Multiple Models from a Single Dataset

(Chapter 12 – Software Project Estimation)

Alain Abran

(Tutorial Contribution: Dr. Monica Villavicencio)

Topics covered

1. Introduction
2. Sensitivity to functional size
3. The empirical study
4. Descriptive analysis
5. Productivity analysis
6. External Benchmarking with the ISBSG Repository
7. Identification of the Adjustment Factors

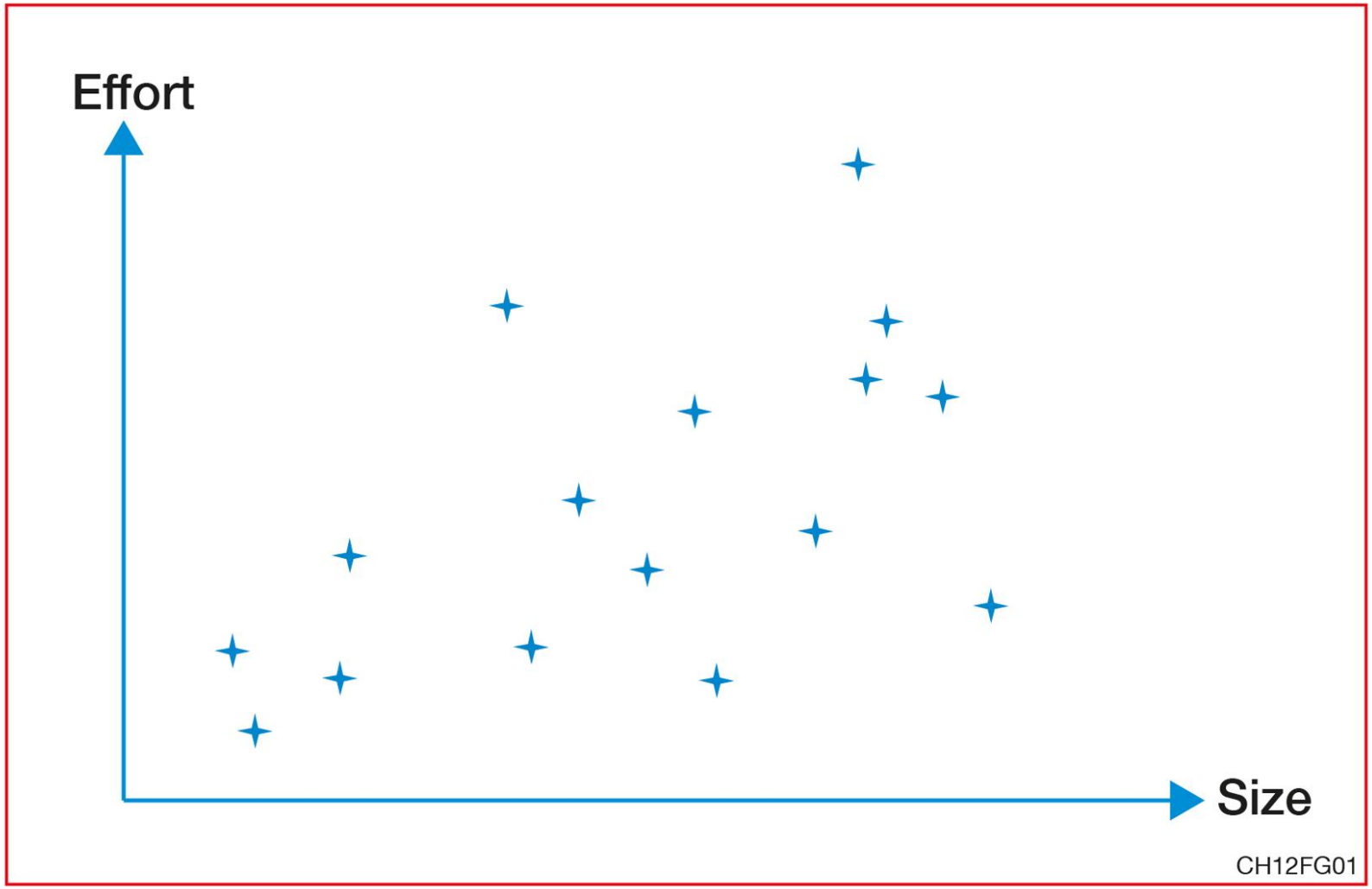
12.1 Introduction

Multiple models

- Classical approach to developing productivity models:
  - single productivity model with many cost factors (i.e. independent variables).
- Alternative approach:
  - A number of simpler models
    - Models that better reflect the major variations in an organization's performance in terms of fixed and variable costs.

12.2 Sensitivity to Functional Size

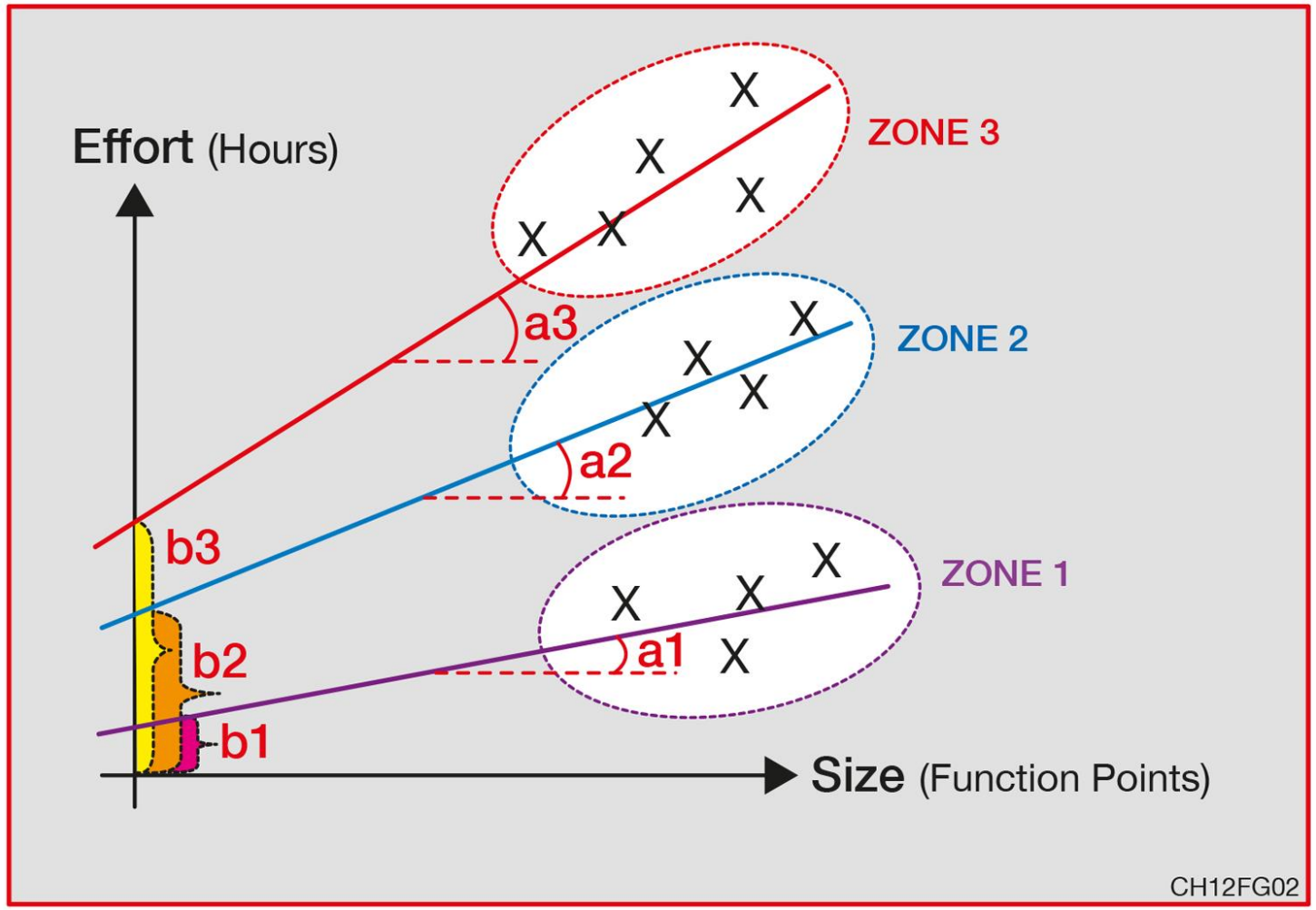
Wedge-shaped dataset



Low & High Sensitivity to Functional Size Increases: Multiple Models

- Zone 1: Populated with a subset of projects demonstrating little sensitivity to increases in size (even large increases in size do not lead to corresponding increases in effort)
  - The effort required is almost insensitive to an increase in the number of functions in the software being developed.
- Zone 3: Populated with a subset of projects demonstrating high sensitivity with respect to functional size.
  - A small increase in size requires a much larger increase in effort – in either fixed or variable costs, or both.
- Zone 2: This lies somewhere in the middle range of the dataset.

Data subsets with distinct sensitivities to increases in functional size



12.3 The Empirical Study

An empirical study (1/2)

- Context:
  - Governmental agency providing specialized financial services to the public.
  - Software applications are similar to those of banking and insurance providers.
  - The productivity was measured on individual projects (models were derived).
- Data collection procedures:
  - Project documentation & related data available to be measured in terms of function point size, effort & duration.

An empirical study (2/2)

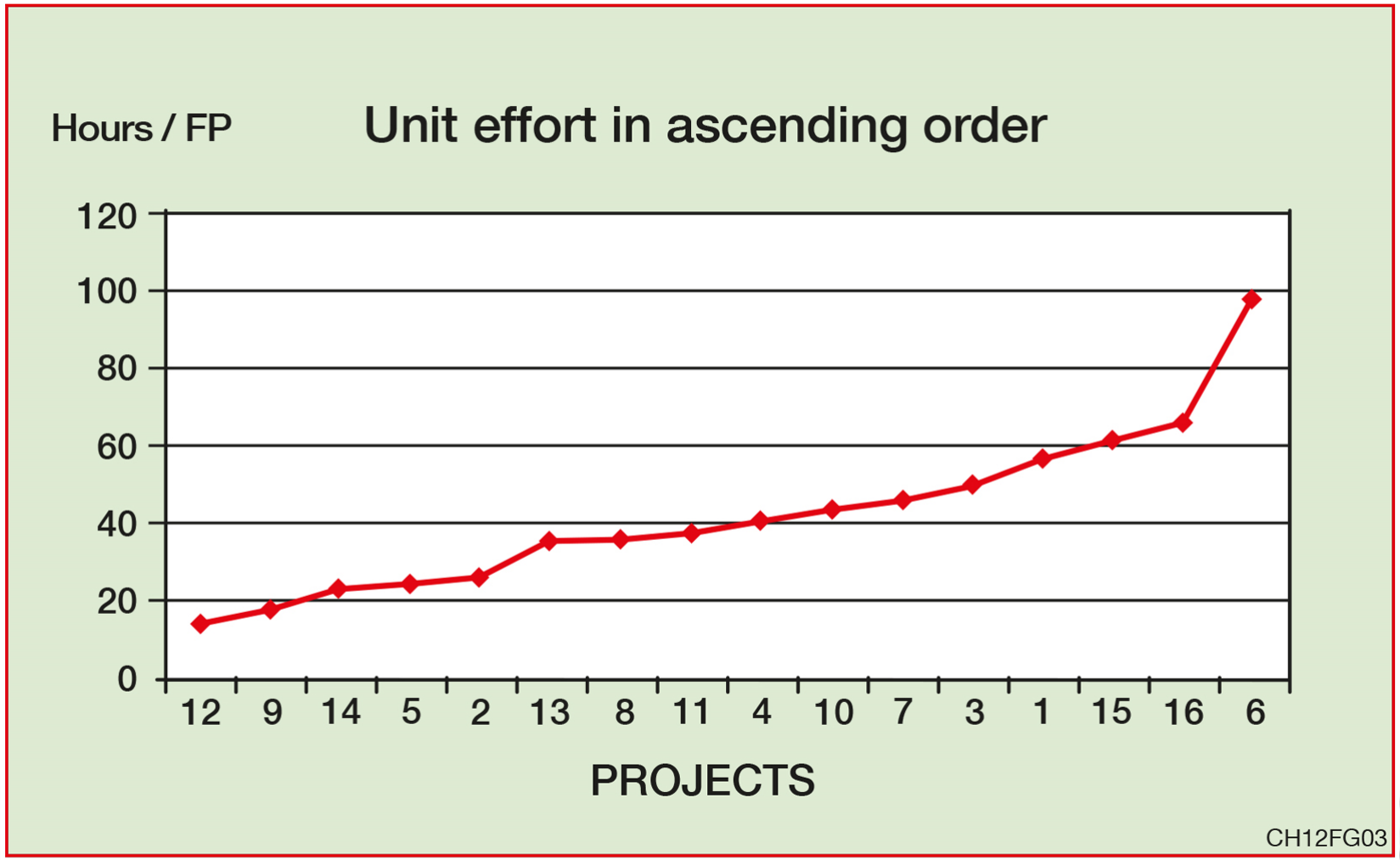
- Data quality control:
  - Critical quantitative variables: effort & functional size.
    - **Effort data:** The time reporting system was considered highly reliable and used for decision making, including payment of invoices when external resources were hired to complement project staffing.
    - **Measurement of functional size:** All functional size measurements were carried out by the same experienced measurer.

12.4 Descriptive Analysis

Project characteristics

- Project sizes vary from a minimum of 111 FP (project 6) to a maximum of 646 FP.
- Effort varies from 4,879 to 29,246 hours.
- Duration varies from 9.6 to 33.6 months.
- Maximum development team sizes from 6 to 35 employees.

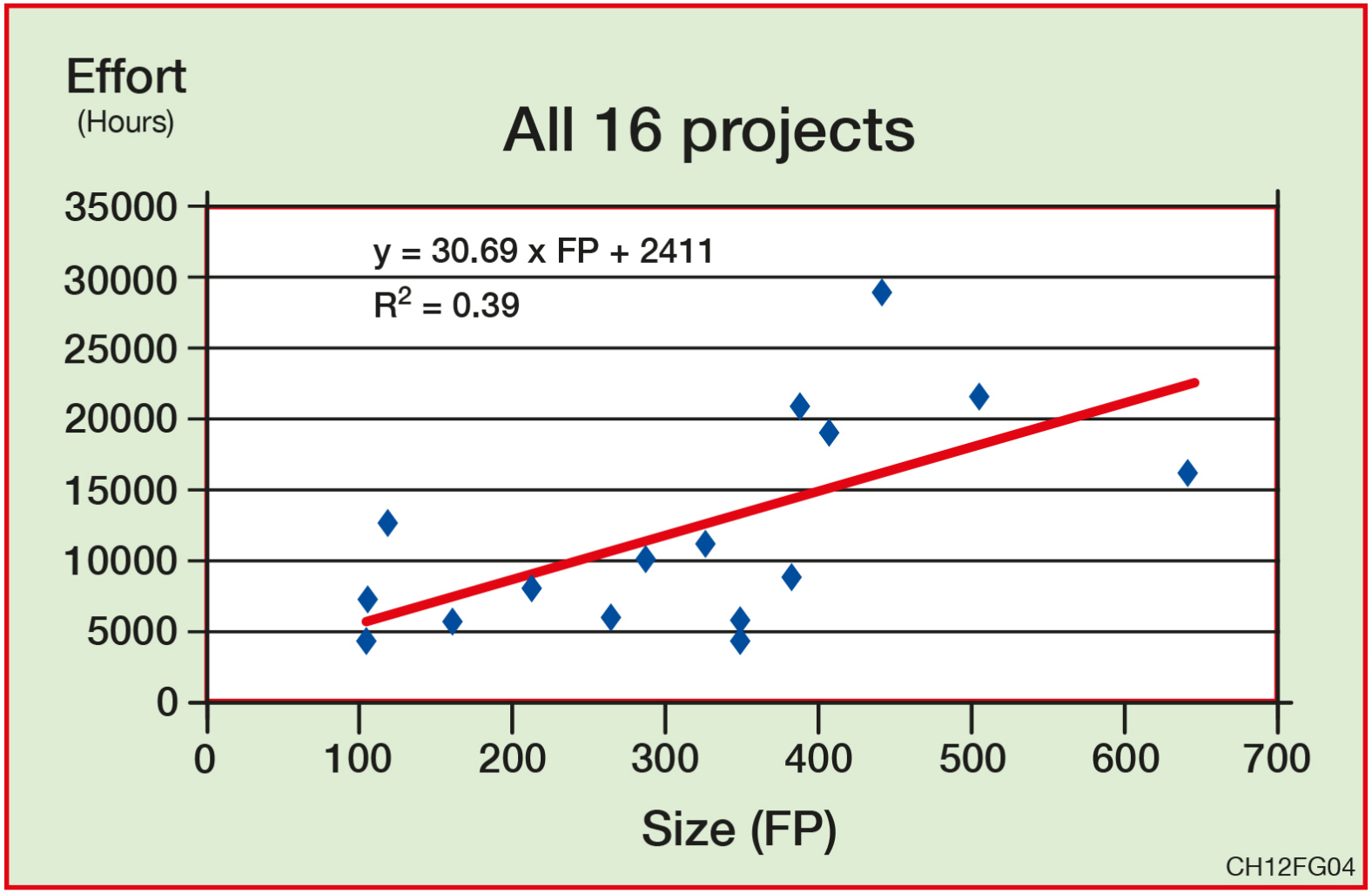
Projects in ascending order of unit effort (hours/FP)



Average unit effort: 42 hours/FP (FPA)

12.5 Productivity Analysis

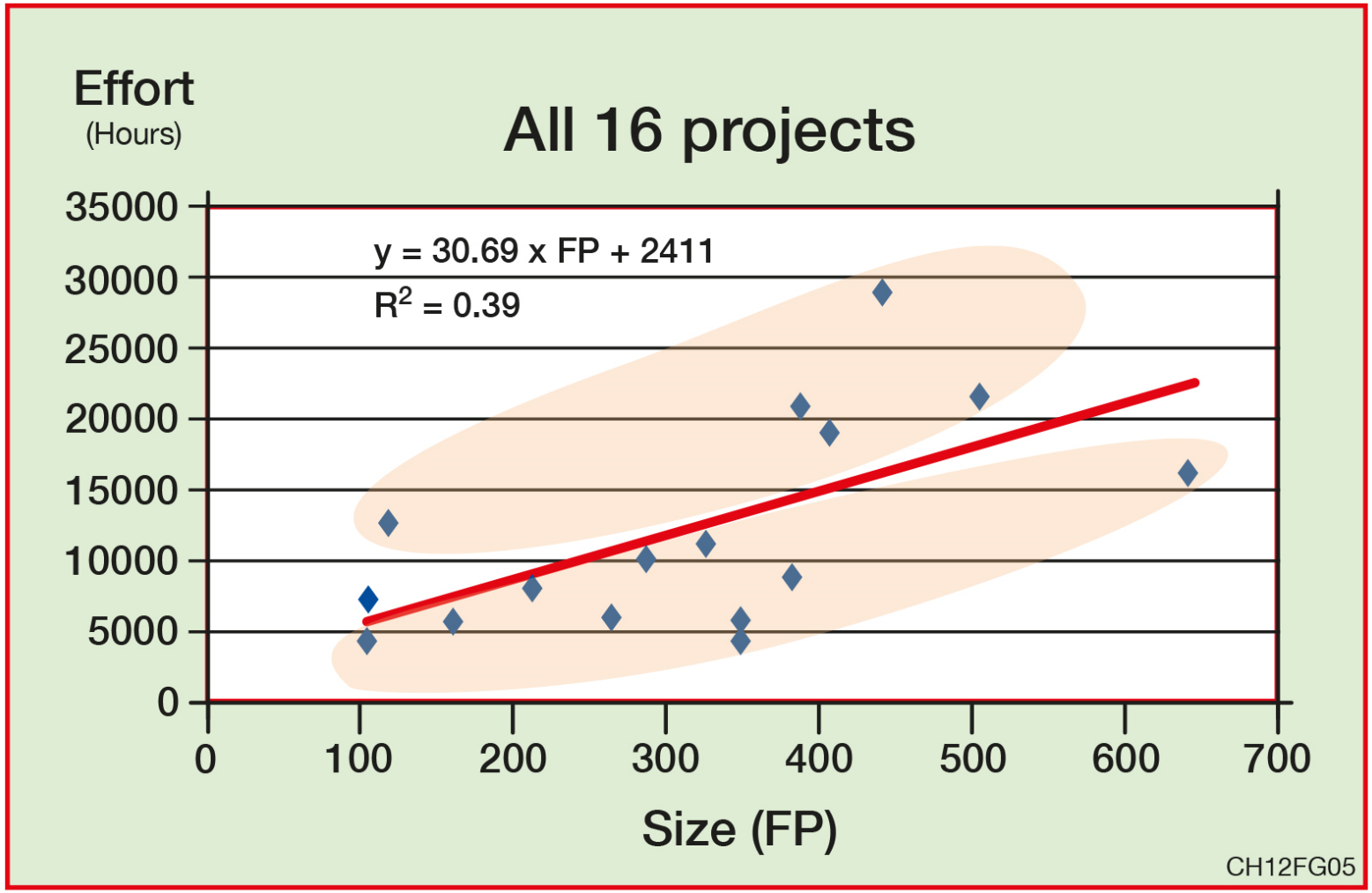
The organization's production model



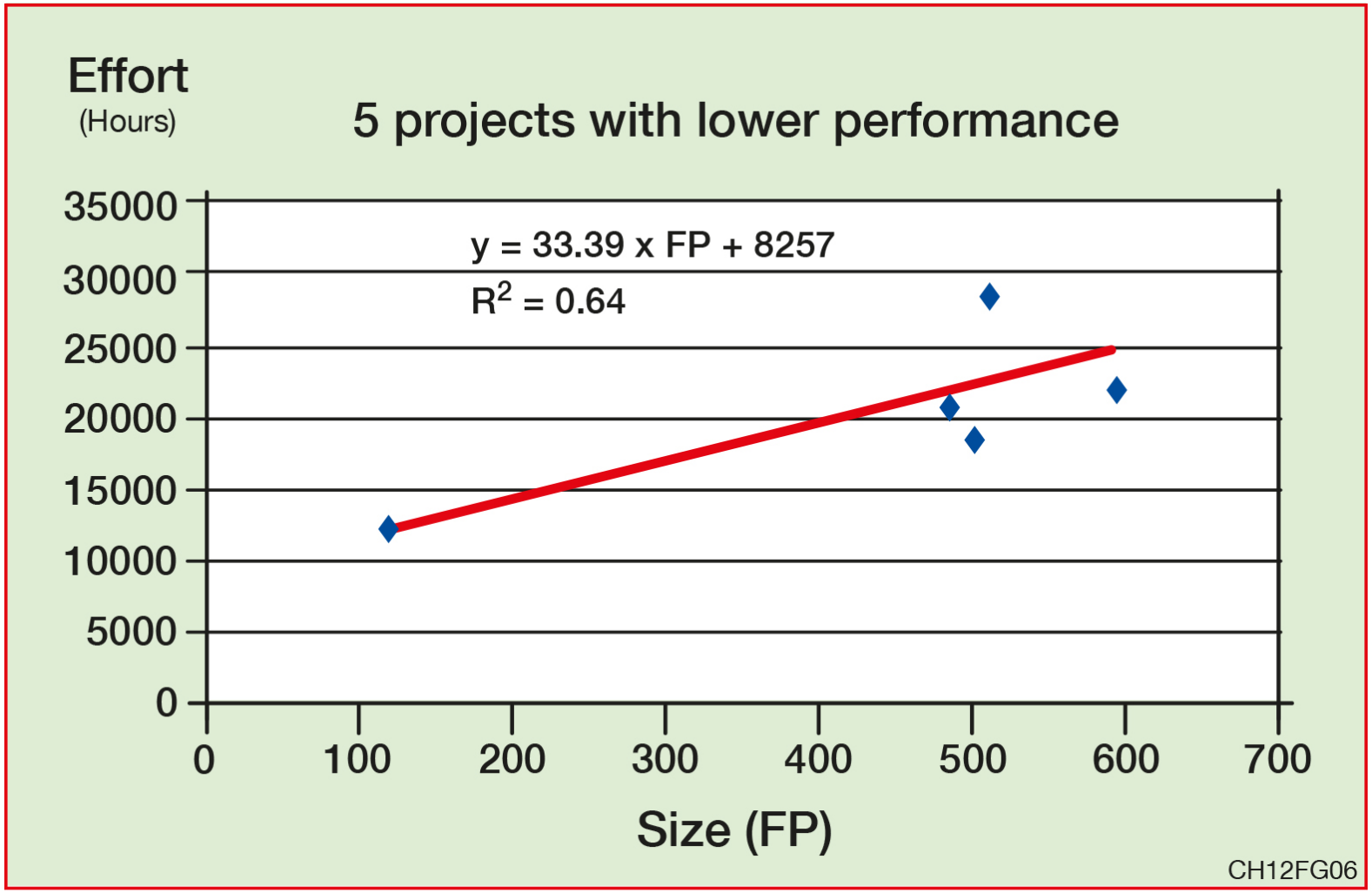
Practical interpretation of the equation

- (Fixed) effort not sensitive to software size = 2,411 hours.
- (Variable) effort sensitive to an increase in software size = 30.7 hours/FP
- The possible reasons for the rather high fixed unit effort values:
  - The acquisition process is highly procedural and time-consuming.
  - The projects have tight constraints and procedural documentation.
  - The projects require lengthy consensus building procedures.
  - The projects require a relatively high number of inspections.

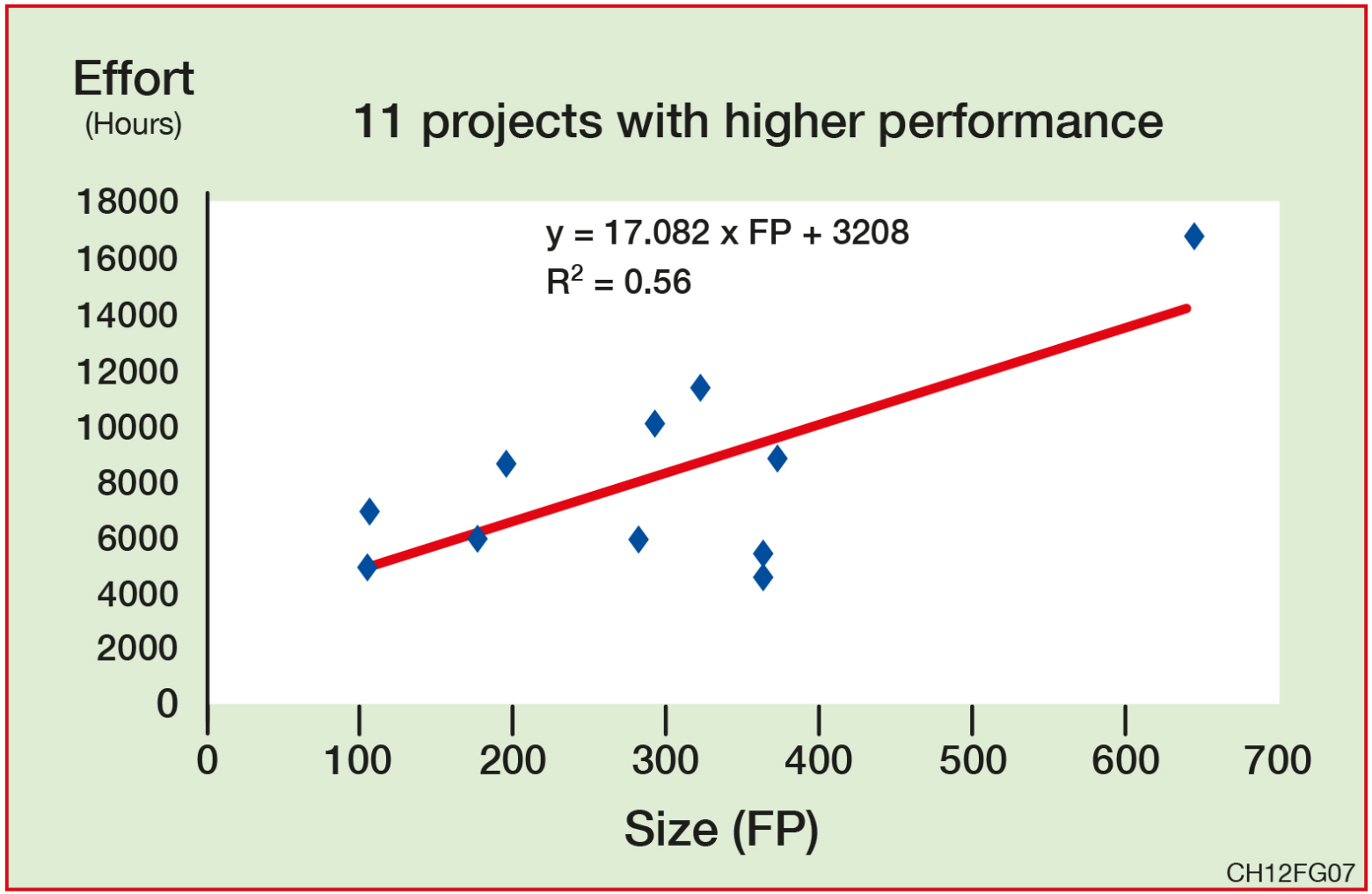
Two subsets of projects within the single dataset



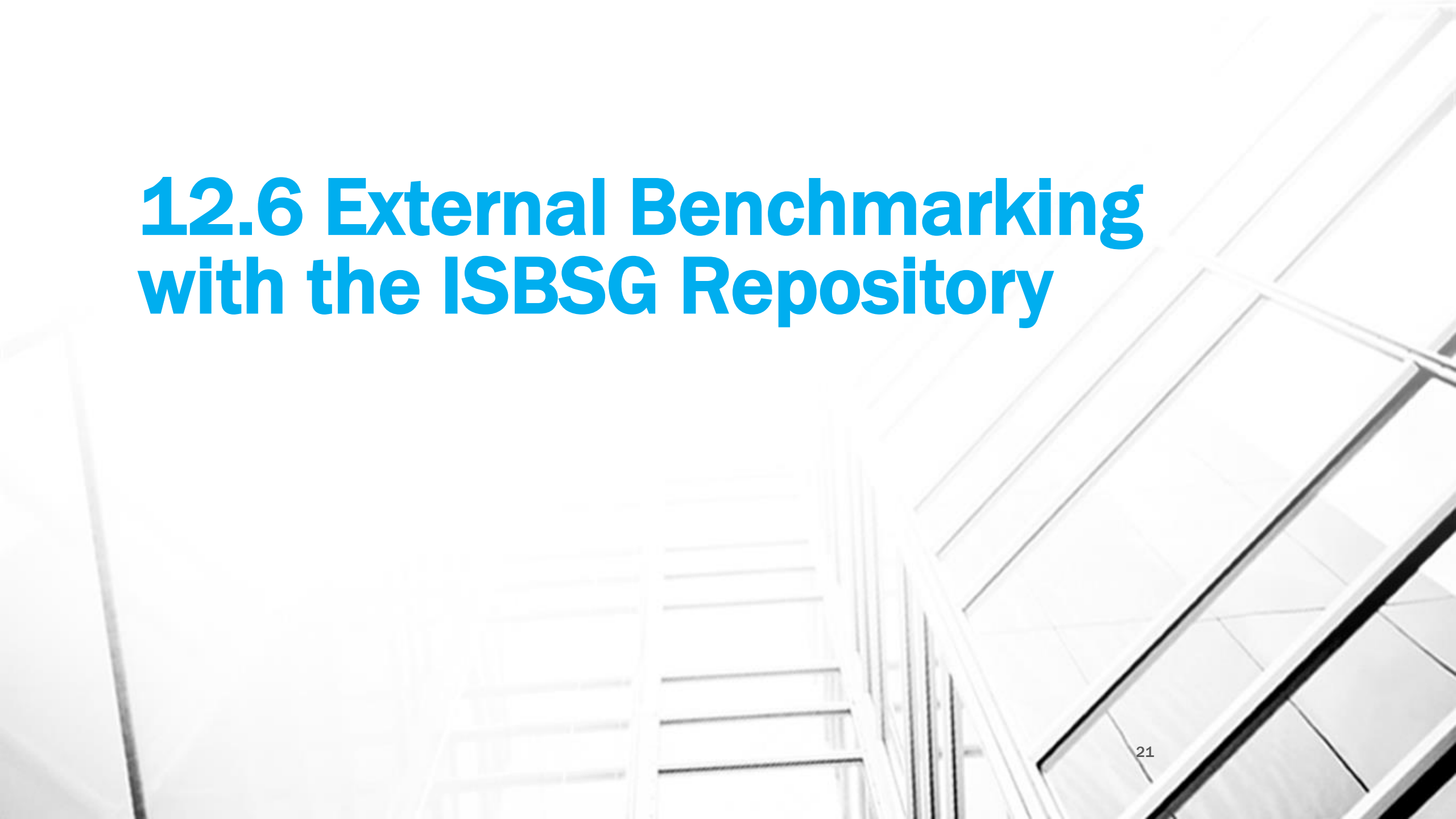
Least productive projects



Most productive projects



12.6 External Benchmarking with the ISBSG Repository



Benchmarking with ISBSG repository

- Benchmarking is the process by which measurement results of a specific entity are compared with the measurement results of similar entities.

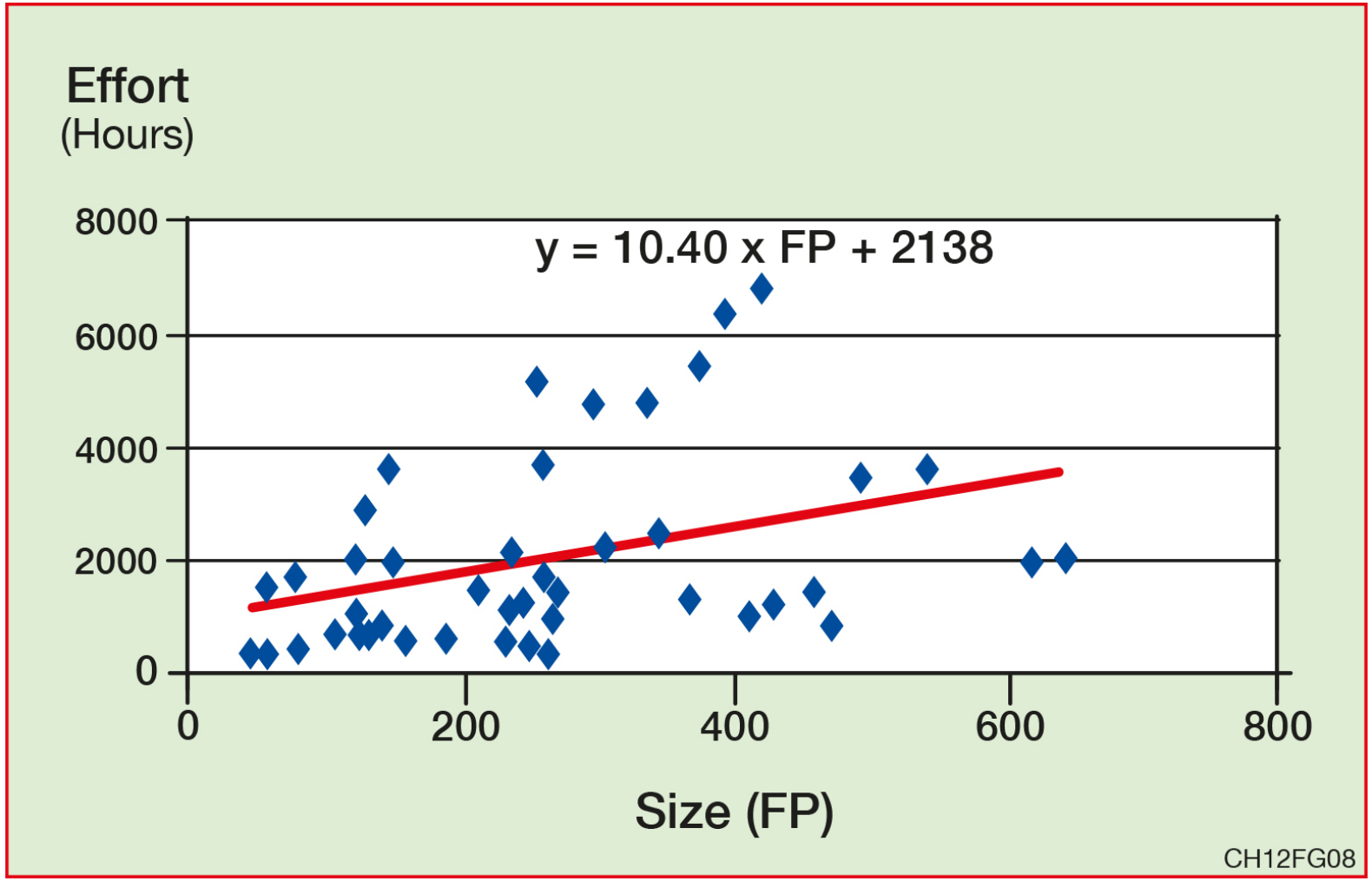
- Criteria used to select a benchmarking repository:

    1. A repository of projects, from both the private and public sectors, representative of software applications providing financial services.

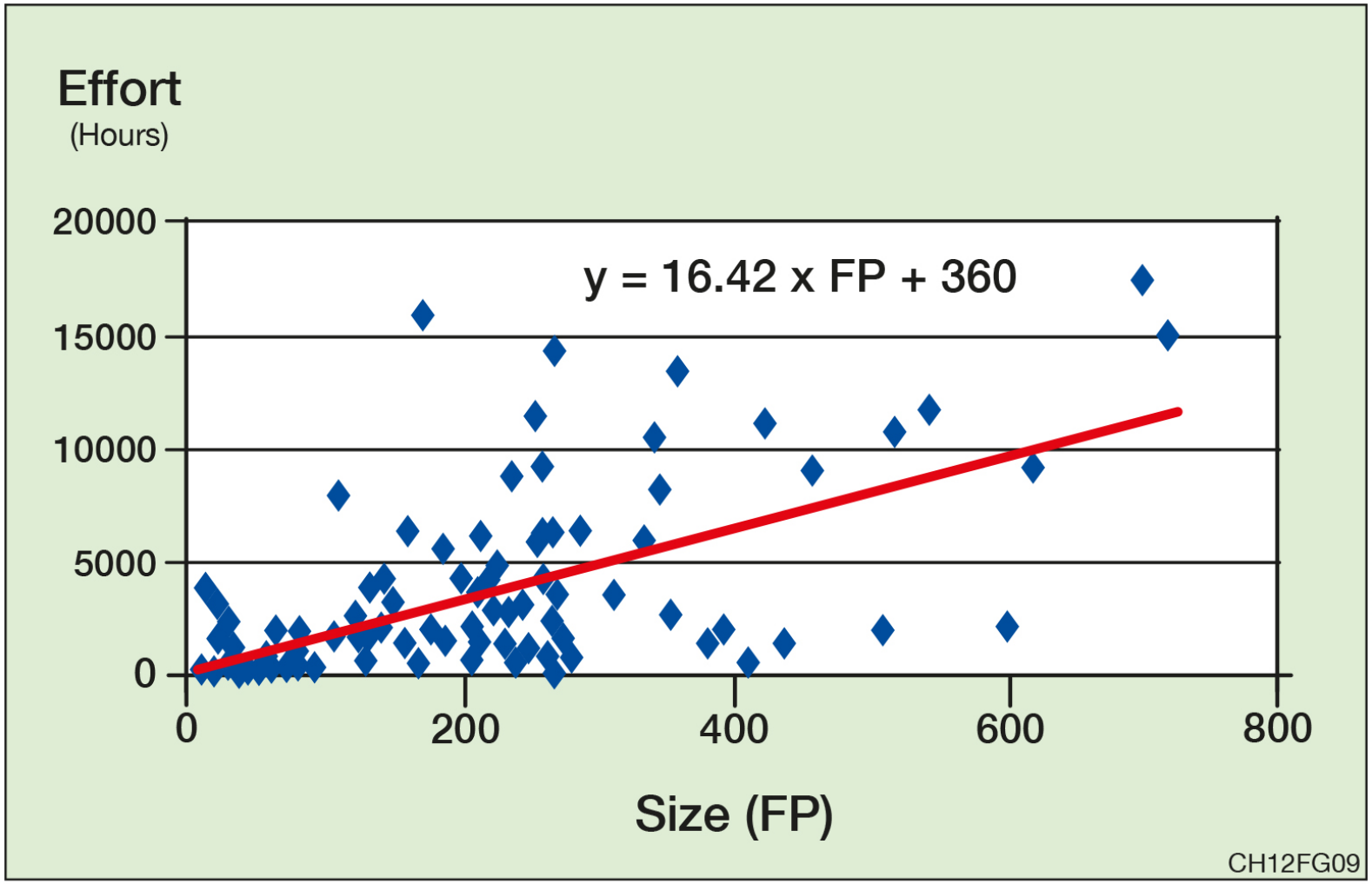
    2. A repository of projects from various countries.

    3. A repository with information available at the data field level (not at the aggregate level).

3GL ISBSG governmental projects



3GL ISBSG financial projects



External benchmarking analysis

	Organization (1)	ISBSG: Government (2)	ISBSG: Financial (3)
Number of projects	16	48	119
Fixed effort (hours)	2,411	2,138	360
Variable effort (hours/FP)	30.7	10.4	16.4

Summary of fixed & variable effort

12.7 Identification of the Adjustment Factors

Projects with the highest unit effort: Factors ideentified in this organization which provided the datga

A. Poorly expressed customer requirements (frequent change requests during a project life cycle).

B. Customers not familiar with the software development process in the organization.

C. High turnover of users involved in the projects, leading to instability in the requirements and delays in decision making.

D. New technologies unknown to the developers.

E. Multiple links with the organization's other software applications.

F. A severely compressed schedule.

Exercises

1. Wedge-shaped datasets often occur in software engineering. Is it always necessary to look for a single productivity model? If not, which economics concepts can help in data analysis and identify models?

2. Identify some criteria to analyze the quality of the documentation available for the measurement of the functional size of the software to be developed.

3. In Table 12.2, are there statistical outliers on the independent variable, functional size?

4. In Table 12.2, are there statistical outliers on the dependent variable, effort?

5. For the dataset in chapter 12, what are the ratios of fixed and variable effort between the two productivity models developed?

6. Compare the performance of the ISBSG projects developed by government organizations and those developed by financial institutions?

7. Compare the performance of the data reported in this chapter with the data from ISBSG governmental organizations.

Term Assignments

1. Collect software project data from your organization, and provide a descriptive analysis of them.
2. Carry out a graphical analysis on size and effort data, and determine whether or not you are looking at candidate multiple models.
3. If you have candidate multiple models, interview the project managers to identify positive and negative productivity factors.
4. Compare your organization's performance with that of similar organizations in the ISBSG repository.