MORS

SCEA
The Society of Cost Estimating and Analysis

Analogies: Techniques for Adjusting Them

R. L. Coleman, J. R. Summerville, TASC; S. S. Gupta, IC CAIG

ASC/Industry Cost & Schedule Workshop Spring 2004 – Destin FL
21 April 2004

SCEA 2004 – Manhattan Beach, CA
17 June 2004

72nd MORSS– Monterey, CA
22 June 2004

Outline

- **Background**
- **The current method**
- **Two new methods**
  - Borrowed slope
  - Relational correlation
- **Conclusion**

richard.coleman@ngc.com (703) 402-3702

Background

richard.coleman@ngc.com (703) 402-3702

Discussion

- **Considerable attention is devoted to techniques in the development of Cost Estimating Relationships (CERs) for parametric estimating**
  - Research on CERs
  - Methods for calibrating
- **Considerable expertise is to be found in buildup techniques**
  - Many Original Equipment Manufacturers (OEMs) have large cost shops which practice buildup
- **Analogy, on the other hand, has been given little attention**
- **Next, some basic definitions …**

richard.coleman@ngc.com (703) 402-3702

Definitions

- **Parametric Estimates: Estimates made by developing statistical "Cost Estimating Relationships" (CERs) based on one or more parameter and cost**
  - Estimates involving parameters but not based on statistical analysis are more properly called either "adjusted analogies" or "adjusted buildups"
- **Analogies: Estimation by assuming that the costs of a new system will be equal to (or similar to) the costs of a system that is similar**
  - "Adjustments" are almost always made
- **Buildups: Physical Bill of Materials (BOMs) and CAD-generated material lists and the like**
  - We do not mean "buildups" consisting entirely of Staffing levels*Duration. Such estimating techniques are little more than "engineering judgment" in fine detail
  - Buildups often include "adjustments" to allow for size differences
- **Composite methods: A method that involves at least two of the three other types**
- **Adjustments: Scaling of a cost by some physical, performance, or other such attribute**
  - Scaling is usually directly proportional to the attribute
  - Scaling parameters are usually countable or measurable and intuitively tied to cost

richard.coleman@ngc.com (703) 402-3702

The Current Method

richard.coleman@ngc.com (703) 402-3702

Typical Adjustments – By Ratio

- **Adjustments, in the analogy or buildup method, typically rely on an "obvious" characteristic**
  - The characteristic is most often weight
  - Sometimes weight of the new system is not known, and so another characteristic is used (often as a proxy for weight)
  - Sometimes a characteristic such as bore diameter of a gun is used
- **Usually the ratio of the values of the characteristic in the new system to the value in the old system is multiplied by the cost of the old system**
  - Sometimes called "j-ing up the estimate"
- **Sometimes the characteristic is transformed in a way that is thought to make it proportional to weight**
  - E.g., the bore diameter of a gun, is cubed
  - In these cases, there may be a presumed relationship to weight,

richard.coleman@ngc.com (703) 402-3702

Implications of the Current Method

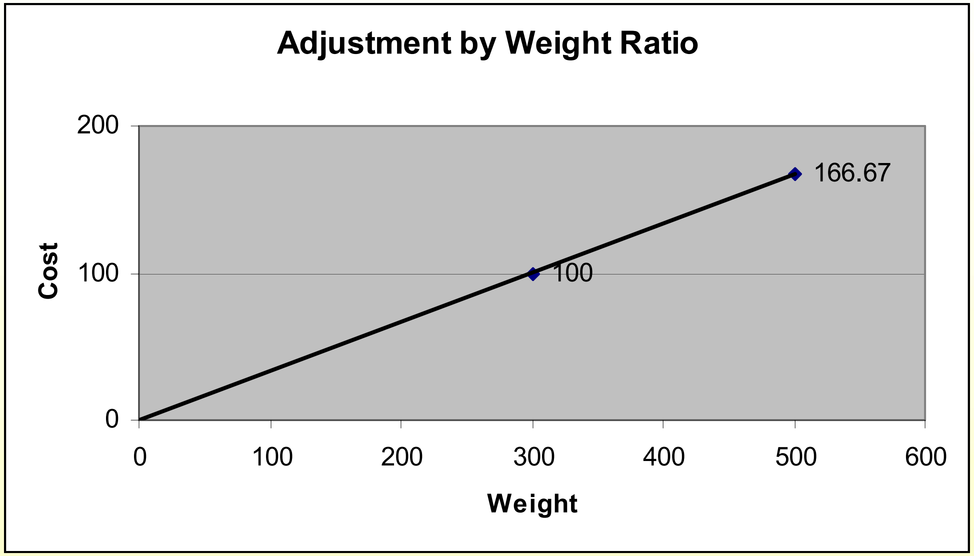
- **An example adjustment by ratio is:**
  – The analogy weighs 300 tons and costs $100M
  – The new system weighs 500 tons and so is assumed to cost (500/300)*$100M = $166.67M
- **This is a typical and familiar adjustment**
  – What is its implication?
  – Should we be inclined to believe it?
  – Is it in accord with what we believe?

  … let's look at a graph to see what it implies … there is a surprise there for most of us … but first, force yourself to <u>predict what the line between the analogy and the prediction looks like</u> … where does it cross the y axis?

richard.coleman@ngc.com (703) 402-3702

Adjustment by Ratio – The Graph

- **The below graph shows the previous adjustment**
  - The analogy weighs 300 tons and costs $100M
  - The new system weighs 500 tons and is assumed to cost (500/300)*$100M = $166.67M
  - Note that the line through the 2 points passes through the origin



<u>Important Observation:</u>
Straight adjustments by ratios *always* pass through the origin! Most observers fail to predict that, even though it is straightforward to show that it must.
<u>Important Question:</u>
Is this reasonable?

richard.coleman@ngc.com (703) 402-3702

Adjustment by Ratio – Reasonable?

- **The y-intercept is a litmus test among cost estimators. There are about three schools of thought:**
    1. CERs “should” pass through the origin
    2. CERs which do not pass through the origin <u>must</u> have an explicable y-intercept
    3. CERs must be statistically derived, and if done properly, the y-intercept is just “what it is”

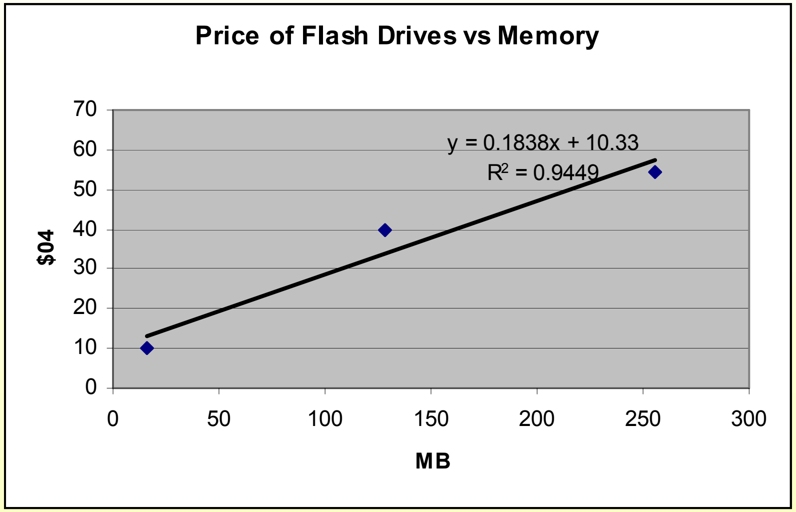
- **We’ll discuss each <u>briefly</u> and then assume you are of school 2 or 3**

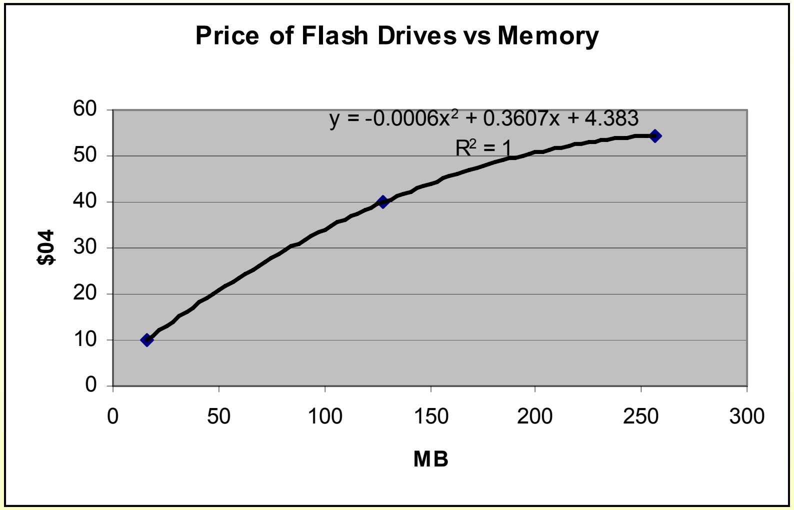
Warnings:
1- Almost anyone is from one of these schools of thought at heart. The writers are no exception.
2- The gulf between these schools is wide.

richard.coleman@ngc.com (703) 402-3702

"Y-Intercepts must pass through the origin"

- **Typical arguments**
  - "If I spend no money, I get no product"
- **Pros:**
  - <u>Sounds</u> good
- **Cons:**
  - Doesn't seem to <u>match the data</u>. E. g., the price of FlashDrives:





richard.coleman@ngc.com (703) 402-3702

"Y-Intercepts must make sense"

- **Typical arguments**
  - "There must be physics-based arguments for CERs"
- **Pros:**
  - Helpful to think about it, within reason
- **Cons:**
  - If practiced to the extreme, good CERs can be rejected just because we do not yet understand them
  - Engineers, who hate cost estimation, can usually talk the analyst to a full stop

richard.coleman@ngc.com (703) 402-3702

"The Y-Intercepts is just what it is"

- **Typical arguments**
  - We are not trying to predict the y-intercept. We are trying to predict the cost of systems of non-zero size.
    - We should take the best advice the data can give us
    - We should extrapolate as little as we can
  - If the data show that the y-intercept is non-zero, we should not reject a CER just because we do not know why
    - Galileo believed the data, even absent a theory of gravity. It took centuries before Isaac Newton knew why – but Isaac Newton wouldn't even have wondered without Galileo showing that there was an explanation missing.
  - This approach is what the practice of statistics currently recommends
- **Pros:**
  - Any existing system (i. e., one of the data points underlying the CER) is well-predicted
- **Cons**
  - If the analysis is not well done, there may be a better CER

richard.coleman@ngc.com (703) 402-3702

Proposal - Two New Methods

- **Borrowed slope[1] – a variant of the methods for calibrating CERs**
  – Adjust a "trusted analogy" by a "trusted slope"
- **Relational Correlation[2] – taking advantage of the geometry of regression**
  – Adjust a "trusted analogy" by a "best guess slope"

1 *A Framework for Costing in a CAIV Environment,* R. L. Coleman, TASC; D. Mannarelli, Navy ARO, ASNE 1996, ADoDCAS 1996

2 *Relational Correlation, What to do when Functional Correlation is Impossible,* ISPA/SCEA 2001, R.L. Coleman, J.R. Summerville, M.E. Dameron, C.L. Pullen; TASC, Inc., S.S.Gupta, IC CAIG

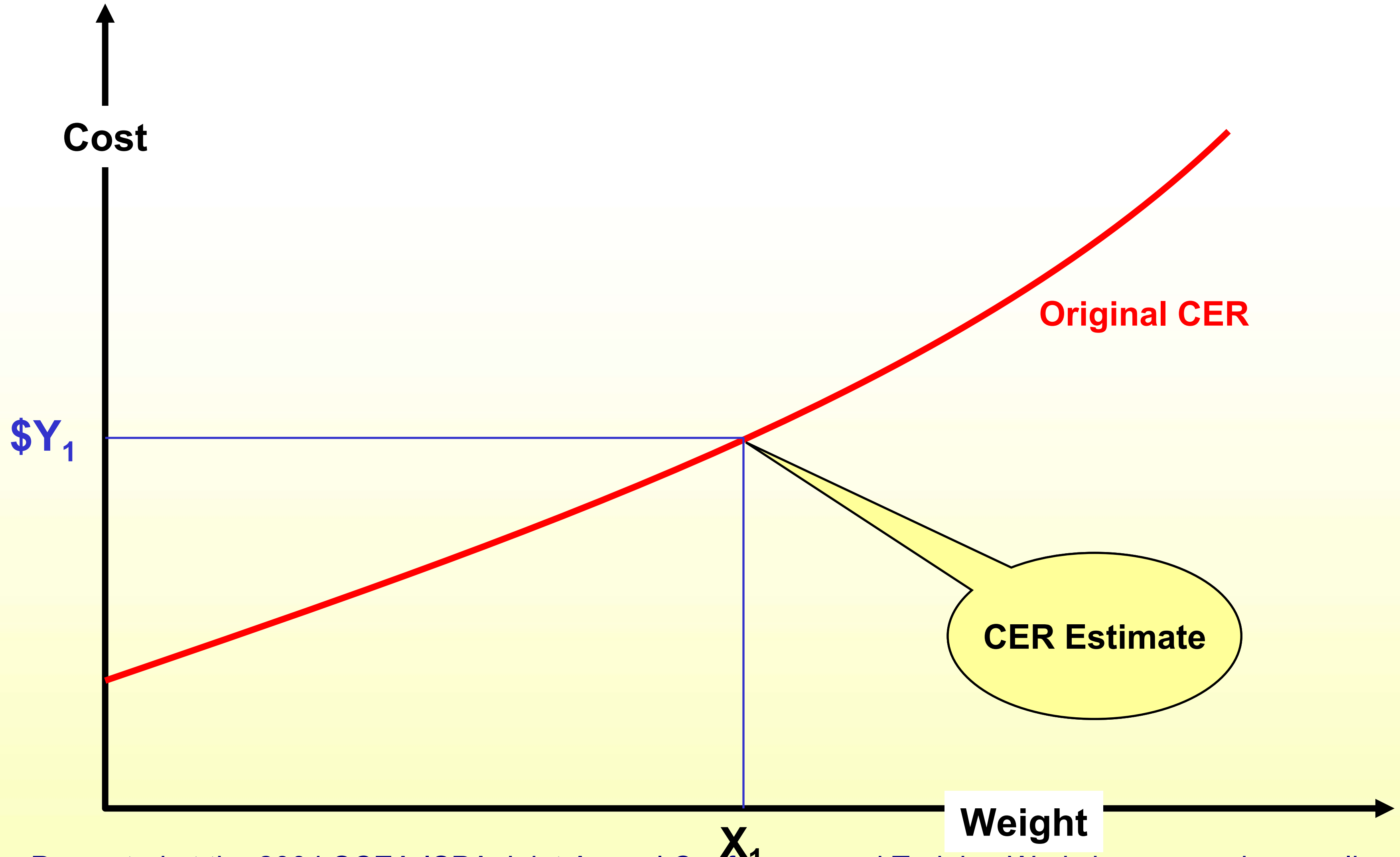
richard.coleman@ngc.com (703) 402-3702

The Borrowed Slope Method

- **Based on "calibrating a CER"**
  - A CER is adjusted to "more trusted," or industry, or company specific data by moving the slope to pass through a point or set of points
  - Picture follows
- **To adjust an analogy, do precisely the same thing**
  - Instead of believing you are adjusting a CER to specific data, think of it as departing from "the most credible point" via "the most credible slope"

richard.coleman@ngc.com (703) 402-3702

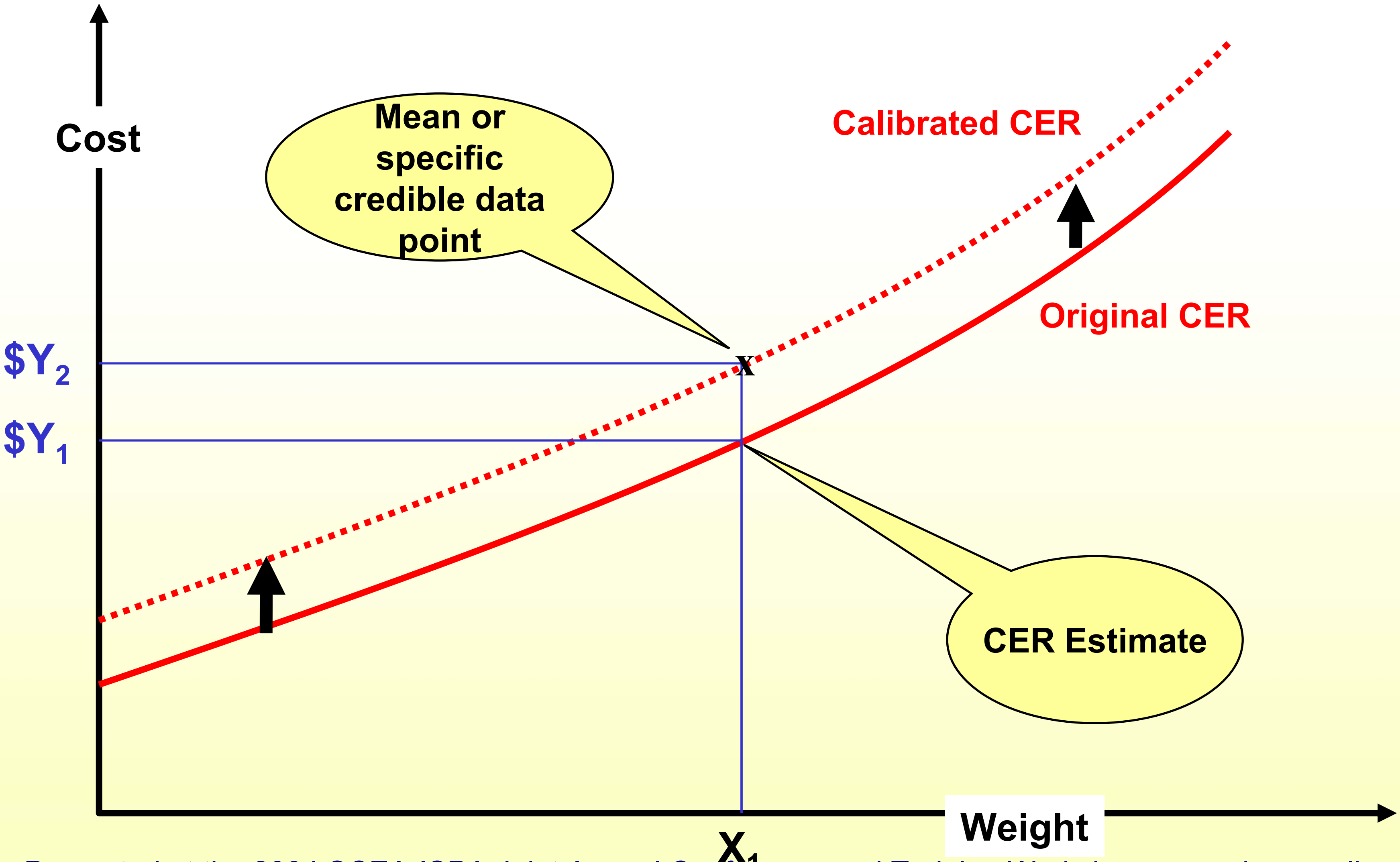
Calibrating CERs



Presented at the 2004 SCEA-ISPA Joint Annual Conference and Training Workshop - www.iceaaonline.com

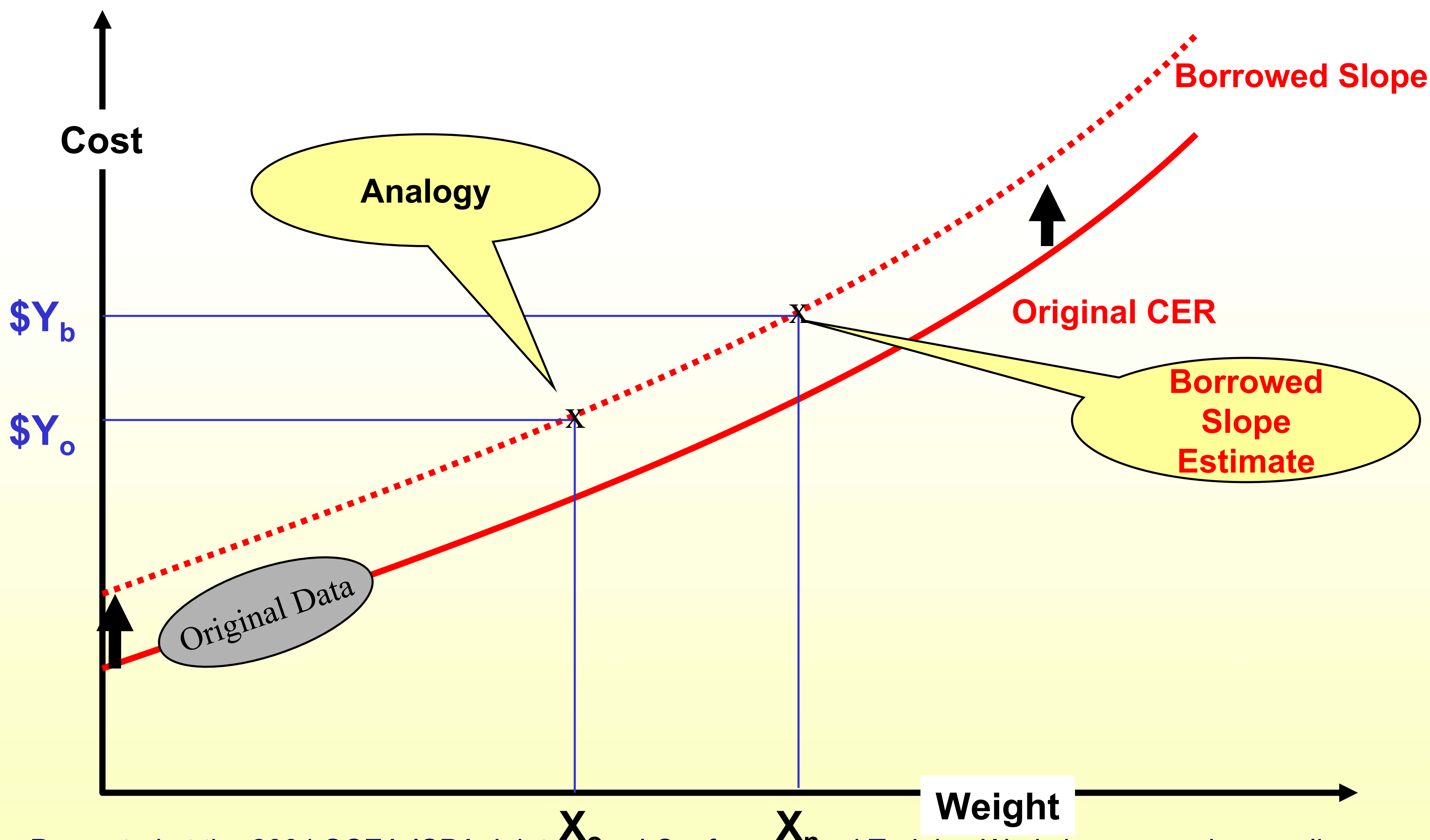
richard.coleman@ngc.com (703) 402-3702

Calibrating CERs



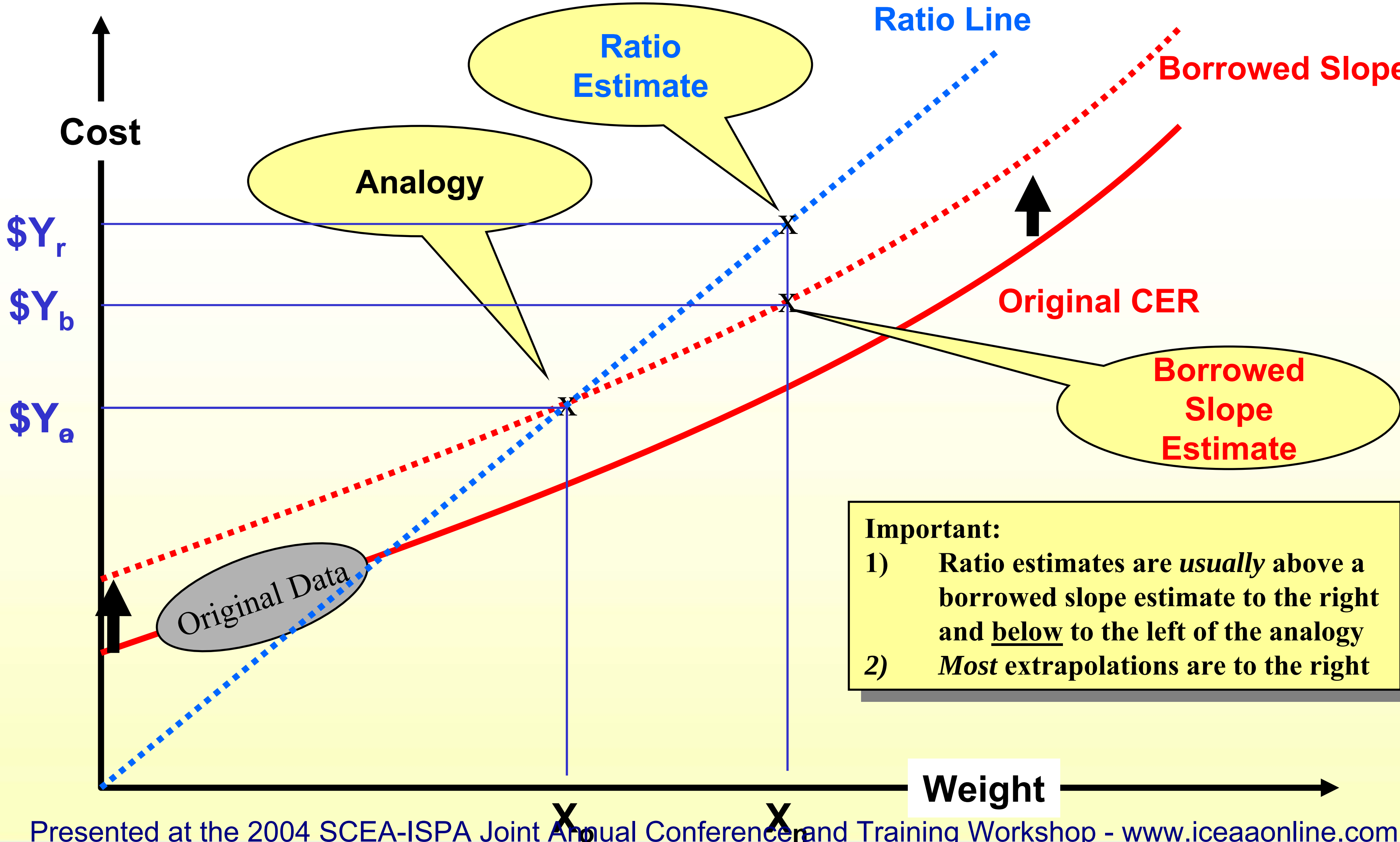
richard.coleman@ngc.com (703) 402-3702

The Borrowed Slope Estimate



richard.coleman@ngc.com (703) 402-3702

Comparison – Borrowed Slope & Ratio Estimates



Important:

1) Ratio estimates are *usually* above a borrowed slope estimate to the right and below to the left of the analogy
2) *Most* extrapolations are to the right

richard.coleman@ngc.com (703) 402-3702

Relational Correlation

- **A *much more esoteric* method is available, which borrows from**
  - Bivariate normality
  - The geometry of regression
- **This method is available when there is no “trusted slope” to borrow**

Warning: Serious math follows.

richard.coleman@ngc.com (703) 402-3702

Bivariate Normality

richard.coleman@ngc.com (703) 402-3702

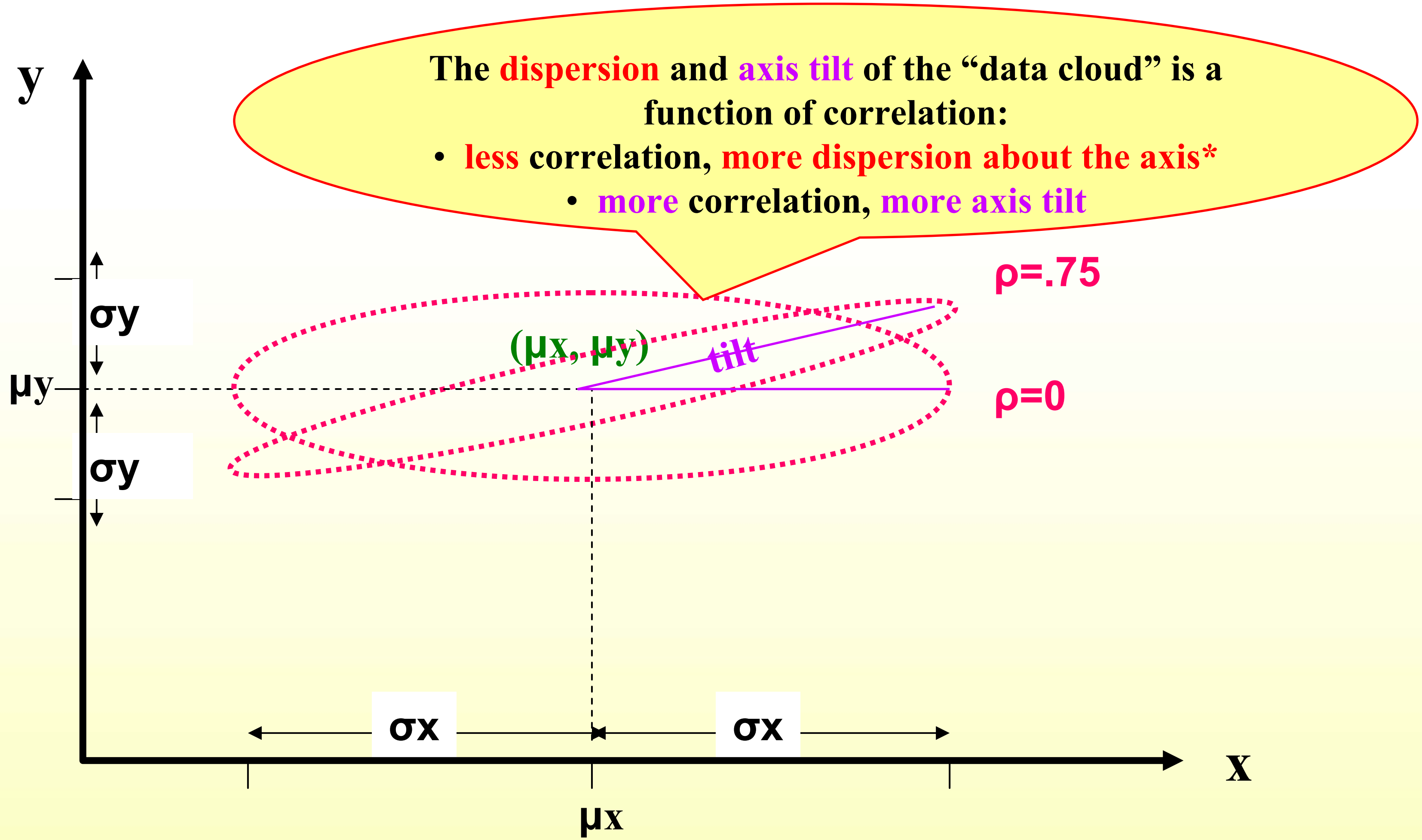
Bivariate Normality

- Suppose X and Y are distributed $N(\mu_x, \sigma_x)$ and $N(\mu_y, \sigma_y)$
- Suppose X and Y are jointly bivariate normal with correlation ρ

…. Then the graph of X and Y will appear as follows

richard.coleman@ngc.com (703) 402-3702

The Bivariate Normal



richard.coleman@ngc.com (703) 402-3702

The Geometry of Regression

richard.coleman@ngc.com (703) 402-3702

Regression

- **The below facts are known to mathematicians, but obscure, and not remembered in cost analysis …**
  - For any two jointly distributed variables, there is a regression line
  - The slope is:

$$m = \rho * (\sigma y / \sigma x)$$

  - The y intercept is:

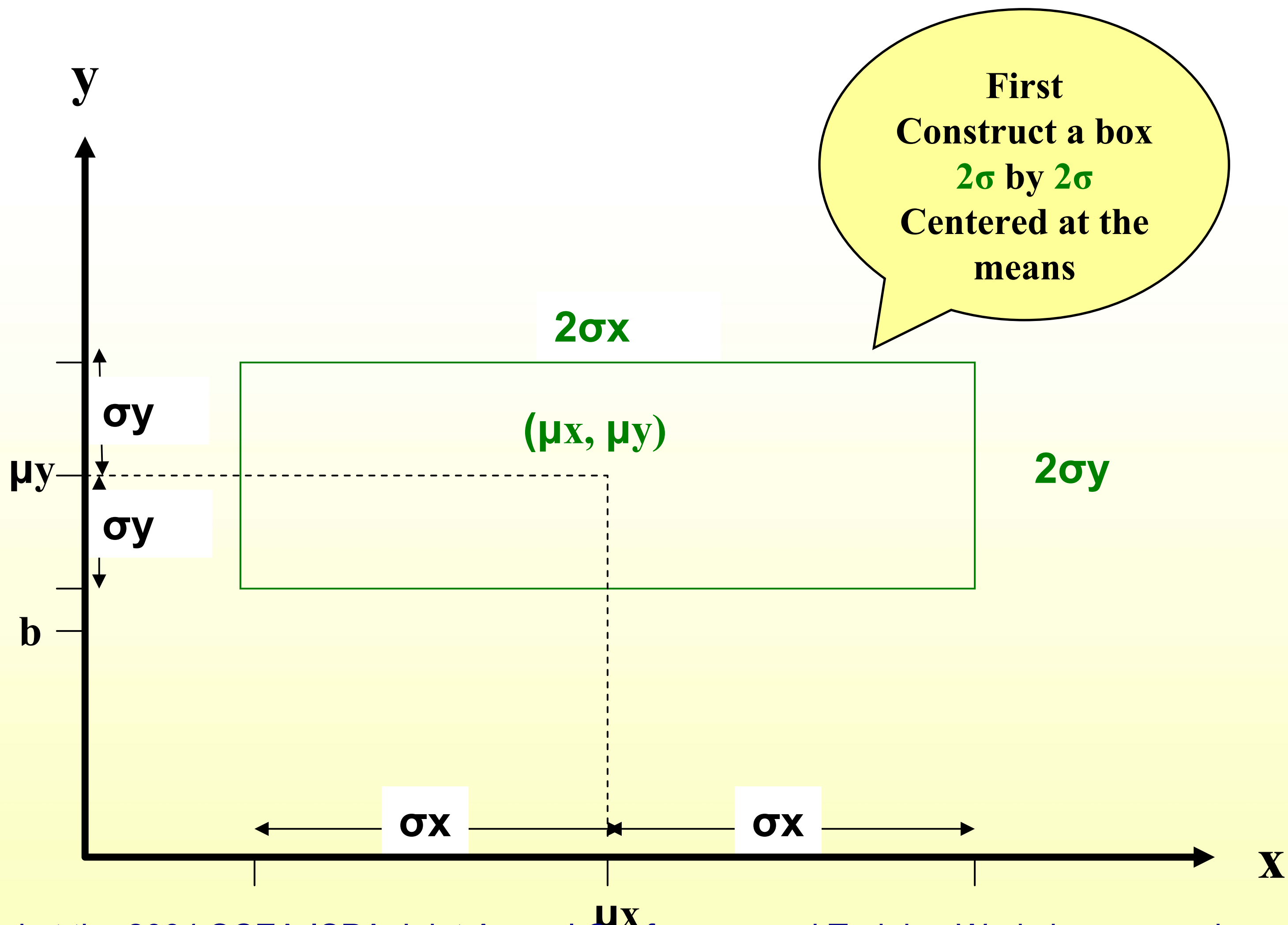
$$b = \mu y - \rho(\sigma y / \sigma x) * \mu x$$

  - If the variables are joint bivariate normal, then ρ is the correlation coefficient

Let's look at the graph…

richard.coleman@ngc.com (703) 402-3702

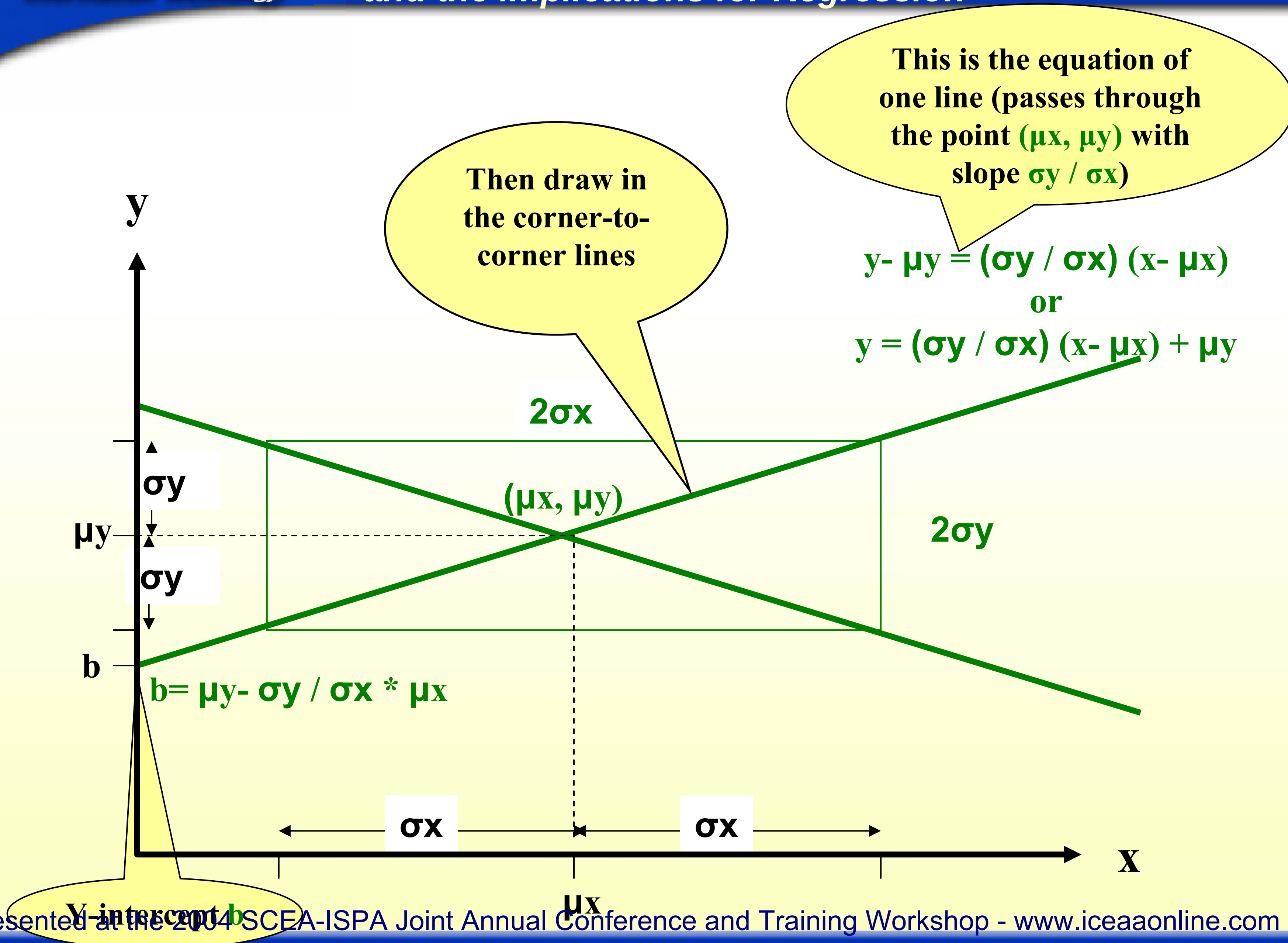
The Geometry of Bivariate Normality
and the implications for Regression



richard.coleman@ngc.com (703) 402-3702

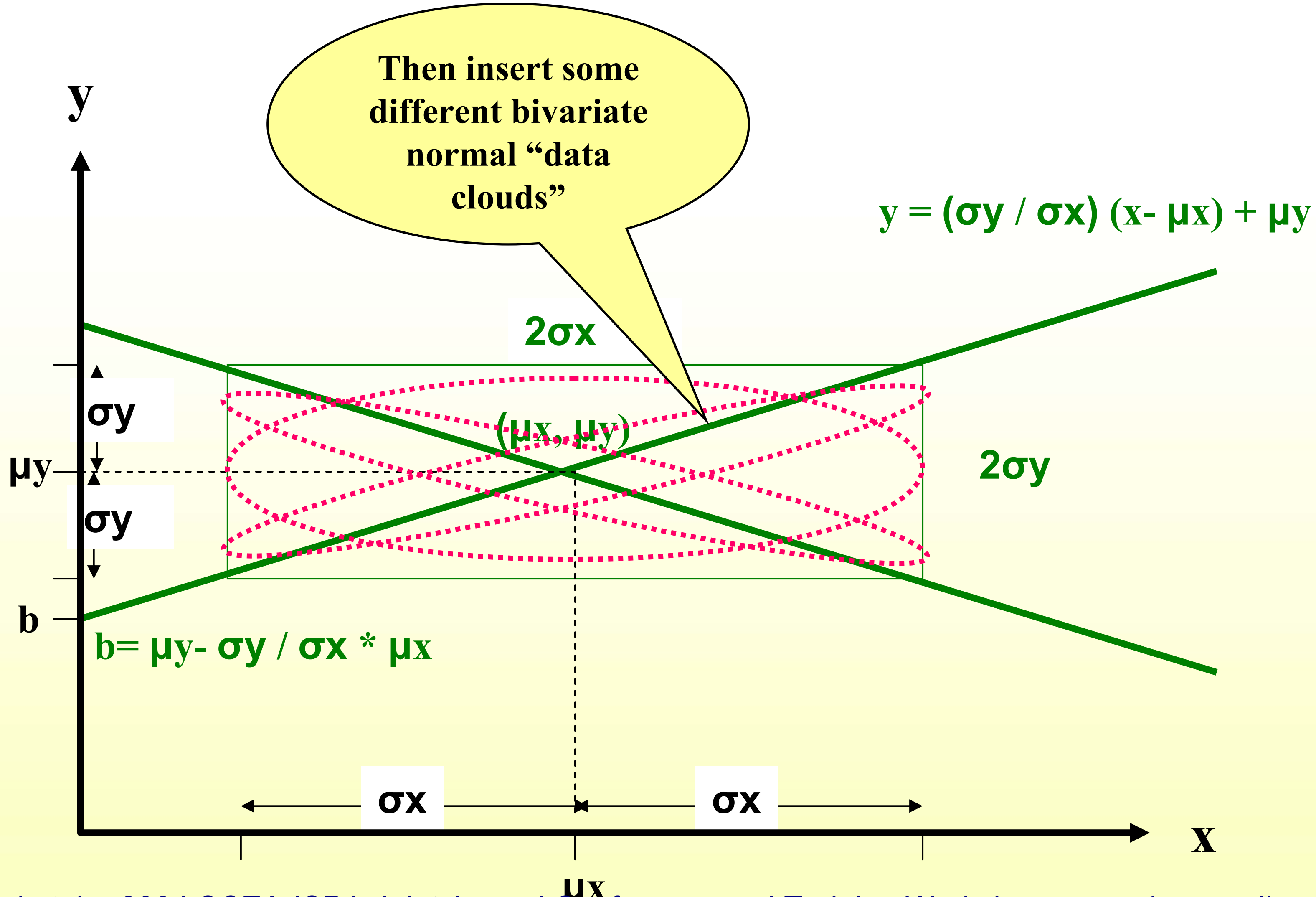
The Geometry of Bivariate Normality

and the implications for Regression



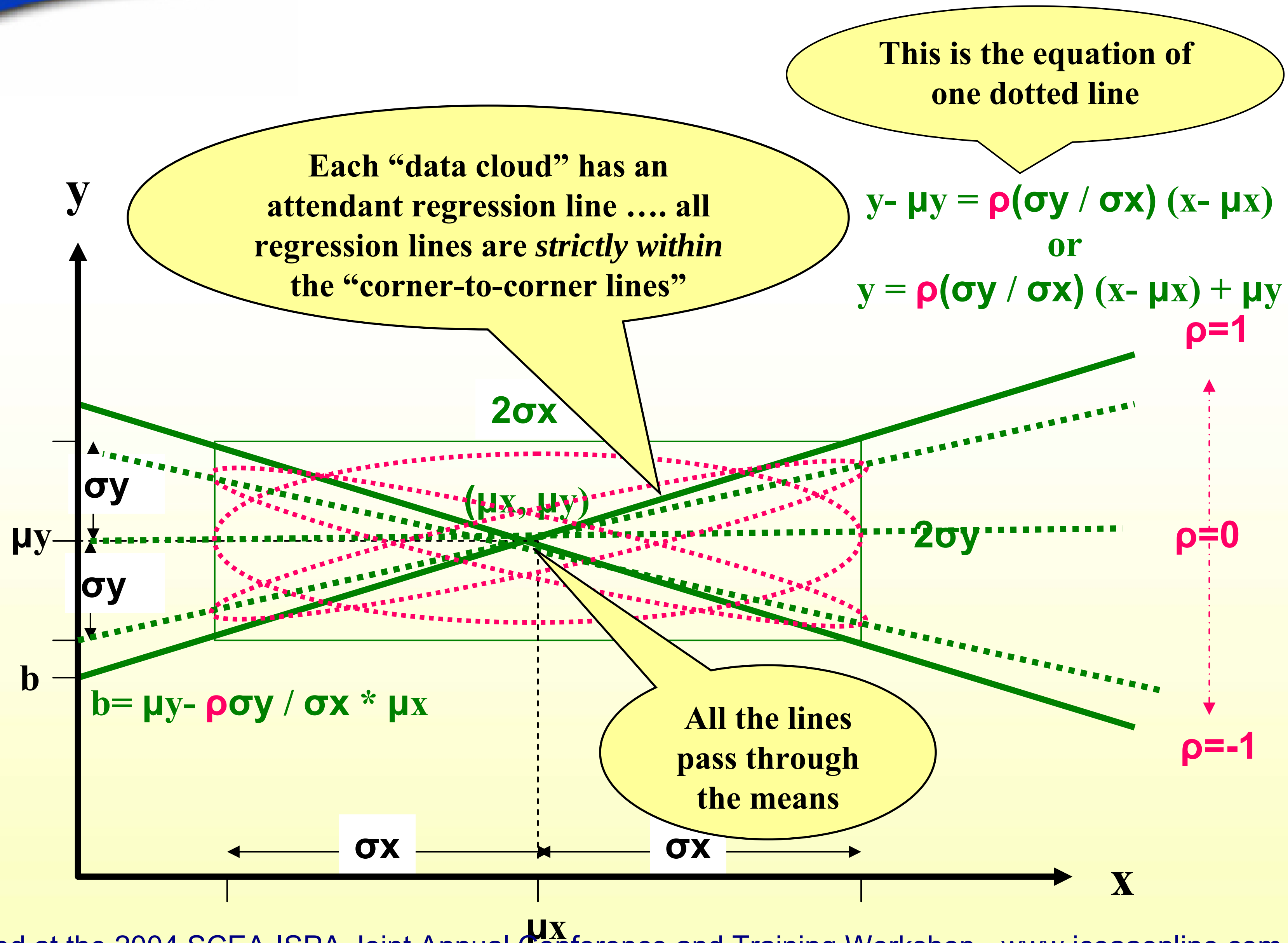
richard.coleman@ngc.com (703) 402-3702

The Geometry of Bivariate Normality
and the implications for Regression



richard.coleman@ngc.com (703) 402-3702

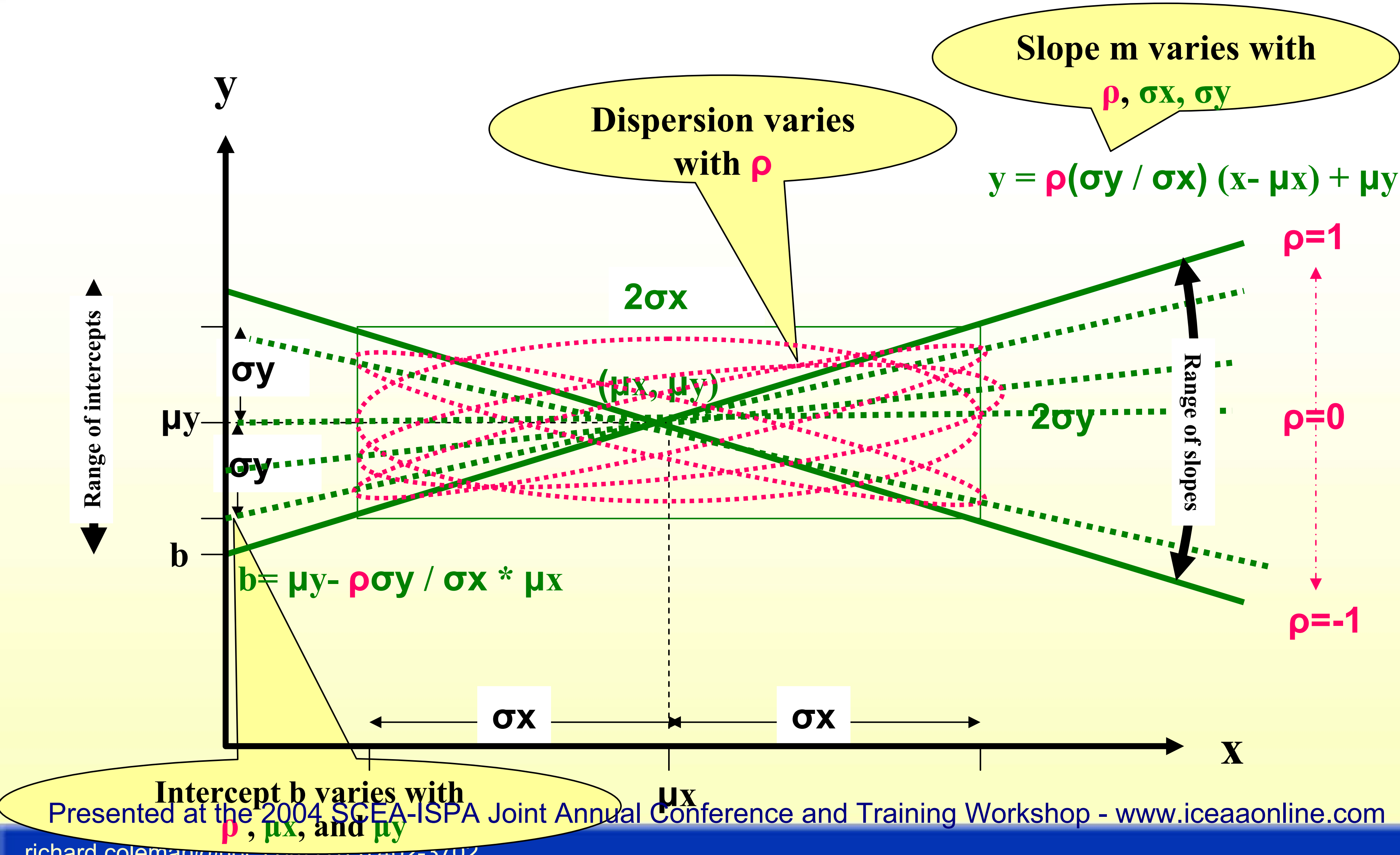
The Geometry of Bivariate Normality
and the implications for Regression



richard.coleman@ngc.com (703) 402-3702

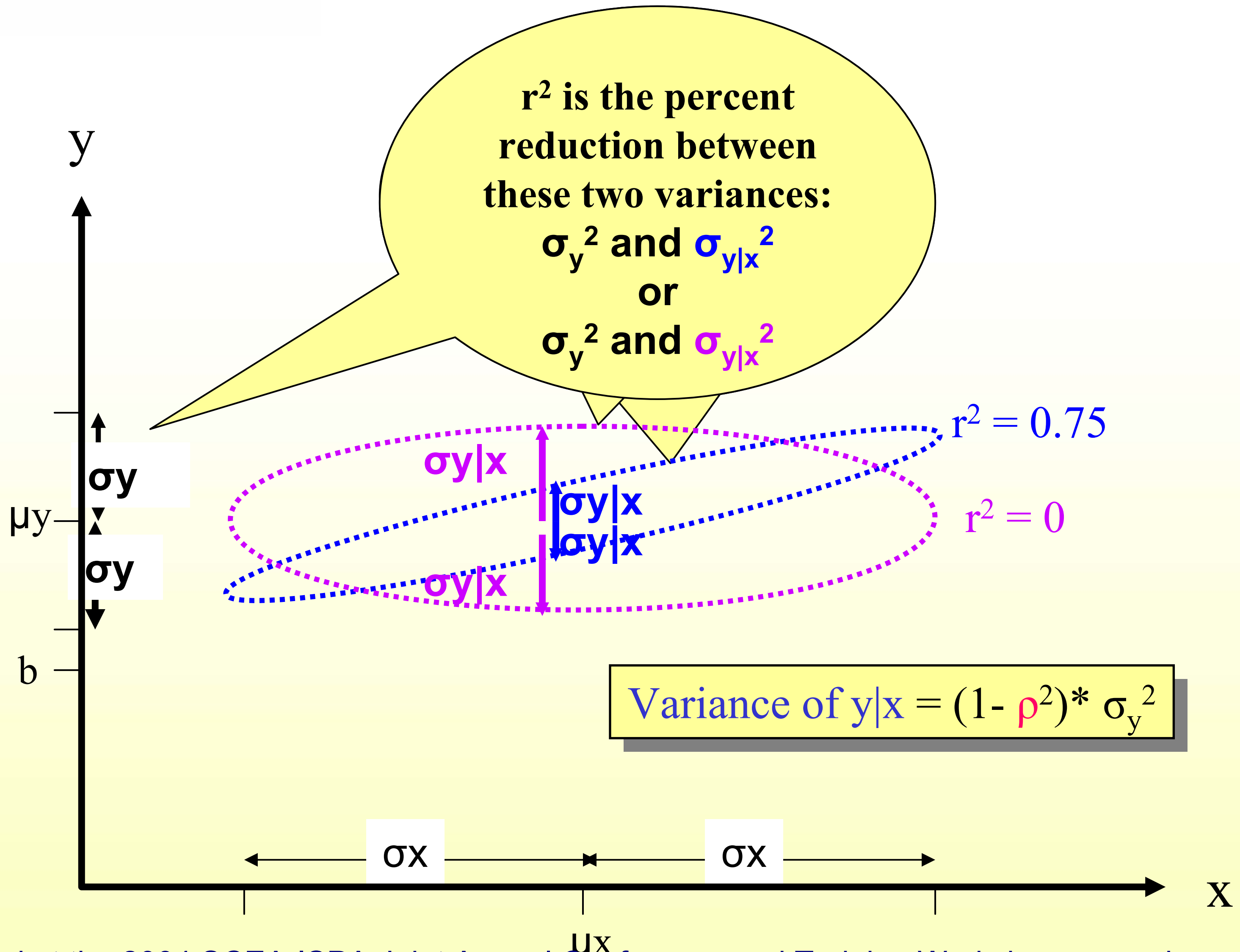
The Geometry of Bivariate Normality

and the implications for Regression



richard.coleman@ngc.com (703) 402-3702

The Geometry of Bivariate Normality
and the implications for Regression



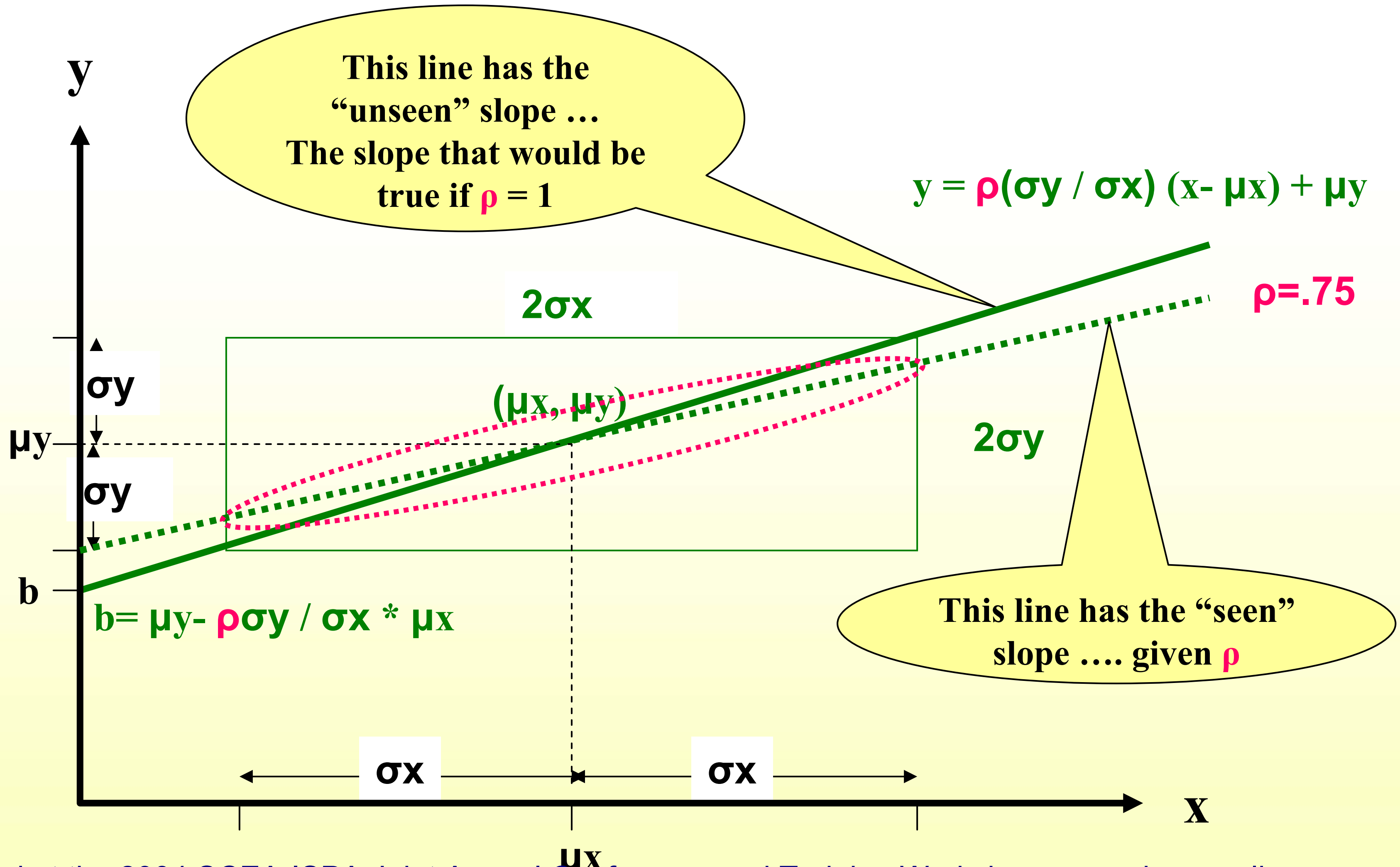
richard.coleman@ngc.com (703) 402-3702

The implications

- **For every regression with apparent slope m, there is an unseen equation**
  - With steeper slope m/ρ which is the *unseen* slope of the two variables
  - With an unseen accompanying y intercept
- **Once we decide upon the means and the variances of x and y, the unseen line is fixed**
  - Once we pick ρ, the regression line is fixed

richard.coleman@ngc.com (703) 402-3702

The Geometry of Bivariate Normality
and the implications for Regression



richard.coleman@ngc.com (703) 402-3702

Implementing Relational Correlation for Analogies

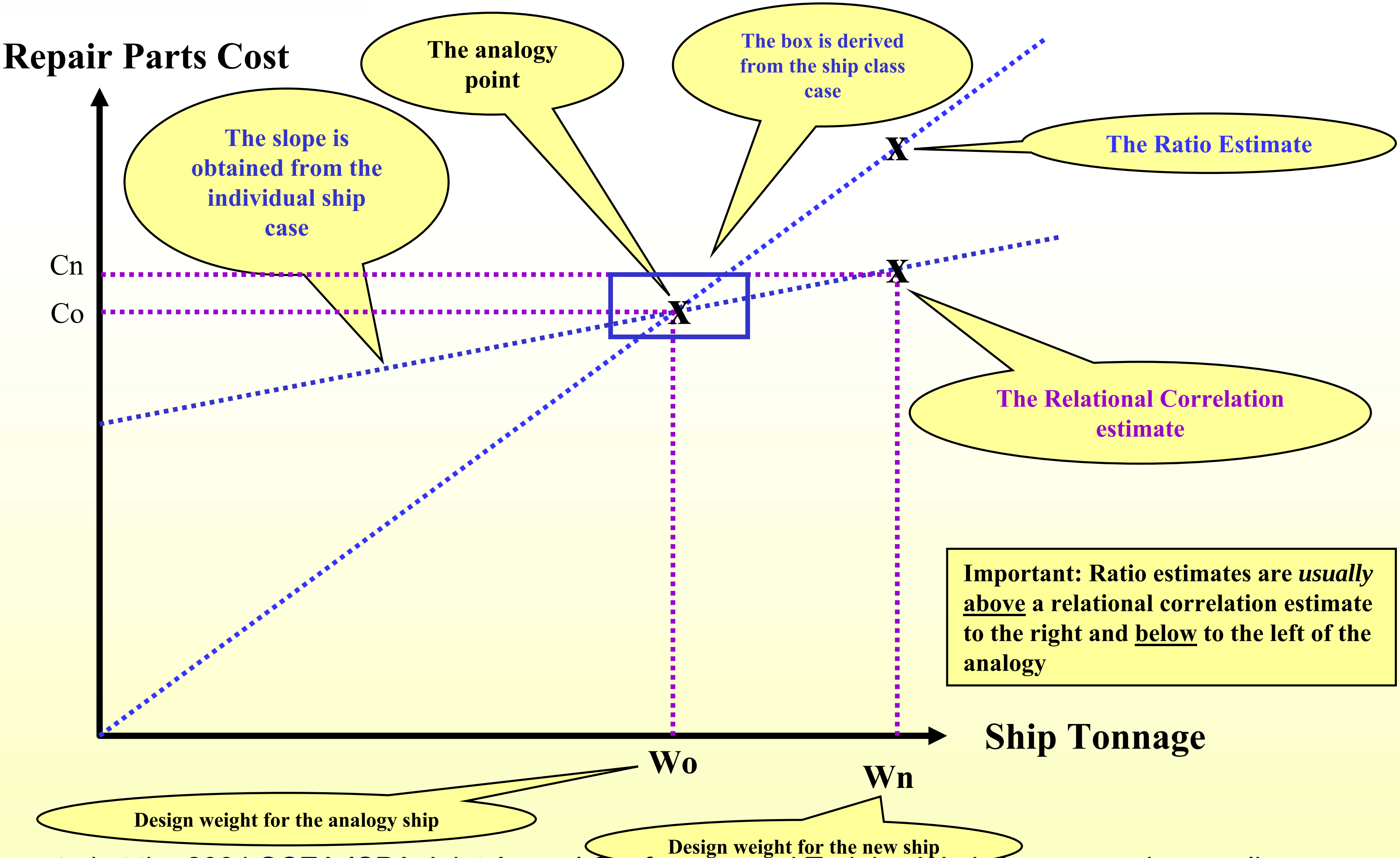
richard.coleman@ngc.com (703) 402-3702

Implementing Relational Correlation for Analogies

- **For Single Point Analogies**
  - Determine a reasonable (preferably historically-based) standard deviation for the x and y variable
  - E.g, to estimate ship repair parts as a function of tonnage you'll need:
    1. The standard deviation for the analogy ship class repair parts cost
    2. The standard deviation for the tonnage within the ship class
    3. The standard deviation of repair parts for a single ship of the class
    - The ratio of 1 and 2 gives you the unseen slope
    - The relationship of 3 and 1 will yield r^2 (Variance of y|x = $(1- \rho^2) * \sigma y^2$)
- **For buildups, do as above, but use an analogy for the values, and apply it to your buildup using percents**

richard.coleman@ngc.com (703) 402-3702

The Relational Correlation Method



Important: Ratio estimates are *usually* above a relational correlation estimate to the right and below to the left of the analogy

richard.coleman@ngc.com (703) 402-3702

Conclusions

- **Adjustments of analogies have received too little attention**
- **Three methods**
  - Ratio adjustments
    - Current practice
    - Overstate above the analogy, and understate below
  - Borrowed slope
    - Needs a CER
  - Relational correlation
    - Esoteric
    - Does not need a CER
- **Hopefully we have convinced you that ratio adjustment is just not good enough!**

richard.coleman@ngc.com (703) 402-3702

Backup & Old

richard.coleman@ngc.com (703) 402-3702

How to Implement Relational Correlation for Expert-Based CERs

The Problem

- **You have two WBS elements**
  - Warhead cost
  - Motor cost
- **You know their historic means and standard deviations – for both cost and the driving parameter, say weight**
  - You know these values from independent data bases
    - So, you cannot get correlation
- **You *do* have a CER to predict warhead cost**
- **You *do not* have a CER to predict motor cost**
  - You believe weight is a driver, but a CER cannot be derived
  - And, the data you have is too far away from your program, it needs to be adjusted … but how?
    - You do not wish to simply factor the cost by the weight change
- **This is a typical problem, and is closely related to the risk problem just described**
- **We will try to predict motor cost as a function of warhead cost … a useful equation as well as a helpful CER**

richard.coleman@ngc.com (703) 402-3702

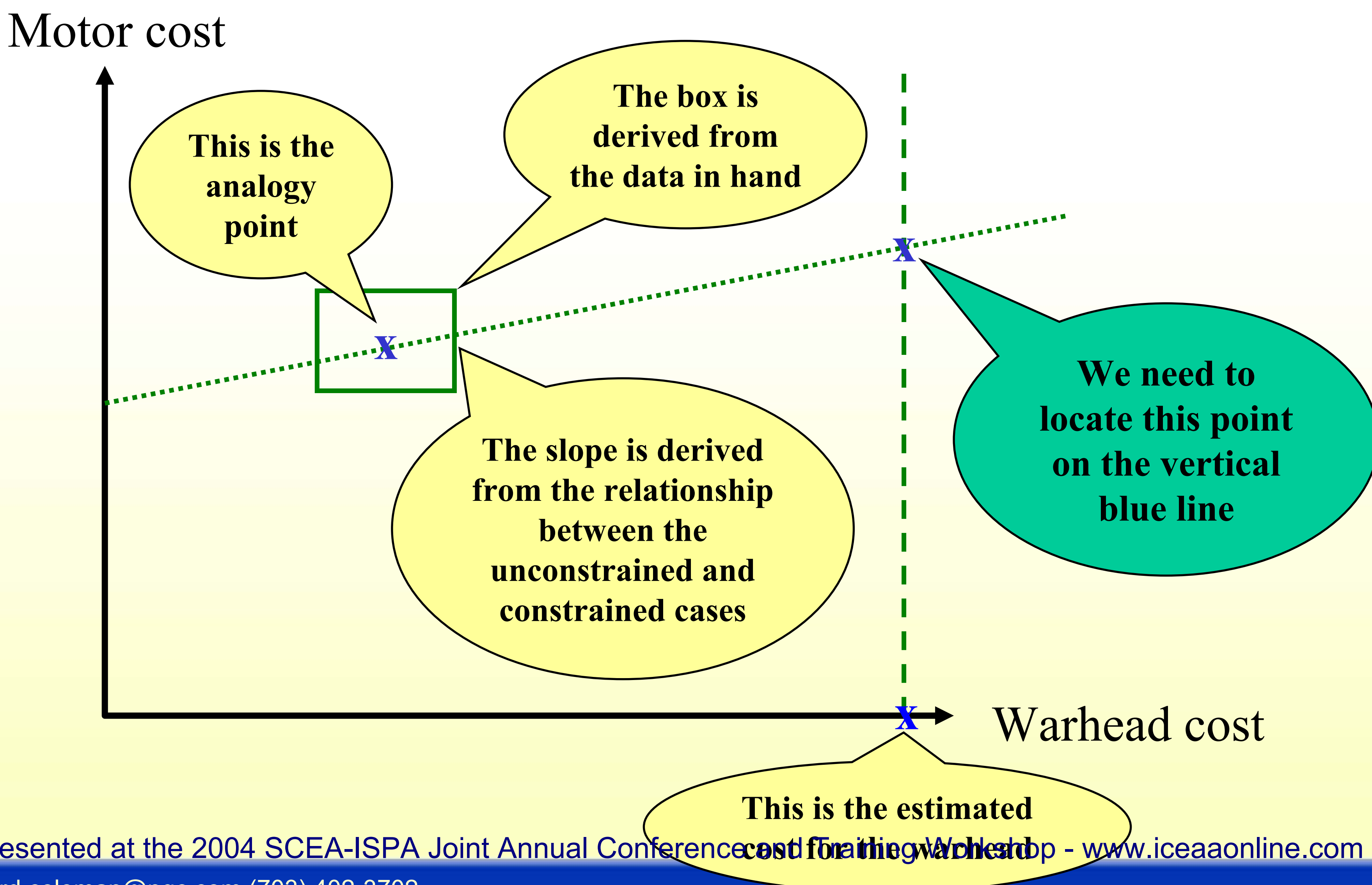
How to Implement Relational Correlation for Expert-Based CERs

1. Ask the engineer: *How much leeway in % do you typically have for weight (or cost) of the motor if design has not yet begun?* (The unconstrained case)
   - *Note: this may differ from the historic variation, but we will use it only in a relative sense*
   - *We will translate the weight fluctuation into cost fluctuation*
2. Ask the engineer: *How much leeway in % do you have for weight (or cost) of the motor, if design of the warhead is complete?* (The constrained case)
3. This will give you r^2:
   - You already knew the "unseen slope", $\sigma y/\sigma x$, now you know the "seen" slope $\rho(\sigma y / \sigma x)$, and you know $b = \mu y - \rho \sigma y/\sigma x * \mu x$
   - The percent reduction in the variance of y is the r^2, and the square root of that is r (Variance of $y|x = (1- \rho^2)* \sigma y^2$)
4. Implement the result as a CER, by passing the slope through the analogy or average of your data.

Note: We do not advocate using such a CER in lieu of a standard CER, only if there is no other recourse

richard.coleman@ngc.com (703) 402-3702

How to Implement Relational Correlation for Expert-Based CERs



richard.coleman@ngc.com (703) 402-3702