Joint UNECE/EUROSTAT
Work Session on
Statistical Metadata

Enhancing User Access to Statistical Metadata
on the Internet

Robert Di Calogero, Paul Austin, Gary Jones, and Michel Le Marois
Statistics Department, International Monetary Fund

Luxembourg
March 6–8, 2002

Opportunities for Enhancing User Access

The IMF aims to:

- Leverage the SDDS format as a global standard for describing international financial and economic statistics

- Leverage the value of the investments made by SDDS subscribers and GDDS participants in developing the metadata posted on the DSBB

- Leverage the wealth of statistical data and metadata that is linked to the DSBB

Improve DSBB Functionality by:

- Enhancing metadata content management
- Developing a common language (XML) for communicating and managing DSBB metadata content
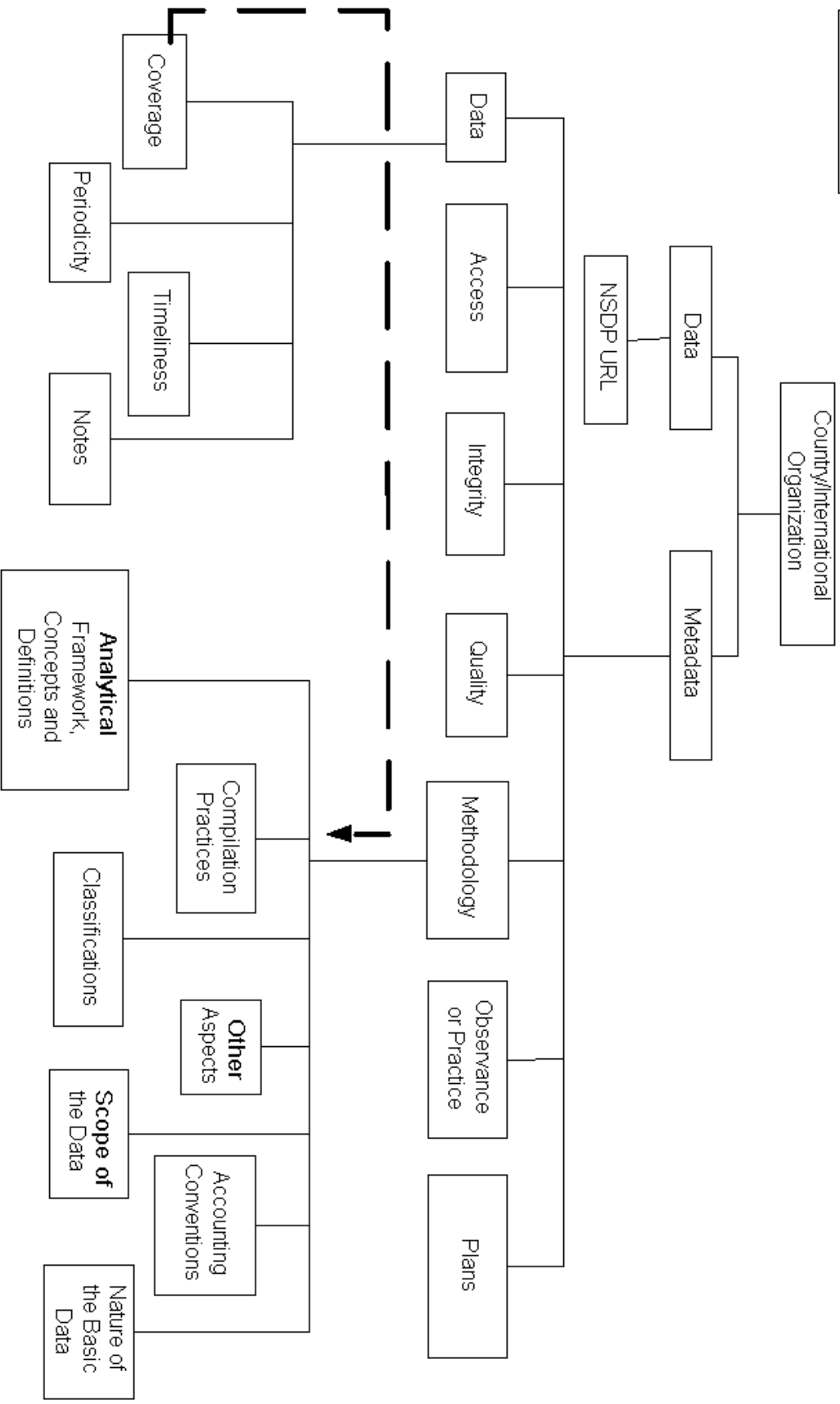- Providing interactivity and data query facilities through improved metadata exchange capabilities

Open Exchange of DSBB Metadata:
Key Elements

- DSBB metadata model
- Rendering of the DSBB model in XML
- Creating a RDBMS that reflects the structure of all DSBB metadata elements
- Deploying a Content Management System to automatically manage DSBB metadata on the web
- Permitting personalized content delivery through content aggregation technologies
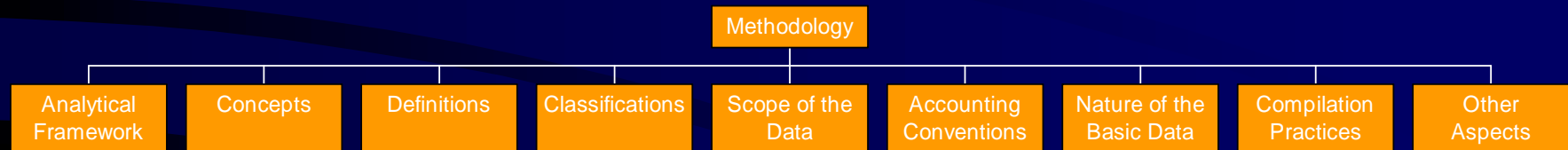
The DSBB Metadata Model

- A unified model that encompasses both SDDS and GDDS metadata, and

- Enables search and query on all DSBB metadata elements

- <u>Diagram A</u>: a schematic breakdown showing each of the four metadata dimensions (data, access, integrity, and quality)

- <u>Diagram B</u>: a schematic breakdown of the metadata elements for summary methodology statements

Diagram A

Country/International Organization
- Data
  - NSDP URL
- Metadata

Data
- Coverage
- Periodicity
- Timeliness
- Notes

Access

Integrity

Quality
- Methodology
  - Analytical Framework, Concepts and Definitions
  - Compilation Practices
  - Classifications
  - Other Aspects
  - Scope of the Data
  - Accounting Conventions
  - Nature of the Basic Data
  - Observance or Practice
  - Plans

Methodology

```
                              ┌──────────────┐
                              │ Methodology  │
                              └──────┬───────┘
   ┌──────┬──────┬──────┬──────┼──────┬──────┬──────┬──────┐
```

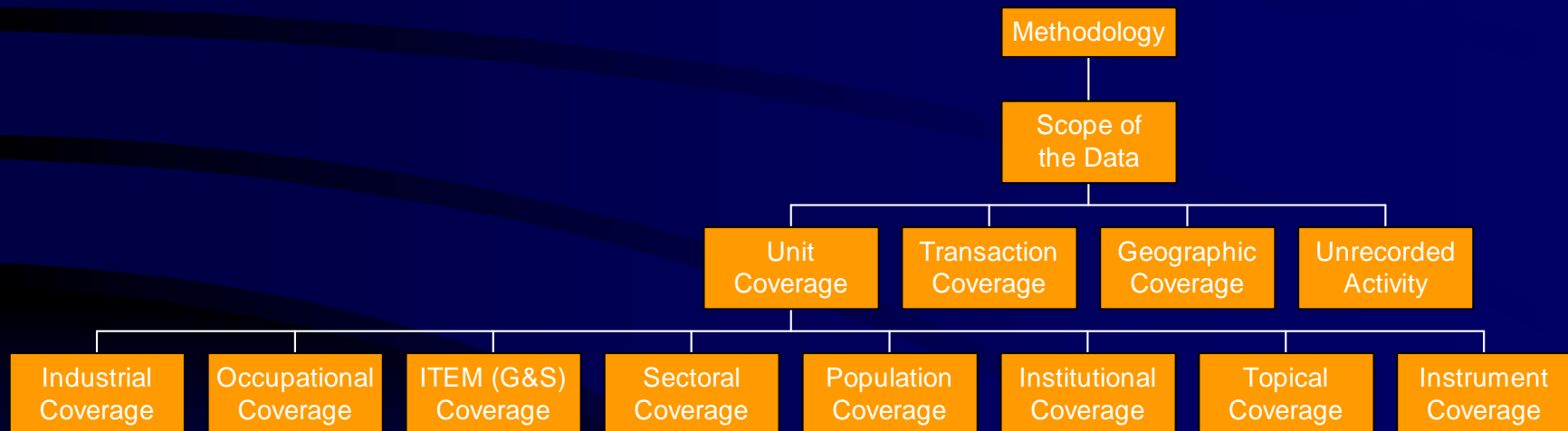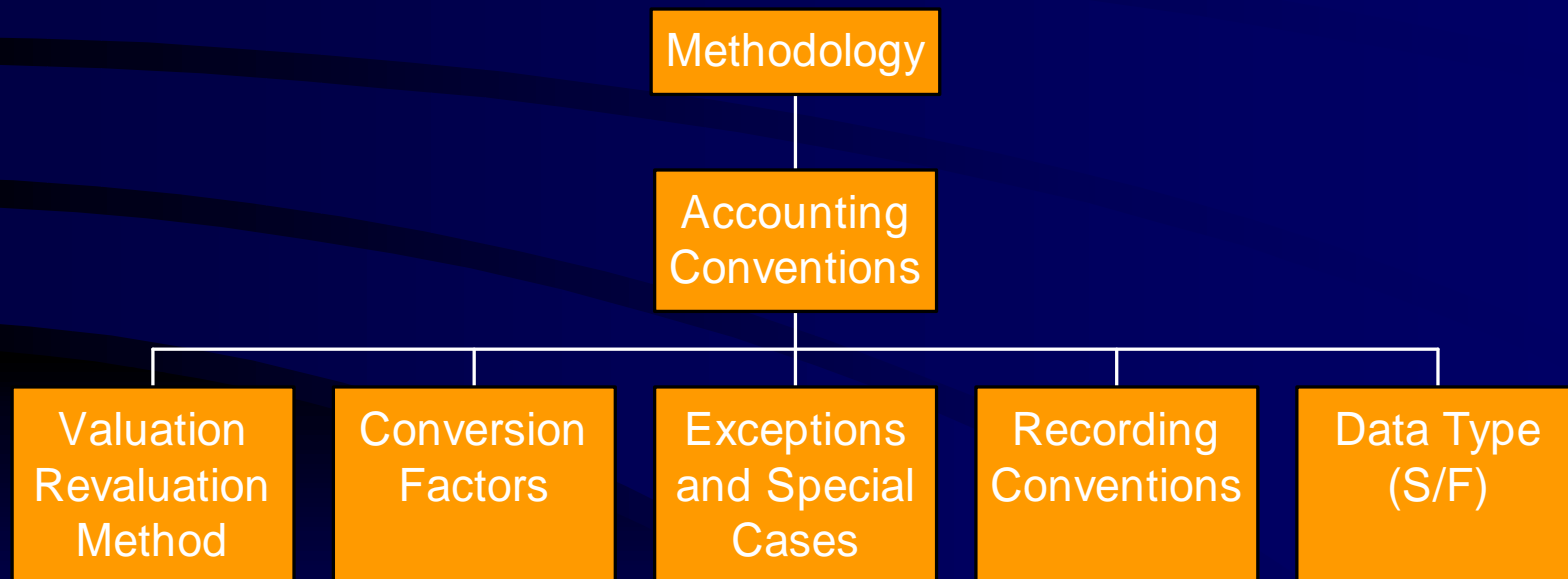| Analytical Framework | Concepts | Definitions | Classifications | Scope of the Data | Accounting Conventions | Nature of the Basic Data | Compilation Practices | Other Aspects |

Methodology

Methodology

Methodology

```
                    ┌─────────────┐
                    │  Coverage   │
                    └──────┬──────┘
                           │
                    ┌──────┴──────┐
                    │ Compilation │
                    │  Practices  │
                    └──────┬──────┘
```

| Assessment and Validation of Intermediate Results | Statistical Adjustments to Basic Data Sources | Estimation Methods | Aggregation Techniques | Consolidation Methods |

Methodology

```
                        ┌──────────┐
                        │ Coverage │
                        └────┬─────┘
                             │
                        ┌────┴─────┐
                        │  Other   │
                        │ Aspects  │
                        └────┬─────┘
     ┌──────────┬───────────┼───────────┬──────────┐
┌─────────┐┌─────────┐┌──────────┐┌─────────┐┌──────────┐
│ Seasonal ││Geographical││Verification and││ Other   ││ Revision │
│Adjustment││   and    ││Consistency ││Measures ││Policy and│
│          ││Commodity ││ of Final  ││         ││Procedures│
│          ││  Detail  ││  Results  ││         ││          │
└─────────┘└─────────┘└──────────┘└─────────┘└──────────┘
```
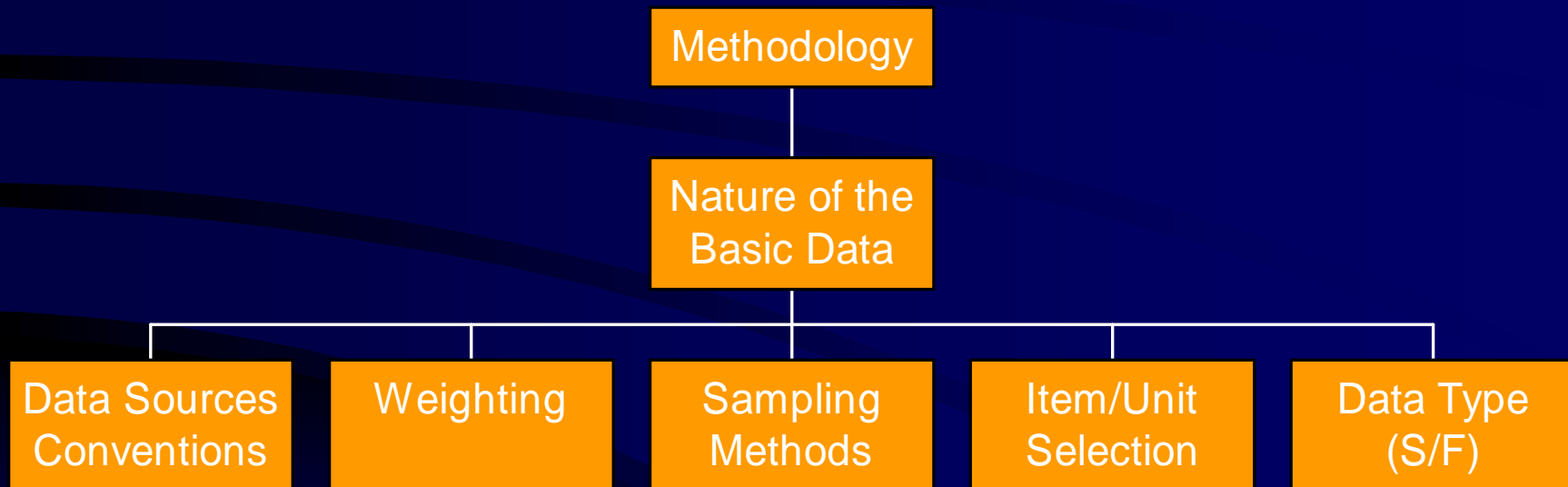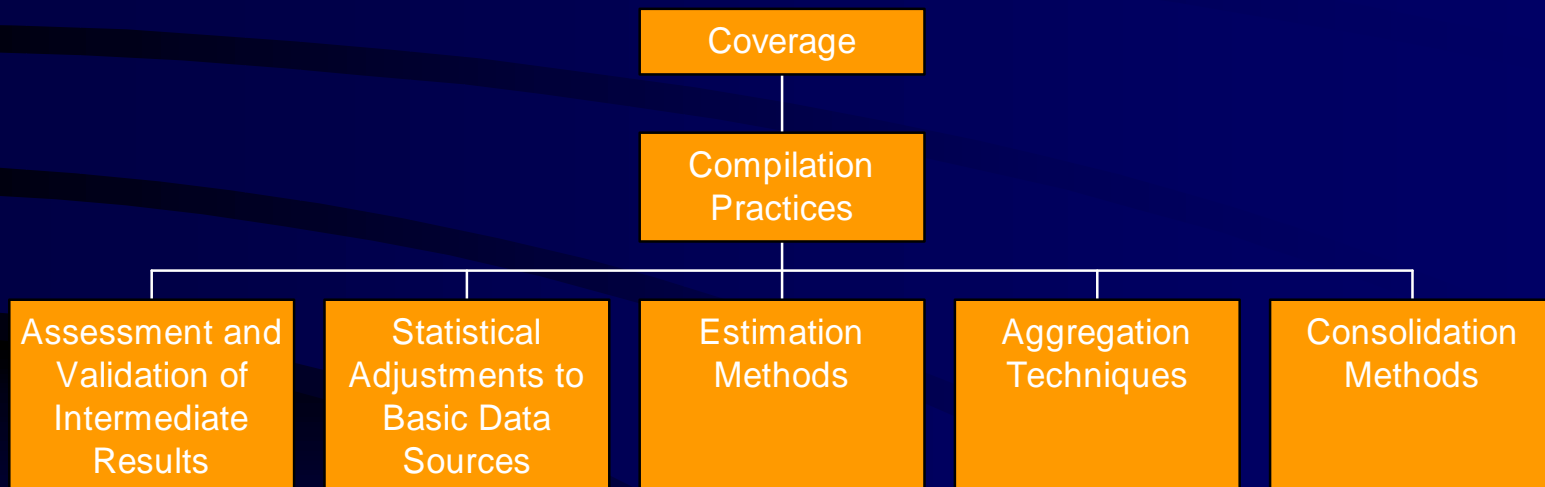
Rendering the DSBB Metadata Model in XML

What is XML (eXtensible Markup Language)?

A document markup language that allows authors to *separate* the content and structure of a document from its presentation

- in XML content/metadata are managed through "tags"
- structure is managed through "Document Type Definitions" (DTDs)

XML: Retaining Structure

- XML provides a standard way to tag text to reflect information structure

- Text may be stored many different ways

- Retaining structure provides a standard way to more easily query, assemble, and disseminate data and metadata

- XML enables the IMF to maximize the informational value inherent in the DSBB's metadata structure

XML Search Engine

- Can review and retrieve XML tags and content in disparate database sources
- XML provides a means for facilitating highly focused metadata searches that can be used to generate comparative queries on particular data on the fly
- XML provides a means for exchanging, editing, managing, validating and publishing SDDS metadata in a more automated and customized fashion

Accessing and Processing Content Elements: XML's Impact upon Users

XML documents are self describing

- data can be managed separately from its presentation
- data can be re-purposed for different media
- user specific data analysis applications (search, query and reporting) can be created through data exchange and database connectivity

Accessing and Processing Content Elements: XML's Impact upon Subscribers

- XML processing will allow subscribers to manage and share information from their own servers/ systems environments

- XML encoding will allow subscribers to more efficiently update metadata

- subscribers can access and retrieve information from other participating agencies

Accessing and Processing Content Elements: XML's Impact upon the IMF

- Develop an XML vocabulary (XML Tags and Document Type Definitions) to describe metadata
- Implement a technology solution that can easily manage, track, store and edit large volumes of XML based metadata
- Obtain a unified metadata repository to maintain archival metadata as *self-described information*

IMF, acting as a portal, will have greater flexibility to disseminate SDDS metadata on the DSBB and link to other Web sites (such as country specific sites or third party information aggregators)

Providing Interactivity and Data Query Facilities through Improved Metadata Interchange

Current Search and Querying Capabilities:

- Click through multiple "static" levels/pages to get information

DSBB users' seek:

- The ability to selectively query DSBB metadata
- To perform comparative metadata queries
- To build on queries that have been executed
- To search and query other websites

Table 1 (Work-in-progress)

Search and Query the DSBB - updated February 12, 2002

It is now possible to search and query the metadata provided by SDDS subscribers and GDDS participants as well as by sector/data category and across sectors/data categories. A large number of pre-defined queries have been established based on frequently requested information. It is also possible to structure queries using the template titles appearing below. As the functionality of this facility is under continuous development, your comments and suggestions are welcome (DDSD@IMF.ORG)

SDDS Subscribers
GDDS Participants
SDDS Subscribers and GDDS Particitpants

GO

Sector

All
Real Sector
Fiscal Sector

Data Category

All
National Accounts
Industrial Production

Metadata Class

Coverage
Periodicity
Timeliness

Metadata Element

Base Year
Frequency of rebasing
Number of establishments surveyed

Search   Clear   Tips

Table 2 (Work-in-progress)

Search and Query the DSBB - updated February 12, 2002

It is now possible to search and query the metadata provided by SDDS subscribers and GDDS participants as well as by sector/data category and across sectors/data categories. A large number of pre-defined queries have been established based on frequently requested information. It is also possible to structure queries using the template titles appearing below. As the functionality of this facility is under continuous development, your comments and suggestions are welcome (DDSD@IMF.ORG)

SDDS Subscribers   GO

Sector
- All
- Real Sector
- Fiscal Sector

Data Category
- Wages and Earnings
- Consumer Price index
- Producer Price Index

Metadata Class
- Coverage
- Periodicity
- Timeliness

Metadata Element
- Base Year
- Number of establishments surveyed
- Number of commodities surveyed

You have selected a query regarding information on the base year, number of establishments covered, and the number of commodities included in the Producer Price Indices of all 50 SDDS subscribers. To modify the the existing search criteria, select from the options above and click on 'Search' below.

Search   Clear   Tips

Table 3 (Work-in-progress)   **AS PROVIDED TO THE IMF**

Results of a PPI query for SDDS Subscribers (partial results) Special Data Dissemination Standard Subscribers

SDDS Subscriber	Base year	Number of establishments surveyed	Number of commodities surveyed
Argentina	1993 = 100	1400	2800
Australia	1988-89 = 100	3200	1400
Austria	2000=100	270	373
Belgium	1980=100	280	1096
Brazil	August 1994=100	1651	462
Canada	1997=100	Approximately 2700	980
Chile	June 1992=100	685	378
Columbia	June 1999=100	2200	730
Costa Rica	December 1999=100	216	1000
Croatia	Previous year=100	320	400
Turkey	1994=100	1300	687
United Kingdom	1995=100	Approximately 3000	980
U.S.A.	1982=100	More than 25,000	Approximately 100,000

Modify Existing Search

The results of your query on the base year, number of establishments covered, and the number of commodities included in the Producer Price Indices of all 50 SDDS subscribers are presented above. To modify the existing search criteria, click on 'Search' below.

Search   Clear   Tips

Table 4 (Work-in-progress)

Modify Existing Search Criteria

Results of a PPI query for SDDS Subscribers (partial results)

SDDS Subscribers   GO

Sector

| All |
| Real Sector |
| Fiscal Sector |

Data Category

| Wages and Earnings |
| Consumer Price index |
| Producer Price Index |

Metadata Class

| Coverage |
| Periodicity |
| Timeliness |

Metadata Element

Base Year	Previous year
	2000 - 1996
	1995 - 1990
Number of establishments	< 500
	500 - 999
	1000 - 1500
Number of commodities	1000 - 1999
	2000 - 2999
	3000 - 3999

You have selected a modified query as follows: base year (2000 - 1996), number of establishments covered (<500), and the number of commodities included (3000 - 3999) in the Producer Price Indices of all 50 SDDS subscribers. To obtain the results of your query click on 'Search' below.

Search   Clear   Tips