# Economics 120A

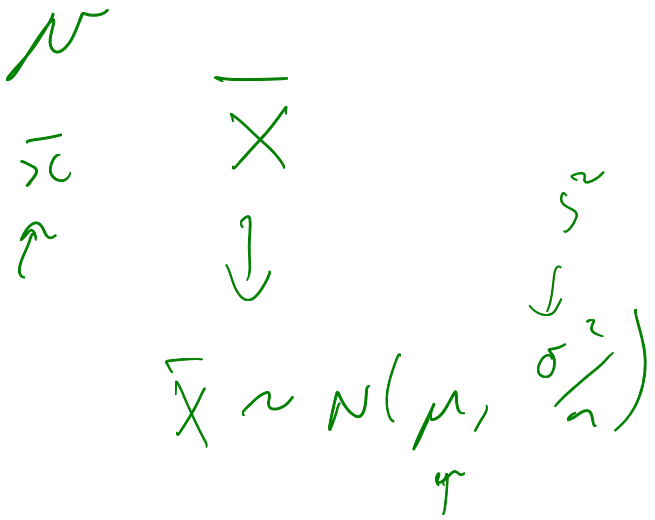
# 10. Case Studies

Graham Elliott

# Plan for Today

Real World examples of the methods and ideas of the course.

1. Equity Premium
2. Alzheimers
3. Forecast errors
4. Behavioural economics



# Equity Premium

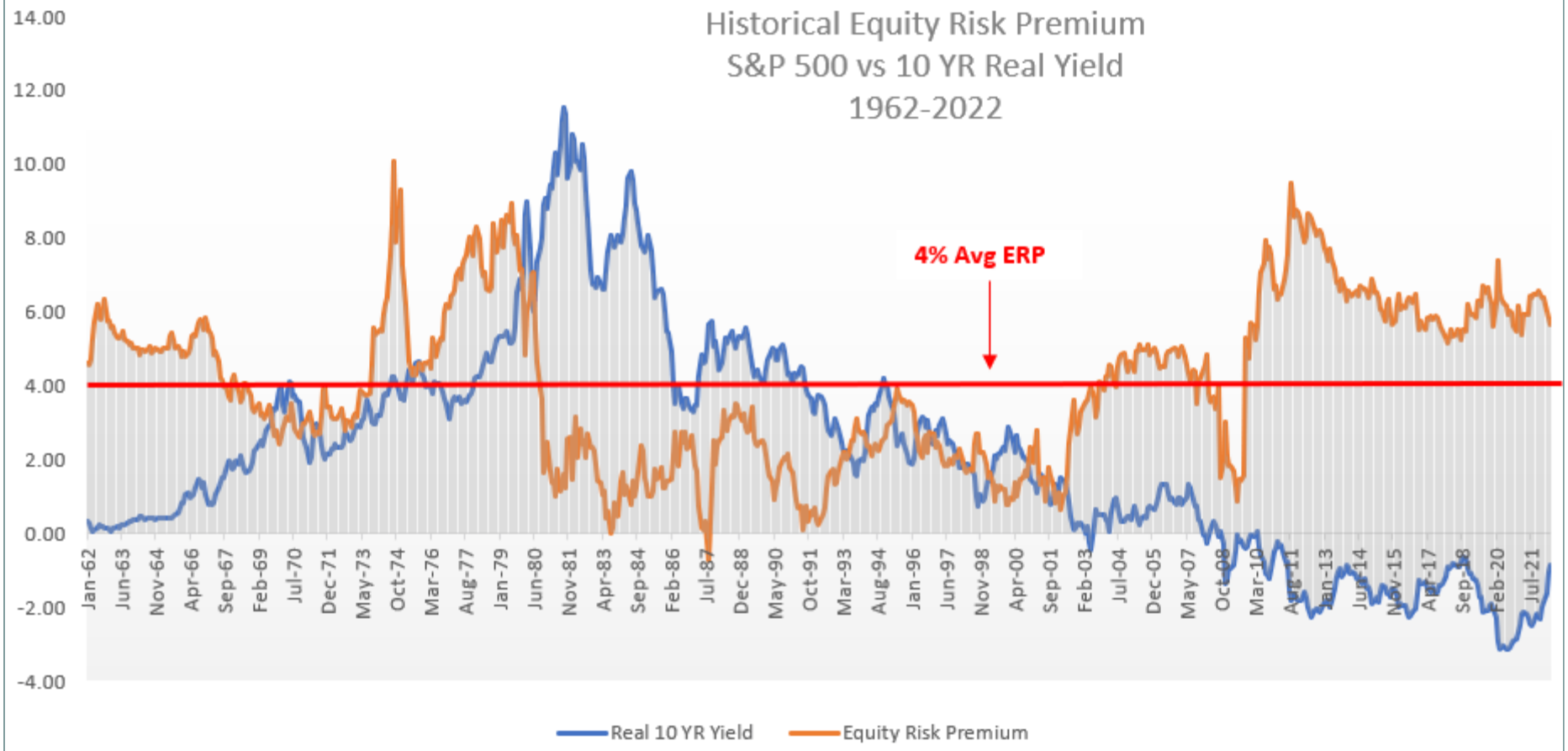
The equity premium is the difference between returns on stocks and the return on a risk free interest rate.

Imagine this is positive, why would you ever lend at the risk free rate instead of investing in stocks?

This should lead to adjustments that make these closer together.

There can still be a gap, which is the return to risk. We might add this to the risk free rate and examine that gap.

# The Data


Historical Equity Risk Premium
S&P 500 vs 10 YR Real Yield
1962-2022
4% Avg ERP
14.00
12.00
10.00
8.00
6.00
4.00
2.00
0.00
-2.00
-4.00
Jan-62
Jun-63
Nov-64
Apr-66
Sep-67
Feb-69
Jul-70
Dec-71
May-73
Oct-74
Mar-76
Aug-77
Jan-79
Jun-80
Nov-81
Apr-83
Sep-84
Feb-86
Jul-87
Dec-88
May-90
Oct-91
Mar-93
Aug-94
Jan-96
Jun-97
Nov-98
Apr-00
Sep-01
Feb-03
Jul-04
Dec-05
May-07
Oct-08
Mar-10
Aug-11
Jan-13
Jun-14
Nov-15
Apr-17
Sep-18
Feb-20
Jul-21
Real 10 YR Yield
Equity Risk Premium

# How would we test for this?

We can define $X_i$ as for year i the excess return adjusted for risk (the adjustment depends on an economic model for measuring risk, standard in the finance literature).

The existence of an "equity premium puzzle" rests on this being positive over time, there is no puzzle if it is zero on average.

We can set up the null and alternative hypotheses to test the hypothesis that there is no puzzle (the economics works).

We have

$EX_i = \mu$

$H_0$: $\mu = 0$ (no equity premium)

$H_1$: $\mu > 0$ (equity premium)

We would need the data to compute the statistics, however the paper reports relevant statistics.

# Some calculated statistics

Goyal + Welch

**Table 1 Descriptive Statistics**

| | Mean | Sdev. | Median | Min | Max | Skew | Kurt | JqBr |
|---|---|---|---|---|---|---|---|---|
| Panel A: Sample 1926–2002 | | | | | | | | |
| Rm($t$) | 9.29 | 19.80 | 13.43 | −58.74 | 45.71 | −0.89 | 3.95 | 13.13 |
| EQP($t$) | 5.57 | 20.00 | 9.37 | −59.82 | 45.41 | −0.75 | 3.66 | 8.59 |

$\bar{x}$ $s$ $q_2$

The data is annual, so we have 76 observations, so n=76.

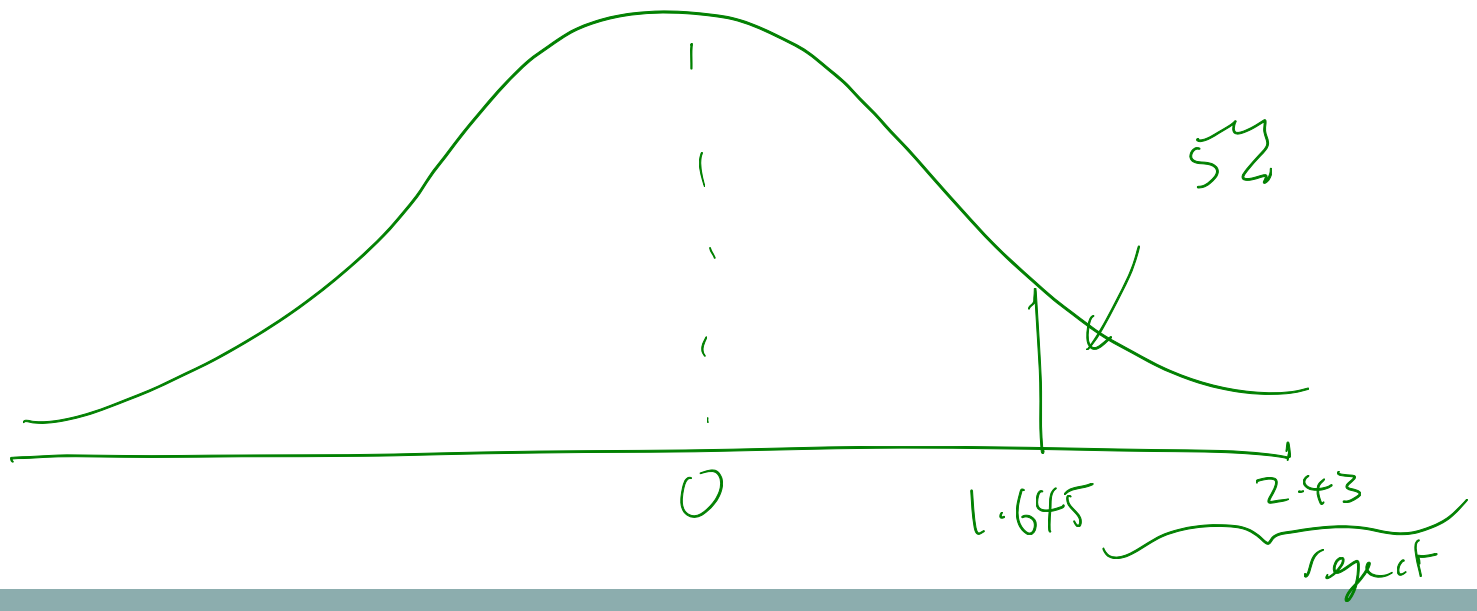
This is enough to do the test or construct confidence intervals.

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

# Tests and Confidence Intervals

We have $t = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}} = \dfrac{5.57 - 0}{20/\sqrt{76}} = 2.43$ ↓

For a size 5% test do we reject? yes, we reject.



# Confidence Interval

If we were interested in the likely additional gains to stocks, we might look at the confidence interval on the equity premium.

Let coverage be 95%.

It is $\bar{x} \pm 1.96 \frac{s}{\sqrt{n}}$

$5.57 \pm 1.96 \frac{20}{\sqrt{76}} = \{1.07, 10.07\}$

# Representative? VSRS?

$\bar{x}$, (s), n

We have data over a large number of years, is it reasonable that each year is representative of say next year (or the years we are hoping to invest over)?

Perhaps not – there might have been some change. In particular, access to the stock market has changed dramatically over the last 20 years, before this it was expensive and to a large extent only available to relatively wealthy investors.

This would require further work!

We also expect some correlation from year to year, so we might want to fix our estimator of the standard error.

Most economists still believe there is something there, although much smaller.

Recent work is in predicting when it might be large (so you have a high weight on stocks in your portfolio) and when it is low (so you have a low weight).

# Alzheimers

WASHINGTON (Reuters) - If both your parents have Alzheimer's disease, you probably are more much likely than other people to get it, researchers said on Monday.

Their study focused on 111 families in which both parents were diagnosed with Alzheimer's disease, the most common form of dementia among the elderly, and assessed the risk for developing it among the offspring.

The parents had 297 children who lived into adulthood. Of the 98 men and women who were at least 70 years old, 41 of them -- about 42 percent -- developed Alzheimer's disease, researchers at the University of Washington in Seattle found.

"That's greater than you would expect in the general population in that age group," Dr. Thomas Bird, one of the researchers, said in a telephone interview.

In the general population, risk for the disease begins to rise at about age 65, with the number of people developing the disease doubling every five years beyond that, experts say.

# Alzheimers

"The numbers will be interesting to follow as they get older and older," Bird said.

Bird said the study is not examining the Alzheimer's risk for people who have one but not two parents who develop the disease.

In order to confirm that both parents actually had Alzheimer's, the researchers reviewed the medical records and in many cases the brain autopsies of those who had died, and tried to meet in person to assess those who still living.

In people with Alzheimer's disease, healthy brain tissue degenerates, causing an inexorable decline in memory and mental abilities. The average length of time from diagnosis to death is about eight years.

# Analysis

$X_i \sim Bernoulli(\pi)$ $(\bar{x}, n)$

First, what is the sample estimate and what did we think it might be?

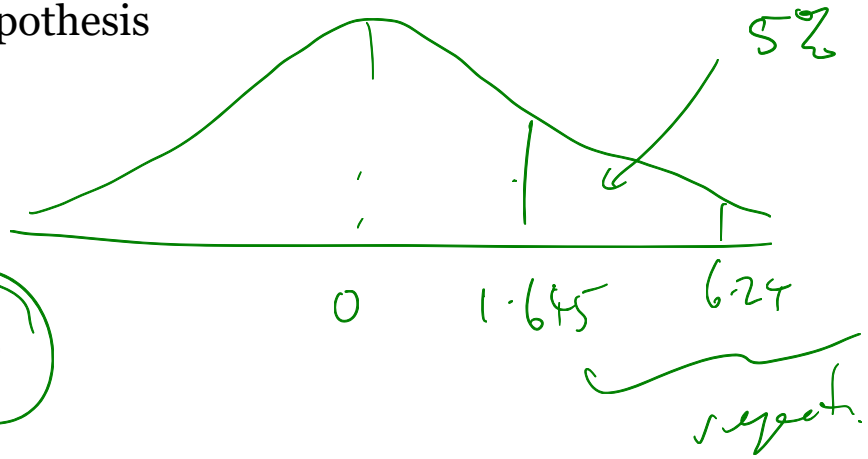
Our null hypothesis might be that there is no effect of multiple parents on the likelihood of getting Alzheimers. This was reported to be 0.1.

The estimated effect was that 41% of the subjects greater than 70 that had both parents with Alzheimers was 41%, which appears much larger.

The sample size (greater than 70yrs old) was n=98.

With the given information, we can test the hypothesis

$$H_0: \mu = 0.1$$
$$H_1: \mu > 0.1$$




We have a t-statistic of

$$t = \frac{0.41 - 0.1}{\sqrt{0.41 * (1 - 0.41)/98}} = 6.24.$$

# Confidence Interval

We have for a 95% confidence interval that the margin of error is

$$MoE = 1.96 * \sqrt{\frac{0.41*(1-0.41)}{98}} = 0.097$$

So the confidence interval is 0.41±0.10={0.31,0.51}.

What type of study is it?

First, classify the study. It is clearly not an experiment (they did not give Alzheimers to anyone).

There is no randomization mentioned in the article, I would presume it is an observational study.

Usually these studies are done where a researcher finds people from medical records that have been diagnosed with the condition, then follows up taking every observation that they can.

Respresentative?

# Conclusions

Clearly there is an effect here, and not a strongly surprising one.

Interpretation?

"I think it confirms that there's a strong genetic component in the disease and that's not a surprise," said Bird, whose study was published in the Archives of Neurology.

Typically brothers/sisters have at least the first 20 years or so of environment in common, not to mention that the habits they obtain in this period will persist.

# Forecast Error Bias

Much of macroeconomics, both policy and individual decision making, depends on forecasts of future conditions.

For example, there is a lot of debate at the moment regarding the likelihood that we are in a depression. How the FOMC responds to this in great part depends on forecasts of output (how big a problem is the decline?) and inflation (how bad will inflation get if they try to use monetary policy).

In practice then, the Fed and businesses etc. depend on the ability to forecast these types of variables for decision making.

Do forecasters do a good job in forecasting?

*An Evaluation of Recent Macroeconomic Forecast Errors*

*Scott Schuh*

*New England Economic Review*

*January/February 2001*

# Basic Idea

Observe forecaster over a period of time, we do not expect that they get the forecast exactly equal to the value of the variable they are forecasting, but we hope that they are close.

One notion of being close is to say that on average, they are equal to the outcome. We call this 'unbiasedness', it is the same notion that we had for the estimator.

Alternatively, they could be biased upwards (picking too high a value on average) or downward (picking too low a value on average).

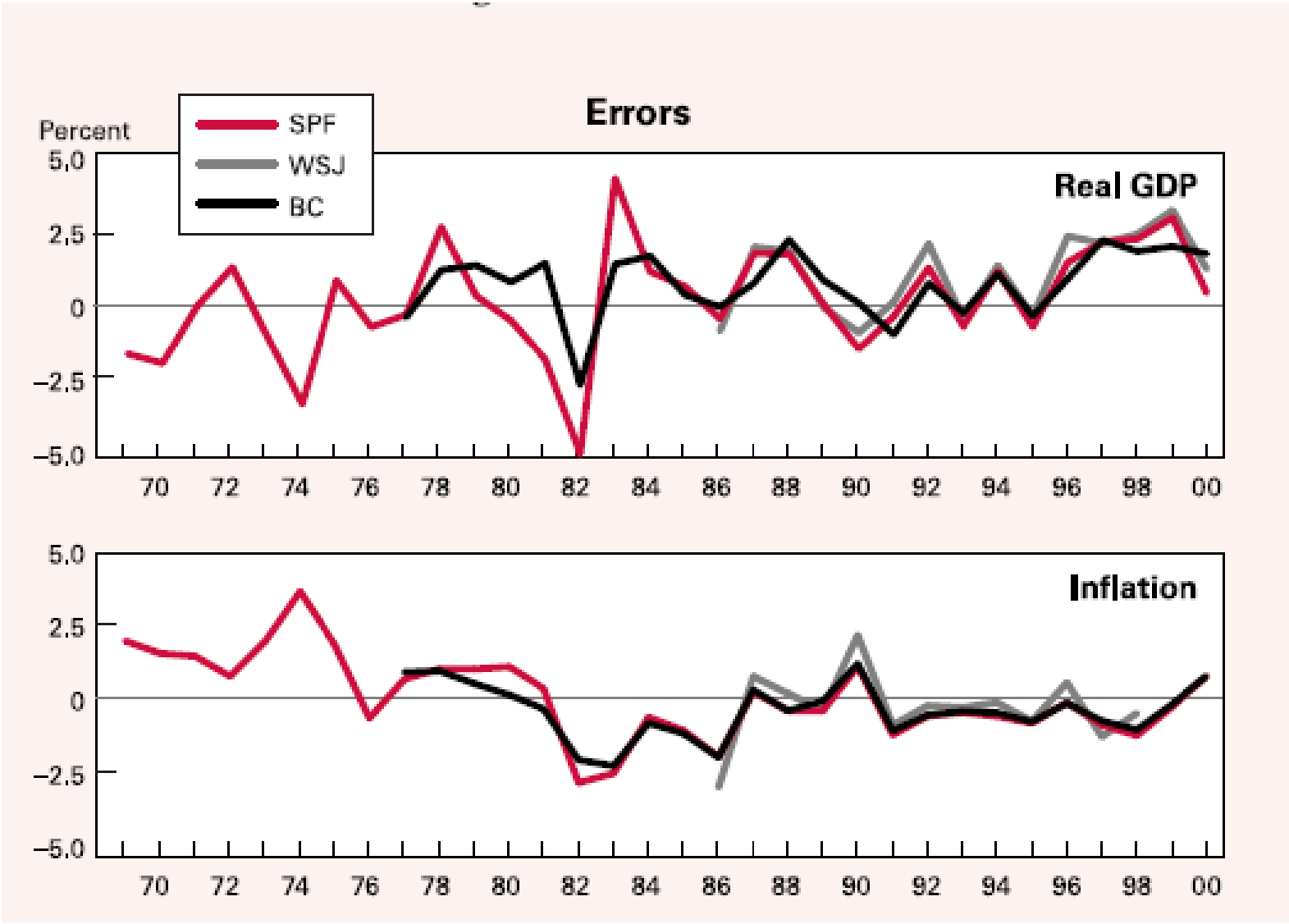
Suppose the forecaster looks ahead one year to forecast inflation, then at the end of the year we can observe the forecasters error.

Let this be $X_i$, where the i represents the particular forecast error. We might over time observe n of these forecast errors.

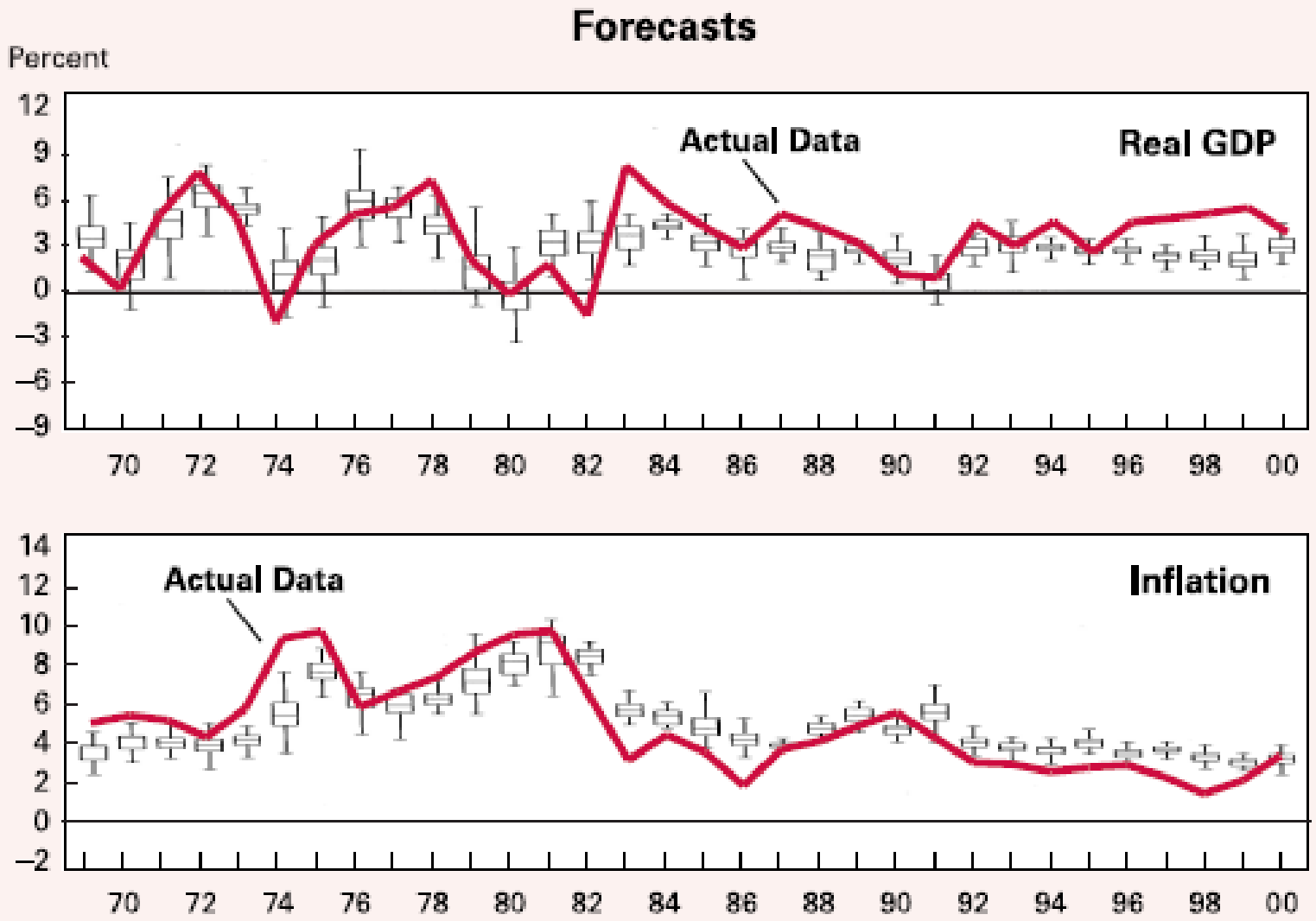
The data we have are for one year ahead, from three groups

1. Survey of Professional Forecasters
2. Wall Street Journal
3. Blue Chip indicators

# Time Series Plots



# *Individual SPF Forecasts and Errors*



# Need formal analysis to answer the question

How would we write our null hypothesis (no bias) and alternative hypothesis (bias) ?

$$H_0: \mu = 0$$
$$H_1: \mu \neq 0$$

We would obtain the variance using our estimator from the data.

Table 1

## *SPF Forecast Error Statistics*

Percent

| | Full Sample | | 1996–2000 | |
|---|---|---|---|---|
| Error | Mean | Std. Dev. | Mean | Std. Dev. |
| $y$ | .30 | 1.82 | **1.93** | .84 |
| $y^r$ | −.14 | 1.96 | **1.68** | .98 |
| $\pi$ | .25 | 1.36 | −.18 | .75 |
| $u$ | −.02 | .84 | **−.53** | .22 |
| $i$ | **−.52** | 1.53 | .10 | .79 |
| $r$ | .03 | 1.29 | .30 | .50 |

Note: Full sample period is 1969–2000 except for $i$ and $r$, which have samples of 1982–2000. **Bold** values indicate that the mean is significantly different from zero at the 5 percent level or better.

# Conclusion

Over the full sample, there is little evidence of bias, although in the recent sample there is more.

Is this evidence against VSRS? Perhaps there has been a change recently, and recent random variables are different from earlier ones.

The actual size of the errors (economic as opposed to statistical significance) is relatively small, from a policy perspective.

# Behavioral Economics Example

**Experimental Tests of the Endowment Effect and the Coase Theorem**

Daniel Kahneman; Jack L. Knetsch; Richard H. Thaler

*The Journal of Political Economy*, Vol. 98, No. 6. (Dec., 1990), pp. 1325-1348.

## The Willingness to Pay–Willingness to Accept Gap, the "Endowment Effect," Subject Misconceptions, and Experimental Procedures for Eliciting Valuations

*By* CHARLES R. PLOTT AND KATHRYN ZEILER*

# Basic Idea

*We conduct experiments to explore the possibility that subject misconceptions, as opposed to a particular theory of preferences referred to as the "endowment effect," account for reported gaps between willingness to pay ("WTP") and willingness to accept ("WTA"). The literature reveals two important facts. First, there is no consensus regarding the nature or robustness of WTP-WTA gaps. Second, while experimenters are careful to control for subject misconceptions, there is no consensus about the fundamental properties of misconceptions or how to avoid them. Instead, by implementing different types of experimental controls, experimenters have revealed notions of how misconceptions arise. Experimenters have applied these controls separately or in different combinations. Such controls include ensuring subject anonymity, using incentive-compatible elicitation mechanisms, and providing subjects with practice and training on the elicitation mechanism before employing it to measure valuations. The pattern of results reported in the literature suggests that the widely differing reports of WTP-WTA gaps could be due to an incomplete science regarding subject misconceptions. We implement a "revealed theory" methodology to compensate for the lack*

# Willingness to Pay

The basic issue here is that subjects in experiments appear to be willing to pay less for the same thing they would only sell for more.

Consider a concert ticket - you might be willing to pay up to $50 for a ticket, but if you own the ticket you might only be willing to accept $60 or more.

This means that as a buyer you prefer $55 to the ticket, as the seller you prefer the ticket to $56, hence you appear to prefer $55 over $56.

This violates rationality notions

The idea that owning something gives it additional value is known as the 'endowment' effect.

# Kahneman et. al (1990)

How is the experiment run?

In Kahneman et al (1990) we take college students, randomize into buyers and sellers. The sellers are given a mug (with the university logo!), matched with a buyer.

The buyer and seller give bids based on what they are willing to pay or accept.

The average difference is the measurement we want to make (WTA-WTP), this will be our $X_i$. $= WTA - WTP$

We could make the null 'no effect' saying that the difference between the WTA and the WTP is zero.

We expect this is a one sided alternative, people on average are unlikely to value stuff they have less than others want to buy it.

So we have

$$H_0: \mu = 0$$
$$H_1: \mu > 0$$

# Representativeness?

The current paper worries about a number of problems --- that the students do not understand what they are supposed to do, even if they understand they might not take it seriously, or that they may influence the answers of other students.

In terms of our general statistical setup, these are problems of representativeness, that the mean of $X_i$ is not what we are trying to measure because it is mixed up with other effects.

For example, confused students might bid stupid numbers because they do not understand the problem, this affects the average measured effect.

They might also just bid randomly, because they do not care about the trade (it is not too important to them, and this is an artificial environment).

Also, if they are allowed to talk amongst themselves, or see what others are doing, this affects the results as well.

# Plott replication of Kahneman

TABLE 2—INDIVIDUAL SUBJECT DATA AND SUMMARY STATISTICS FROM KKT REPLICATION

| Treatment | Individual responses (in U.S. dollars) | Mean | Median | Std. dev. |
|---|---|---|---|---|
| WTP | 0, 0, 0, 0, 0.50, 0.50, 0.50, 0.50, 0.50, 1, 1, 1, 1, 1, 1.50 | 1.74 | 1.50 | 1.46 |
| ($n$ = 29) | 2, 2, 2, 2, 2, 2.50, 2.50, 2.50, 3, 3, 3.50, 4.50, 5, 5 | | | |
| WTA | 0, 1.50, 2, 2, 2.50, 2.50, 3, 3.50, 3.50, 3.50, 3.50, 3.50, 4, 4.50 | 4.72 | 4.50 | 2.17 |
| ($n$ = 29) | 4.50, 5.50, 5.50, 5.50, 6, 6, 6, 6.50, 7, 7, 7, 7.50, 7.50, 7.50, 8.50 | | | |

# t-test

The data are not paired, we can still do a t-test but I had to estimate it using a formula we have not seen (you will be able to do this easily in 120b).

My estimated standard error for the difference in the means was 0.485, and the estimated difference is 2.98.

The t statistic is $t = \frac{2.98-0}{0.485} = 6.14$.

Many theories (mostly in psychology literature) to explain these findings (remember, data leads to implications and further theorizing).

Endowment effect: We are more averse to losing things we have than we are excited by getting things we do not have.

Legal ramifications: Value of a loss is not the purchase price of a similar object.

# Plott response

The objective of this paper was not simply to argue that the data are not representative, but to show it by doing the experiment well and making the effect smaller in the process.

The paper makes some effort to make the data representative was employed.

-The students had trial runs (for which they made money if they did it well) to familiarize themselves with the problem.
-The calculations were done on data from the final (with the mug, instead of tokens) round.
-Bids and payouts were anonymous, so that the results did not get affected by people worried about what others might think.

Basically though the argument is that the experimental design is responsible for these large effects.

# Improved design

| Experiment | Treatment | Individual responses (in U.S. dollars) | Mean | Median | Std. dev. |
|---|---|---|---|---|---|
| Experiment 1: (USC/practice) | WTP ($n$ = 15) | 0, 1, 1.62, 3.50, 4, 4, 4.17, 5, 6, 6, 6.50, 8, 8.75, 9.50, 10 | 5.20 | 5.00 | 3.04 |
| | WTA ($n$ = 16) | 0, 0.01, 3, 3.75, 3.75, 3.75, 5, 5, 5, 6, 6, 6, 7, 11, 12, 13.75 | 5.69 | 5.00 | 3.83 |
| Experiment 2: (USC/no practice) | WTP ($n$ = 12) | 1, 2, 3.50, 5, 5, 5, 8, 8.50, 9, 11.50, 13, 23 | 7.88 | 6.50 | 6.00 |
| | WTA ($n$ = 14) | 0.50, 1, 2, 2.50, 2.50, 4.50, 4.50, 5.70, 6.25, 8, 8, 8.95, 12, 13.50 | 5.71 | 5.10 | 4.00 |

# Conclusion

The difference between willingness to pay and willingness to accept basically goes away, it is statistically small and could easily be due to randomness.

The findings in the literature may not be real, just an artifact of the way in which the experiments were run.