How Economists Analyze Data: An Application to Demand for Cigarettes 1960-2008

Ted Juhl [1]

[1] Department of Economics
University of Kansas

May 27 KU Mini College

Outline

Introduction

Introduce Yourself:

- State your name.
- What did you of study in college?
- Is your job in same area? (NO)

Introduction

Without education, we are in horrible and deadly danger of taking educated people seriously.

G.K. Chesterton

Introduction

College teaches you how to figure out stuff.

Mark Juhl

Examples of Empirical Studies in the News

Diet Soda Causes Weight Gain!

- So I should just drink the real stuff?
- Theory: fake sweeteners cause more eating.
- My theory: you start drinking diet soda because you gained weight!

Examples of Empirical Studies in the News

Elderly People Who Walk Slowly Will Die of Heart Disease!

- Researchers followed elderly subjects and recorded pace.
- They found that slower walkers had more heart problems.
- Researchers conclude that walking slower is bad for heart.
- My theory: if you have heart problems, you walk slowly!

Economists and Panel Data

- Suppose we want to explore the effect of education on wages.
- We could observe many people and see what happens.
- Problem: It may be that we confuse “gifted" people with educated people.
- We want to see what happens when you “educate" a random person.

Economists and Panel Data

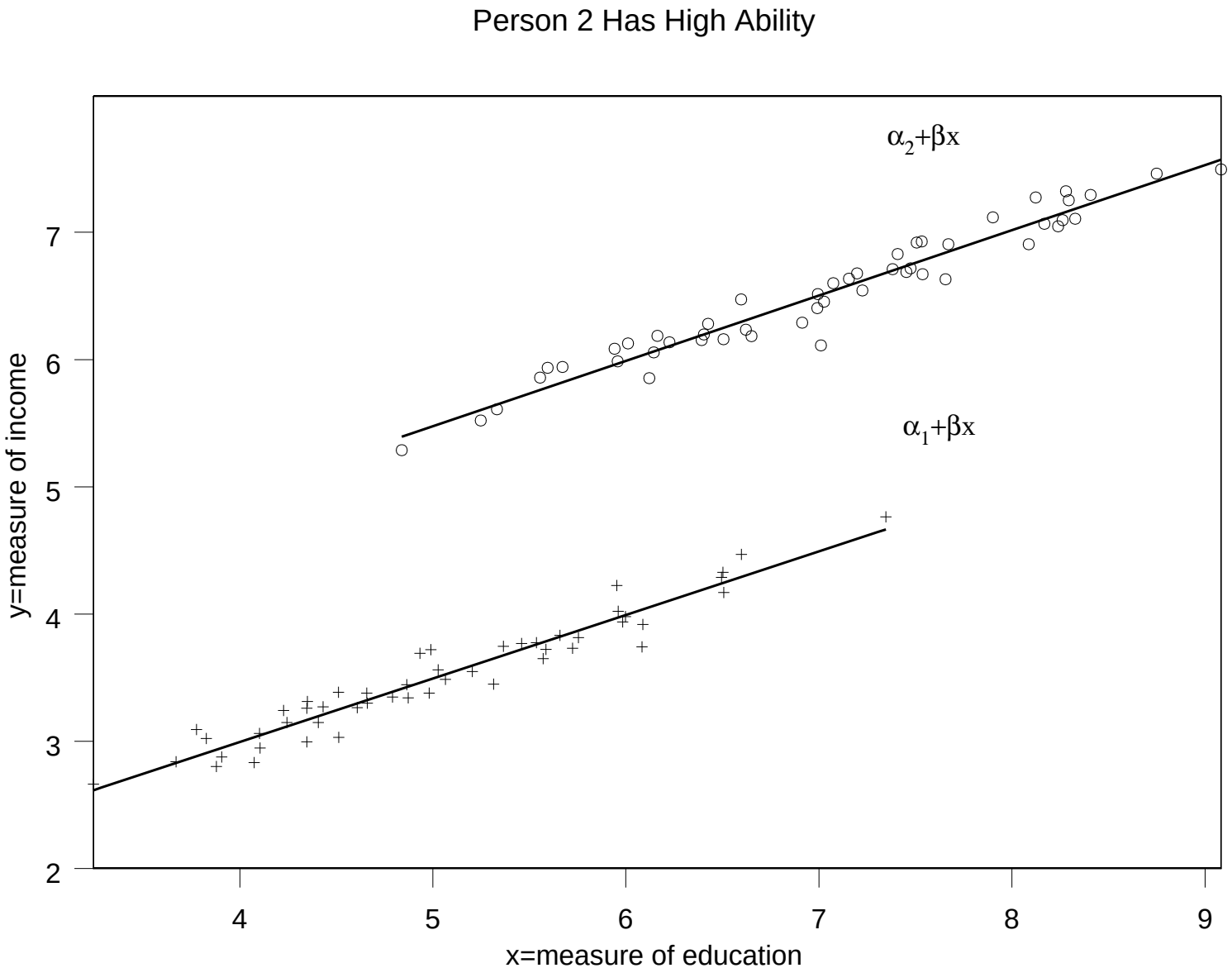
- A panel data set observes many individuals.
- Each individual is followed over time.
- An individual could also be a firm, country, state, etc.
- This is the way much of current data is collected.

Economists and Panel Data

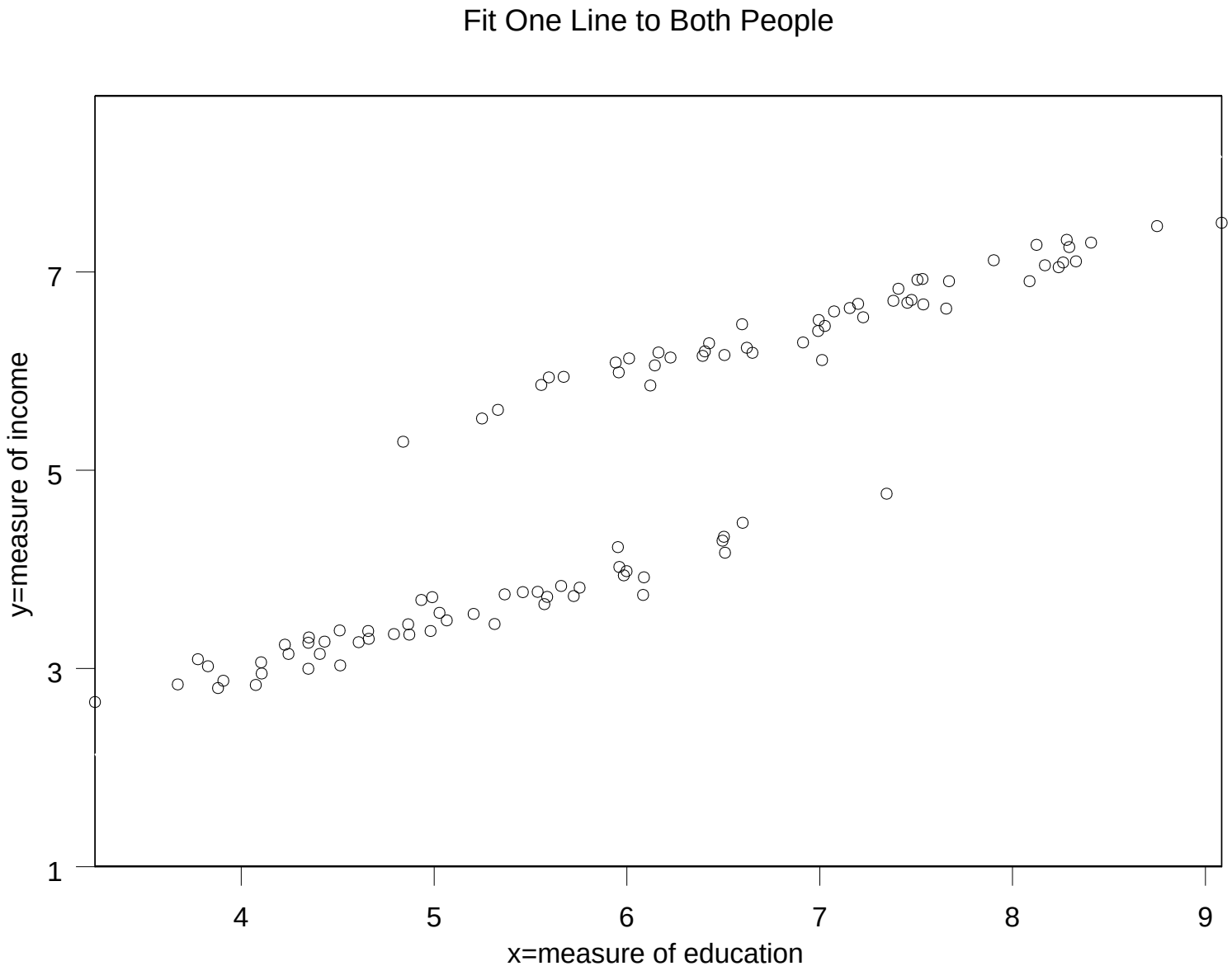
Recent Studies Using Panels

- Do CEO's get a raise for firing workers? (No)
- Does a common currency increase trade between countries? (Yes, a lot)
- Is there evidence for discrimination based on race and gender? (Yes, no, and sometimes)

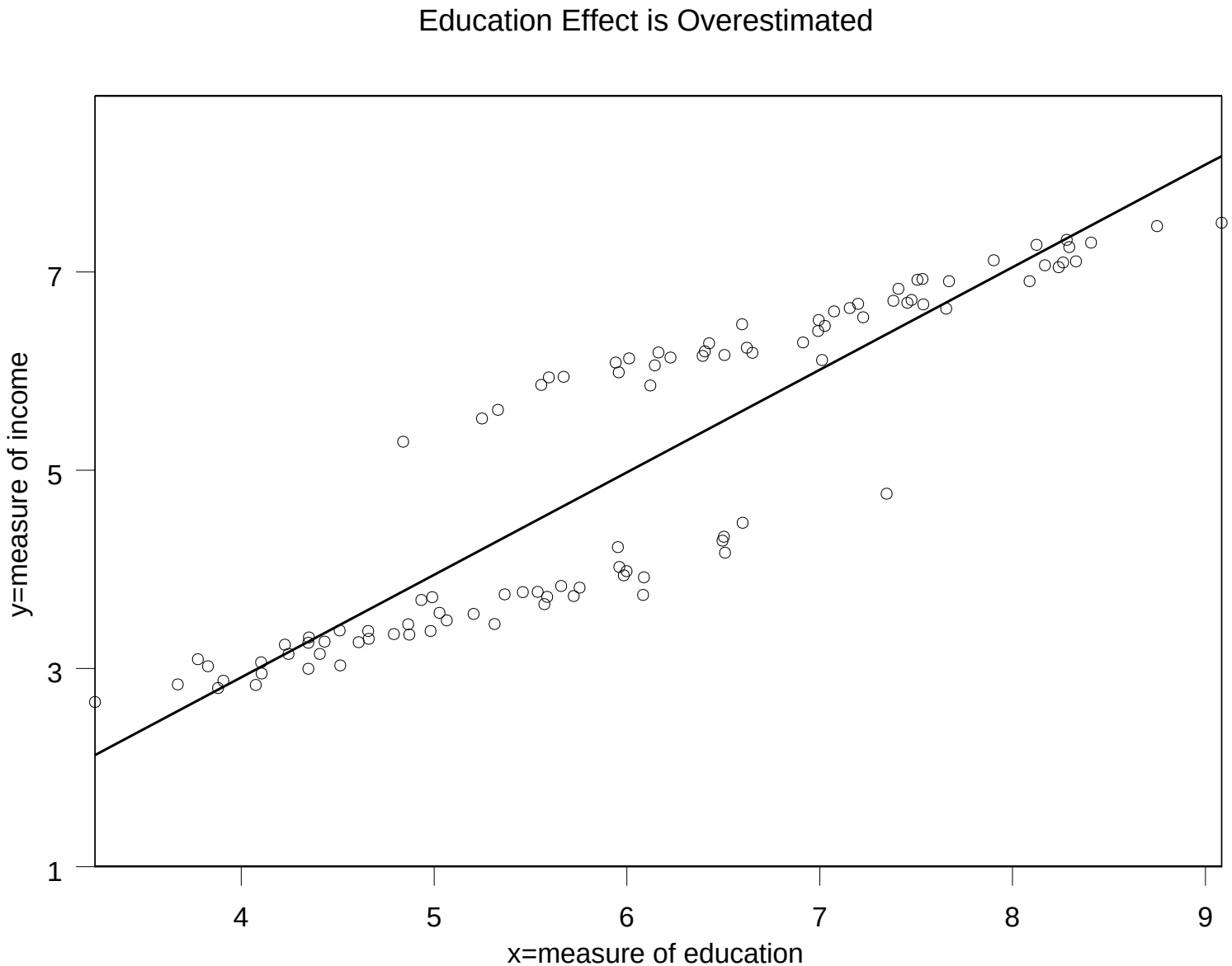
Panel Data and Education



Panel Data and Education



Panel Data and Education

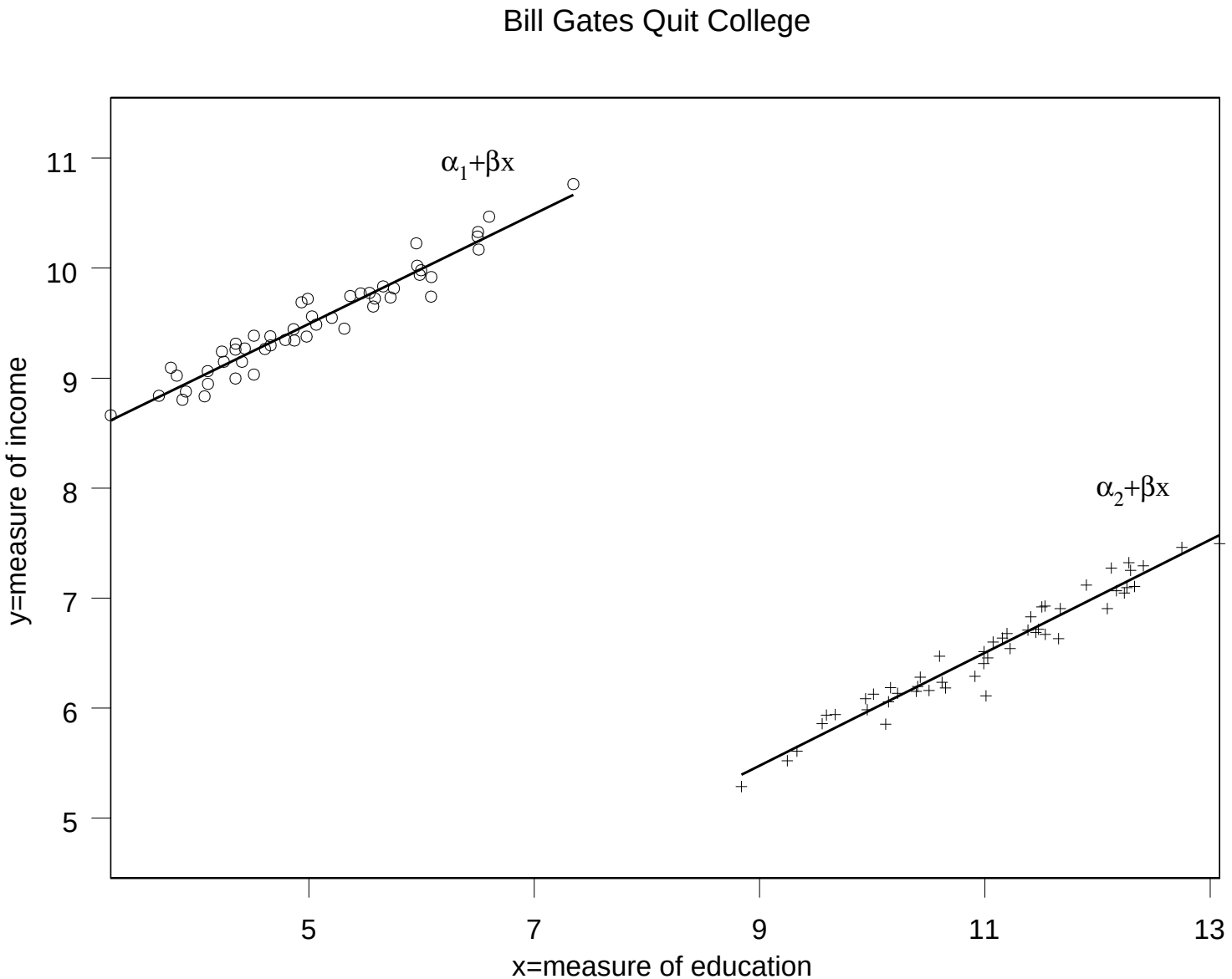


Panel Data and Education

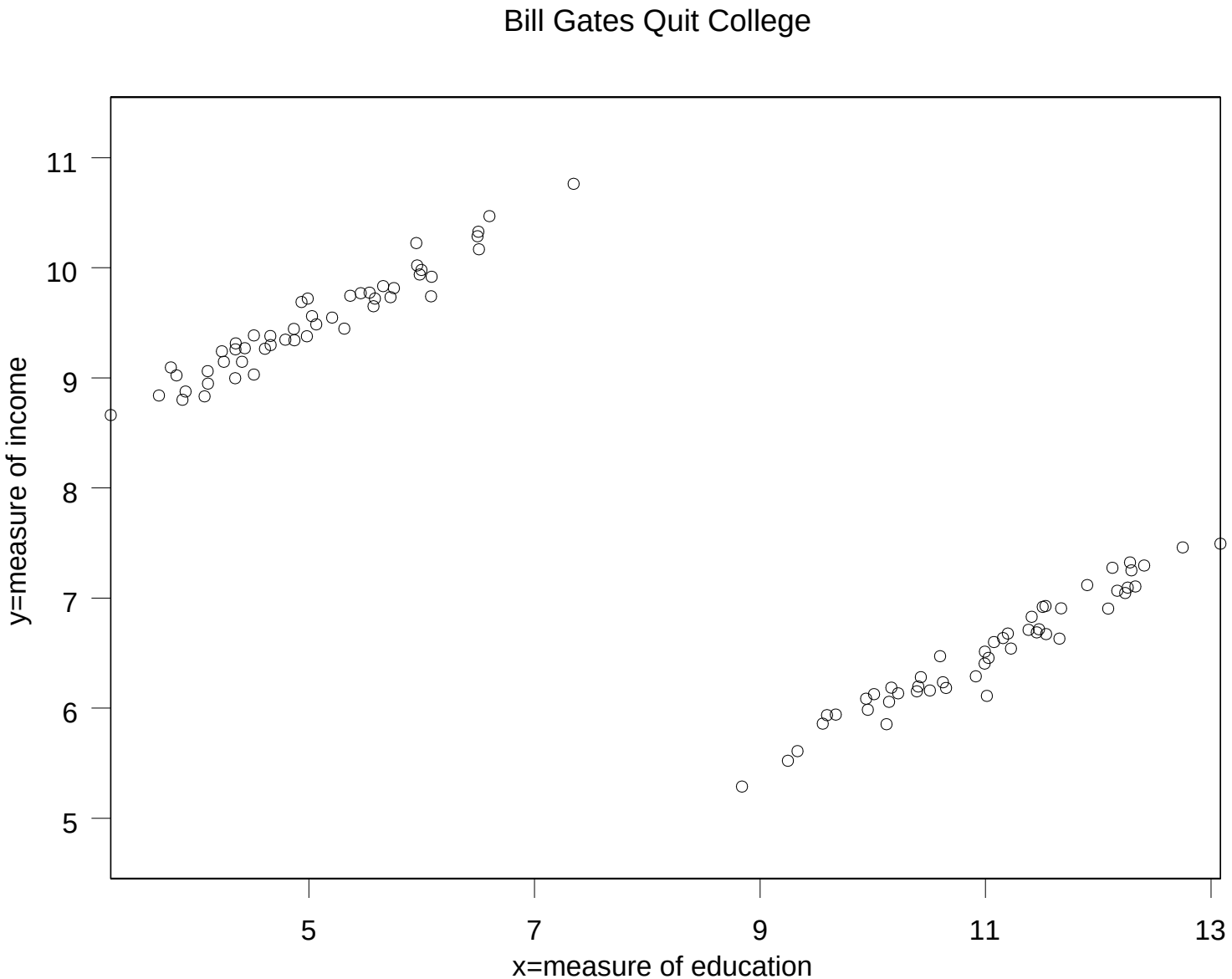
If we don't take individuals into account

- we can get poor estimates of effects of education (too large in this example)
- but ... it could be worse

Panel Data and Education

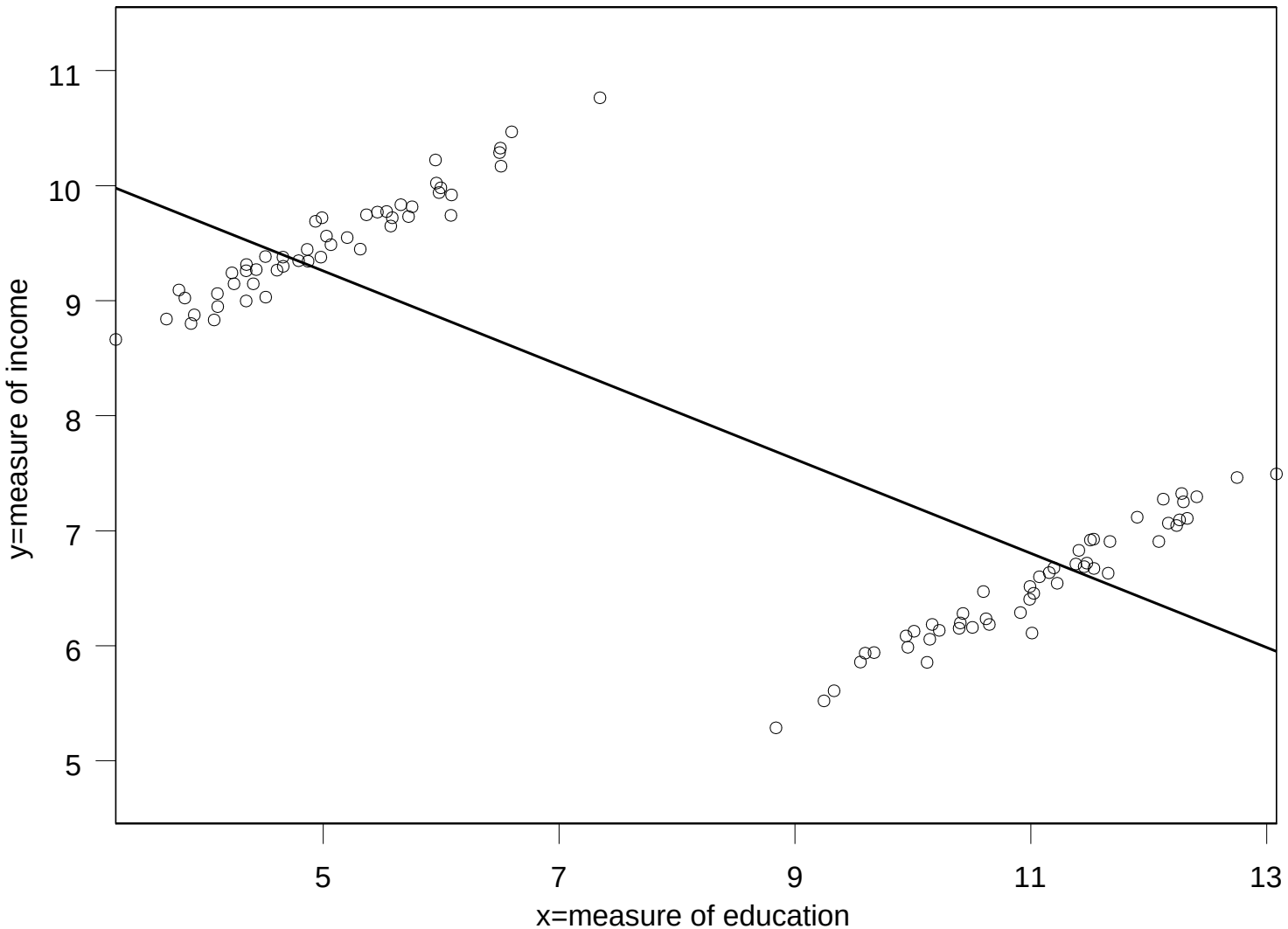


Panel Data and Education



Panel Data and Education

Conclusion: I Should've Quit College!



Data Description for Cigarettes

I'll quit smoking when cigarettes cost $\times$ *per pack.*

Billboard on 435 near KCI

Things We Can See to Measure Cigarette Consumption

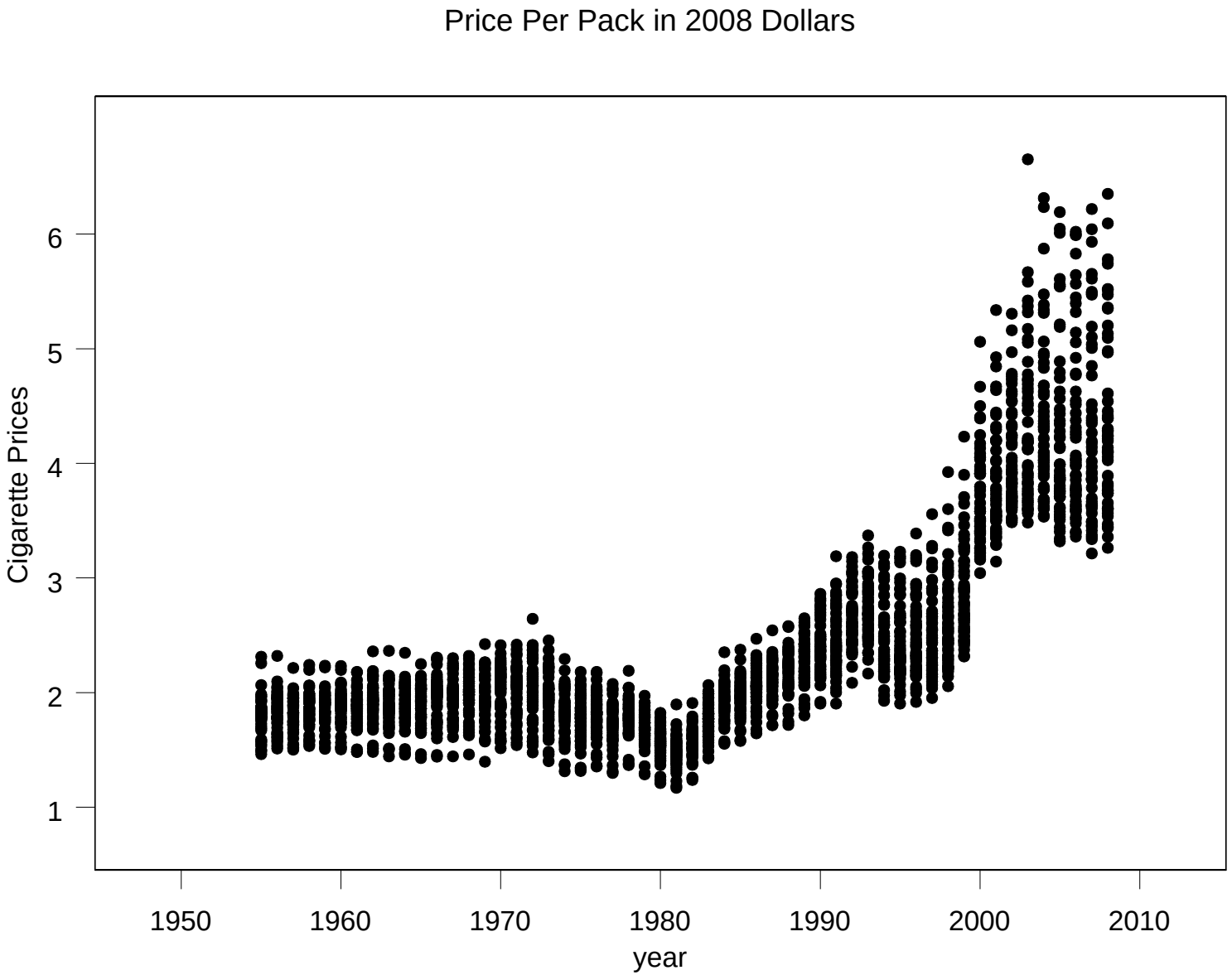
- Packs per capita
- Prices (in real dollars, $ 2008)
- State Income per Capita (in real dollars)
- Prices in neighboring states

Data

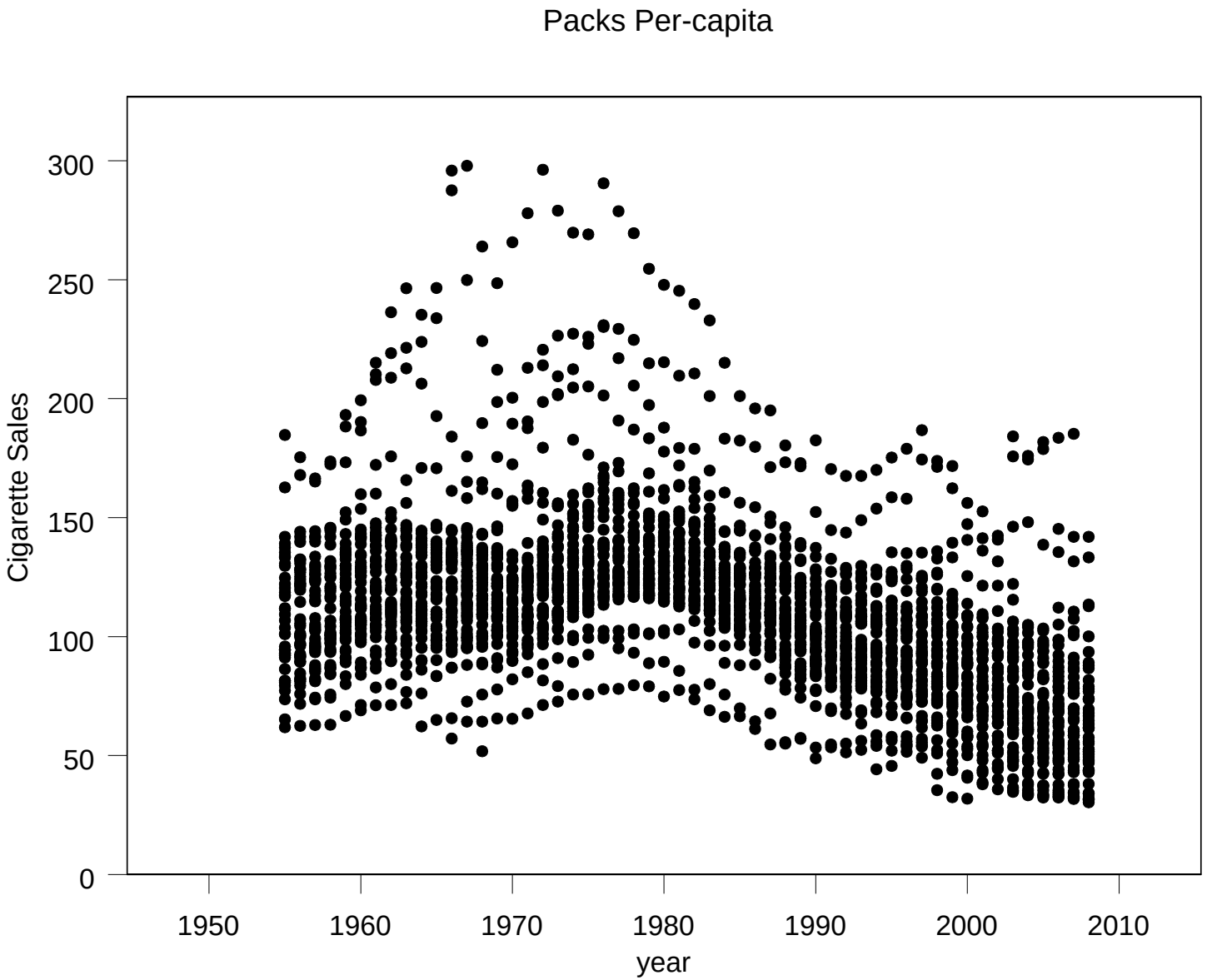
We observe data for

- 50 states $+$ D.C.
- from 1955 to 2008
- This is Panel Data where each state is the individual.

Cigarettes


Price Per Pack in 2008 Dollars
Cigarette Prices
1
2
3
4
5
6
1950
1960
1970
1980
1990
2000
2010
year

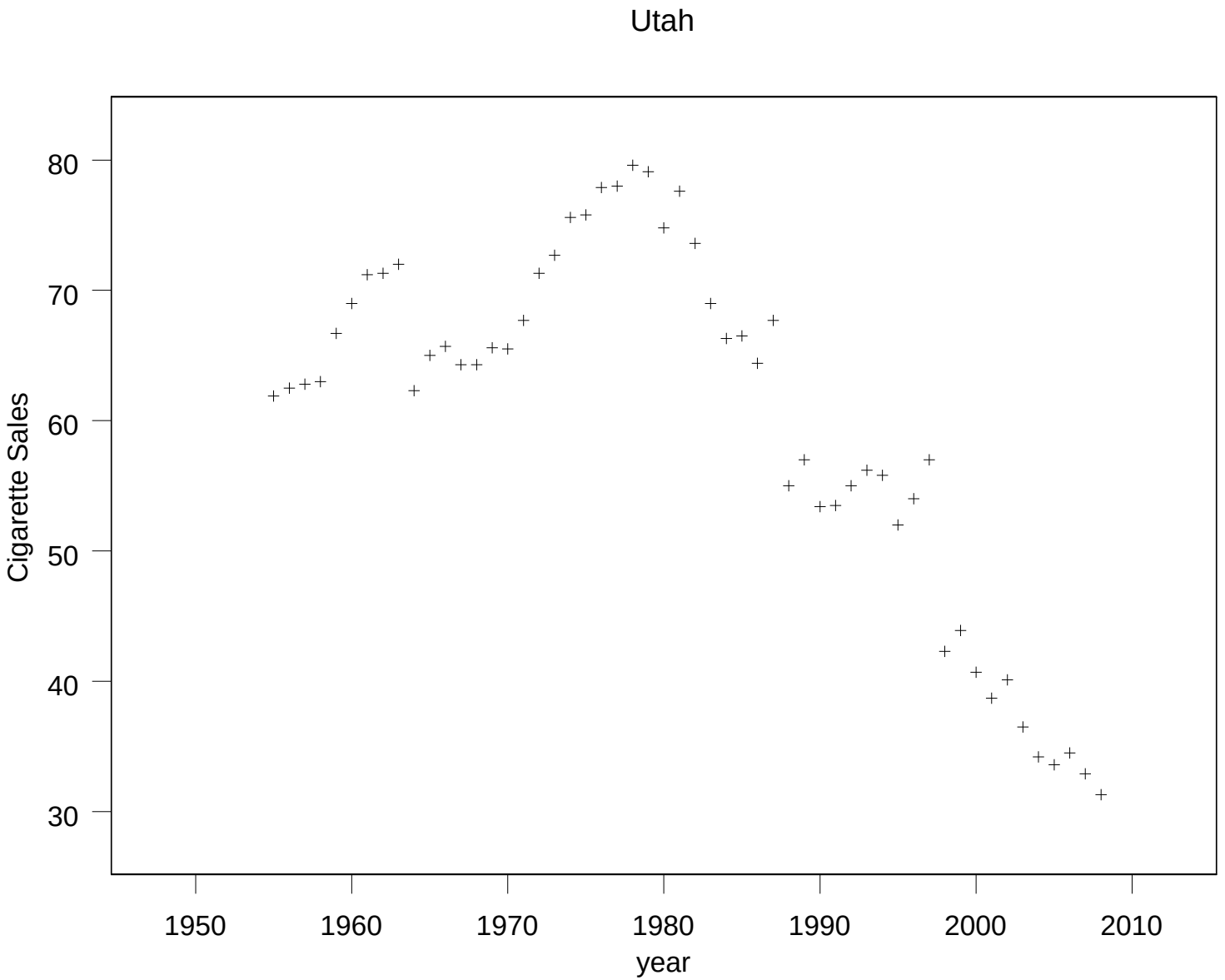
Cigarettes


Packs Per-capita
Cigarette Sales
300
250
200
150
100
50
0
1950
1960
1970
1980
1990
2000
2010
year

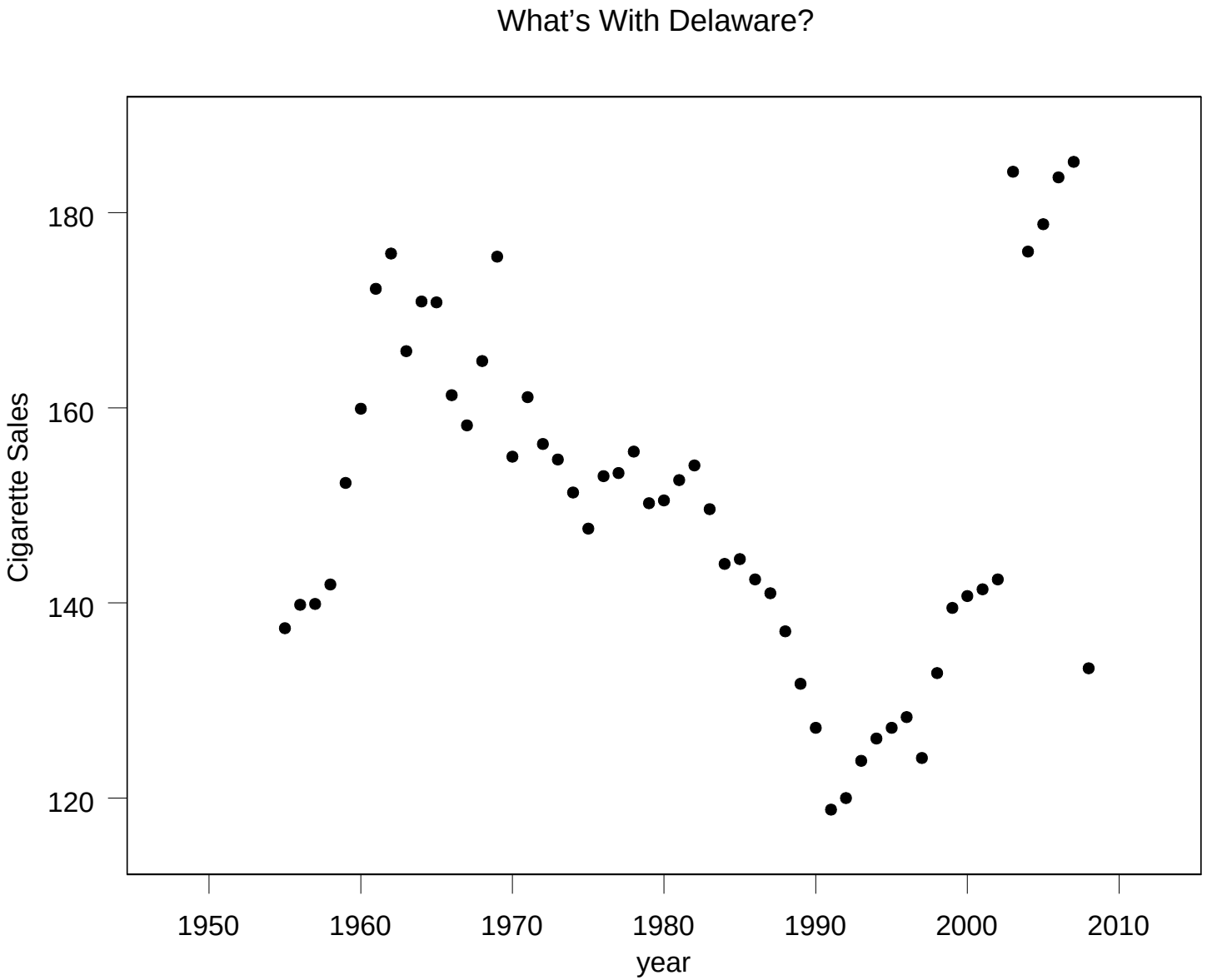
Cigarettes


What State is This?
Cigarette Sales
80
70
60
50
40
30
1950
1960
1970
1980
1990
2000
2010
year

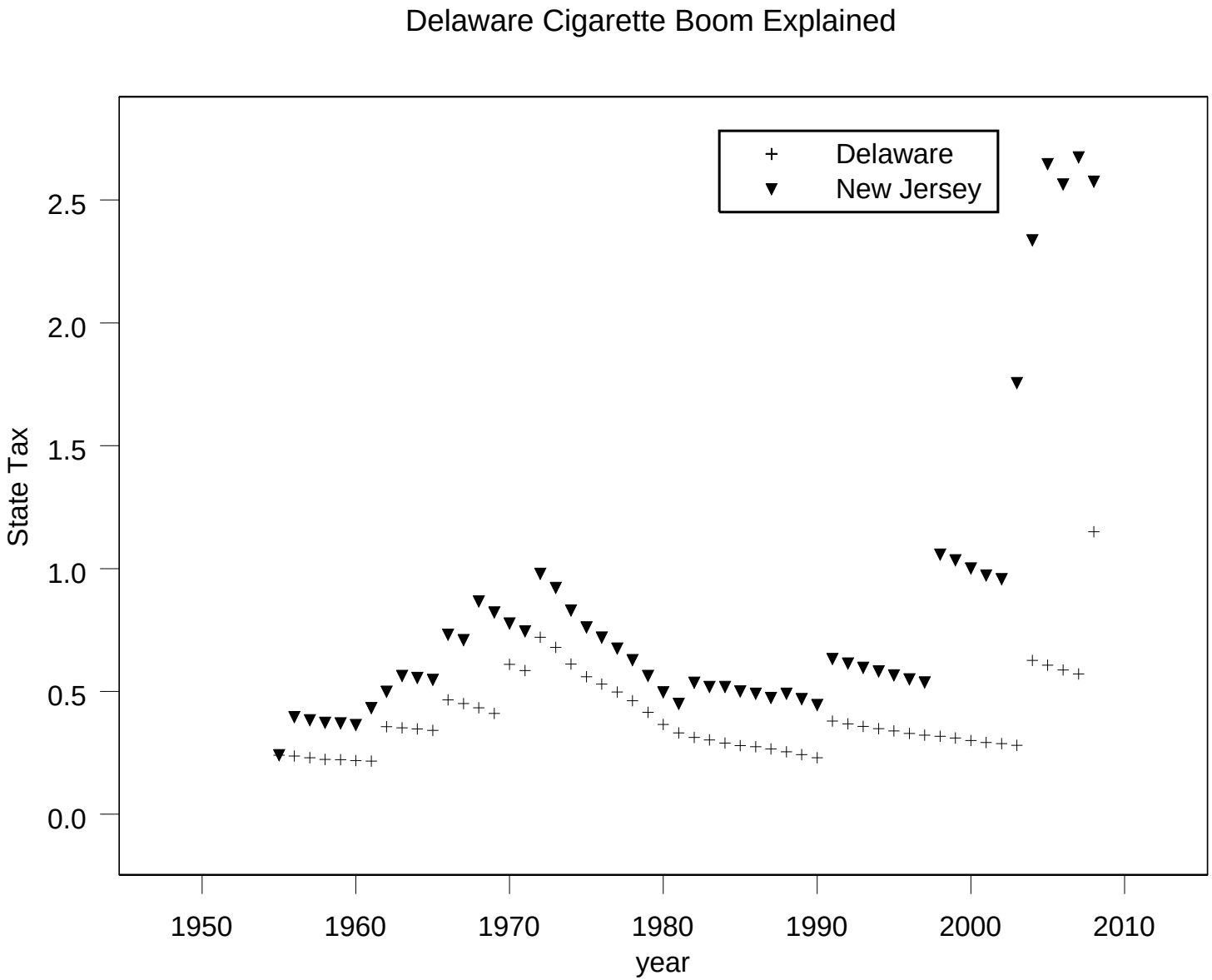
Cigarettes


Utah
Cigarette Sales
80
70
60
50
40
30
1950
1960
1970
1980
1990
2000
2010
year

Cigarettes


What's With Delaware?
Cigarette Sales
180
160
140
120
1950
1960
1970
1980
1990
2000
2010
year

Cigarettes

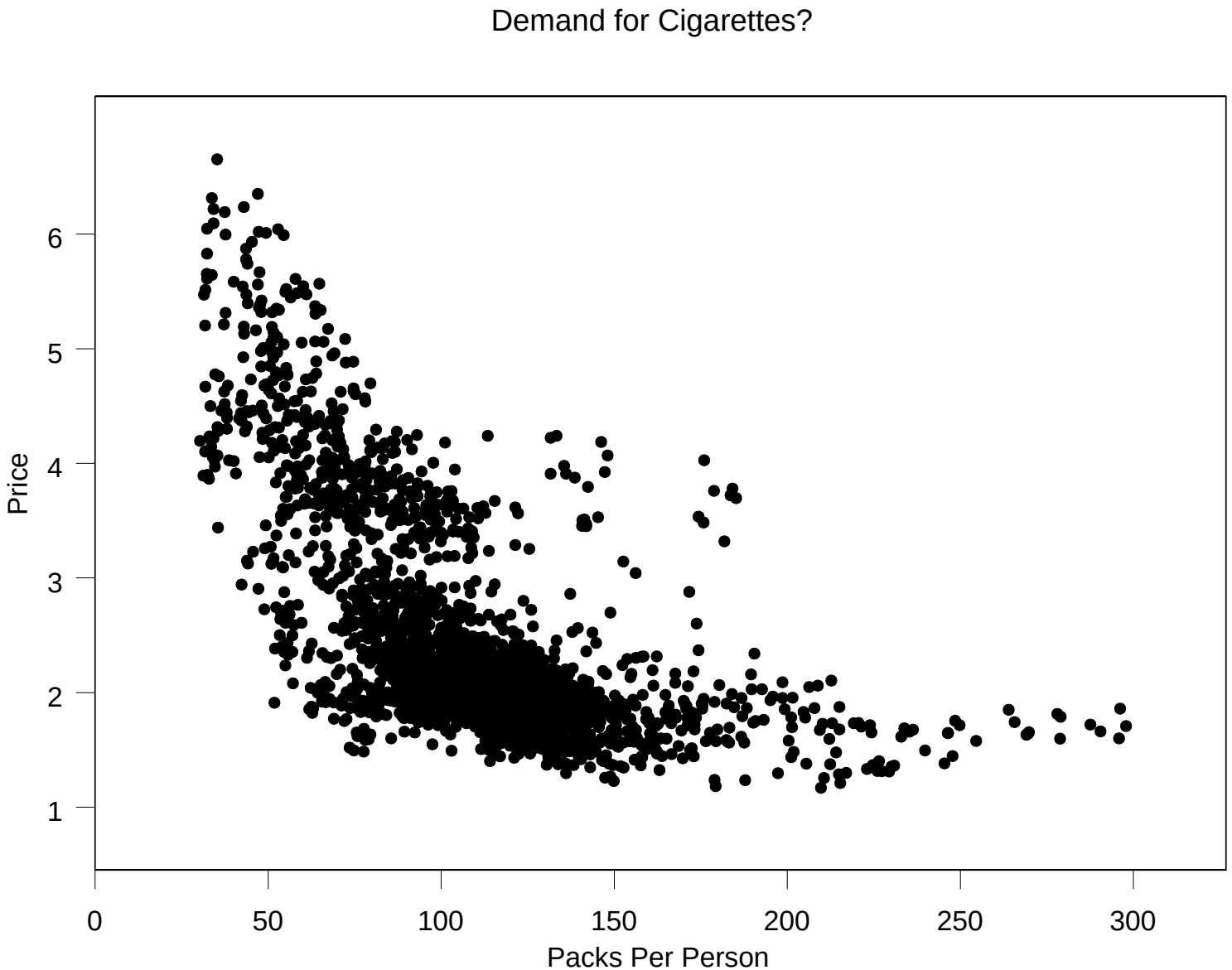

Delaware Cigarette Boom Explained
Delaware
New Jersey
State Tax
0.0
0.5
1.0
1.5
2.0
2.5
1950
1960
1970
1980
1990
2000
2010
year

Demand for Cigarettes

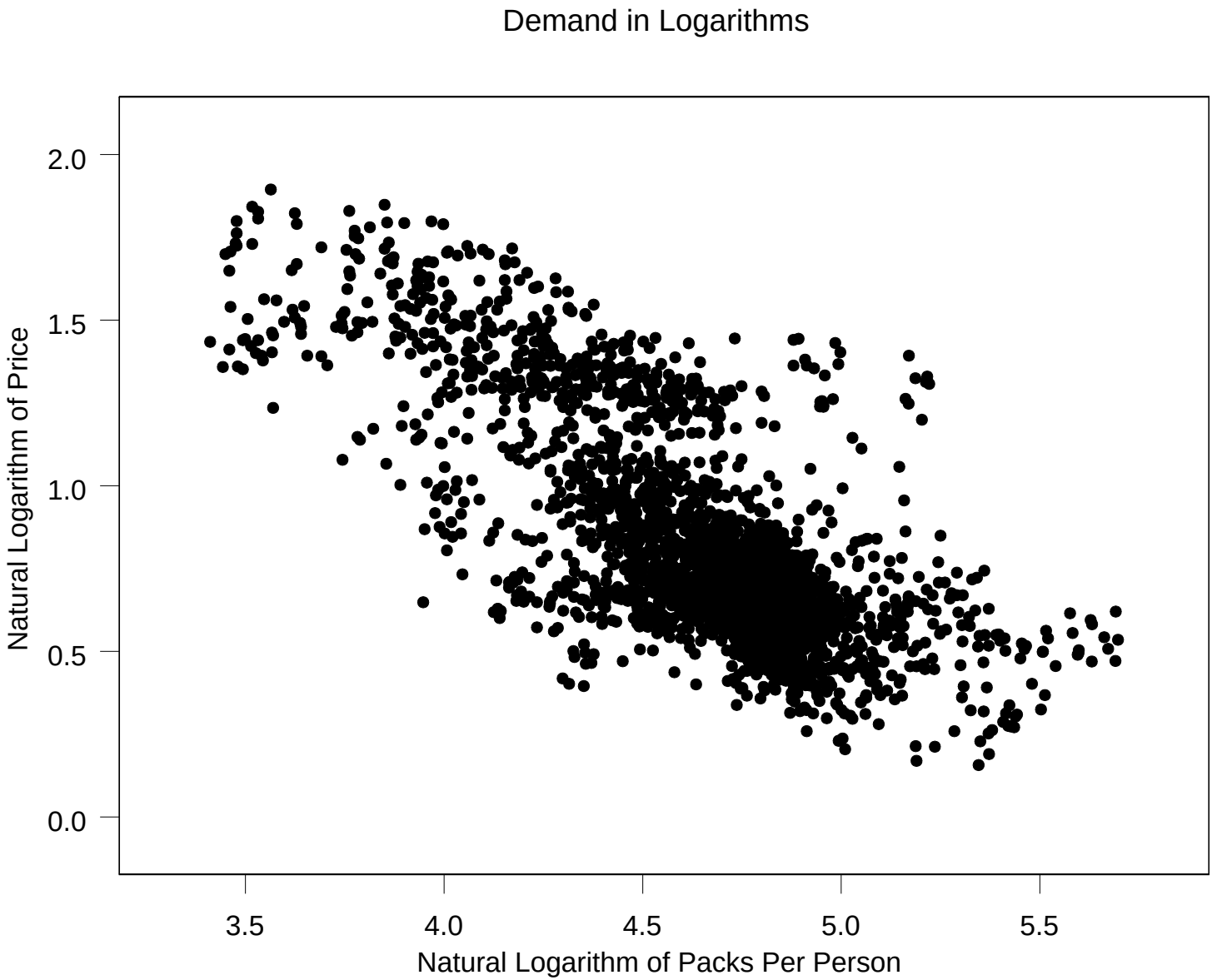
Demand for cigarettes in each state depends on:

- Price of cigarettes. As price $\uparrow$, quantity demanded $\downarrow$.
- State income. The effect could be positive or negative.
- Prices in other states. As prices in other states $\uparrow$, quantity demanded in this state $\uparrow$.

Cigarettes


Demand for Cigarettes?
Price
1
2
3
4
5
6
0
50
100
150
200
250
300
Packs Per Person

Cigarettes

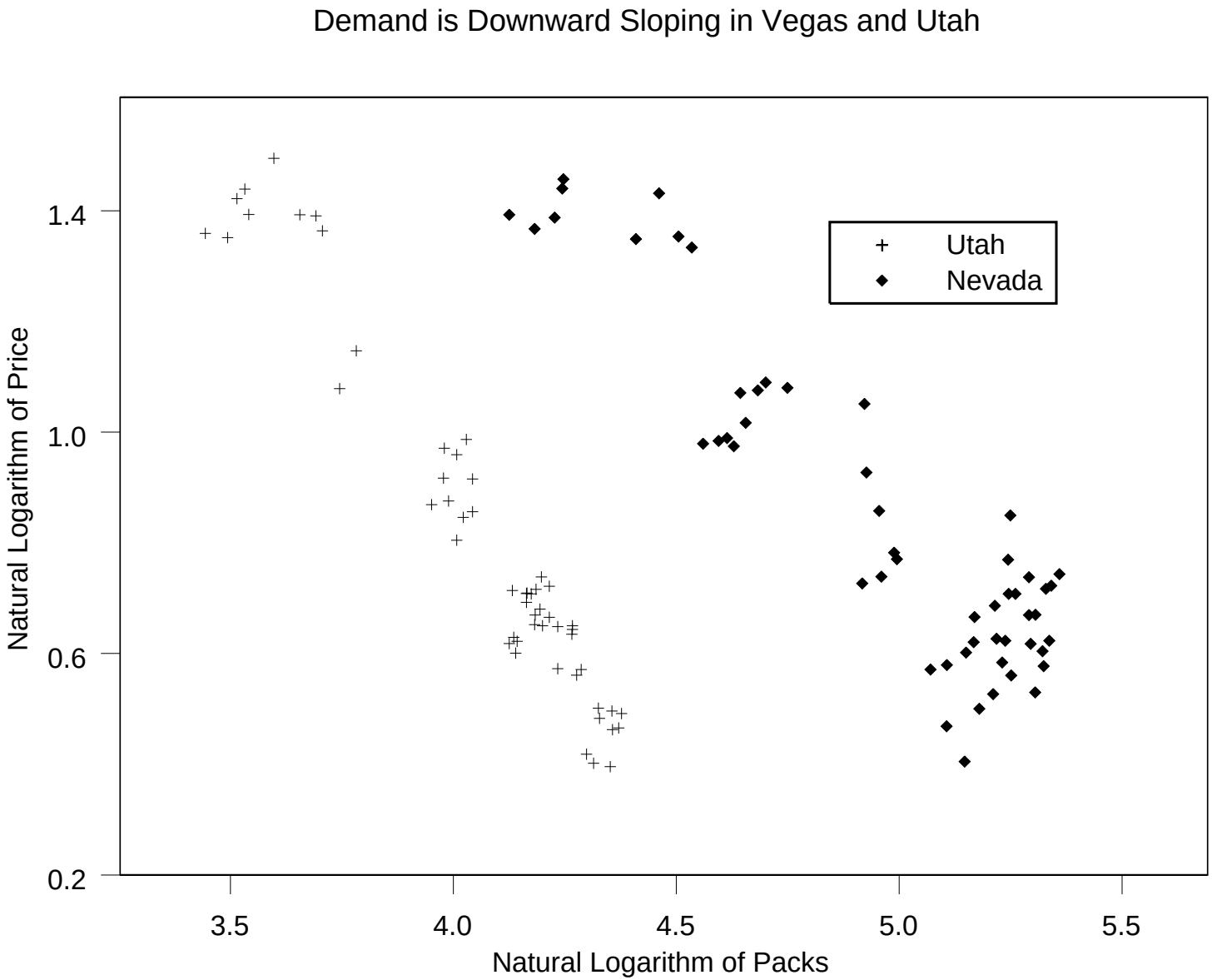

Demand in Logarithms
Natural Logarithm of Price
2.0
1.5
1.0
0.5
0.0
3.5
4.0
4.5
5.0
5.5
Natural Logarithm of Packs Per Person

Cigarette Demand and Panel Data

The data we have shown combines all states

- Do all the states have the same demand?
- Do all states respond to prices the same way?

Cigarettes


Demand is Downward Sloping in Vegas and Utah
Natural Logarithm of Price
1.4
1.0
0.6
0.2
3.5
4.0
4.5
5.0
5.5
Natural Logarithm of Packs
Utah
Nevada

Cigarette Demand and Panel Data

Summary:

1. It looks like different states have different demand curves.
2. They seem to have very similar slopes.
3. Panel data is an appropriate technique to estimate the slopes.

Models of Demand

All models are wrong, but some are useful.

George Box

Models of Demand

English: Demand for cigarettes in each state depends on

- price of cigarettes
- income of state
- prices in neighboring states
- amount of cigarettes purchased last year

Models of Demand

Math:

$$\begin{aligned} lncigs_{it} &= \alpha_i + \beta_1 lncigs_{i,t-1} + \beta_2 lnprice_{it} + \beta_3 lnincome_{it} \\ &\quad + \beta_4 lnpriceother_{it} + \lambda_t + v_{it} \\ i &= 1, \ldots, 51 \\ t &= 1955, \ldots, 2008 \end{aligned}$$

$lncigs_{it}$ = natural log of cigarettes for state i at time t

$lncigs_{i,t-1}$ = same thing but for the previous period

$lnprice_{it}$ = natural log of real price

$lnpriceother_{it}$ = natural log of neighbor price

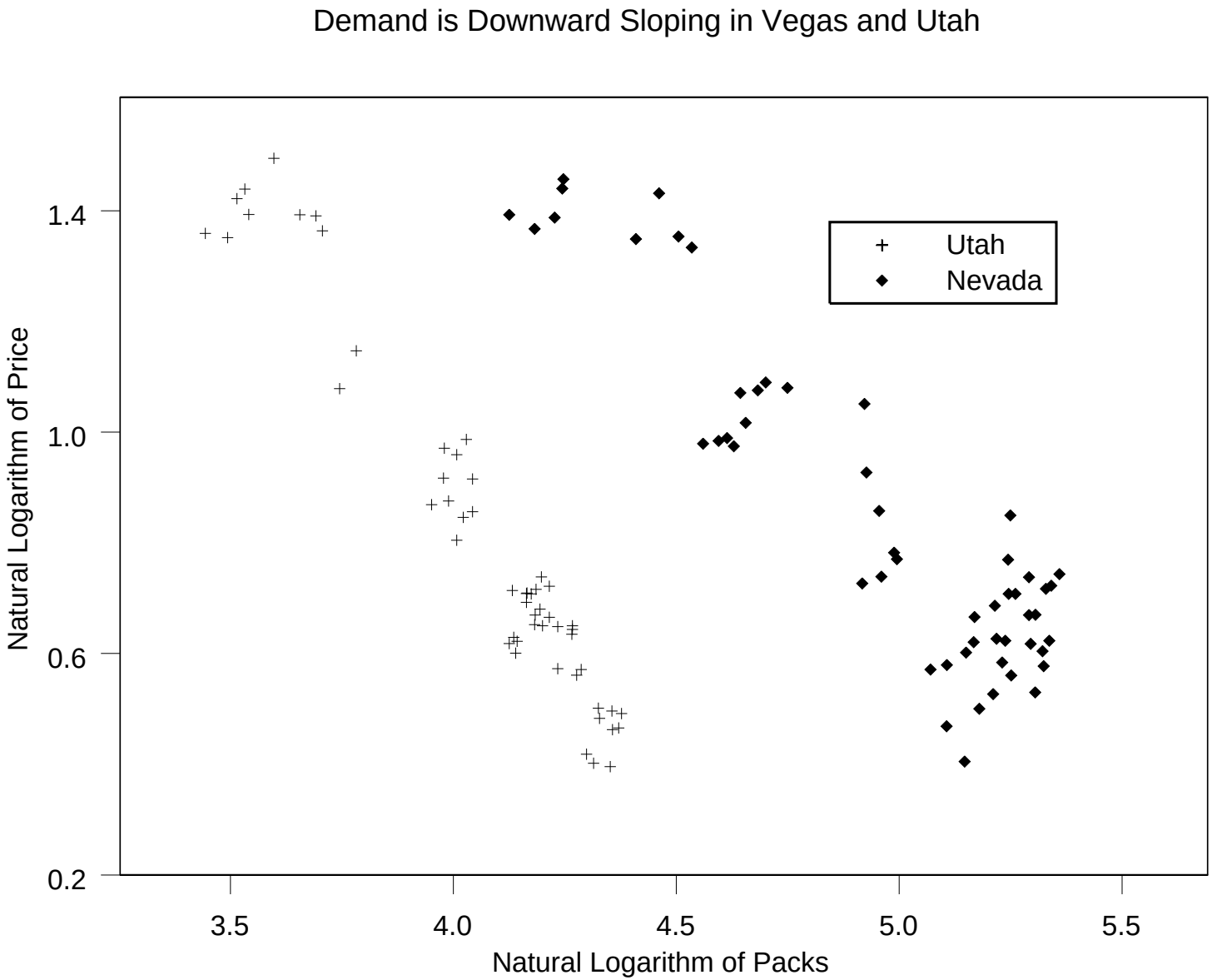
$lnincome_{it}$ = natural log of per capita real income

α_i = each state can have a shift but same slope

λ_t = each period can have events like advertising bans etc

v_{it} = error

Models of Demand



Models of Demand

Interpretation:

- Dang! I should have gone to the session on spelling bees in Korea!
- Each coefficient, β_j, represents an elasticity.

Models of Demand

What is an elasticity?

$$\begin{aligned} lncigs_{it} = \alpha_i + \beta_1 lncigs_{i,t-1} + \beta_2 lnprice_{it} + \beta_3 lnincome_{it} \\ + \beta_4 lnpriceother_{it} + \lambda_t + v_{it} \\ i = 1, \ldots, 51 \\ t = 1955, \ldots, 2008 \end{aligned}$$

- Elasticity of price, β_2 represents the percentage change in demand if we change the price by 1 percent.
- If $\beta_2 \leq -1$, demand is elastic.
- When demand is elastic and taxes per pack increase, tax revenues *decrease*!
- If $-1 < \beta_2 < 0$, demand is inelastic.
- When demand is inelastic and taxes per pack increase, tax revenues *increase*.

Models of Demand

What about β_1?

$$\begin{aligned} lncigs_{it} = \alpha_i + \beta_1 lncigs_{i,t-1} + \beta_2 lnprice_{it} + \beta_3 lnincome_{it} \\ + \beta_4 lnpriceother_{it} + \lambda_t + v_{it} \\ i = 1, \ldots, 51 \\ t = 1955, \ldots, 2008 \end{aligned}$$

- β_1 is a measure of habit formation.
- The closer β_1 is to one, the more persistent.

Estimation

$$\begin{aligned} lncigs_{it} &= \alpha_i + \beta_1 lncigs_{i,t-1} + \beta_2 lnprice_{it} + \beta_3 lnincome_{it} \\ &\quad + \beta_4 lnpriceother_{it} + \lambda_t + v_{it} \\ i &= 1, \ldots, 51 \\ t &= 1955, \ldots, 2008 \end{aligned}$$

- We don't know any of the β values.
- We must use our data and many pages of matrices to find estimated values.

Estimation

For 1960-1970

$$\begin{aligned} lncigs_{it} = \alpha_i + \beta_1 lncigs_{i,t-1} + \beta_2 lnprice_{it} + \beta_3 lnincome_{it} \\ + \beta_4 lnpriceother_{it} + \lambda_t + v_{it} \end{aligned}$$

- $\hat{\beta}_2 = -0.475$. If price increases one percent, demand declines by 0.475 percent.
- $\hat{\beta}_3 = 0.365$. If income increases one percent, demand increases by 0.365 percent.
- $\hat{\beta}_4 = 0.039$. If neighbor state raises cigs by one percent, demand in our state increases by 0.039 percent. This estimate is not statistically significant.
- $\hat{\beta}_1 = 0.305$. Demand is not very persistent.

Estimation

For 1970-1980

$$lncigs_{it} = \alpha_i + \beta_1 lncigs_{i,t-1} + \beta_2 lnprice_{it} + \beta_3 lnincome_{it} + \beta_4 lnpriceother_{it} + \lambda_t + v_{it}$$

- $\hat{\beta}_2 = -0.483$. If price increases one percent, demand declines by 0.483 percent.
- $\hat{\beta}_3$ is not statistically significant
- $\hat{\beta}_4 = 0.091$. If neighbor state raises cigs by one percent, demand in our state increases by 0.091 percent.
- $\hat{\beta}_1 = 0.547$. Demand is more persistent.

Estimation

For 1980-1990

$$lncigs_{it} = \alpha_i + \beta_1 lncigs_{i,t-1} + \beta_2 lnprice_{it} + \beta_3 lnincome_{it} + \beta_4 lnpriceother_{it} + \lambda_t + v_{it}$$

- $\hat{\beta}_2 = -0.218$. If price increases one percent, demand declines by 0.218 percent. This is less elastic than previous decades.
- $\hat{\beta}_3 = 0.129$ If income increases one percent, demand increases by 0.129 percent.
- $\hat{\beta}_4$ is not significant.
- $\hat{\beta}_1 = 0.677$. Demand is more persistent.

Estimation

For 1990-2000

$$\begin{aligned} lncigs_{it} = \alpha_i + \beta_1 lncigs_{i,t-1} + \beta_2 lnprice_{it} + \beta_3 lnincome_{it} \\ + \beta_4 lnpriceother_{it} + \lambda_t + v_{it} \end{aligned}$$

- $\hat{\beta}_2 = -0.405$. If price increases one percent, demand declines by 0.405 percent. This is similar to 1960's and 1970's.
- $\hat{\beta}_3 = 0.026$ If income increases one percent, demand increases by 0.026 percent.
- $\hat{\beta}_4 = 0.075$. If neighbor state raises cigs by one percent, demand in our state increases by 0.075 percent.
- $\hat{\beta}_1 = 0.661$. Demand is roughly same persistent as 1980's.

Estimation

For 2000-2008

$$lncigs_{it} = \alpha_i + \beta_1 lncigs_{i,t-1} + \beta_2 lnprice_{it} + \beta_3 lnincome_{it} + \beta_4 lnpriceother_{it} + \lambda_t + v_{it}$$

- $\hat{\beta}_2 = -0.781$. If price increases one percent, demand declines by 0.781 percent. This is most elastic time period.
- $\hat{\beta}_3 = 0.208$ If income increases one percent, demand increases by 0.208 percent.
- $\hat{\beta}_4 = 0.143$. If neighbor state raises cigs by one percent, demand in our state increases by 0.143 percent. This time period is most sensitive to neighbor price changes.
- $\hat{\beta}_1 = 0.382$. Demand is less persistent.

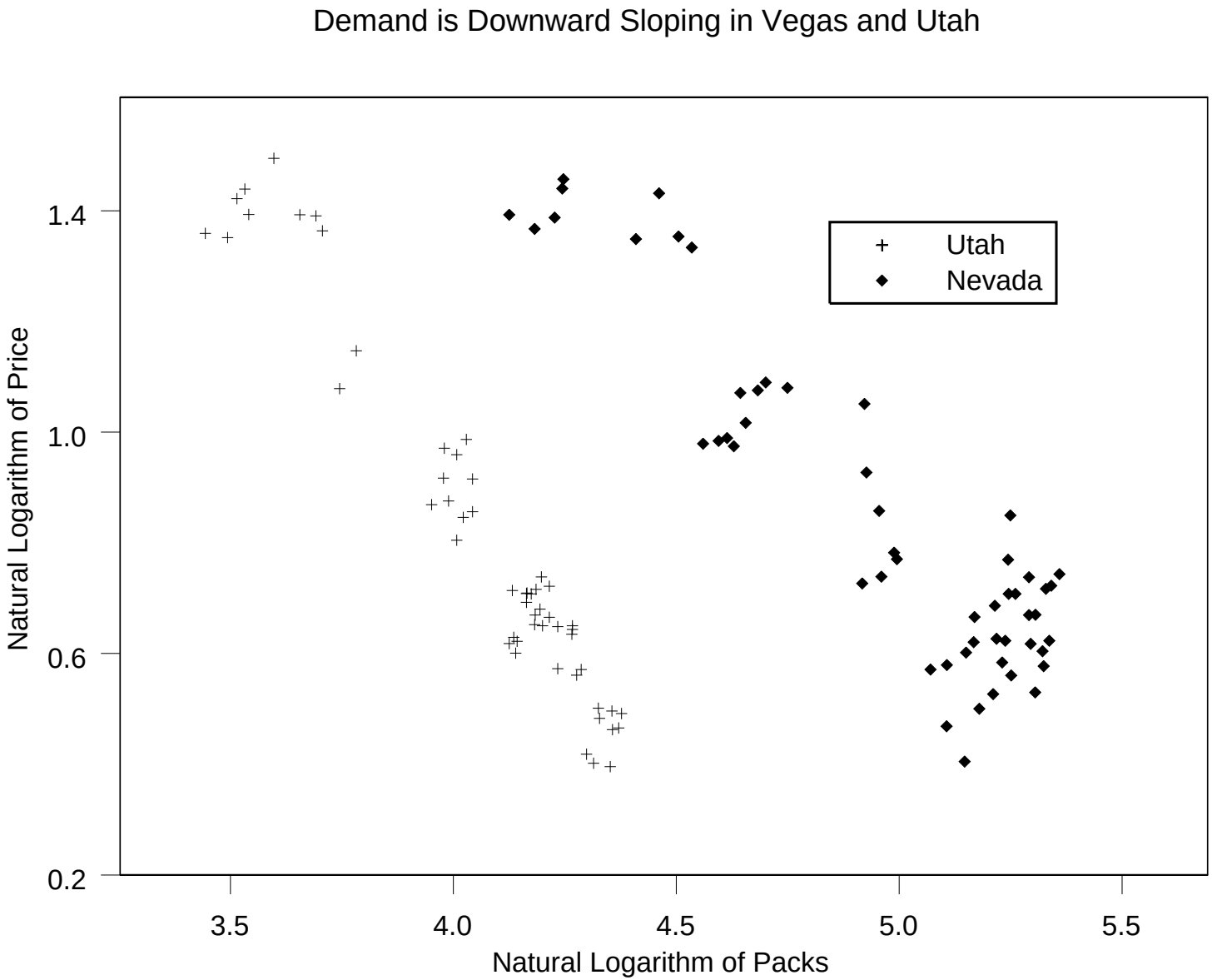
Are States Equal in Elasticity?

We should test if

$$\beta_{1i} = \beta_1$$
$$\beta_{2i} = \beta_2$$
$$\beta_{3i} = \beta_3$$
$$\beta_{4i} = \beta_4.$$

English: All the states have same *slope*

Cigarettes


Demand is Downward Sloping in Vegas and Utah
Natural Logarithm of Price
1.4
1.0
0.6
0.2
3.5
4.0
4.5
5.0
5.5
Natural Logarithm of Packs
Utah
Nevada

Are States Equal in Elasticity?

$$
\begin{aligned}
&-\frac{1}{2}\sum_{i=1}^{N}\left(vec(\Sigma_i^{-1})^\top-\frac{\iota_T^\top\Sigma_i^{-1}\otimes\iota_T^\top\Sigma_i^{-1}}{\iota_T^\top\Sigma_i^{-1}\iota_T}\right)(X_i\otimes X_i)vec(I_k)\\
&+\frac{1}{2}\sum_{i=1}^{N}\left(w_i^\top\Sigma_i^{-1}\otimes w_i^\top\Sigma_i^{-1}\right)(X_i\otimes X_i)vec(I_k)\\
&+\frac{1}{2}\sum_{i=1}^{N}\left[\left(w_i^\top\Sigma_i^{-1}\iota_T\iota_T^\top\Sigma_i^{-1}w_i\right)\frac{\iota_T^\top\Sigma_i^{-1}\otimes\iota_T^\top\Sigma_i^{-1}}{(\iota_T^\top\Sigma_i^{-1}\iota_T)^2}\right](X_i\otimes X_i)vec(I_k)\\
&-\frac{1}{2}\sum_{i=1}^{N}\left(\frac{w_i^\top\Sigma_i^{-1}\iota_T\iota_T^\top\Sigma_i^{-1}\otimes w_i^\top\Sigma_i^{-1}}{\iota_T^\top\Sigma_i^{-1}\iota_T}\right)(X_i\otimes X_i)vec(I_k)\\
&-\frac{1}{2}\sum_{i=1}^{N}\left(\frac{w_i^\top\Sigma_i^{-1}\otimes w_i^\top\Sigma_i^{-1}\iota_T\iota_T^\top\Sigma_i^{-1}}{\iota_T^\top\Sigma_i^{-1}\iota_T}\right)(X_i\otimes X_i)vec(I_k)
\end{aligned}
$$

Future Directions

There are several key features that could be explored:

- Current model assumes constant elasticity. This appears to work well but we should check.
- We need a more careful model of neighbor state behavior.
- If Kansas increases cigarette taxes per pack, we need to know if tax revenue may actually decrease.