ANALYSIS GROUP
ECONOMIC, FINANCIAL and STRATEGY CONSULTANTS

WHITE & CASE

Econometric Analysis in Litigation

February 21, 2006

Pierre Y. Cremieux, Ph.D., pcremieux@analysisgroup.com
Matthew S. Wild, Esq., mwild@whitecase.com

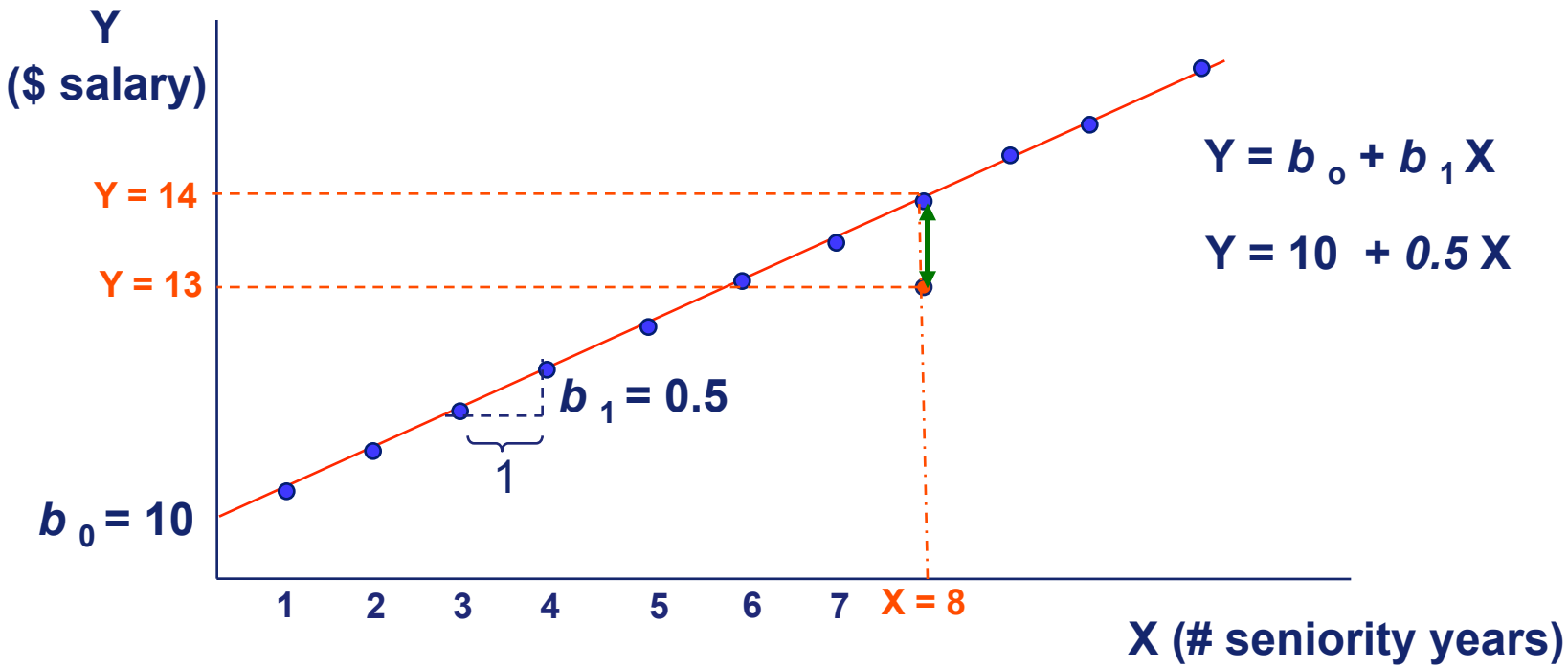
What We Want to Achieve Today

- **Enable you to work confidently with economic experts**
  - Push back on your experts
  - Understand how to review an opposing expert's analysis
- **Encourage you to learn more about econometrics and regression analysis**
  - Sources provided in bibliography

What is Regression Analysis?

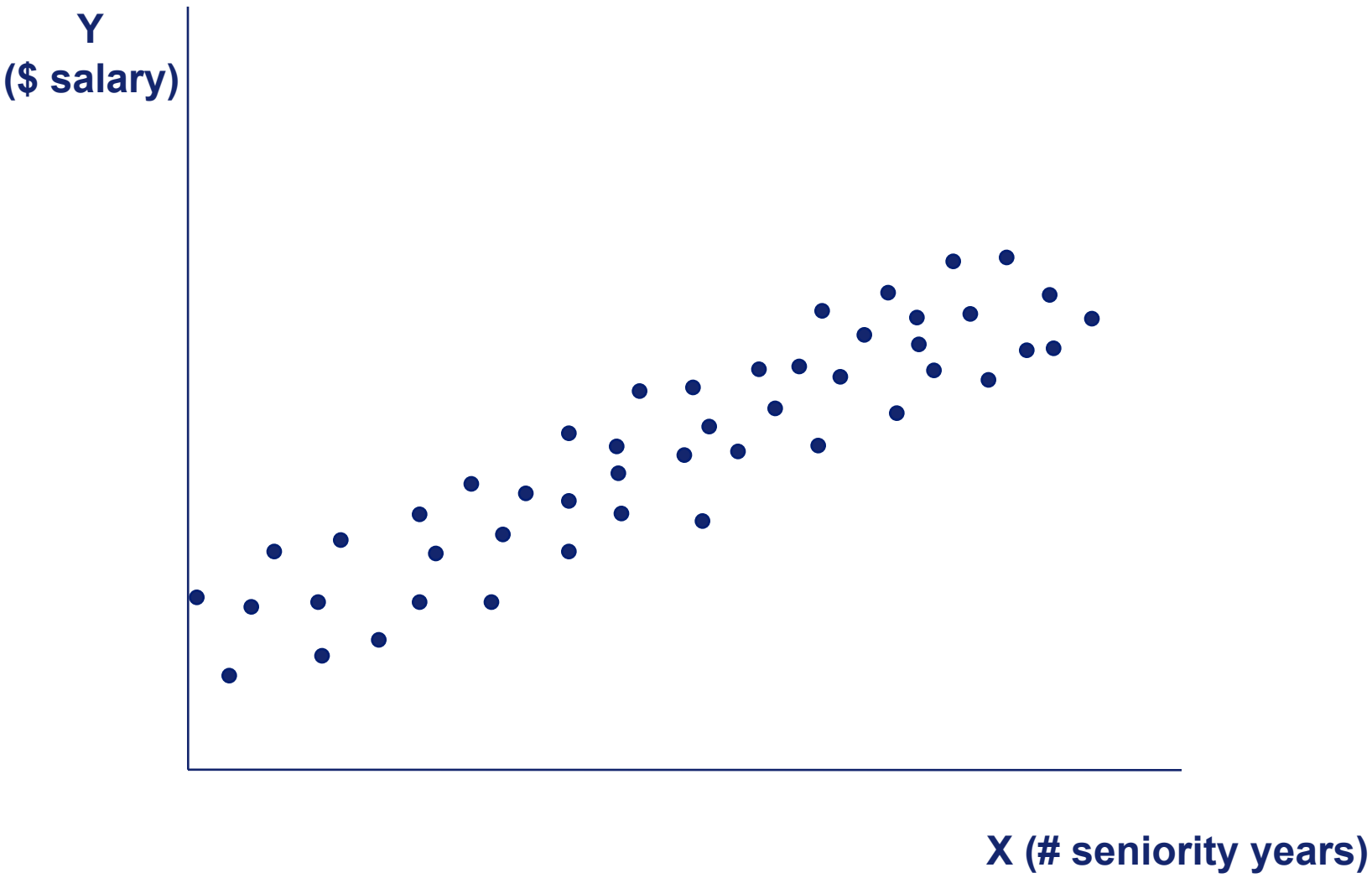
- **Regression analysis is**
  - A statistical approach developed to investigate the relationship between two variables
  - A statistical model of reality that simplifies the real world without losing the essence of a complex relationship
- **Regression analysis mixes art and science**
  - The art is in developing a statistical model simple enough to be manageable but not so simple that it becomes meaningless
  - The science is in the application of the right mathematical tools at the right time and under the right circumstances

Regression Model: A Simple Example

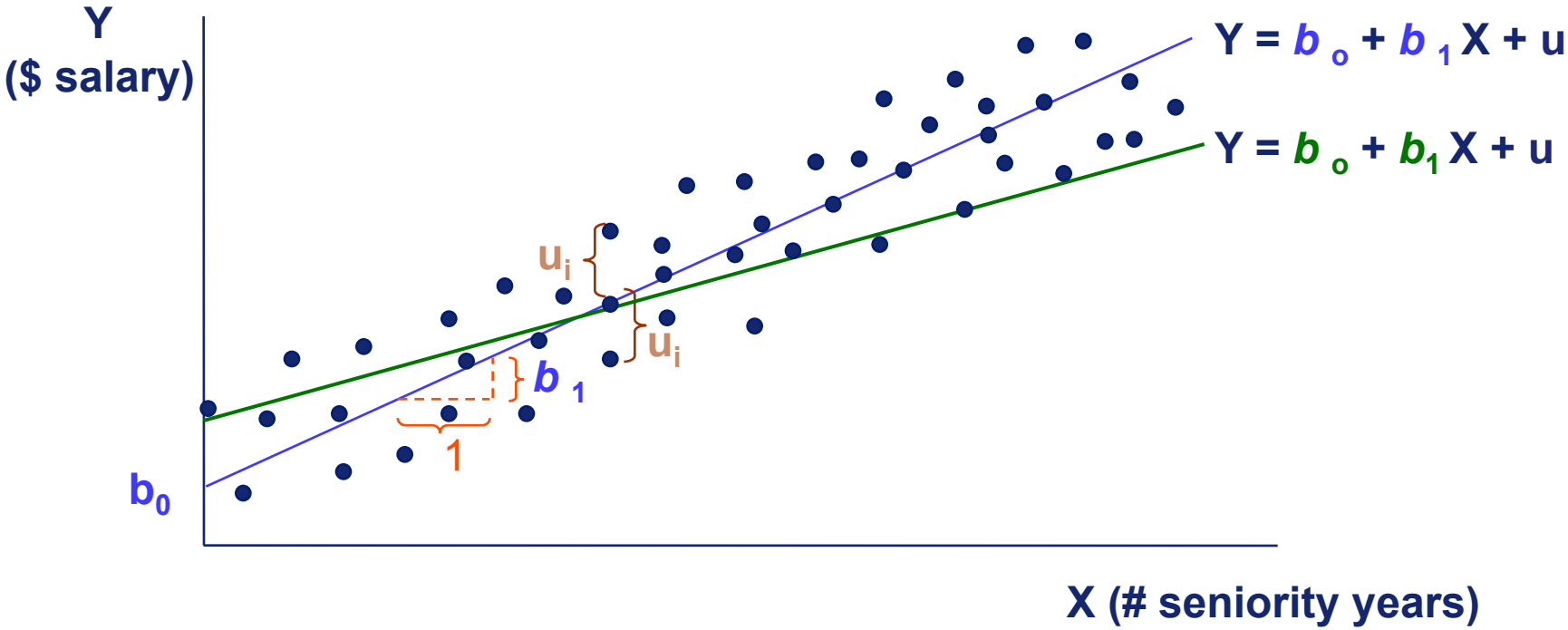


- The blue points illustrate salary at yearly intervals
- The red line illustrates a deterministic relationship
- A deterministic relationship is exact

Regression Model: A Simple Example *(cont.)*



Regression Model: A Simple Example *(cont.)*



- What line can be drawn that is closest to most of the points (blue line)?
- Regression analysis estimates b_0 and b_1 by "minimizing" the square distance between the line and the scatter points ("unexplained residuals" = u_i)
- Since not all the observations are on the line, Adj-R^2 <100%
- The blue line has a higher Adj-R^2 than the green line

Common Uses of Regression

- **Antitrust damages – calculating price effects**
  - A regression is a "generally accepted method for proving antitrust damages."[1]
- **Securities damages – event studies**
  - "An event study is a statistical regression analysis that examines the effect of an event on dependent variable, such as a corporation's stock [and] … is an accepted method for the evaluation of materiality damages" in a securities fraud litigation.[2]
- **Employment economics – discrimination claims**
  - "[I]t is clear that a regression analysis … may serve to prove a plaintiff's [racial discrimination] case."[3]
- **Other commercial disputes**
  - "[M]ultiple regression analysis … has often been used to calculate lost profits."[4]

Regression Statistics as Evidence

- **Use of data to support theories or estimate unknown values:**
  - Was the price of a stock sensitive to misrepresentation by firm management?
    - If so, by how much?
  - Have anticompetitive acts increased prices?
    - If so, by how much?
- **Financial / economic relationships can be described by a regression equation**
- **Unknown values can be predicted using known values**
- **Accuracy and reliability of results can be tested**
- **But: Results are only as good as the data and the model**

Regression Models: The Basics

Basic linear model: $Y_i = a + b_i * X_i + u_i$

a = the intercept – the value of Y when X is zero

Y = the **dependent** variable – what is being explained

X = the **independent** variables – what does the explaining

b = the coefficients – what we will estimate

u = the disturbance term – unexplained randomness

- **Salary (Y) is partially explained by**
  - Seniority (X)
  - Other factors
  - Randomness (u)

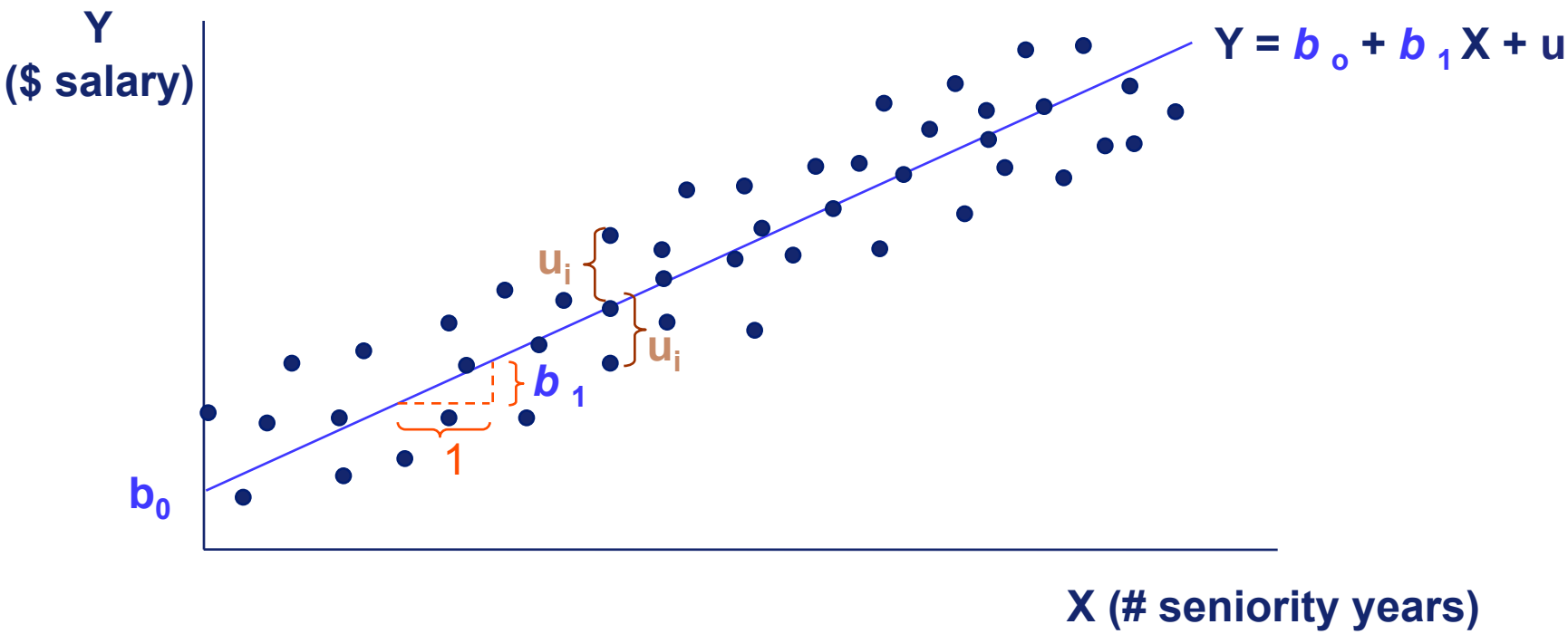
Interpreting Regression Output

- **R^2**
  - Measures how well the blue line traces the scatter of points. The percentage of variation in the dependent variable explained by the relationship established with the independent variables. "Adjusted R^2" (another statistic) used to remove apparent "explanatory" effects of additional variables
  - The higher the R^2 the stronger the correlation
- **Coefficients**
  - Measure the relationship between the dependent variable and the independent variable. Look at the sign and the size
  - Do they make sense (conform with theory)?
- **Standard error**
  - Measures the precision with which the independent variable predicts the dependent variable. It measures the spread of the coefficient estimates that would be generated by repeated sampling

Interpreting Regression Output *(cont.)*

- **T-statistic**
  - Coefficient divided by standard error
  - Measures the "statistical significance" of coefficient
  - A 95% confidence interval for each coefficient estimate can be calculated based on its standard error
  - At least one court has insisted upon a 95% degree of confidence rejecting results based upon a 85% degree of confidence.[5]
- **P-value**
  - "Significance level" of the t-statistic
  - Shows the probability of obtaining the estimated β if the true coefficient is, in fact, zero

Regression Model: A Simple Example *(cont.)*



- What line can be drawn that is closest to most of the points (blue line)?
- Regression analysis estimates b_0 and b_1 by "minimizing" the square distance between the line and the scatter points ("unexplained residuals" = u_i)
- Since not all the observations are on the line, Adj-R^2 <100%
- The blue line has a higher Adj-R^2 than the green line

Interpreting Regression Output *(cont.)*

Secretary in the Business Services Industry

Dependant Variable: Yearly Salary

Variable	Coef. [A]	Std Err. [B]	t stat [C]=[A]/[B]	P>|t| [D]
Intercept	15527	2402	6.47	<.0001
age	264	63	4.17	<.0001
Number of Observations	196			
R-Squared	0.03			

Sample year: 2002

Data source: Current Population Survey, Bureau of Labor Statistics

Interpreting Regression Output *(cont.)*

Dependant Variable: Yearly Salary

Variable	Coef.	Std Err.	t stat	P>|t|
Intercept	-12139	12058	-1.010	0.316
Female (ref: Male)	-1749	7119	-0.250	0.806
Black (Ref: Caucasian)	-1905	3560	-0.540	0.593
American Indian (Ref: Caucasian)	812	4191	0.190	0.847
Asian (Ref: Caucasian)	-4573	3796	-1.200	0.231
Unionized (ref: non union)	3613	5013	0.720	0.472
Reference group: <10th Grade				
10th Grade	4360	9966	0.440	0.662
11th Grade	2324	1775	1.310	0.193
12th Grade no diploma	4250	8881	0.480	0.633
High school graduate	12479	4458	2.800	0.006
Some college but no degree	17039	5703	2.990	0.003
Associate degree (Vocational)	14439	5575	2.590	0.011
Associate degree (Academic)	27378	9266	2.950	0.004
College degree	23511	5767	4.080	0.000
Master's degree	23562	7322	3.220	0.002
age	240	86	2.810	0.006

n	196
R-Squared	0.33

What Regression Analysis Cannot Do:

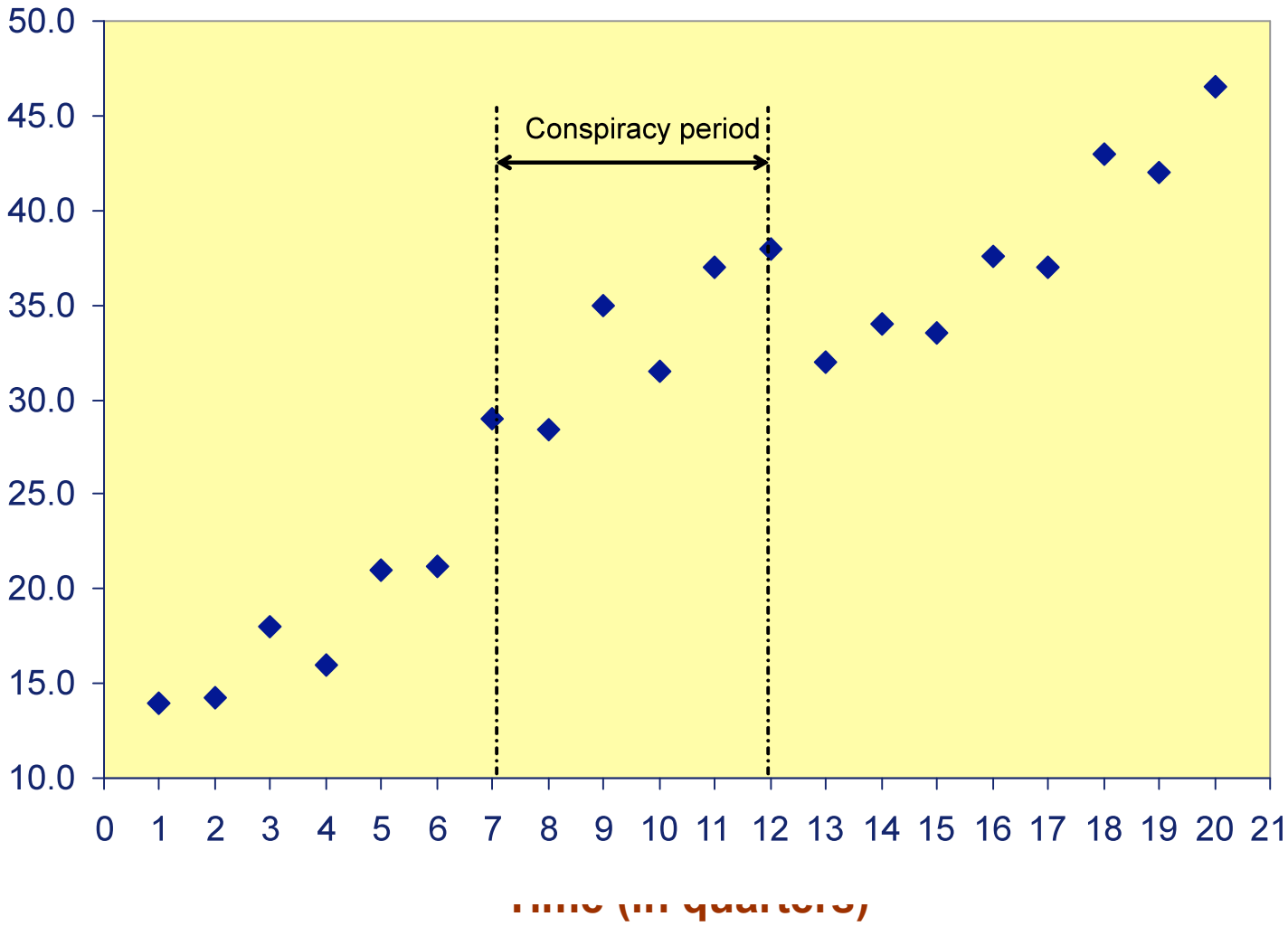
- **Regression analysis cannot "prove" anything with certainty**
  - We are dealing with stochastic relationships
  - We can show whether a relationship is sufficiently likely (at a specified level of confidence)
- **Regression analysis cannot independently prove causality**
  - Need theory to help here
  - Regression can provide empirical support for a theory of causality

Common Pitfalls of Regression Analysis

- **Data may be incomplete or biased**
  - Courts generally admit models with incomplete data,[6] but often find that they lack credibility.[7]
  - Data are contaminated (no "clean" period). Courts often exclude models based upon biased data.[8]
- **Methods used may be inappropriate (model is mis-specified)**
  - The model omits an important explanatory variable
    - Courts generally do not exclude models with this flaw,[9] unless the movant can demonstrate that the omitted variable(s) affected the outcome of the analysis.[10] In antitrust cases, however, Judge Posner as well as the Eighth Circuit en banc have rejected econometric models because they failed to account for non-conspiratorial causes of price changes. These opinions did not explain the showing that was required.[11]
  - Model fails to account for a change over time in the relationship between the dependent and independent variables
    - At least one court has discredited an analysis with this flaw where a party demonstrated that the flaw affected the outcome.[12] Another court noted, however, that no "authority would require use of a Chow test in this case."[13]
- **Results are misinterpreted**
- **Analysis is not properly presented to the trier of fact**

Example 1: Antitrust Case Where Price Fixing is Alleged

Claim: Collusion Between 7th and 12th Quarter Increased Prices



Data

Time (Quarters)	Price ($ 000)	Conspiracy Period	Conspiracy Variable
1	14.0	No	0
2	14.2	No	0
3	18.0	No	0
4	16.0	No	0
5	21.0	No	0
6	21.2	No	0
7	29.0	Yes	1
8	28.4	Yes	1
9	35.0	Yes	1
10	31.5	Yes	1
11	37.0	Yes	1
12	38.0	Yes	1
13	32.0	No	0
14	34.0	No	0
15	33.5	No	0
16	37.6	No	0
17	37.0	No	0
18	43.0	No	0
19	42.0	No	0
20	46.5	No	0

Measuring an Alleged Anticompetitive Price Increase

- **Example: Anticompetitive price increase**
- **Hypothetical regression results (with no conspiration dummy included):**
  - Price = 13.9 + 1.6 * Time + u

    (8.7) **(11.9)**

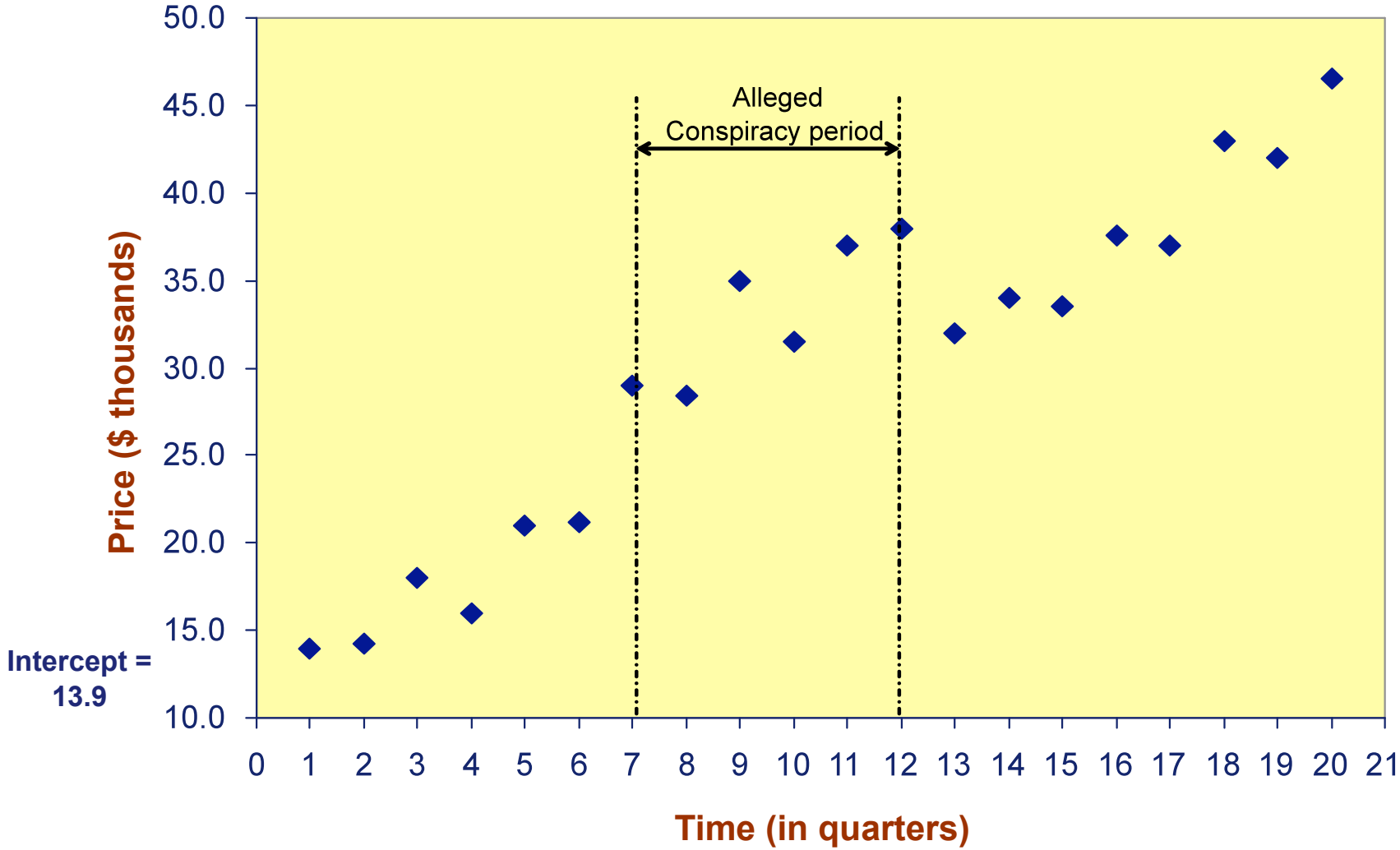
    $R^2 = 0.89$

    Sample size = 20

Note: T statistics are reported in parentheses

Measuring an Alleged Anticompetitive Price Increase *(cont.)*

Estimated line when no conspiration period dummy variable is included as a regressor



Measuring an Alleged Anticompetitive Price Increase *(cont.)*

- **Example: Anticompetitive price increase**
- **Hypothetical regression results: During the anticompetitive pricing years:**
  - (Estimated) Price = 11.4 + 1.6 * Time + 6.2 * (Conspiracy period)

    (12.9) **(23.8)** **(7.2)**

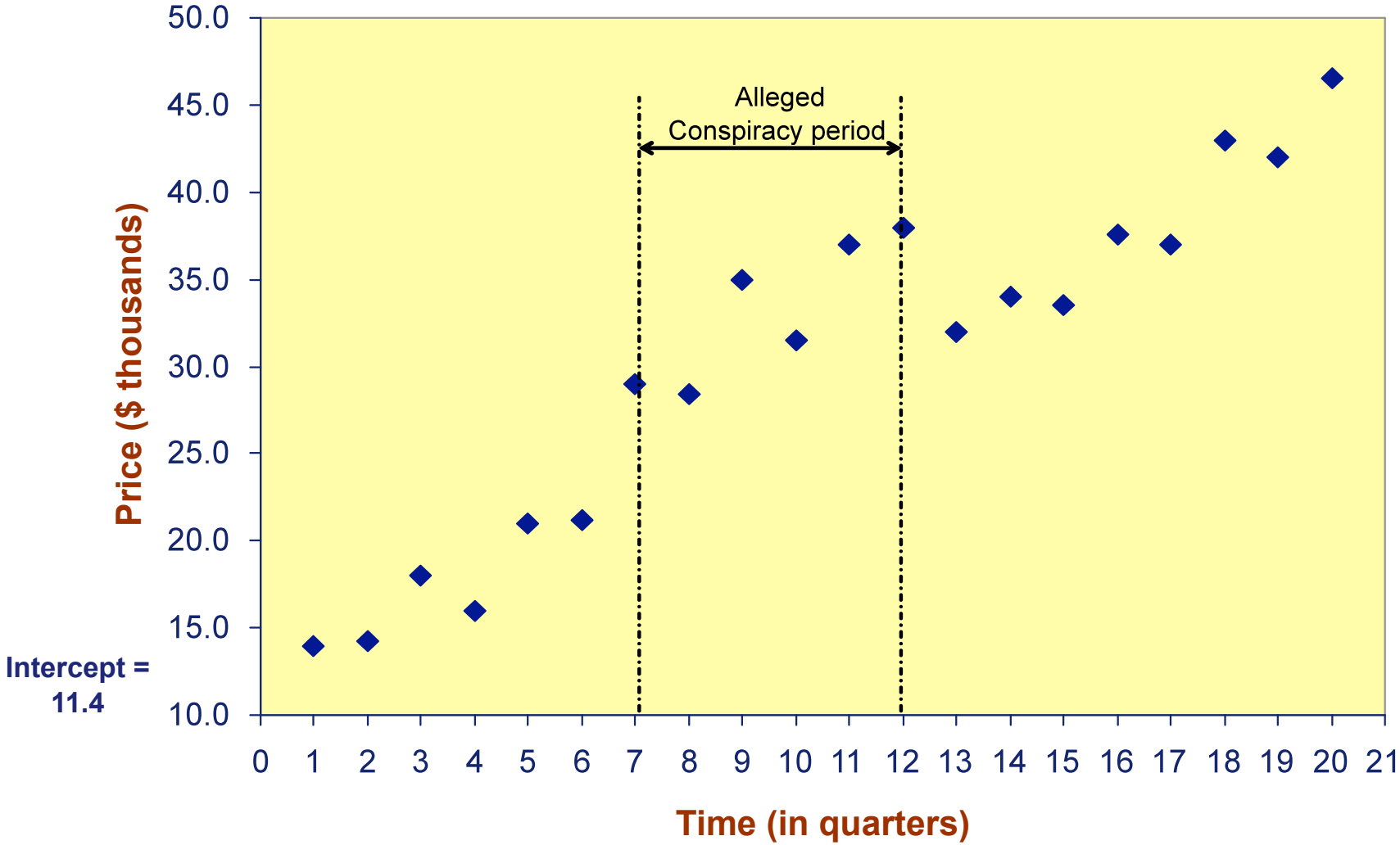
    $R^2 = 0.97$

    Sample size = 20

Note: T statistics are reported in parentheses

Measuring an Alleged Anticompetitive Price Increase *(cont.)*

If a dummy variable for the conspiracy period is statistically significant, regression may be consistent with collusion



Measuring an Alleged Anticompetitive Price Increase *(cont.)*

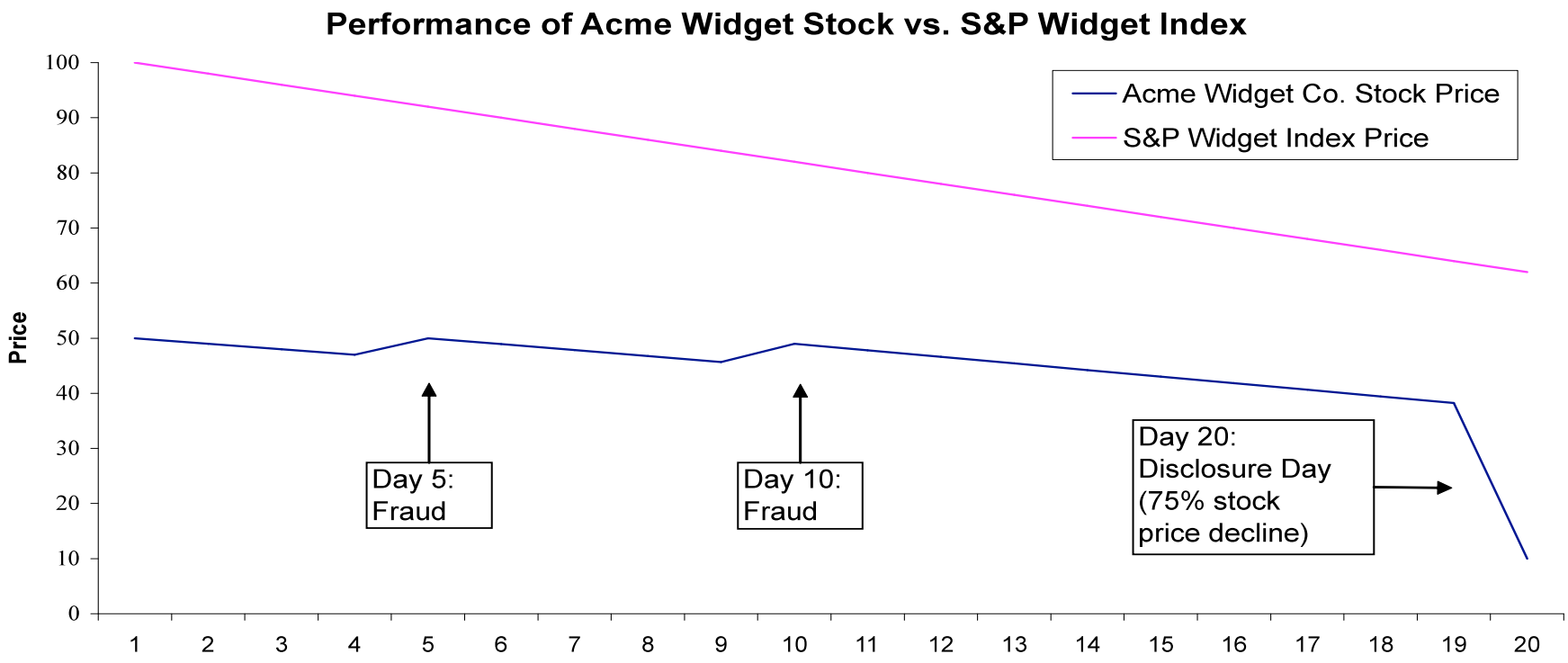
Have the following issues been considered?

- **Market factors:**
  - Input prices. One court observed that "[w]ith respect to changes in price, changes in cost are one obvious explanatory variable."[14]
  - Changes in demand. One court observed that "the effect of … demand … might impact price levels."[15]
  - Capacity constraints. That same court also observed that "the effect of supply … might impact price levels."[16]
  - Level of competition and market structure. Judge Posner held that where the expert "failed to correct for the effect of market share on [plaintiff's] prices, … no reasonable jury could estimate damages."[17]
- **Other issues**
  - Is the model used found in the literature?
  - Has the model been commented on in legal cases?
  - Would the approach and interpretation pass peer-review scrutiny?
  - Has the expert ever used this approach in a refereed publication?

Example 2: 10b-5 Case Where Fraud is Alleged

Assessing Securities Fraud

- Acme Widget Co. is alleged to have made fraudulent announcements on day 5 and day 10 of the class period
- In the absence of fraud, Acme Widget Co.'s stock price performance may resemble an industry index such as the S&P Widget Index
- When the fraud is disclosed on day 20, Acme's stock price drops 75%



Assessing Securities Fraud *(cont.)*

- **In the absence of fraud, price behavior during the class period could be modeled using regression analysis as follows:**

  $$\text{Acme stock return}_t = a + b^*(\text{Industry index return})_t + u_t$$

- **If fraud-related announcements were made on days 5, 10, and 20, then we can test the effect of these announcements using the following model:**

  $$\text{Acme stock return}_t = a + b^*(\text{Industry index return})^t + d_1^*(\text{day } 5) + d_2^*(\text{day } 10) + d_3^*(\text{day } 20) + u_t$$

  - The coefficients d_1 through d_3 will be informative on the likelihood that the announcement had an effect on the stock price

Assessing Securities Fraud (cont.)

- **In the presence of material fraud-related announcements, we would expect d_1 , d_2 , and d_3 to have values that reflect the effect of those announcements after accounting for the industry trend:**

$$\text{Acme stock return}_t = 0 + 1*(\text{Industry index return})_t + 0.09*(\text{day } 5) + 0.10*(\text{day } 10) - 0.70*(\text{day } 20) + u_t$$

- **In this example, Acme Widget Co.'s daily returns follow the industry index daily returns with a 1-to-1 relationship**
- **The disclosure of fraud on day 20 accounts for 70% of Acme's stock price decline on that day**
- **The remaining 5% drop is the industry effect**

Assessing Securities Fraud *(cont.)*

- **In the absence of announcements, we would expect $d_1=d_2=d_3=0$. The stock but-for price can be calculated in the following way:**

$$\text{Acme stock return}_t = a + b*(\text{Industry index return})_t + 0*(\text{day } 5) + 0*(\text{day } 10) + 0*(\text{day } 20) + u_t$$

- **Given the actual and but-for prices, it is possible to calculate damages**

Assessing Securities Fraud *(cont.)*

- **In the real world, using regression in securities fraud event studies is more complex:**
  - Stock price behavior may reflect general market trends as well as industry trends
    - Often multiple indices are used as independent variables in the regression (e.g., S&P 500 market index and Healthcare industry index)
  - If the class period is “tainted” with many fraud-related days, a separate “clean” period may be used for estimating the regression
  - Other non-fraud company announcements may be driving the stock price on fraud-related days
    - The effect of fraud must be separated from the effect of non-fraud-related news
  - Stock price data is noisy
    - The relationship of the company with the index and the effect of news on given days should be tested for statistical significance

Reviewing an Opposing Expert's Analysis

Reviewing an Opposing Expert's Analysis

- **Data:**
  - Have they been provided?
  - Can they be duplicated (if public)?
  - What cleaning process was followed (What is the sample? What errors were corrected? How were extreme values treated?)
  - Are alternative sources available?
  - Is sampling choice appropriate?
- **Analysis:**
  - Does the model depart from established peer-reviewed models, or experts own position in other cases or publications?
  - Does any case law comment on the model used?
  - Are there econometric issues to be aware of?
  - Are the results sensitive to alternative specifications?
  - Are results inconsistent with testimony, documents, or common sense?

Conclusion

- Regressions can help identify a relationship in a variety of litigations such as securities, antitrust, discrimination, commercial damages, and intellectual property
- Regression analysis alone cannot establish causation
- A regression expresses the stochastic relationship between two or more variables
- A good theoretical framework is useful to develop a robust statistical analysis
- Regressions are only as good as the data that are used in the analysis
- Regressions analysis requires great care, particularly when specifying the relationship that is the subject of study

ANALYSIS GROUP
ECONOMIC, FINANCIAL and STRATEGY CONSULTANTS

WHITE & CASE

Econometric Analysis in Litigation

February 21, 2006

Pierre Y. Cremieux, Ph.D., pcremieux@analysisgroup.com
Matthew S. Wild, Esq., mwild@whitecase.com