and Green 2012.

Table 1: Question 6 Table

Study	Number in the control group	Number voting in the control group	Number in the treatment group	Number voting in the treatment group
2006 Michigan	26481	8755	5310	2123
2007 Michigan	348277	88960	12391	3791
2009 Illinois	15676	2600	9326	1936

a) Estimate the ATE for each study.

Answer:

```
In [1]: clear
        scalar p_c_MI06 = 8755/26481
        scalar p_t_MI06 = 2123/5310
        global ATE_MI06 = p_t_MI06 - p_c_MI06
        disp $ATE_MI06
```

.06919727

```
In [2]: scalar p_c_MI07 = 88960/348227
        scalar p_t_MI07 = 3791/12391
        global ATE_MI07 = p_t_MI07 - p_c_MI07
        disp $ATE_MI07
```

.05048232

```
In [3]: scalar p_c_IL09 = 2600/15676
        scalar p_t_IL09 = 1936/9326
        global ATE_IL09 = p_t_IL09 - p_c_IL09
        disp $ATE_IL09
```

.04173304

b) Estimate the standard error for each study. Use the standard errors (squared) to calculate the precision of each study.

Answer:

```
In [4]: global SE_MI06 = sqrt((p_t_MI06 * (1-p_t_MI06))/5310 + ///
        (p_c_MI06 * (1-p_c_MI06))/26481)
        disp $SE_MI06
```

```
.00731764
```

```
In [5]: global SE_MI07 = sqrt((p_t_MI07 * (1-p_t_MI07))/12391 + ///  
      (p_c_MI07 * (1-p_c_MI07))/348227)  
      disp $SE_MI07
```

```
.00420513
```

```
In [6]: global SE_IL09 = sqrt((p_t_IL09 * (1-p_t_IL09))/9326 + ///  
      (p_c_IL09 * (1-p_c_IL09))/15676)  
      disp $SE_IL09
```

```
.00514433
```

```
In [7]: global prec_MI06 = 1/$SE_MI06^2  
      global prec_MI07 = 1/$SE_MI07^2  
      global prec_IL09 = 1/$SE_IL09^2
```

```
In [8]: disp $prec_MI06  
      disp $prec_MI07  
      disp $prec_IL09
```

```
18674.87
```

```
56551.044
```

```
37786.98
```



- c) Assuming that these three samples are random draws from the same population, calculate a precision-weighted average of the three studies. (Hint: weight each estimate by the inverse of its squared standard error.)

Answer:

```
In [9]: scalar weighted_mean = ($ATE_MI06*$prec_MI06 + ///  
      $ATE_MI07*$prec_MI07 + ///  
      $ATE_IL09*$prec_IL09)/($prec_MI06+$prec_MI07+$prec_IL09)  
      disp weighted_mean
```

```
.05064947
```

- d) Show that this estimate is identical to what one obtains by using the Bayesian updating formula recursively: use the results from the 2006 study as your priors, and update them based on the 2007 study to form a posterior mean and posterior variance based on equations (11.2) and (11.3); then update this posterior using the results from the 2009 study.

Answer:

Using the MI2006 study as the prior and updating based on MI2007 implies the following parameters:

$$g = 0.069, \sigma_g^2 = 1/18675, \beta = 0, \sigma_\beta^2 = 0, x_e = 0.042, \text{ and } \sigma_{x_e}^2 = 1/56550$$

Plugging these numbers into equation (11.3) gives:

$$\sigma_{\bar{\tau}|x_e}^2 = \frac{1}{\frac{1}{\sigma_g^2} + \frac{1}{\sigma_\beta^2 + \sigma_{x_e}^2}} = \frac{1}{18675 + 56550} + \frac{1}{75225} = 0.000032934$$

Plugging these numbers into equation (11.4) gives:

$$p_1 = \frac{\sigma_{\bar{\tau}|x_e}^2}{\sigma_g^2} = \frac{\sigma_\beta^2 + \sigma_{x_e}^2}{\sigma_g^2 + \sigma_\beta^2 + \sigma_{x_e}^2} = \frac{0 + 1/56550}{1/18675 + 0 + 1/56550} = 0.248$$

$$p_2 = \frac{\sigma_{\bar{\tau}|x_e}^2}{\sigma_\beta^2 + \sigma_{x_e}^2} = \frac{\sigma_g^2}{\sigma_g^2 + \sigma_\beta^2 + \sigma_{x_e}^2} = 1 - p_1 = 0.752$$

Finally, plugging these numbers into equation (11.2) gives the posterior estimate based on the first two studies:

$$E[\bar{\tau}|X_e = x_e] = p_1 * g + p_2(x_e - \beta)$$

$$= 0.248 * 0.069 + 0.752 * (0.051 - 0) = 0.055$$

The next step uses this estimate as the prior and updates based on IL2009. That exercise uses the following parameters:

$$g = 0.055, \sigma_g^2 = 1/75225, \beta = 0, \sigma_\beta^2 = 0, x_e = 0.042, \text{ and } \sigma_{x_e}^2 = 1/37787$$

Plugging these numbers into equation (11.3) gives:

$$\sigma_{\bar{\tau}|x_e}^2 = \frac{1}{\frac{1}{\sigma_g^2} + \frac{1}{\sigma_\beta^2 + \sigma_{x_e}^2}} = \frac{1}{75225 + 37787} = 0.00000884859$$

Plugging these numbers into equation (11.4) gives:

$$p_1 = \frac{\sigma_{\bar{\tau}|x_e}^2}{\sigma_g^2} = \frac{\sigma_\beta^2 + \sigma_{x_e}^2}{\sigma_g^2 + \sigma_\beta^2 + \sigma_{x_e}^2} = \frac{0 + 1/37787}{1/75225 + 0 + 1/37787} = 0.666$$

$$p_2 = \frac{\sigma_{\bar{\tau}|x_e}^2}{\sigma_\beta^2 + \sigma_{x_e}^2} = \frac{\sigma_g^2}{\sigma_g^2 + \sigma_\beta^2 + \sigma_{x_e}^2} = 1 - p_1 = 0.334$$

Finally, plugging these numbers into equation (11.2) gives the posterior estimate based on all three studies:

$$\begin{aligned} E[\bar{\tau}|X_e = x_e] &= p_1 * g + p_2(x_e - \beta) \\ &= 0.666 * 0.055 + 0.334 * (0.042 - 0) = 0.0507 \end{aligned}$$

This estimate matches what we obtained earlier based on a precision-weighted average.

- e) Use equation (11.3) to estimate the variance of the precision-weighted average. Take the square root of the variance in order to obtain the standard error. In order to estimate the 95% confidence interval, use the following procedure, which is based on a large-sample approximation. Obtain the lower bound of the interval by subtracting 1.96 times the standard error from the precision-weighted average; obtain the upper bound of the interval by adding 1.96 times the standard error to the precision-weighted average.

Answer:

$$\begin{aligned} \sqrt{\sigma_{\bar{\tau}}^2|x_e} &= \sqrt{\frac{1}{\sigma_g^2} + \frac{1}{\sigma_{\beta}^2 + \sigma_{x_e}^2}} \\ &= \sqrt{\frac{1}{75225 + 37787}} \\ &= 0.00297 \end{aligned}$$

The Lower bound is $0.051 - (1.96)(.00297) = 0.0448$. The upper bound is $0.051 + (1.96)(.00297) = 0.0565$.

These results are confirmed by the `rmeta` package:

```
In [10]: clear
         set obs 3
         gen ate = .
         replace ate = $ATE_MI06 if _n ==1
         replace ate = $ATE_MI07 if _n ==2
         replace ate = $ATE_IL09 if _n ==3

         gen ate_se = .
         replace ate_se = $SE_MI06 if _n ==1
         replace ate_se = $SE_MI07 if _n ==2
         replace ate_se = $SE_IL09 if _n ==3

In [11]: qui metaan ate ate_se, fe

In [12]: // inverse-variance weighted effect
         disp r(eff)

.05064947
```



```
In [13]: // inverse-variance weighted se
         disp sqrt(r(effvar))

.00297465
```

- f) Explain why the confidence interval formed in part (e) is likely to understate the true amount of uncertainty associated with the estimate of the population ATE.

Answer:

The formulas used to calculate the estimates and standard errors assume that the studies represent independent random samples from the same population. If the studies are biased samples from the population of citizens and campaign contexts (perhaps because they all involve North-erners and low-salience political contests) or if the studies are non-independent (because they were all conducted or analyzed by some of the same authors), the true degree of uncertainty may be understated.

Question 7

According to the logistic regression coefficients reported in Table 11.2, the intercept in Region 1 is 8.531 and the slope is -1.978. Based on these numbers, what proportion of those offered a price of 100 shillings is expected to buy a bed net? How does this compare to the actual rate of purchases at this price?

Answer:

The logistic model for Region 1 is

$$\begin{aligned} Pr[Y_i = 1] &= \frac{1}{1 + e^{-(\beta_0 + \beta_1 \ln(D-i))}} \\ &= \frac{1}{1 + e^{-(\beta_0 + \beta_1 \ln[100])}} \\ &= 0.359 \end{aligned}$$

The corresponding empirical value from this region is 0.340.

Question 8

Table 11.3 presents observed and predicted values for each of the experimental sites in the Kenya bed nets study. Data for this experiment may be found at <http://isps.research.yale.edu/FEDAI>.

- a) Verify that mispredictions mostly occur in Region 4 by calculating this region's contribution to the total chi-square statistics for each model.

Answer:

```
In [1]: import delimited ./data/chapter11/GerberGreenBook_Chapter11_Table_11_3, clear

In [2]: tabstat m1_x2 m2_x2, statistics(sum) by(region)
```

Summary statistics: sum
by categories of: region (Region)

region	m1_x2	m2_x2
-----+-----		
1	.67	1.14
2	2.17	2.53
3	.45	.5
4	7.1	13.36
5	2.14	2.53
6	1.91	2.05
-----+-----		
Total	14.44	22.11

For Model I, Region 4's mispredictions contribute 7.10 to the overall chi-square of 14.44; For Model II, Region 4's mispredictions contribute 13.36 to the overall chi-square of 22.10. In other words, the mispredictions in Region 4 are greater (using the chi-square metric) than the other regions' combined.

- b) Re-estimate the two logistic regression models presented in Table 11.3, this time excluding Region 4. Reproduce Table 11.3.
Answer:

In [3]: `clear`

```
import delim ./data/chapter11/Chapter_11_Dupas_2010_Dataset, clear
//import delim Dupas_QJE_2010, clear

gen log_price = log(price)

gen purchased =.
replace purchased = 1 if purchasednet=="yes"
replace purchased = 0 if purchasednet=="no"

rename cfw_id region

drop if (price ==0 | region==4)

qui glm purchased log_price i.region c.log_price#i.region, ///
family(binomial) link(logit)
predict double preds_1
```

```

qui glm purchased log_price i.region, family(binomial) link(logit)
predict double preds_2

```

```

generate non_p =.
replace non_p=1 if purchased==0
replace non_p=0 if purchased!=0

generate pred_nonp_1 = 1-preds_1
generate pred_nonp_2 = 1-preds_2

```

```

In [4]: collapse (sum) purchases=purchased ///
        non_purchases=non_p ///
        pred_purchases_1=preds_1 ///
        pred_nonpurchases_1=pred_nonp_1 ///
        pred_purchases_2=preds_2 ///
        pred_nonpurchases_2=pred_nonp_2, by(region price)

```

```

In [5]: gen chi_square_1 = (purchases- pred_purchases_1)^2/ pred_purchases_1 + ///
        (non_purchases- pred_nonpurchases_1)^2/ pred_nonpurchases_1

gen chi_square_2 = (purchases- pred_purchases_2)^2/ pred_purchases_2 + ///
( non_purchases- pred_nonpurchases_2)^2/ pred_nonpurchases_2

```

```

In [6]: // print table
list region price purchases non_purchases pred_purchases_1 ///
pred_nonpurchases_1 chi_square_1
list region price purchases non_purchases pred_purchases_2 ///
pred_nonpurchases_2 chi_square_2

```

	region	price	purcha~s	non_pu~s	pred_pu~1	pred_n~1	chi_sq~1
1.	1	70	16	13	15.441598	13.5584	.0431908
2.	1	100	16	31	16.920757	30.07924	.078289
3.	1	130	12	37	12.291133	36.70887	.0092048
4.	1	190	5	23	3.8221953	24.1778	.420315
5.	1	250	2	28	2.5243177	27.47568	.1189098
6.	2	40	46	15	47.582604	13.4174	.2393082
7.	2	80	40	30	35.429156	34.57085	1.194044
8.	2	120	18	46	21.212681	42.78732	.7277873
9.	2	200	10	49	9.7755597	49.22444	.0061763
10.	3	50	42	16	43.139894	14.86011	.1175591
11.	3	90	33	27	31.052283	28.94772	.2532181
12.	3	150	18	40	18.042949	39.95705	.0001484
13.	3	210	9	39	9.7648732	38.23513	.0752127
14.	5	60	27	10	26.083854	10.91615	.1090662
15.	5	110	12	25	15.218787	21.78121	1.156442

16.		5	140	11	18	8.6973594	20.30264	.8707839	
17.		6	50	14	5	14.837853	4.162147	.2159737	
18.		6	100	11	7	8.7298279	9.270172	1.146296	
19.		6	150	4	14	5.4323191	12.56768	.5408933	
+-----+									
+-----+									
		region	price	purcha~s	non_pu~s	pred_pu~2	pred_n~2	chi_sq~2	
+-----+									
1.		1	70	16	13	14.85299	14.14701	.181574	
2.		1	100	16	31	16.687555	30.31244	.0439238	
3.		1	130	12	37	12.498146	36.50185	.0266531	
4.		1	190	5	23	4.1147946	23.88521	.2232385	
5.		1	250	2	28	2.8465156	27.15348	.2781327	
+-----+									
6.		2	40	46	15	47.744037	13.25596	.293164	
7.		2	80	40	30	35.468777	34.53122	1.173467	
8.		2	120	18	46	21.135136	42.86486	.6943625	
9.		2	200	10	49	9.652051	49.34795	.0149967	
10.		3	50	42	16	44.054578	13.94542	.3985204	
+-----+									
11.		3	90	33	27	31.294685	28.70532	.1942352	
12.		3	150	18	40	17.510472	40.48953	.0196039	
13.		3	210	9	39	9.1402666	38.85973	.0026588	
14.		5	60	27	10	25.287202	11.7128	.366482	
15.		5	110	12	25	15.49734	21.50266	1.358088	
+-----+									
16.		5	140	11	18	9.215458	19.78454	.506534	
17.		6	50	14	5	14.607437	4.392563	.1092606	
18.		6	100	11	7	8.7615558	9.238444	1.114256	
19.		6	150	4	14	5.6310073	12.36899	.687486	
+-----+									

```
In [7]: qui tabstat chi_square_1, statistics(sum) save
        scalar model_1_chi_sq=el(r(StatTotal),1,1)
```

```
In [8]: scalar pvalue_1=chiprob(9, model_1_chi_sq)
        disp %8.5f model_1_chi_sq
        disp %8.6f pvalue_1
```

7.32282

0.603548

```
In [9]: qui tabstat chi_square_2, statistics(sum) save
        scalar model_2_chi_sq=el(r(StatTotal),1,1)
        scalar pvalue_2=chiprob(13, model_2_chi_sq)
        disp %8.5f model_2_chi_sq
        disp %8.6f pvalue_2
```

7.68664

0.863399

- c) Although excluding Region 4 from the analysis improves the apparent fit of both models, why might it be considered a questionable practice to exclude Region 4?

Answer:

The mispredictions in Region 4 were not expected before the results became known; the adjustments to the model are therefore post hoc. Since the overall degree of fit is not worse than would be expected by chance when Region 4 is included, it may be that the anomalies in that region are due to random sampling variability.

Question 9

Because the log transformation of price is undefined when price is zero, we excluded the zero price condition from the analysis of bed net purchases in Tables 11.2 and 11.3.

- a) If we exclude zero prices from our experimental analysis, will our estimate of the causal effect of price be biased?

Answer:

No. The exclusion is based on the treatment not on the results. Because those receiving the zero price are a random subset of all subjects, excluding these observations does not lead to biased estimates of the ATE. Thinking back to Chapter 7, missingness here is unrelated to potential outcomes.

- b) Suppose we reasoned that a nominal price of zero nevertheless involves some transaction cost, as villagers have to make the effort to redeem their vouchers. For a given subject, we may model the probability of making a purchase as:

$$Pr[Y = 1] = \frac{1}{1 + e^{-(\beta_0 + \beta_1 \ln[V_i + \gamma])}}$$

where γ represents the transaction cost of redeeming the voucher. In order to estimate γ , insert a positive value of γ , and use logistic regression to estimate the revised model; note the value of the log-likelihood for this model. Repeat this exercise for different values of γ . Obtain the “maximum likelihood estimate” of γ by finding the value of γ that maximizes the log-likelihood.

Answer:

We tried different values of γ until we came upon the value 19, which maximized the log-likelihood:

```
In [1]: import delim ./data/chapter11/Chapter_11_Dupas_2010_Dataset,clear

In [2]: gen purchased =.
         replace purchased = 1 if purchasednet=="yes"
         replace purchased = 0 if purchasednet=="no"
         rename cfw_id region
```

```

matrix t=J(100, 2, .)
matrix colnames t=gammas lls

forvalues i = 1/100 {

    gen log_price_star_`i' = log(price + `i')
    qui glm purchased log_price_star_`i' i.region,
        family(binomial) link(logit)

    matrix t[`i', 1] = `i'
    matrix t[`i', 2] = `e(ll)'

}

```

```

In [3]: svmat double t, names(col)
        qui sum lls
        list gammas if lls==r(max)

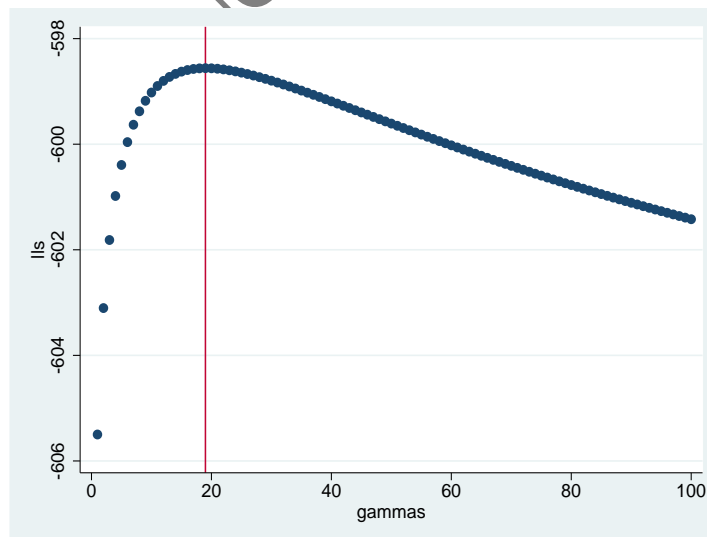
        scatter lls gammas, xline(19)
        graph export ../results/chapter11/exercise_11_9_graph.pdf

```

```

+-----+
| gammas |
|-----|
19. |      19 |
+-----+

```



- c) What is the substantive interpretation of the maximum likelihood estimates of γ and β_1 ? (Note that the standard errors using this method understate the true sampling variability because they are conditional on a particular choice of γ . Ignore the reported standard errors, and just interpret the estimates.)

Answer:

The maximum likelihood estimate of γ is 19 shillings, which suggests that redeeming the voucher involves some transaction cost even when the nominal price is very small. The coefficient on the treatment variable, $\ln(\text{price} + 19)$, is -1.98, which suggests that for every one unit change in this rescaled version of the treatment, the log-odds of purchase declines by -1.98. For example, suppose that the offer price rises from 0 to 100 shillings. The $\ln(\text{price} + 19)$ would change from 2.94 to 4.78, and this change would reduce the log-odds of purchase by 3.63. To illustrate what this means in terms of percentage points, suppose that a person has a 50% chance of making a purchase at a price of zero (50% implies a log-odds of 0). If that person were to be offered a price of 100, the predicted probability of a purchase is $1/(1 + e^{(-3.63)}) = 0.026$, or 2.6%.

Question 10

This chapter discussed the modeling issues that arise when randomly assigned treatments vary in intensity or duration. The table below considers a different case, where subjects are randomly assigned treatments but then choose to take different dosages. In this experiment, subjects were randomly assigned to receive one of two pre-recorded phone messages.⁴ The treatment script encouraged people to vote and revealed whether members of the household voted in the past two elections. The control script encouraged people to recycle. Both calls were made the day before the election. Both calls were answered at similar rates. The table below presents voting rates for households that answered the phone call. Voting rates are broken down by how long the person answering the phone took to hang up after the recorded message began.

Table 2: Question 10 Table

Treatment group	Duration of call before respondent disconnected				Total
	1-10 seconds	11-20 seconds	21-30 seconds	31-40 seconds	
Call encouraged voting	16.6 (187)	17.4 (784)	19.7 (983)	24.3 (2,032)	21.4 (3,986)
Call encouraged recycling	17.5 (143)	18.3 (619)	18.9 (1,132)	19.8 (2,012)	19.2 (3,906)

Entries are percent voting, with Ns in parentheses. The sample is restricted to one-voter households. Both scripts were approximately 35 seconds long.

- a) Focusing only on households that answered the phone, estimate the apparent average effect of assignment to the script that encouraged voting?

Answer:

The estimated ATE is $21.4 - 19.2 = 2.2$ percentage points.

- b) Does this table provide convincing evidence that “the longer a person listens to a recorded message that encourages voting, the more effective that message will be in terms of boosting voter turnout?” Why or why not?

Answer:

No. The length of time one listens before hanging up is not randomly assigned, and the people who listen for a given length of time to one script are not necessarily comparable in terms of potential outcomes to those who listen the same length of time to the other script. For example, it may be that people who are very interested in politics (and very likely to vote)

⁴Gerber et al. 2010.

listen to the entire voting script but do not listen to the entire recycling script. We cannot infer anything about the effect of listening duration from these results unless we impose some strong assumptions that do not follow from the experimental design.

DO NOT DISTRIBUTE