



Normalizing the ISBGS Software Benchmark

Lee Fischman & Karen McRitchie

Galorath Incorporated

100 North Sepulveda Blvd, Suite 1801

El Segundo, CA 90245 USA

Overview

- **ISBSG Background**
- **How Galorath uses the ISBSG data set**
- **Data mapping**
- **Normalization methods**
- **Error checking**
- **Results**

ISBSG Data Set Background

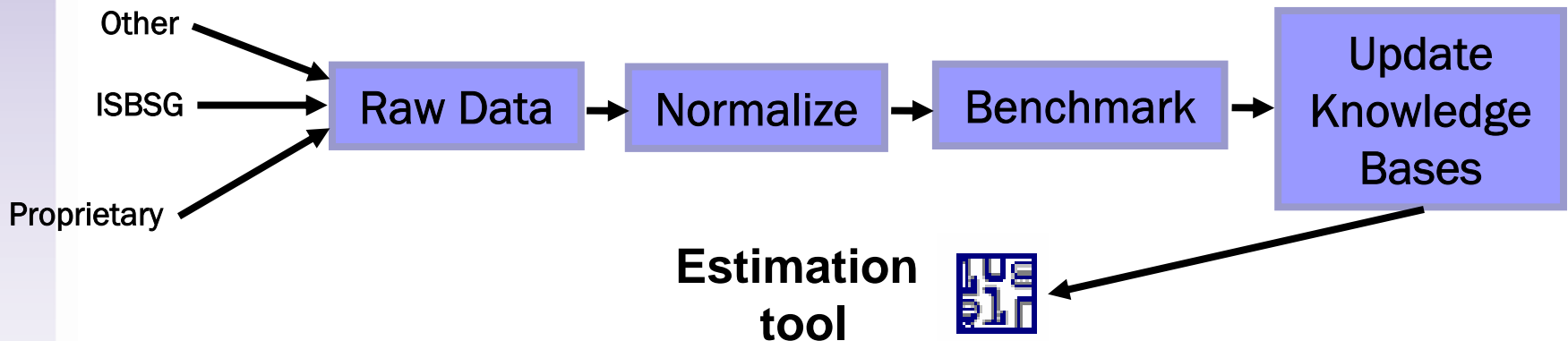
- The International Software Benchmarking Standards Group collects data for software development and enhancement projects
- They more recently added a repository for maintenance projects
- ISBSG offers data sets for purchase which are based on actual software projects
- The subject of this study was ISBSG release 9 which contains 3023 observations



see www.isbsg.org for more information

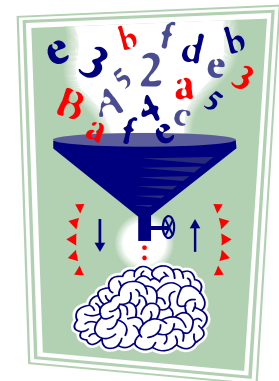
Galorath's Use of ISBSG

- Galorath Incorporated offers the a suite of software project estimation, planning and control tools
- The SEER-SEM is the core estimation tool of this suite and provides knowledge based estimation capability
- The estimation tool includes knowledge bases which capture parameter settings that target productivity rates for different types of projects and scenarios
- The knowledge bases are developed by Galorath based on analysis of various data sets,
 - ISBSG is one of them



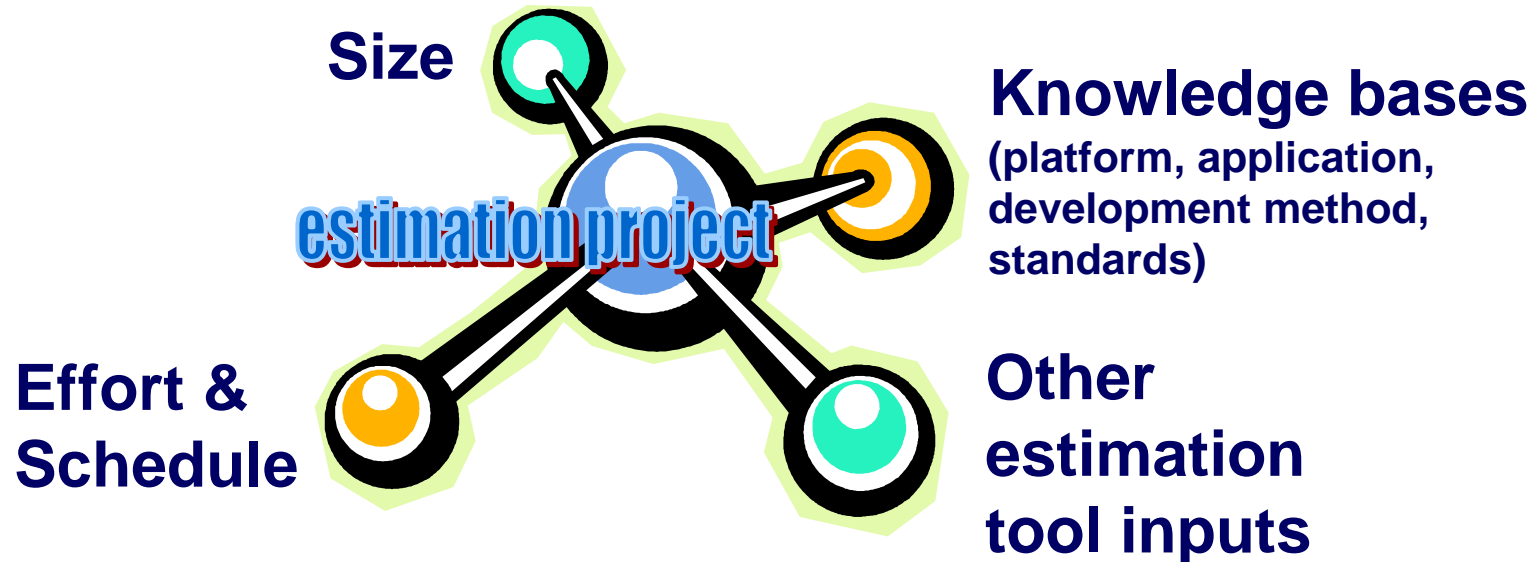
How Data Is Incorporated Into “Knowledge Bases”

- The estimating tool does not ship with raw data
- Knowledge bases include parameter settings based on analysis of historical data
- How is this analysis done?
 - For each observation, model the project data in the estimating tool
 - Use the estimating tool’s calibration mode to get an actual vs. estimate
 - Group & aggregate observations into categories
 - Look at estimate ratios, “technology ratings” and productivity rates
 - Identify trends
 - Absolute productivity for broad groups
 - Relative productivity, from category to category
 - Make adjustments
 - Repeat & retest



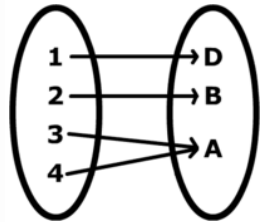
For Each Observation...

- Modeling a project in the estimation tool requires certain information



- Needless to say, the ISBSG data set needs to have descriptive and quantitative information mapped to the estimation tool inputs
- A data mapping process has been developed and tailored to the ISBSG data set

Normalization



Mapping ISBSGs
to tool categories



Acceptable
as-is



Focus of
our work

■ Knowledge bases

- Mapping ISBSG qualitative descriptors into estimation tool attribute names
 - platform, application, method, standard
- Repeatable process developed

■ Size normalization

- Our analysis utilizes unadjusted function points
- ISBSG provides a well defined measure of size, requiring no special normalization or analysis

■ Effort normalization

- The estimation tool assumes effort for certain phases
- Effort needs to be normalized so that all projects are comparable

Why our own normalization?

Although ISBSG provides normalized work effort, we ourselves normalized to achieve the following:

- Full transparency into normalization method
- Ability to remove implementation phase
- Removal of resource level 4

How Did ISBGS Do Normalization?

Effort was added or removed depending on whether these phases/resources were recorded in each record.

| | Plan | Spec | Build | Test | Impl |
|---|-------------|-------------|--------------|-------------|-------------|
| Developers | 10 | 19 | 46 | 18 | 7 |
| Developers + Support | 7 | 14 | 55 | 13 | 11 |
| Developers + Support + Users | 8 | 25 | 33 | 23 | 11 |

George Ansell & Chris Lohan, Normalising Effort in the ISBSG Repository, Australian Conference on Software Measurement (2002)

Galorath's Normalization Steps

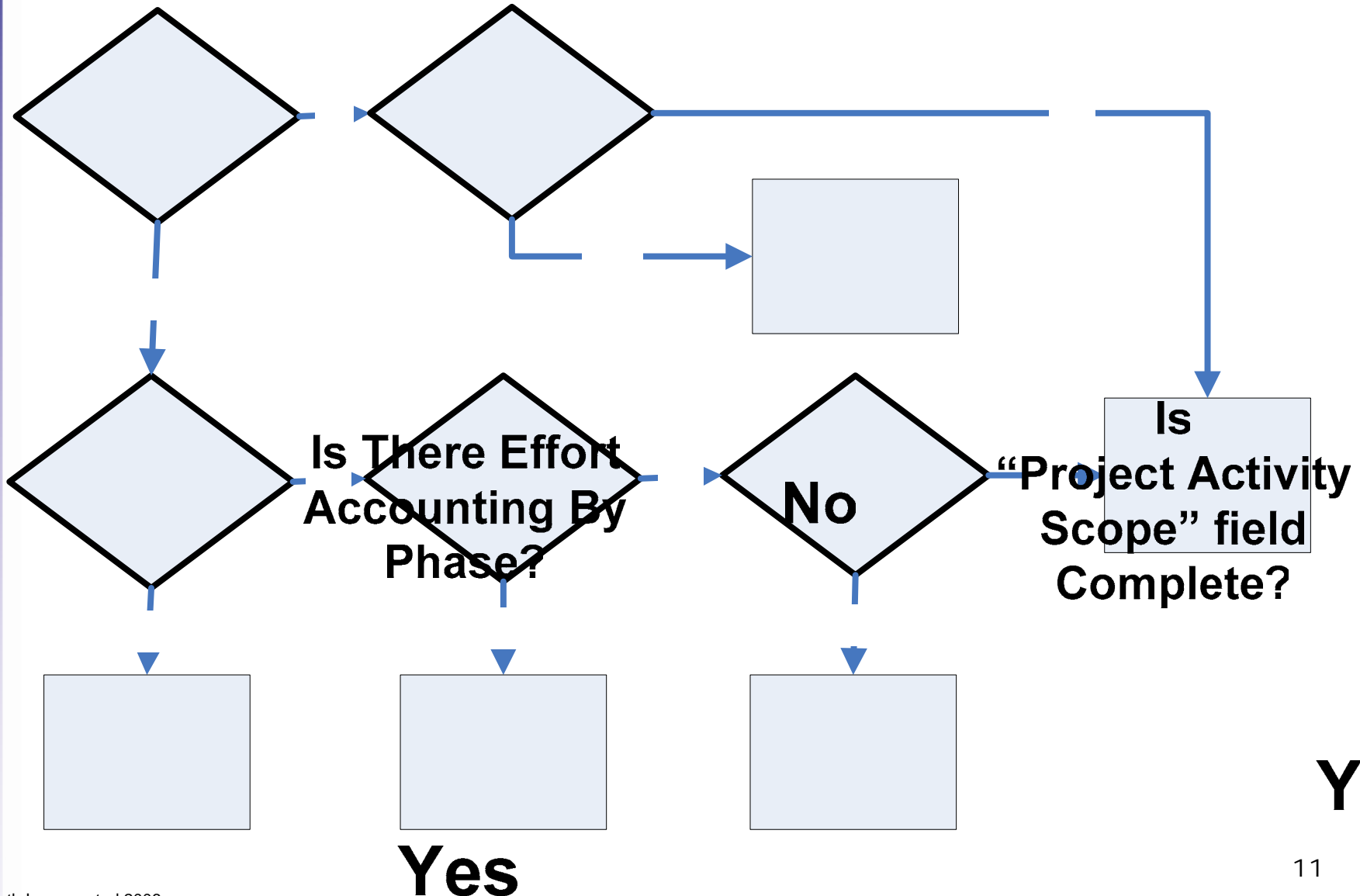
Phase normalization

- Fill in potentially missing phases in “Project Activity Scope” field
- Estimate potentially omitted work effort
- Choose best normalization method given available info
- Remove effort due to Implementation

Resource normalization

- Remove effort due to resource level 4

Phase Normalization Decision Process



Yes

Phase Reporting In ISBSG v9

Respondent may have provided list of phases included

Respondent may have provided a detailed effort accounting by phase

| Project Activity Scope | Effort Plan | Effort Specify | Effort Design | Effort Build | Effort Test | Effort Implement |
|--|-------------|----------------|---------------|--------------|-------------|------------------|
| Planning;Specification;Build;Test;Implement; Planning;Specification;Build;Test; | 100 | 400 | | 1000 | 200 | 150 |
| <i>x</i> | | 281 | | 359 | 156 | |
| Specification;Build;Test;Implement; Planning;Specification;Build;Test; | 62 | 17172 | | 136 | 230 | |
| Planning;Specification;Build;Test; | 3 | 459 | | 2 | 1 | |
| | | 75 | <i>x</i> | 248 | 136 | |
| Specification;Build;Test; | | | | | | |
| Planning;Specification;Build;Test;Implement; | 320 | 480 | | 800 | 960 | |
| Planning;Specification;Build; | | 967 | | 905 | | |
| Planning;Build;Test; | 427 | | | 882 | 358 | |
| Planning;Build;Test;Implement; Specification;Build;Test; | 1248 | 0 | | 2504 | 8768 | 12520 |
| Specification;Build;Test;Implement; | <i>x</i> | 153 | | 1961 | 510 | 255 |
| | 90 | 720 | | 1043 | 1682 | |
| Planning;Specification;Build;Test; | | 640 | | 811 | 936 | |
| Planning;Specification;Build; | | 102 | | 4 | | |
| | 39 | 164 | | 315 | 65 | 65 |

Backup slide: ISBSG's definition of phases

“Project Activity Scope” Field - Did People Get Their Answers Right?

Fill in missing phases

| | | | | |
|------|------|------|------|------|
| Plan | Spec | Buil | Test | Impl |
| | Spec | Buil | Test | |

‘Design’ is rarely included – is that because most people consider it part of Spec?

Did this respondent believe Plan included Spec and Design?

| | | | |
|------|------|------|------|
| Plan | Spec | Buil | Test |
| Plan | Spec | Buil | Test |
| | Spec | Buil | Test |

| | | | |
|------|------|------|------|
| Plan | Spec | Buil | Test |
|------|------|------|------|

| | | | |
|------|------|------|------|
| Plan | Spec | Buil | Test |
| | | Buil | |

Figure out whether some phases were inadvertently combined

| | | | |
|------|------|------|------|
| Plan | Buil | Test | Impl |
|------|------|------|------|

| | | | | |
|--|------|------|------|------|
| | Spec | Buil | Test | Impl |
| | | | Test | |

| | | | | |
|------|------|------|------|------|
| Plan | Spec | Buil | Test | Impl |
|------|------|------|------|------|

| | | | | |
|--|------|------|------|--|
| | Spec | Buil | Test | |
|--|------|------|------|--|

| | | | | |
|------|------|------|------|------|
| Plan | Spec | Buil | Test | Impl |
|------|------|------|------|------|

Was Test only what was intended?

Similar questions for detailed effort accounting

“Project Activity Scope” Replacement Table – Judging When Phases Are Missing

‘High’ ratings are given in obvious instances of phase exclusion.

continued...

| Phases Included | Probability Requiring Interpolation? | Phases Included | Probability Requiring Interpolation? |
|----------------------|--|--------------------------|--|
| Impl | High | Plan | High |
| Test | High | Plan Impl | Low |
| Buil | High | Plan TestImpl | High |
| BuilTest | Low | Plan Buil | Low |
| BuilTestImpl | Low | Plan Buil Impl | Low |
| DesiBuil | Low | Plan BuilTest | Low |
| DesiBuilTestImpl | Low | Plan BuilTestImpl | Low |
| Spec | High | Plan DesiBuilTestImpl | Low |
| Spec Impl | Low | PlanSpec | High |
| Spec Test | High | PlanSpec TestImpl | High |
| Spec TestImpl | High | PlanSpec Buil | Low |
| Spec Buil | Low | PlanSpec Buil Impl | Low |
| Spec Buil Impl | Low | PlanSpec BuilTest | Low |
| Spec BuilTest | Low | PlanSpec BuilTestImpl | Low |
| Spec BuilTestImpl | Low | PlanSpecDesiBuil Impl | Low |
| SpecDesiBuil | Low | PlanSpecDesiBuilTest | Low |
| SpecDesiBuil Impl | Low | PlanSpecDesiBuilTestImpl | Low |
| SpecDesiBuilTest | Low | | |
| SpecDesiBuilTestImpl | Low | | |

Normalization Steps

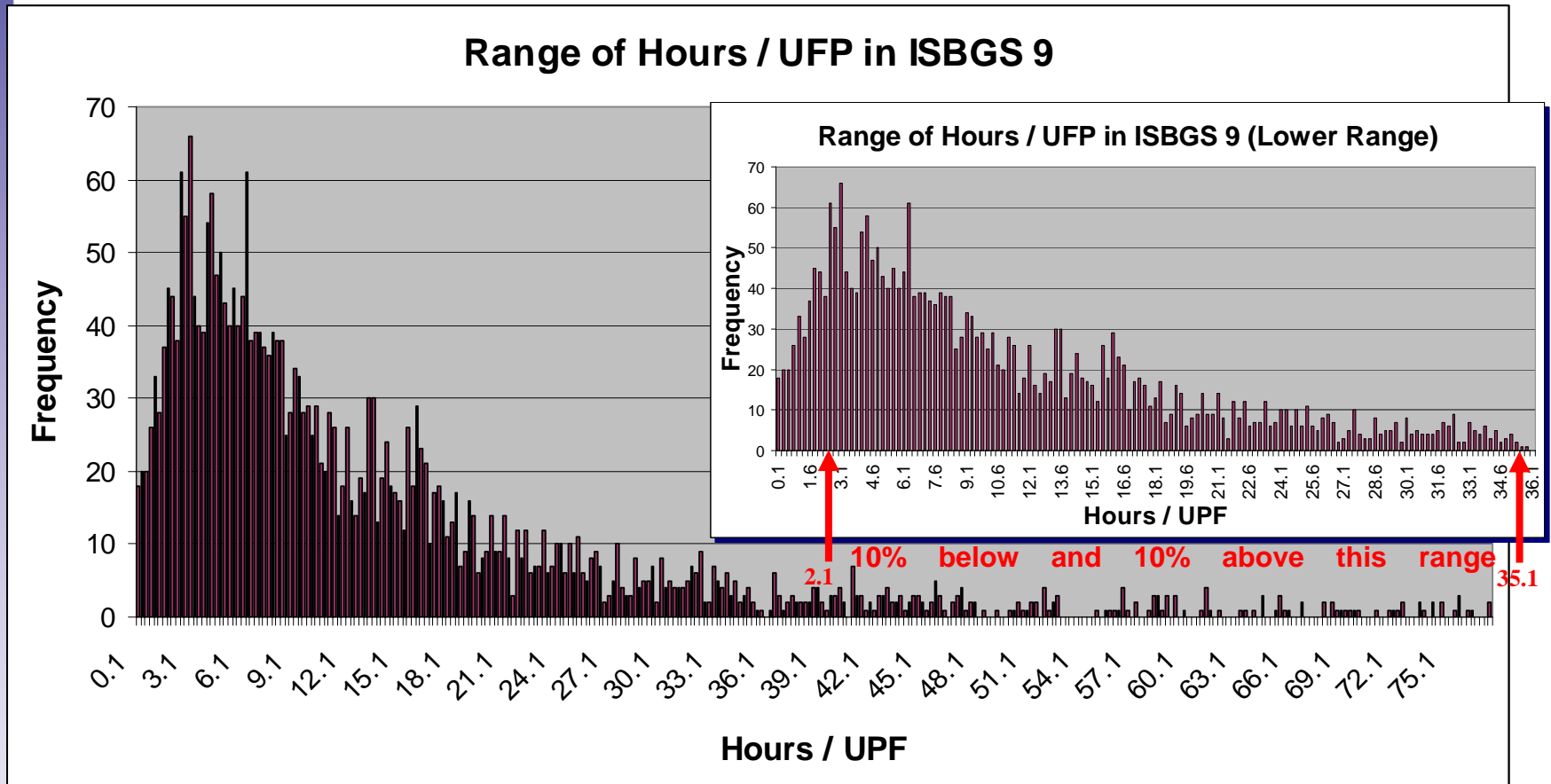
Phase normalization

- Fill in potentially missing phases in “Project Activity Scope” field
- Estimate potentially omitted work effort
- Choose best normalization method given available info
- Remove effort due to Implementation

Resource normalization

- Remove effort due to resource level 4

Productivity In ISBSG Hints At Omitted Work Effort



Lower hours/UFP may hint at projects where not all effort was accounted for

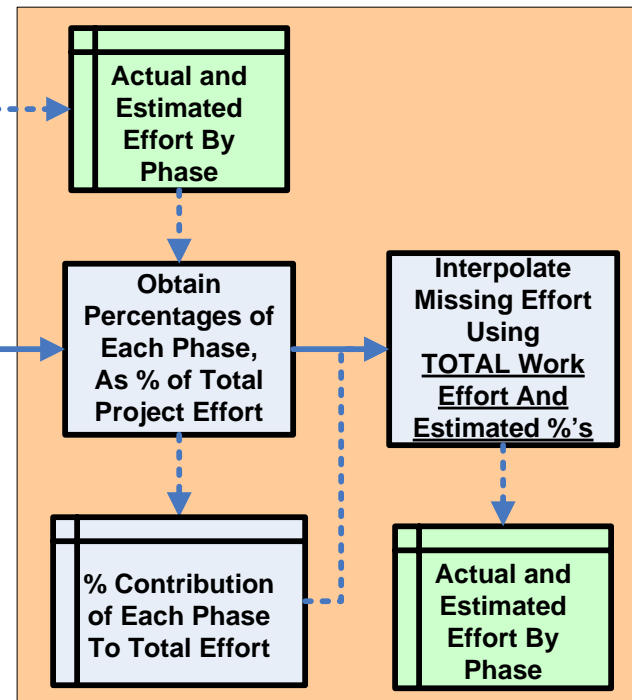
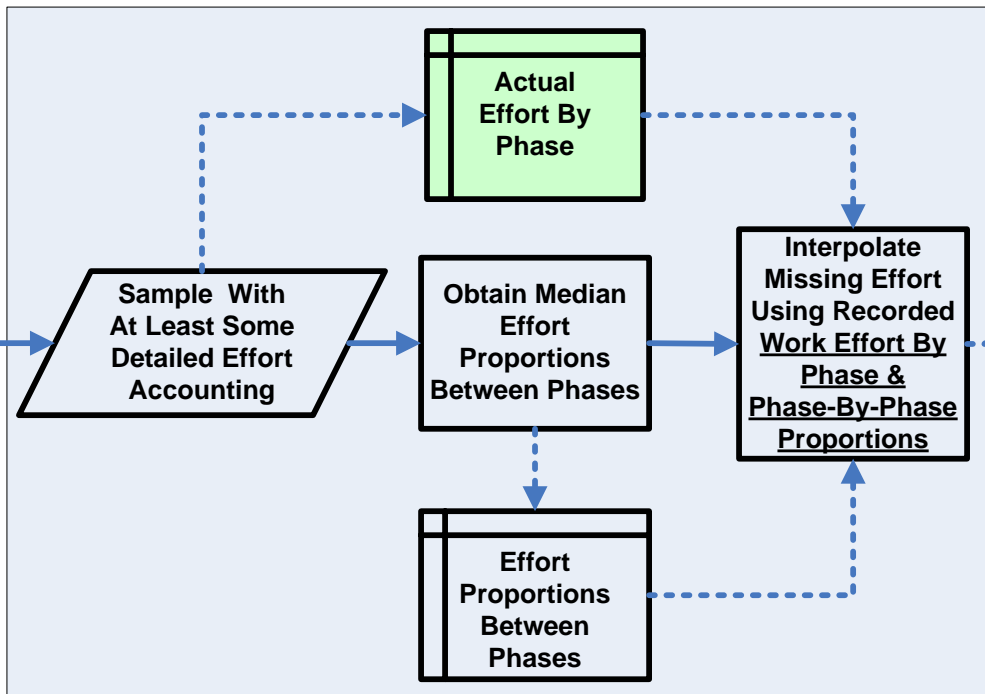
Calibration Process – Different Options Developed Depending on Circumstance



“Final” use in 43 instances

“Final” use in 81 instances

Work Effort By Phase (Partial Or Better)



“Phase Proportions” Method
Used when effort accounting by phase is present but incomplete

“Phase Percentages” Method
Used when Project Activity Scope field is present but list of phases is incomplete

Phase Proportions Calibration Method

“Missing from Detailed Effort By Phase”

Fill in effort for phases missing from the detailed phase accounting....

| Project Activity Scope | Effort Plan | Effort Specify | Effort Design | Effort Build | Effort Test | Effort Implement |
|--|-------------|----------------|---------------|--------------|-------------|------------------|
| Planning; Specification; Build; Test; Implement; | 100 | 400 | | 1000 | 200 | 150 |
| Planning; Specification; Build; Test; | | 281 | | 359 | 156 | |

....using the proportions established from reported phase-level work effort.

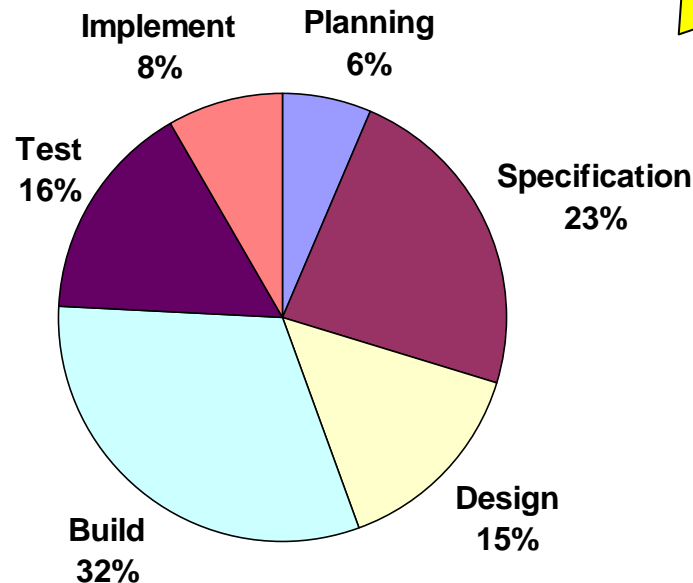
| Matrix summarizing combinations to left | | | | | | |
|---|---------|---------|-----------|----------|---------|---------|
| Numbers below are A / B | | | | | | |
| A | WE Plan | WE Spec | WE Design | WE Build | WE Test | WE Impl |
| WE Plan | 0.354 | 0.391 | 0.185 | 0.400 | 0.884 | |
| WE Spec | | 0.659 | 0.598 | 1.157 | 3.133 | |
| WE Design | | | 0.271 | 0.932 | 2.688 | |
| WE Build | | | | 2.000 | 6.182 | |
| WE Test | | | | | 2.613 | |
| WE Impl | | | | | | |
| B | WE Plan | WE Spec | WE Design | WE Build | WE Test | WE Impl |

Phase Proportions Calibration Method Resulting Effort Percentages By Phase

Actual, reported
effort by phase

Effort by phase
estimated from this
calibration method

Combined so that each record has either
actual or estimated effort by phase

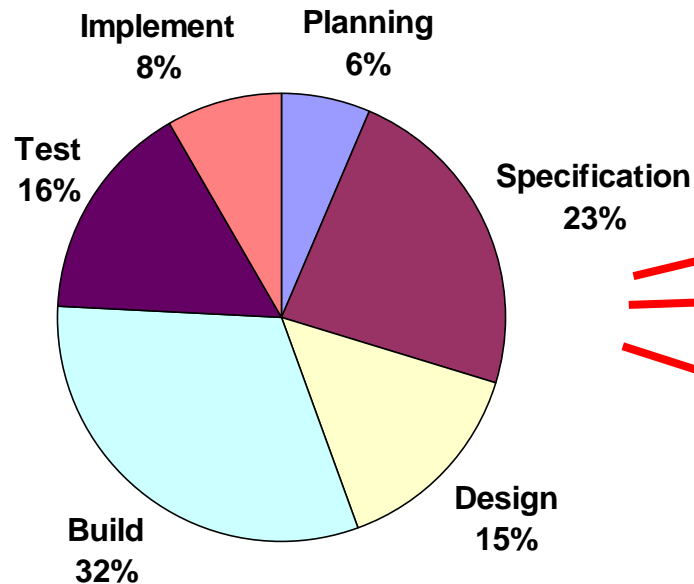


Phase Percentages Calibration Method

“Missing from Project Activity Scope list”

Using average percentages established from last method...

...*add back in* estimated effort for phases not listed as included.



| Planning | Specification | Design | Build | Test | Implement |
|----------|---------------|----------|-------|------|-----------|
| | | 273.2297 | | | 65.32259 |
| | | 117.5626 | | | |
| 1396.617 | | 3249.218 | | | 1453.017 |
| | | 2615.03 | | | 38.15955 |
| | | 68.67665 | | | |
| 17.20378 | | 40.02446 | | | 22.23922 |

Compare results of last page with other studies

| | ISBSG 02 | MIS 96 | NASA 93 | Cocom 00 | Grady 93 | Rubin 99 | Canada 94 |
|----------|-------------|-----------|------------|-------------|-------------|-------------|--------------|
| P | 10 | 13 | | | | | |
| S | 19 | 18 | 20 | 17 | 31 | 30 | 30 |
| B | 46 | 43 | 37 | 43 | 28 | 28 | 29 |
| T | 18 | 19 | 26 | 22 | 24 | 15 | 17 |
| I | 7 | 9 | | | | 14 | 14 |



6/8

23/17
(15/23 design)

32/30

16/13

8/9

**(proportion/
percentage
methods)**

Source: Ansell & Lokan, ACOSM 2002

Normalization Steps

Phase normalization

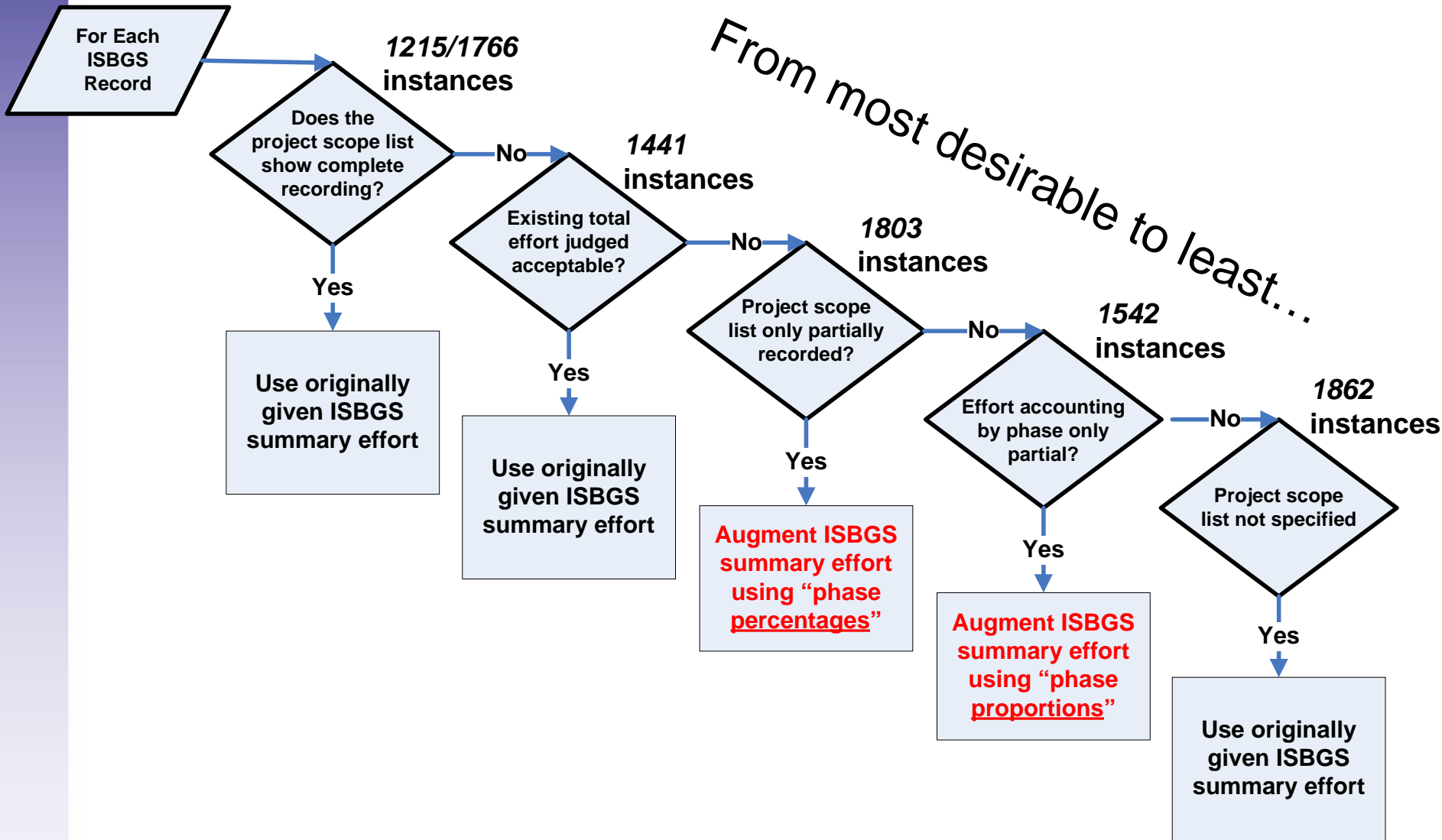
- Fill in potentially missing phases in “Project Activity Scope” field
- Estimate potentially omitted work effort
- Choose best normalization method given available info
- Remove effort due to Implementation

Resource normalization

- Remove effort due to resource level 4

The Normalization Measure Taken Depended On The Info Available

From most desirable to least...



Normalization Steps

Phase normalization

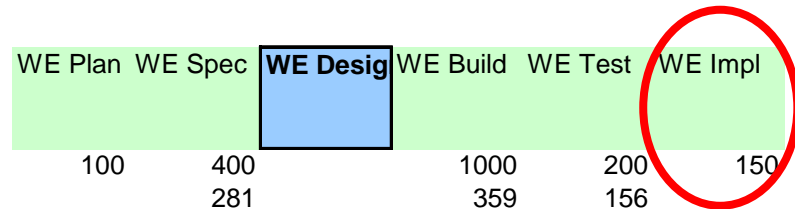
- Fill in potentially missing phases in “Project Activity Scope” field
- Estimate potentially omitted work effort
- Choose best normalization method given available info
- Remove effort due to Implementation

Resource normalization

- Remove effort due to resource level 4

Implementation Effort Removal Approaches

Best case: Explicitly given.



Otherwise: Use an estimated value.

Removal of Implementation phase effort is consistent with the normalization method used:

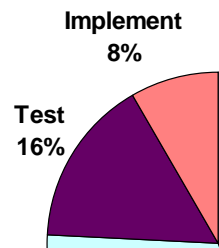
- Proportions between phases – Implementation removed based on its historically proportionate share, given the balance of other phases.

Matrix summarizing combinations to left
Numbers below are A / B

| A | WE Plan | WE Spec | WE Design | WE Build | WE Test | WE Impl |
|-----------|---------|---------|-----------|----------|---------|---------|
| WE Plan | 0.354 | 0.391 | 0.185 | 0.400 | 0.884 | 0.884 |
| WE Spec | | 0.659 | 0.598 | 1.157 | 3.133 | 3.133 |
| WE Design | | | 0.271 | 0.932 | 2.688 | 2.688 |
| WE Build | | | | 2.000 | 6.182 | 6.182 |
| WE Test | | | | | 2.613 | 2.613 |
| WE Impl | | | | | | 1.000 |
| B | WE Plan | WE Spec | WE Design | WE Build | WE Test | WE Impl |

- Percentages of each phase - relative to percentage given for Implementation

Backup slide: ISBSG's definition of phases



Normalization Steps Completed Thus Far

Phase normalization

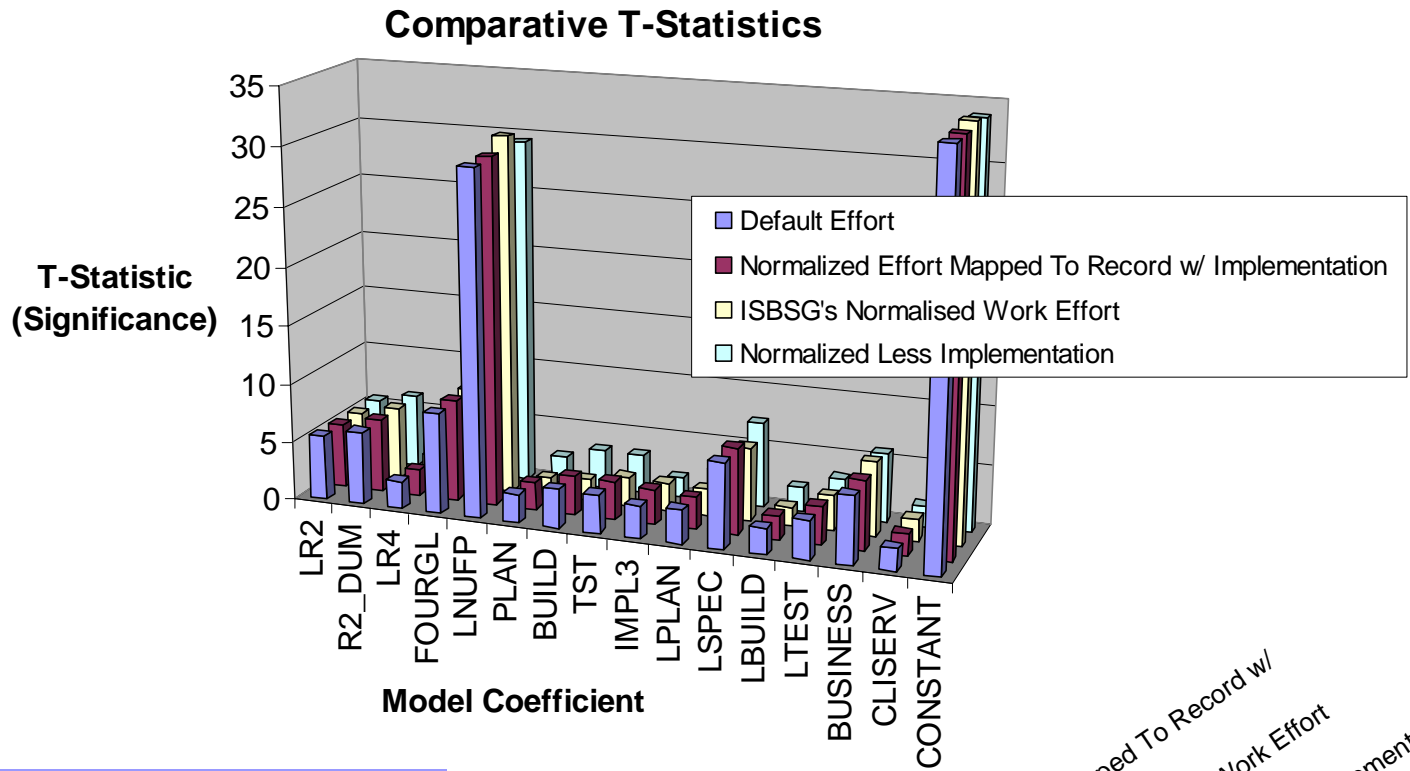
- Fill in potentially missing phases in “Project Activity Scope” field
- Estimate potentially omitted work
- Choose the most appropriate given available info
- Remove effort due to implementation

Time to review!

Resource normalization

- Remove effort due to resource level 4

Comparative Performance of the Different Effort Measures – Very Similar



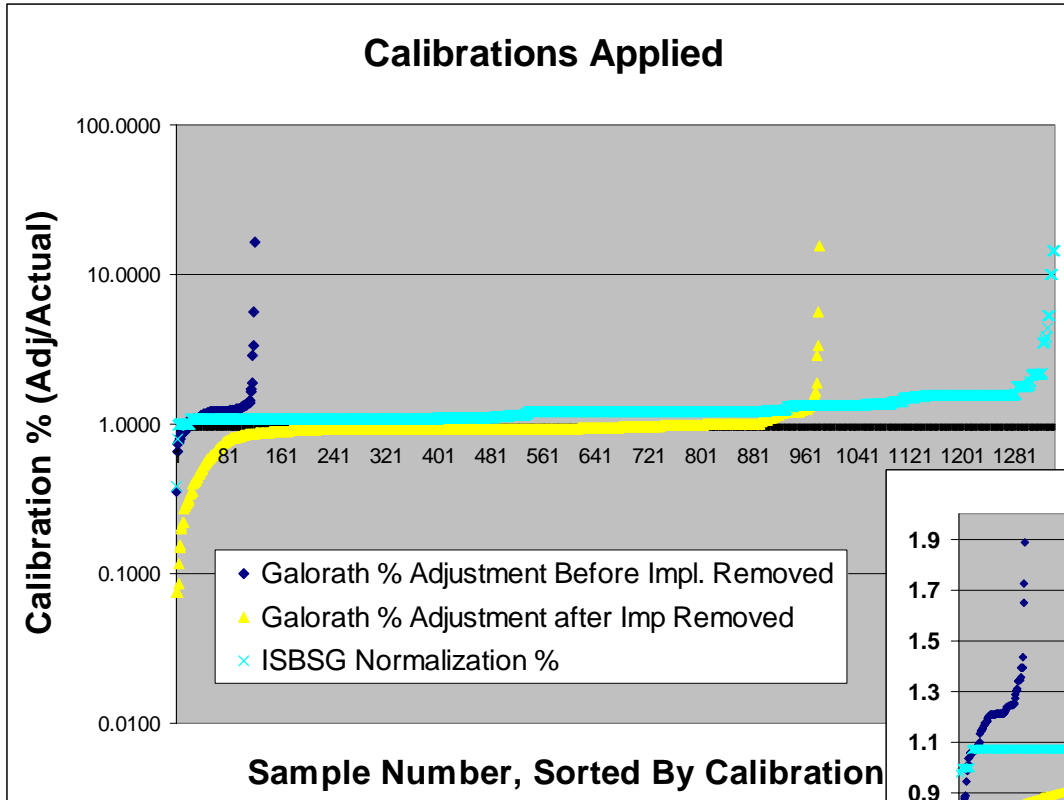
Probable reason:

Explainable portion of relationship is unchanged (all adj. R²'s are about .5)

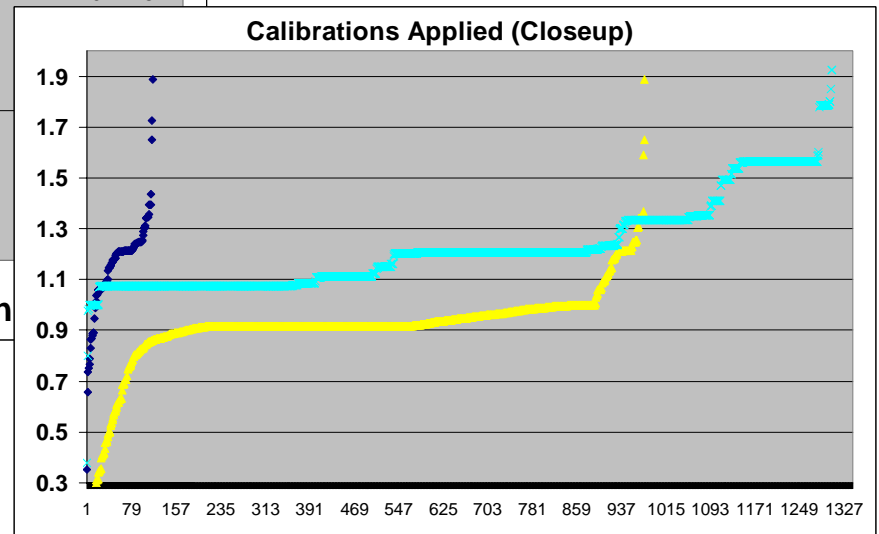
| # Times Best or Second Best | Default Effort | Normalized Effort Mapped To Record w/ Implementation | ISBSG's Normalised Work Effort | Normalized Less Implementation |
|-----------------------------|----------------|--|--------------------------------|--------------------------------|
| # Times Best or Second Best | 10 | 7 | 6 | 10 |
| # Times Best | 7 | 1 | 4 | 4 |

Normalization Statistics

ISBSG v9 Compared To Galorath



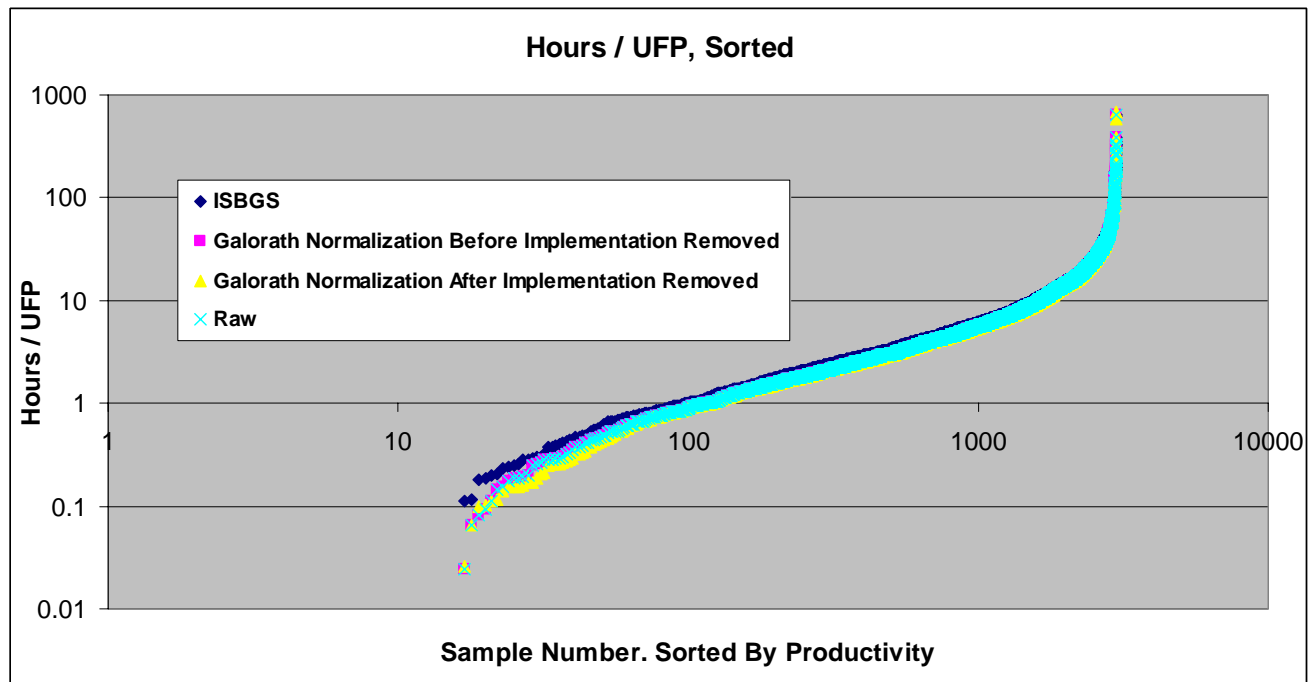
Out of 3024 records
(those not shown
received no calibration)



of adjustments
ISBSG Galorath
1359 998
Correlation
-9.7%

Impact of Normalization On Productivity

| | Galorath Normalization | | | |
|----------------|------------------------|--------|-------------------------------|------------------------------|
| | Raw | ISBGS | Before Implementation Removed | After Implementation Removed |
| Average | 16.377 | 18.179 | 16.377 | 16.153 |
| Median | 8.890 | 9.711 | 8.890 | 8.559 |
| Stdev | 28.439 | 31.992 | 28.439 | 30.461 |



Normalization Steps

Phase normalization

- Fill in potentially missing phases in “Project Activity Scope” field
- Estimate potentially omitted work effort
- Choose best normalization method given available info
- Remove effort due to Implementation

Resource normalization

- Remove effort due to resource level 4

Method Used For Resource Normalization: Regression

| VARIABLE NAME | ESTIMATED COEFFICIENT | T-RATIO | P-VALUE | DEFINITION OF VARIABLE |
|---------------|-----------------------|---------|---------|---|
| LR2 | 0.2842 | 5.52 | 0 | 1 when resource level is 2 or above * log UFPs |
| R2_DUM | -1.7423 | -6.25 | 0 | 1 when resource level is 2 or above |
| LR4 | 0.0436 | 2.15 | 0.032 | 1 when resource level is 4 * log UFPs |
| FOURGL | -0.4255 | -8.72 | 0 | 1 when 4GL language is used |
| LNUFP | 0.6960 | 29.36 | 0 | log of UFPs |
| PLAN | 0.6842 | 2.38 | 0.017 | 1 when the Plan phase is included, according to revised list of phases |
| BUILD | -1.3634 | -3.31 | 0.001 | 1 when the Build phase is included, according to revised list of phases |
| TST | 1.1291 | 3.26 | 0.001 | 1 when the Test phase is included, according to revised list of phases |
| IMPL2 | 0.0907 | 1.68 | 0.094 | 1 when the Implementation phase is included, according to ORIGINAL list |
| LPLAN | -0.1511 | -2.79 | 0.005 | PLAN * log of UFPs |
| LSPEC | 0.1249 | 7.33 | 0 | SPEC * log of UFPs |
| LBUILD | 0.1721 | 2.10 | 0.036 | BUILD * log of UFPs |
| LTEST | -0.2344 | -3.29 | 0.001 | LTEST * log of UFPs |
| BUSINESS | -0.2778 | -5.87 | 0 | 1 for non-critical Business projects |
| CLISERV | -0.1593 | -1.76 | 0.078 | 1 for non-critical Client-Server projects |
| CONSTANT | 4.0837 | 33.85 | 0 | |

Notes

Dependant variable was normalized effort, less implementation

Regressions were scale-adjusted (i.e., heteroskedasticity).

Coefficient values were highly stable given numerous alternative models.

Adjusted for all R² models hovered just below .5 in all cases (using weighted least squares)

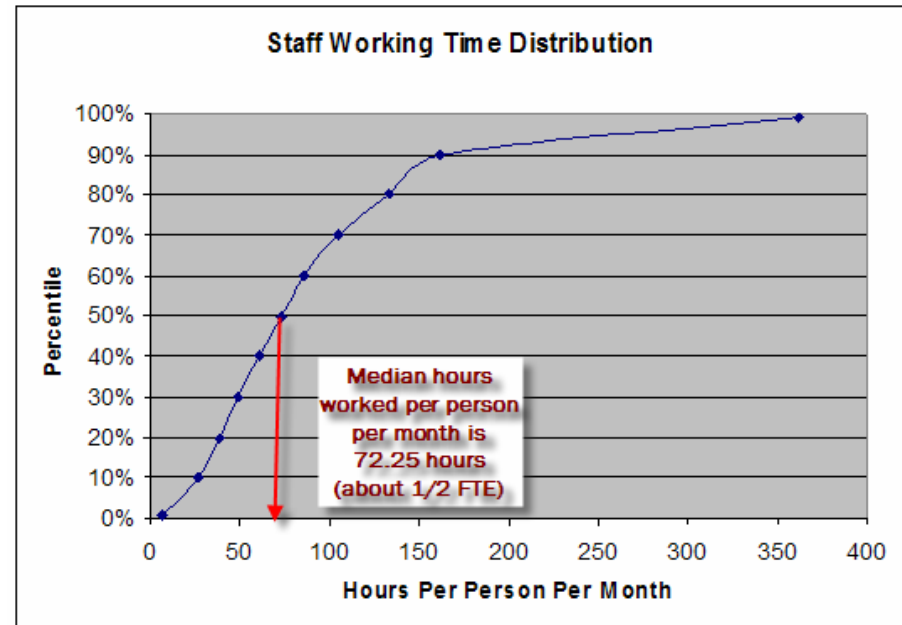
Adjustments to R4 were applied on a % basis (estimate without R4 / with R4)

Refining reported schedule

- It is desirable to use schedule data when calibrating
- 85% of project report schedule information
- **Inactive schedule check – removing unproductive time**
 - Elapsed schedule – inactive schedule should give a good measure of productive time spent
 - However, several projects (44) show inactive time within ½ calendar month of elapsed time
 - For those observations, elapsed time was used
- **Staff level check – if information provided, compute a schedule**
 - 41 observations reported no schedule, but did have staff (max=average) levels reported
 - Derived schedule = (Summary Work Effort/152)/ Staff Level

Checking staff levels

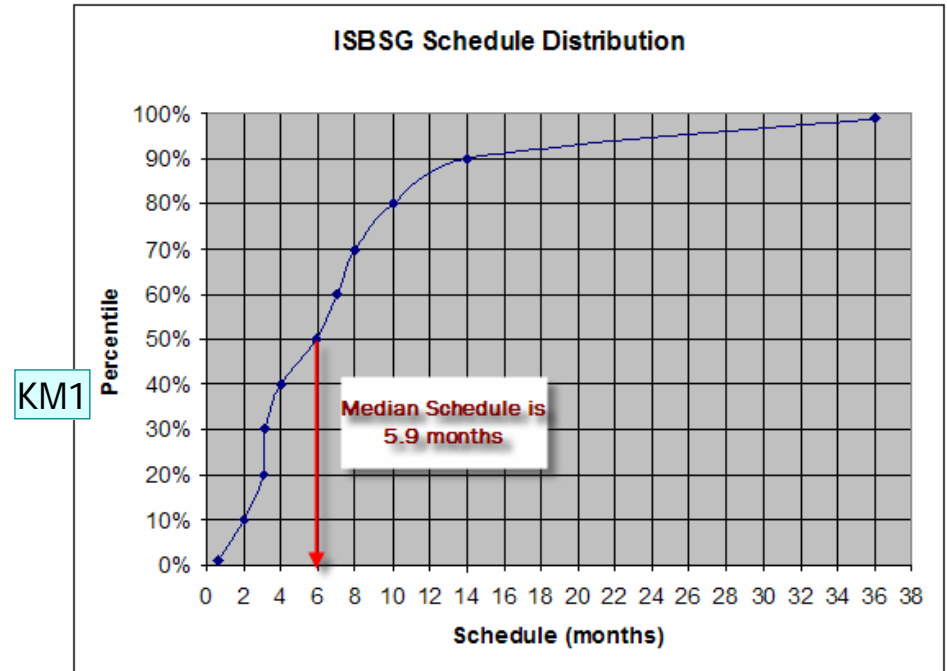
- **Staff levels can provide insight when calibrating**
 - ISBSG reports average and max staff levels for 1106 (37%) observations
- **Needed to determine if this data is appropriate for use in calibration**
- **Analysis of monthly work effort shows**
 - Reported staff levels were typically less than a FTE
 - Some extremes show average staff work effort to be greater than 2x an FTE
- **Reported staff levels were not used for calibration because:**
 - Inconsistent work effort by reported staff levels (some part time, some full time, some more)
 - Reported staff levels may not correlate with normalized work effort used for calibration
- **Instead, derived average staff levels were computed based on normalized effort and schedule**



Hours Per Person Per Month = Summary Work Effort / Average Team Size / Schedule

Checking extremes

- **98% of observations were between .6 months and 36 months**
- **Short bursts**
 - Checked for schedules < 1 month which reported more than 10 people
 - Detected 11 of these cases, did not use schedule
- **Long slogs**
 - Checked for schedules > 2 years and less than 1 FTE
 - Detected 18 of these cases, schedule data deemed not reliable for calibration



KM1

There were 15 of these using the ISGSG normalized work effort.

Karen McRitchie, 7/12/2006

Conclusion / Wrap Up

- We found anomalies in the ISBGS data set, although not many
- We have a consistent set of qualitative descriptors from which to categorize data trends
- Missing phases may not have been omitted, but rather reported in other phases
- Phase proportions after normalization are consistent with other studies
- Fewer calibrations, or none, can achieve similar results in estimating (probably because the majority of records are well specified)
- Observations calibrated did not match observations adjusted by the ISBSG normalized work effort
- The data set is prepared for analysis for estimation tool knowledge base update (along with other datasets)



Backup Slides

ISBSG's Definition of Phases

| Project Phase | Possible Phase Components | |
|------------------|--|---|
| Plan | Preliminary Investigations Overall Project Planning Feasibility Study | Cost Benefit Study Project Initiation Report Terms of Reference |
| Specify | Systems Analysis Requirements Specification Review & Rework Requirements | Spec. Architecture Design/Specification Review & Rework Architecture Spec |
| Design | Functional / External Design Create Physical / Internal Design(s) | Review and Rework Design(s) |
| Build | Package Selection Construct Code & Program Software Review or Inspect & Rework Code | Package customisation / interfaces Unit Test Integrate Software |
| Test | Plan System or Performance Testing System Testing Performance Testing | Create & Run Automated Tests Acceptance Testing |
| Implement | Prepare Releases for Delivery Install Software Releases for Users Prepare User Documentation | Prepare & Deliver User Training Provide User Support |

ISBSG Description of Resource Levels

ISBSG Description of Each Resource Level

1 = development team effort

project team, project management, project administration

2 = development team support

database administration, data administration, quality assurance, data security, standards support, audit & control, tech support

3 = computer operations involvement

software support, hardware support, information centre support, computer operators, network administration

4 = end users or clients

user liaisons, user training time, application users and/or clients

What Does “Implementation” Include?

- **Prepare Releases for Delivery**
- **Install Software Releases for Users**
- **Prepare User Documentation**
- **Prepare & Deliver User Training**
- **Provide User Support**



Driven by factors beyond functional size (users, locations, business units..)

Acknowledgement

- **Thanks to Peter Hill of ISBSG for answering several questions about the ISBSG data set and providing detail on the ISBSG normalization process**

Platform knowledge base

- **SEER-SEM platform knowledge base generically describes the operational platform**
- **ISBSG fields used to determine platform**
 - Development Platform (Mainframe, Midrange, PC)
 - Architecture (Standalone, Client-Server, Multi-Tier, Web)
 - Business Area Type
 - Application Type

Application knowledge base

- **SEER-SEM application knowledge base describes the general application type**
- **ISBSG fields used to determine application**
 - Business Area Type (used to set defaults for those that did not have details)
 - Application Type
- **Most applications readily mapped to existing SEER-SEM categories**

Acquisition Method knowledge base

- **SEER-SEM acquisition method knowledge base reflects reuse scenarios**
- **ISBSG fields used to determine application**
 - Development type (new, enhancement)
 - Function counts (added vs. changed vs. deleted)

Development method knowledge base

- **SEER-SEM development method knowledge base describes the development method or paradigm**
- **ISBSG fields used to determine application**
 - Development techniques
 - Case tool used
 - Package Customization and Degree of Customization