Education PPP: an Output-Oriented Approach

$$Exp = PPP \times Quantity \times Quality$$

Tertiary

Vocational

Secondary

Primary

No schooling
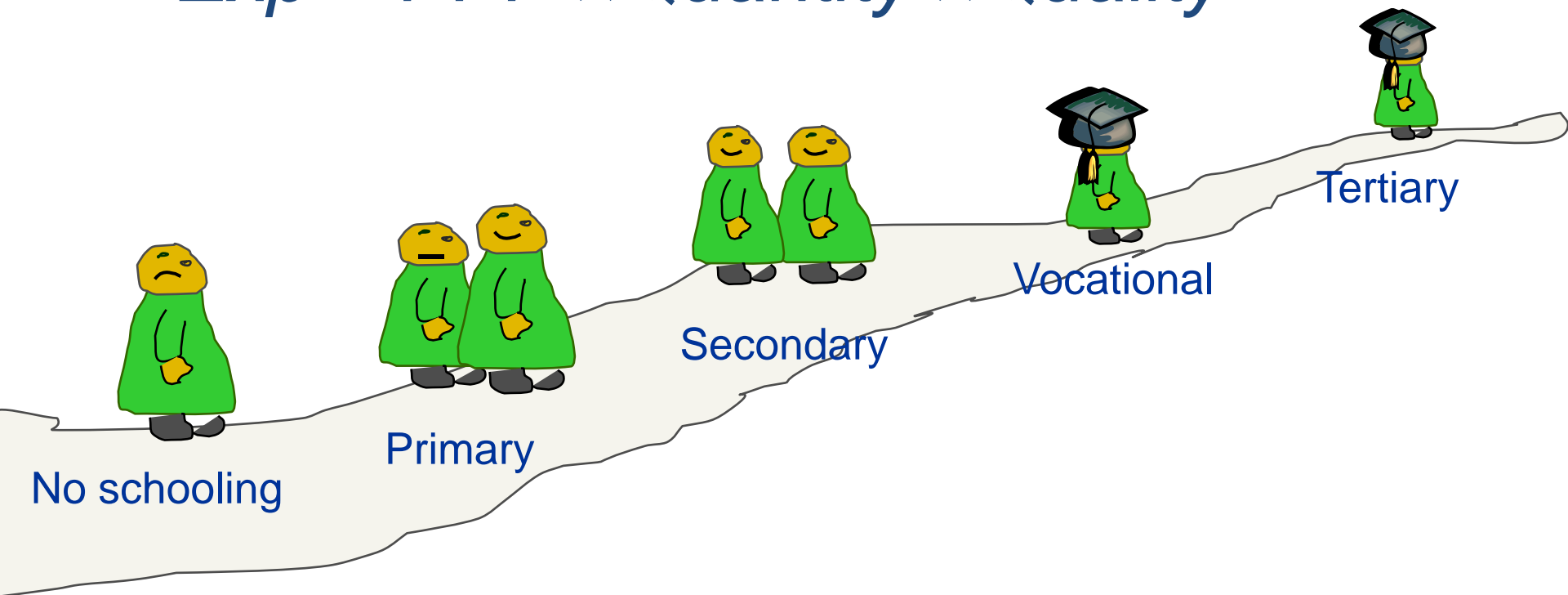
Outline: 1. Quantity
2. Quality
3. Conceptual model

No schooling

Primary

Secondary

Vocational

Tertiary
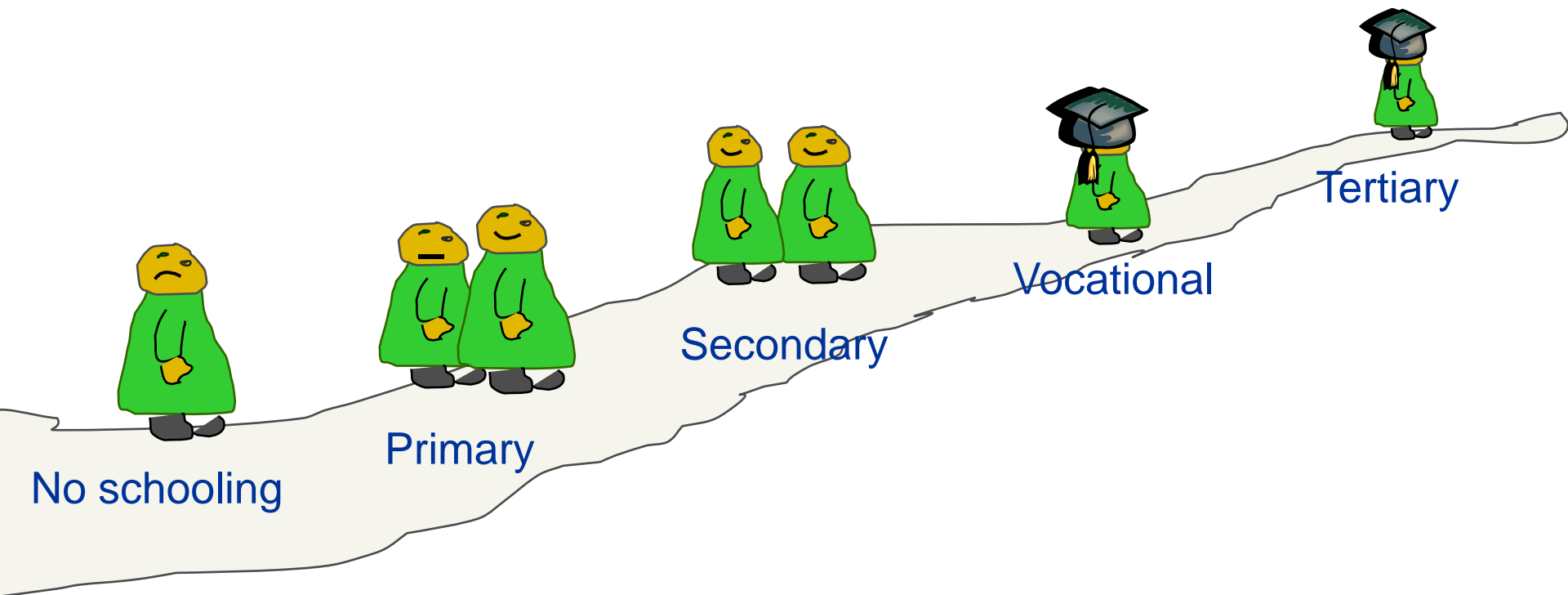
Measuring the **Quantity** of Education

- Basic element is **pupil counts**: how many pupils consumed education services?

- Enrollment figures generally available from national Education Management Information Systems (EMIS), but must be validated

- **Actual services** consumed by pupils calculated using a time-on-task measure of the pupil-hours volume of instructional activity

Validating enrollment figures (method 1)

- Adjust for school coverage (are all schools included in National EMIS systems?)

- Sample survey of schools

- Compare to time series trends

- Adjust for pupil expenditures

- Ensure all sub-sectors report and look for consistencies

Validating Enrollment Figures (method 2)

- Compare Enrollments from EMIS and Household Surveys, such as DHS

- Differentiating between Enrollment and Attendance (EMIS captures enrollment, household surveys – attendance)

- Use Verification Protocols developed by AED, UIS, UNICEF and UNESCO

Method 2 verification protocols

- Use UIS data when no survey MICS/DHS data are available

- Use MICS/DHS when no UIS data available

- Use UIS when data is < 5% difference,

- Use expert when data is > 5 % difference

Adjust Quantity by Estimating Pupil Hours

- Subtract Days Lost from Official School Year

- Collect Teacher Absence Data and Roll-up

- Collect Student Absence Data and Roll-up

- Calculate Daily Time Loss

- Survey Time on Task

- Sample Schools and Estimate Absence Data across Stages/Levels/Organization

Stratification

- Disaggregation by ISCED levels (some national education systems do not align neatly w ISCED)

- Possibly, disaggregation by:
  - Public
  - Private
  - Complementary models (religious, community schools)

Measuring the <u>Quality</u> of Education

- Moving towards a common learning metric

- Composite Learning Scores in Education PPP

- **Data imputation** will help address the problem of missing quality scores
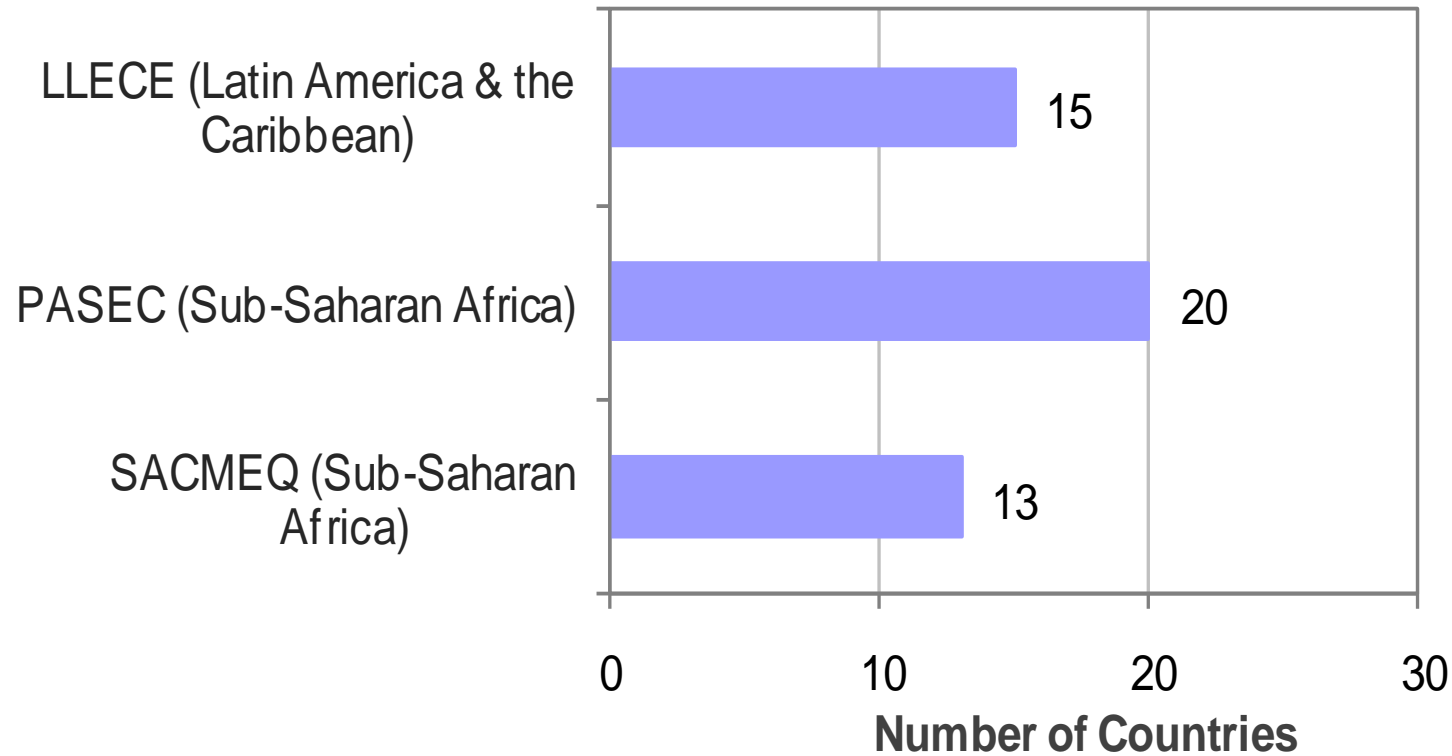
Why impute learning scores in education PPP?

- A fair assessment of value is inevitably linked to outputs
- In education, outputs are learning outcomes, measured by assessments of student achievement
- Participation in assessments is growing, but far from global
- Imputation of missing values is a way of accounting for quality in a broader group of countries than those currently available
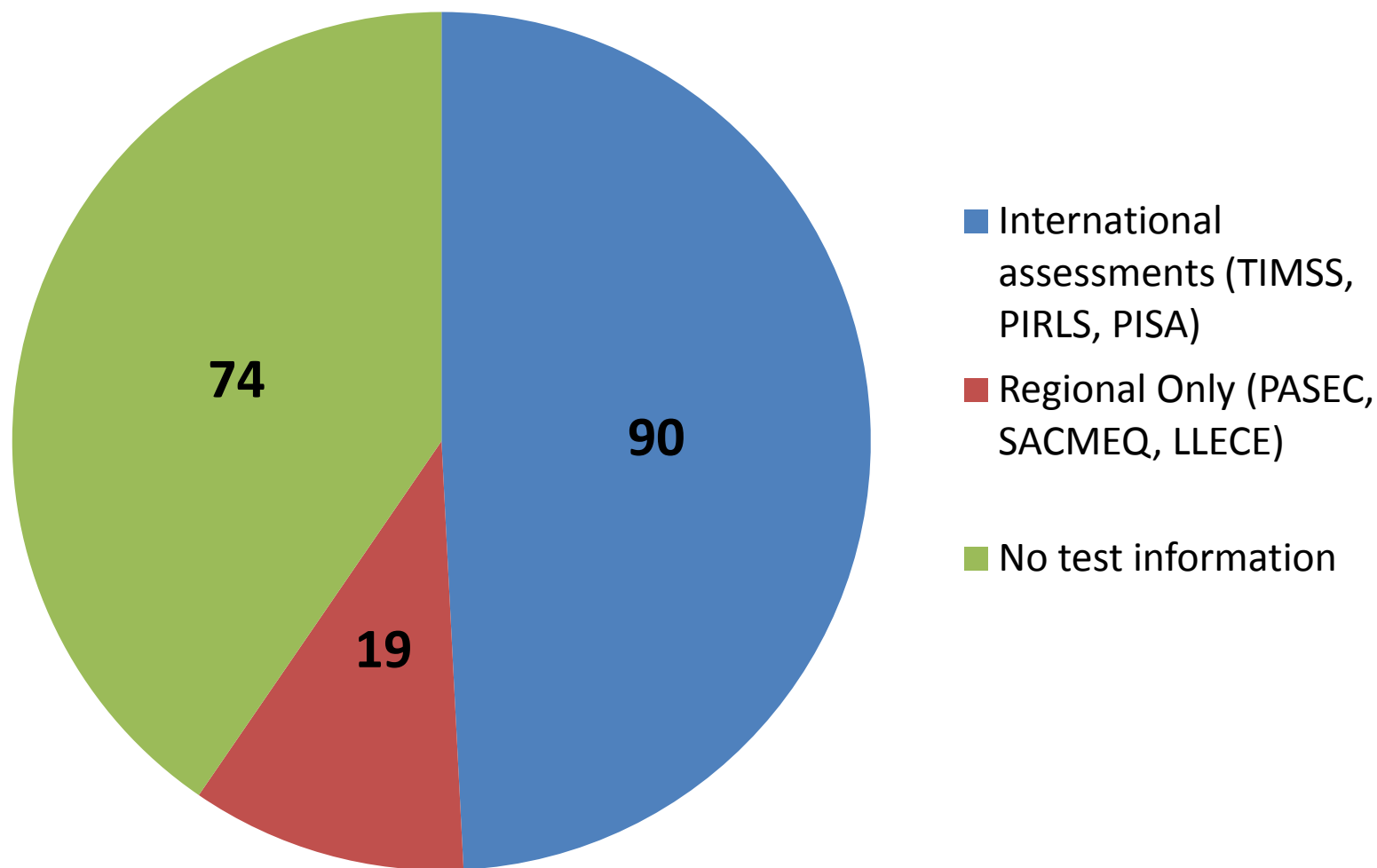- International and regional assessments will be used.

International Achievement Studies: Regional Coverage
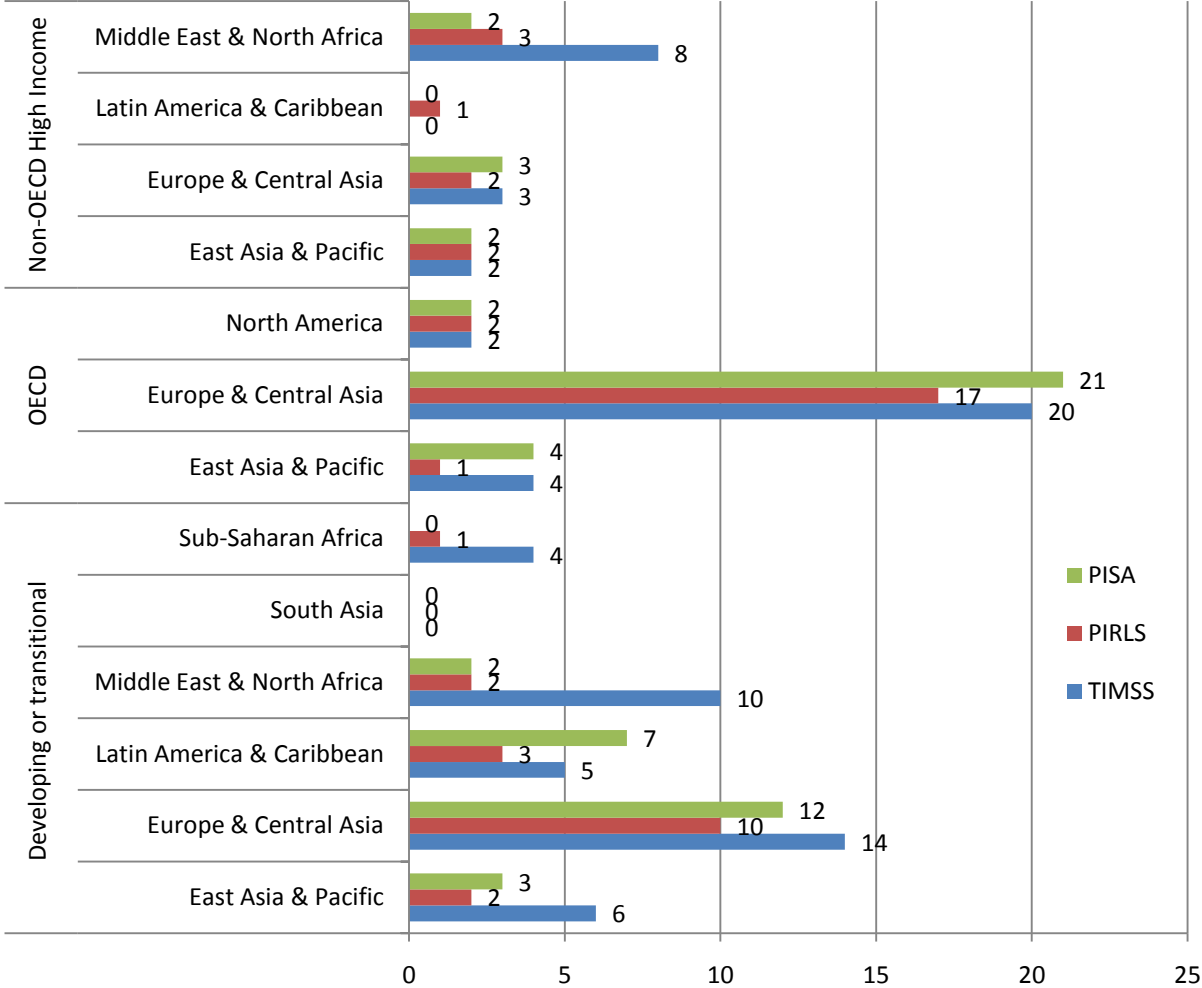
Additional Sources of Data: Regional Studies

Availability of Student Achievement Scores: ALL ICP PARCIPATING COUNTRIES (N=182)

■ International assessments (TIMSS, PIRLS, PISA)

■ Regional Only (PASEC, SACMEQ, LLECE)

■ No test information

Education Policy & Data Center

Towards a Common Quality Adjustment

- Data imputation is necessary to include all countries in a quality adjustment process
- All available information must be used
- Reliability and validity will be tested

Imputation Methods: Some Examples

- Crouch (Crouch and Vinjevold, 2006)
  - Recursive regression imputation in several steps, based on correlation with TIMSS 1999 scores
  - Univariate regressions with target learning score (TIMSS 1999) as dependent variable, one assessment at a time as sole predictor
  - Result: several predicted scores for some countries, weighted average taken as imputed value
  - Weights based on correlation with TIMSS 1999

- Hanushek and Woessmann (2009, 2010)
  - "Calibration" of major international achievement studies into a new scale.
  - Based on U.S. performance in international studies, and NAEP as reference point.
  - Mean and standard deviation of 13 OECD countries

Multiple Imputation (MI)

- Long-established methodology for treating missing data (for overview see Rubin, 1996; Allison, 2002)

- Allows to use all available information

- Accounts for uncertainty associated with imputation (multiple estimates generated)

- Imputes more than one variable with missing data at a time

- Drawback: demands on technology/ time to produce reliable estimates.

Imputation of Learning Scores: the Process

- Define dataset: three "core" assessments (state-of-the-art) and predictors of education quality

  – GDP per capita

  – Age structure (percent youth)

  – Per pupil expenditure

  – Teacher-student ratio

  – Other variables known or theorized to predict quality

- Missing values on predictor variables will also be imputed

Multiple Imputation (MI): the Mechanics

- Conditional prediction models are specified for each of the variables with missing data

- MI runs iteration cycles "filling in the blanks" in all of the variables across 3-5 chains, based on the specified models

- The goal is to achieve convergence of the distributions of imputed values across the chains, thereby increasing reliability

- Estimates are then pooled across the chains

- The result is 3-5 completed (imputed) datasets

The Composite Learning Score: Stages

- Because of the high proportion of missing data on learning assessments, the construction of the learning scores will be done in stages

- Stage 1: Countries with international assessments: Impute all three "core" assessments, construct single score from weighted average (90 out of 183)

- Stage 2: Countries with regional assessments: impute one composite learning score (19 countries)

- Stage 3: Countries with no assessments, but complete data on predictors of quality: impute one composite learning score (TBD: total is 74)

- Actual or imputed learning score will be transformed to a specified scale (currently Mean= 500; SD=100)

Sources of Error and Threats to Validity

- Unequal, nonrandom geographic representation in international assessments

- Regional assessments and EGRA lacking in standardization, hence lower reliability

- Large proportion of ICP countries (74) never participated in learning assessments, hence quality scores will be derived solely based on inputs and proxies.

- MI will account for some of the uncertainty (averaging out across multiple estimates)

- Crude methods (e.g. single imputation using regression or predictive mean matching) may be applied for difficult cases.

After the Imputation: Adjustments

- Once in the PPP model, learning scores will be adjusted for student SES and length of schooling

- **SES Index**: variables from available international achievement databases, as well as household surveys

- **Length of schooling**: weighted average of years spent through graduation from secondary, based on the proportion attending preschool

After the Imputation: Adjustments

- Once in the PPP model, learning scores will be adjusted for student SES and length of schooling

- **SES Index**: variables from available international achievement databases, as well as household surveys

- **Length of schooling**: weighted average of years spent through graduation from secondary, based on the proportion attending preschool
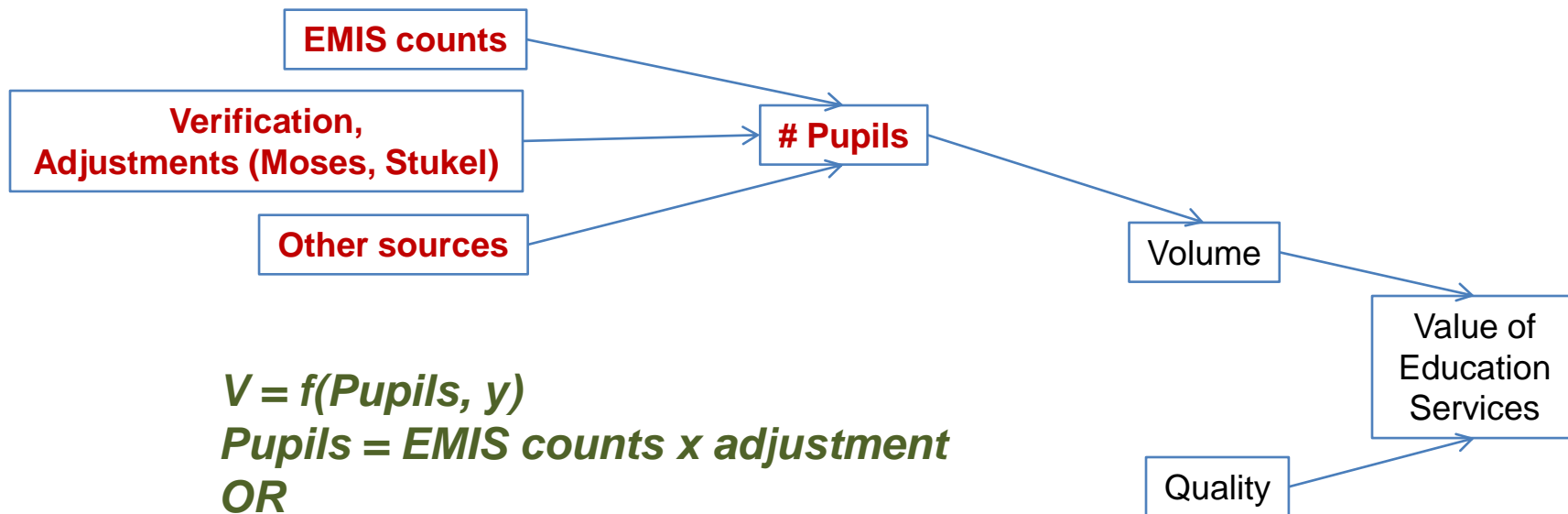
Conceptual model

Quantity

Value of Education Services
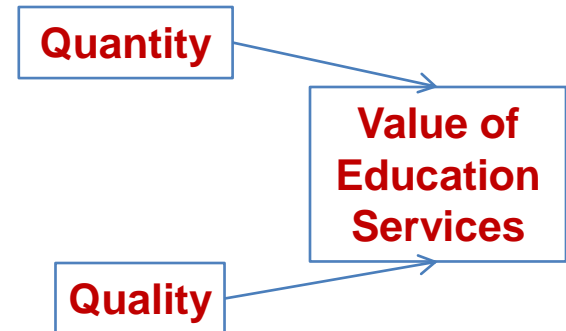
Quality

$$Value\ of\ Education\ Services = \ Quantity \times Quality$$

$$Expenditures = PPPadjustment \times Quantity \times Quality$$

EMIS counts

Verification,
Adjustments (Moses, Stukel)

Other sources

Pupils

Volume

Value of
Education
Services

Quality

$V = f(Pupils, y)$
$Pupils = EMIS\ counts\ x\ adjustment$
OR
$V = HH\ survey\ attendance$

EMIS counts → # Pupils

Verification, Adjustments (Moses, Stukel) → # Pupils

Other sources → # Pupils

Pupils → Volume

Time on Task → Volume

Volume → Value of Education Services

Quality → Value of Education Services
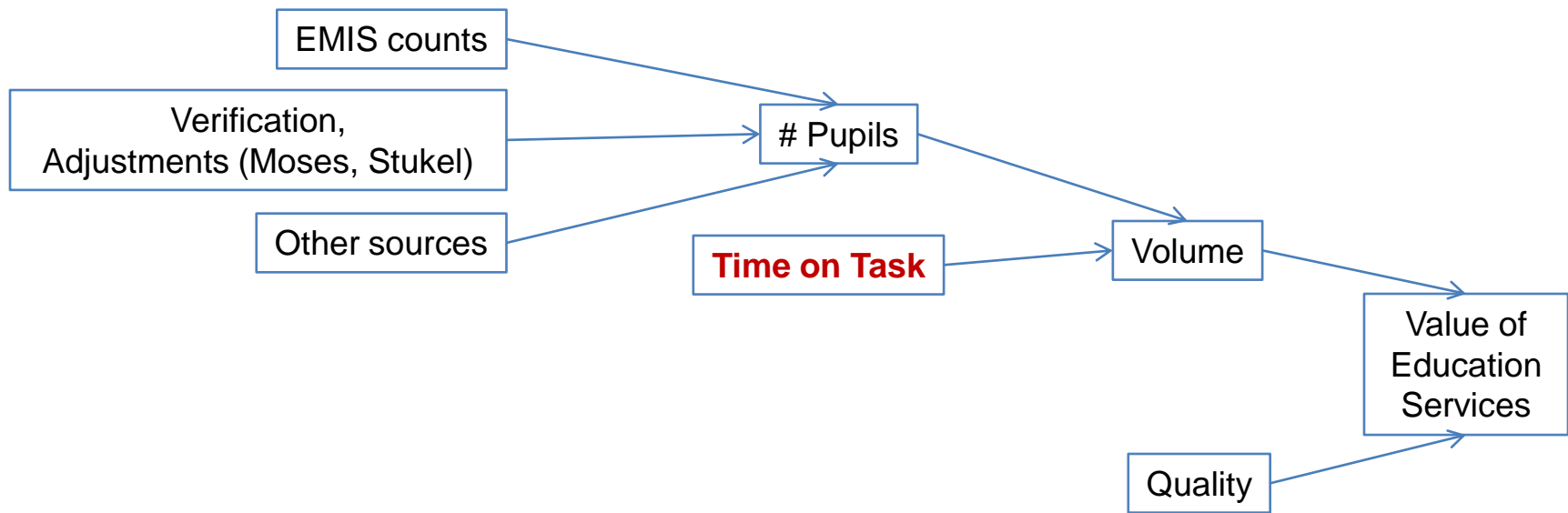
$$Quantity = fulltime\ equivalent\ of\ Pupils(ftP)$$

$$\text{Refined}: \quad Quantity'$$
$$= ftP \times \frac{actual\ instruction\ time\ (time\ on\ task)}{benchmark\ annual\ instruction\ time}$$
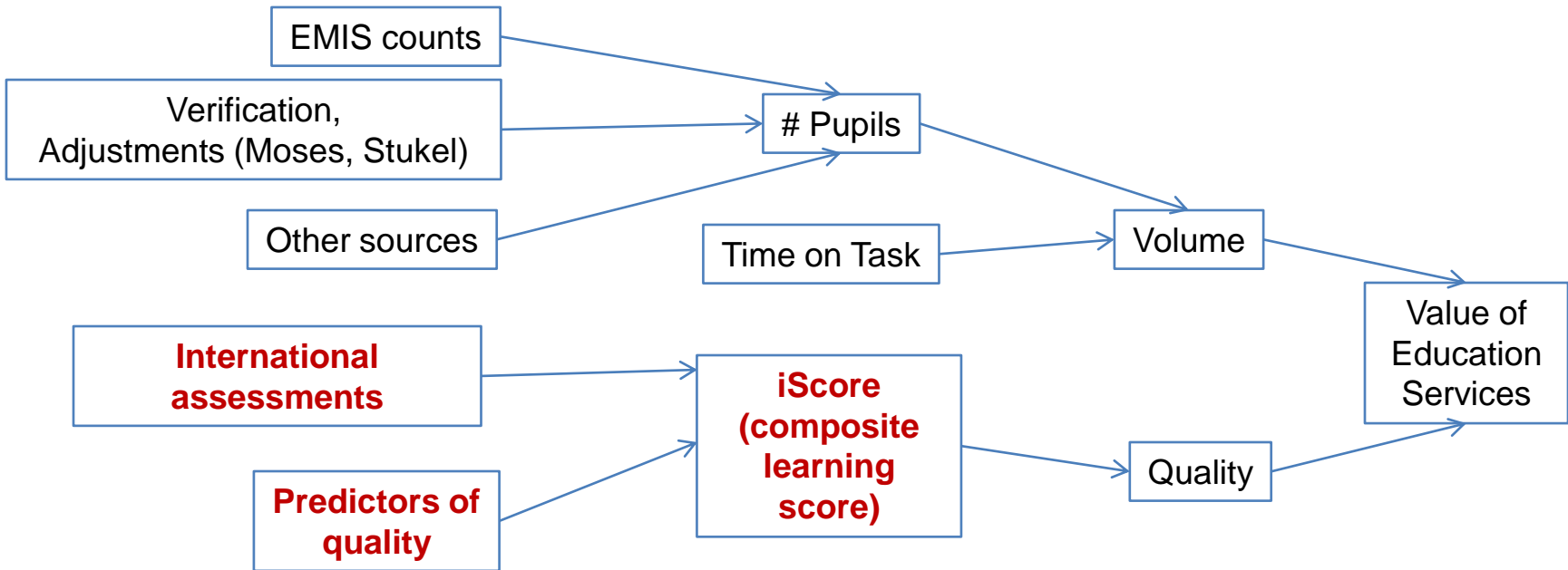
$$iScore = f(learning\ scores, inputs, context)$$

$$Quality(0) = (\partial \times iScore)$$

EMIS counts

Verification,
Adjustments (Moses, Stukel)

Other sources

Pupils

Time on Task

Volume

International assessments

iScore
(Crouch, MI)

GDP, other

Quality

Value of
Education
Services

HH surveys

SES adjustment
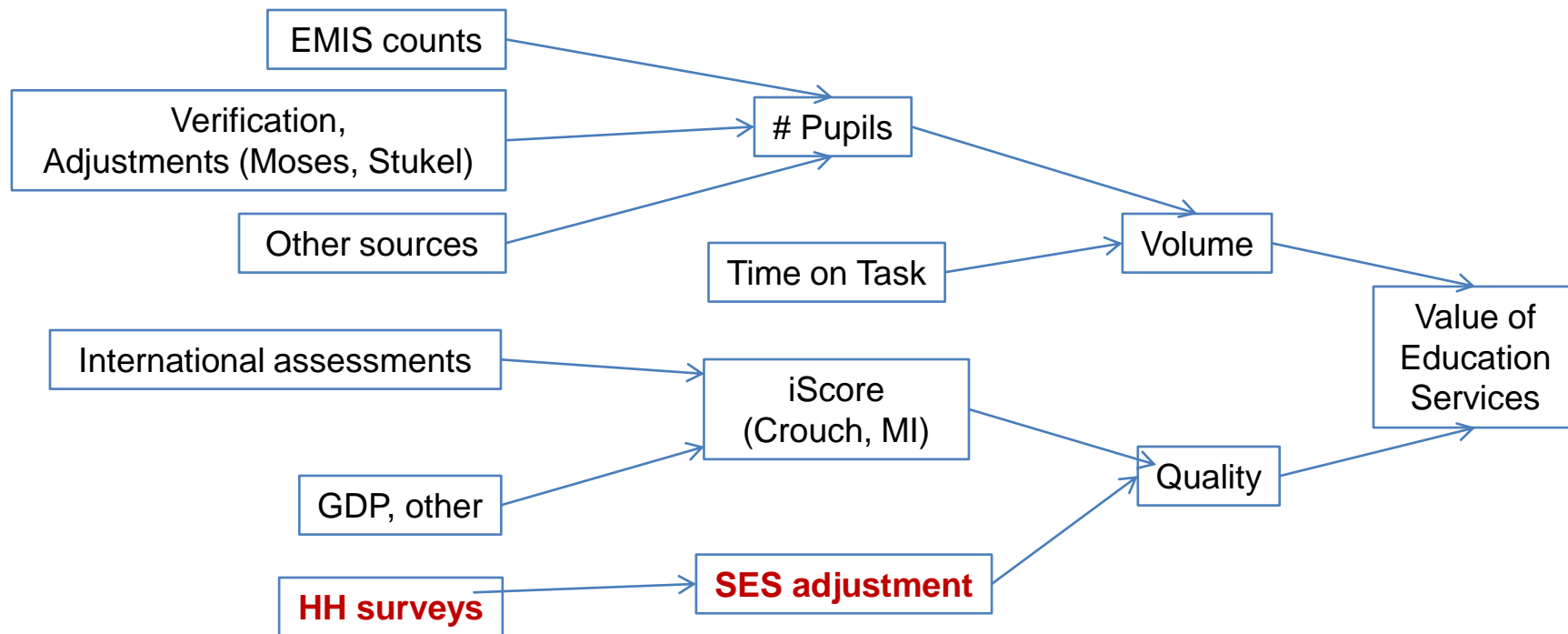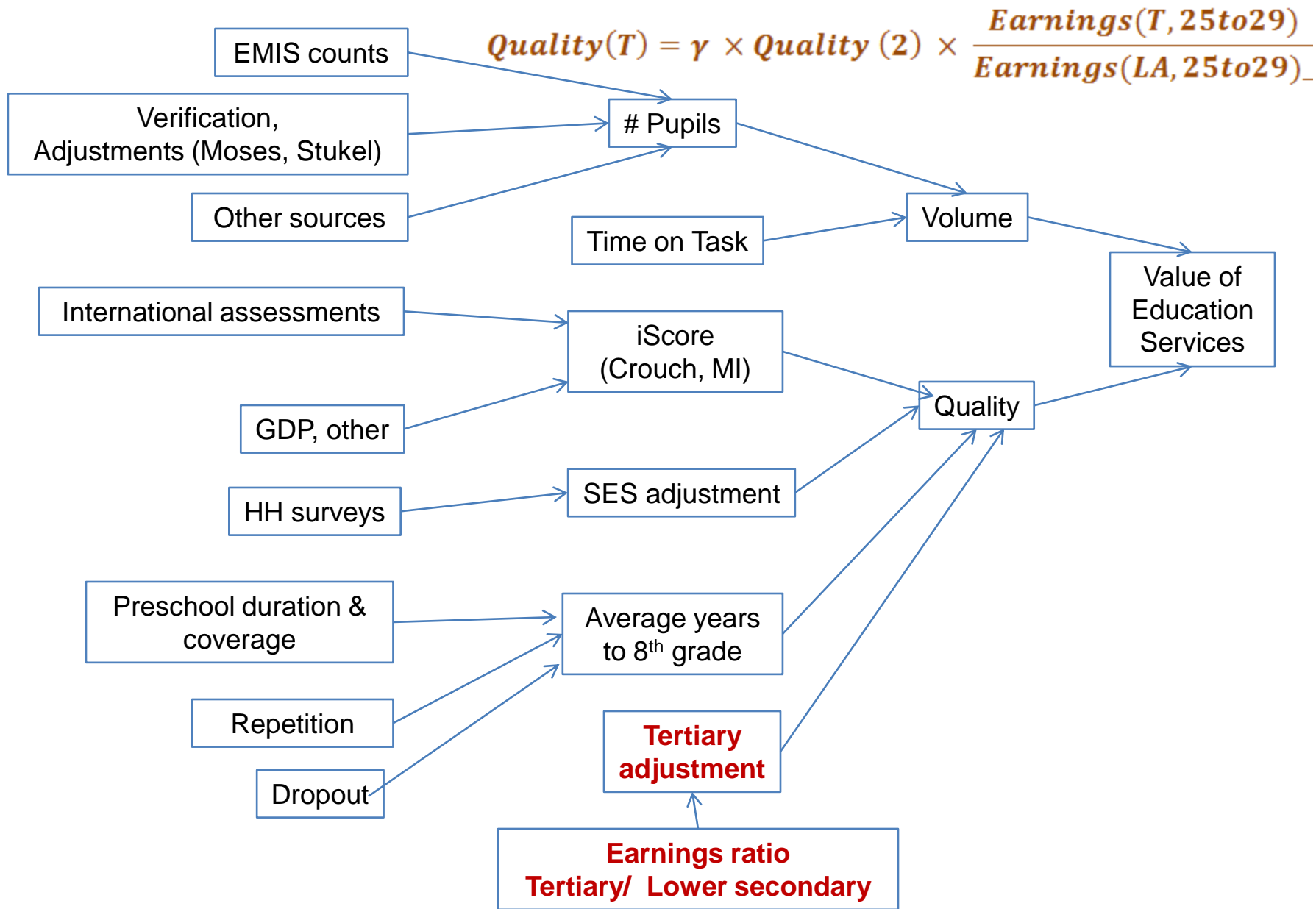
$$Quality(1) = (\partial \times iScore) \times (\propto \times HC)$$
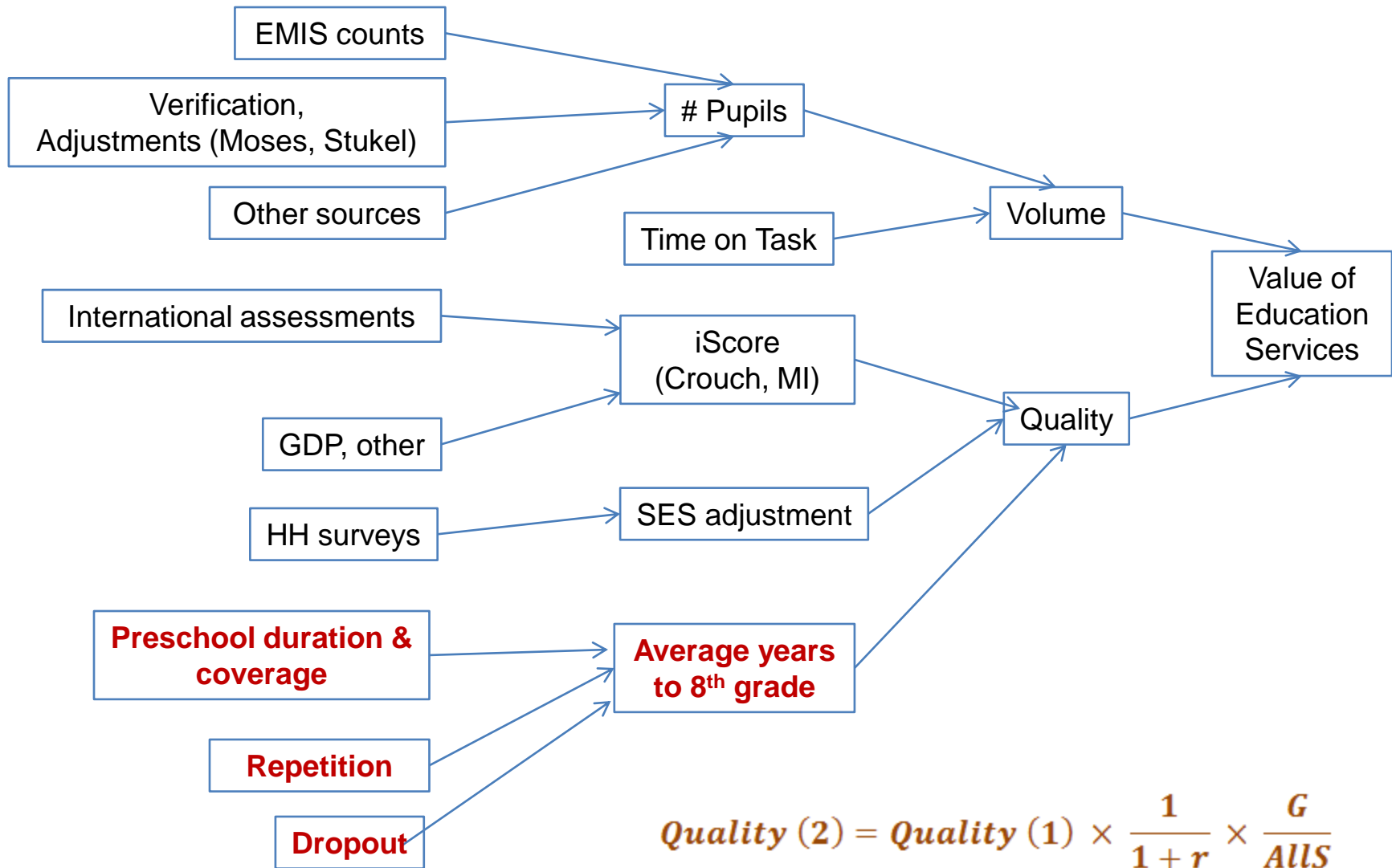
$$Quality\ (2) = Quality\ (1) \times \frac{1}{1+r} \times \frac{G}{AllS}$$

$$Quality(T) = \gamma \times Quality\ (2) \times \frac{Earnings(T, 25 to 29)}{Earnings(LA, 25 to 29)}$$

EMIS counts

Verification,
Adjustments (Moses, Stukel)

Other sources

Pupils

Time on Task

Volume

International assessments

iScore
(Crouch, MI)

GDP, other

Value of
Education
Services

SES adjustment

HH surveys

Quality

Preschool duration &
coverage

Average years
to 8th grade

Repetition

**Tertiary
adjustment**

Dropout

**Earnings ratio
Tertiary/ Lower secondary**

Further notes

All calculations by ISCED level
Private, public, religious may be aggregated
Q applied to preschool – secondary
Q(t) separate for tertiary
Data on key inputs available for most countries

Pilot to include:
EMIS count verifications
Time on task estimates
Learning assessments (?)

The Roadmap

1. (ICP/Regional and External) Refine proposed conceptual model of output methodology into an implementable instrument

2. (External) First-level secondary data assessment and collection (for model-testing phase)

3. (External) iScore construction: imputation, adjustment.

4. (External) Identify insufficient data coverage on important predictors, develop plan to collect missing data.

The Roadmap, cont'd

5. (Local/regional and external) Train local or regional technical teams to verify pupil counts, implement the sample survey for EMIS information verification.

6. (Local/regional and external) Train local or regional technical teams to assess actual instruction time, implement survey to collect missing information on instruction time.

7. (Local/regional) Assess whether EMIS counts or household survey estimates are the more reliable source of pupil count information; verify EMIS counts

8. (Local/regional) Implement assessment and surveys of actual instruction time

The Roadmap, cont'd

9. (External, but can train regional experts) Collect SES, repetition, dropout, and preschool attendance data for all ICP countries where such information is available.

10. (External, but can train regional experts) Calculate Quality and Quantity for all countries with complete dataset.

11. (External) Rerun the imputation models with newly collected information, in particular, time-on-task.

The Roadmap, cont'd

12. (External) Evaluate the value added of time of task, and the difference from validated pupil enrollment numbers.

13. Analyze and process the information into PPP adjustment values for the piloted countries, to evaluate the process, and to propose adjustments where necessary for a full-scale implementation of the output approach to education PPPs.