DESTATIS
Statistisches Bundesamt

Considerations to deal with the frozen cell problem in τ-Argus Modular

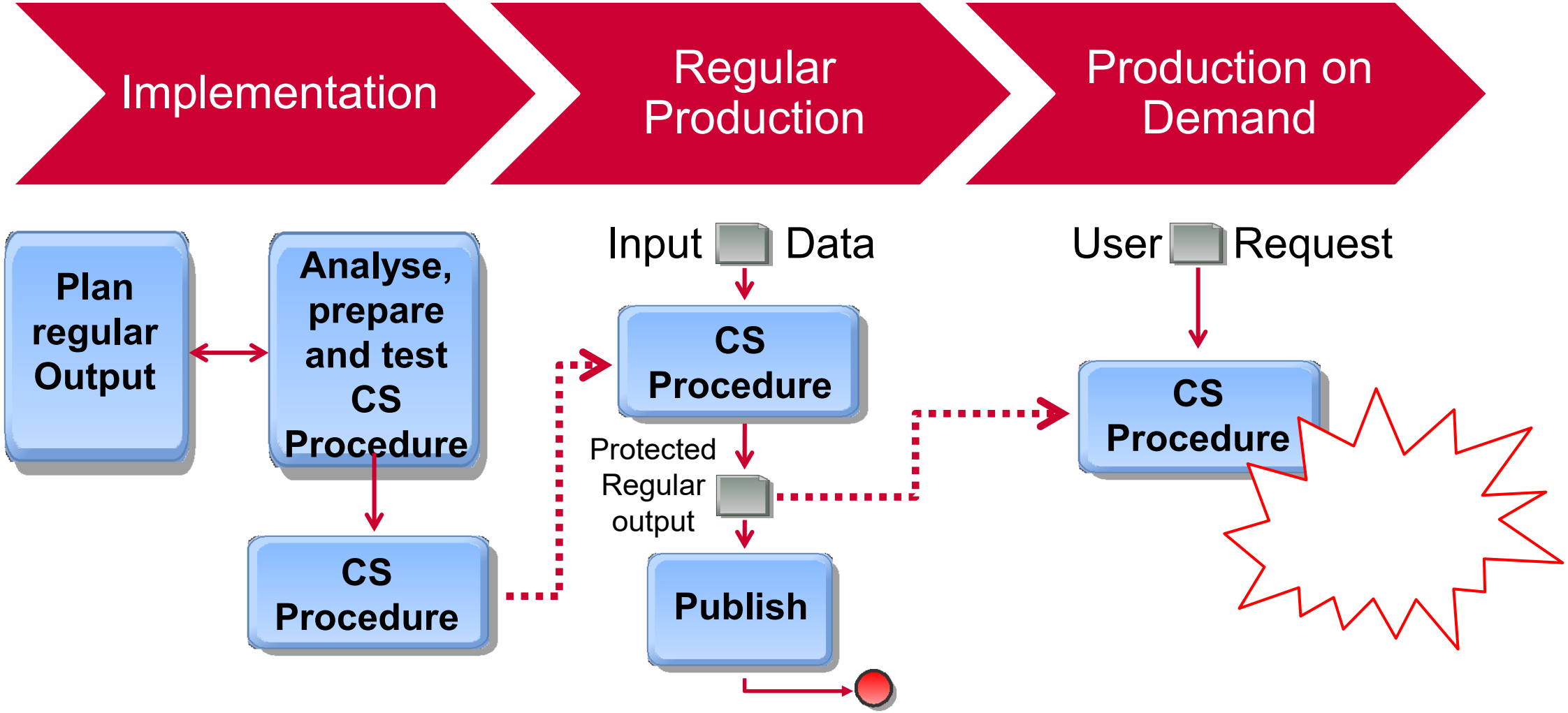
Joint UNECE/Eurostat Work Session on Statistical Data Confidentiality
1 to 3 December 2021, Poznań

Sarah GIESSING*, Peter-Paul de Wolf, Michel Reiffert*, Felix Geyer***

* Destatis

** CBS

Cell Suppression (CS) – Process Model



„Frozen“ cells can lead to infeasible instances

Example

- Margin cells published and must not be suppressed
- Unsafe interior cell
- Option A: Set margin cells to „protected“
  - Modular quits with error

Region x NACE	- A	A1	A2
- R1	18	10	8
1A	4	-	4
1B	14	10	4

Fig. 1 Simple table, infeasible due to frozen cells

„Frozen“ cells can lead to infeasible instances

Example

- Margin cells published and must not be suppressed
- Unsafe interior cell
- Option A: Set margin cells to „protected“
  - Modular quits with error
- Option B: Assign penalty cost to margin cells
  - Modular will suppress them anyway
  - → Inconsistent to published table → unsafe

Region x NACE			
	- A	A1	A2
- R1	18	X	X
1A	4	-	4
1B	14	X	X

Fig. 1 Simple table, infeasible due to frozen cells

Need pragmatic solution…

…for hierarchical / linked tables

- Accept hidden risks from phase 2
  - i.e. suppress all interior cells of an infeasible subtable, including the zero cells

Region x NACE			
	- A	A1	A2
- R1	18	10	8
1A	4	X	X
1B	14	X	X

Fig. 1 Simple table, infeasible due to frozen cells

- Identify subtables that can be released safely, or to
  - somewhat relaxed protection standards

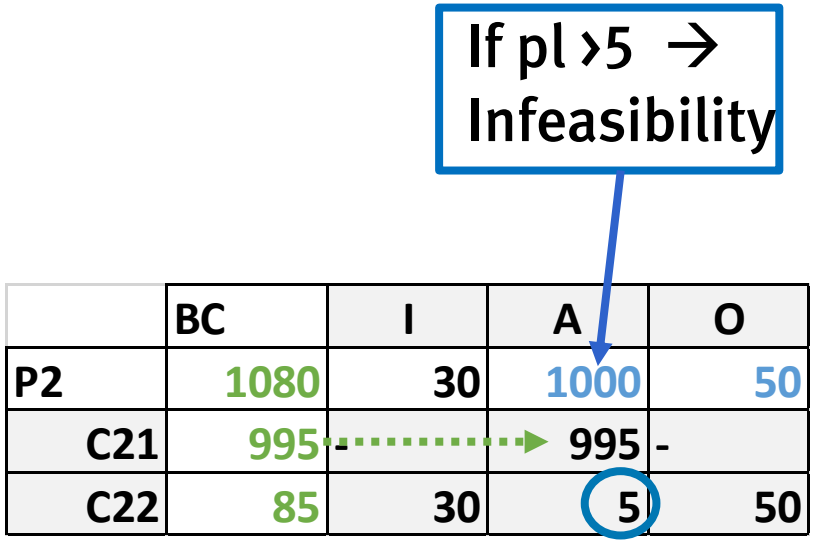
Need pragmatic solution…

…for hierarchical / linked tables

- Identify subtables that can be released safely, or to
  - somewhat relaxed protection standards

Example:

- Subtable with temporary primary unsafe cells



	BC	I	A	O
P2	1080	30	1000	50
C21	995	-	995	-
C22	85	30	5	50

Instance where feasibility depends on the protection level of a cell, here: cell P2/A

Example: Heuristic protection levels of temporary primary cells in Modular

primary with pl :=4

primary with pl :=20

Max pl=20

	BC	I	A	O
R	1645	90	1000	555
P1	550	50	-	500
P2	1080	30	1000	50
P3	15	10	-	5

Parent subtable (R,BC) to subtable (P2,BC).
Protection levels: (P1,O): 30 (R,A): 4

	BC	I	A	O
P2	1080	30	1000	50
C21	995	-	995	-
C22	85	30	5	50

Instance where feasibility depends on the protection level of a cell, here: cell P2/A

Example: Heuristic protection levels of temporary primary cells in Modular

primary with pl :=4

primary with pl :=20

Max pl=20

temporary primary with pl :=min(20,q*1000)

temporary primary with pl :=min(20,q*50)

	BC	I	A	O
R	1645	90	1000	555
P1	550	50	-	500
P2	1080	30	1000	50
P3	15	10	-	5

Parent subtable (R,BC) to subtable (P2,BC).
Protection levels: (P1,O): 30 (R,A): 4

	BC	I	A	O
P2	1080	30	1000	50
C21	995	-	995	-
C22	85	30	5	50

Instance where feasibility depends on the protection level of a cell, here: cell P2/A

Example: Heuristic protection levels of temporary primary cells in Modular

primary with pl :=4

primary with pl :=20

Max pl=20

	BC	I	A	O
R	1645	90	1000	555
P1	550	50	-	500
P2	1080	30	1000	50
P3	15	10	-	5

Parent subtable (R,BC) to subtable (P2,BC).
Protection levels: (P1,O): 30 (R,A): 4

temporary primary with pl :=min(20,q*1000)

Recall:
If pl >5 → Infeasibility

temporary primary with pl :=min(20,q*50)

	BC	I	A	O
P2	1080	30	1000	50
C21	995	-	995	-
C22	85	30	5	50

Instance where feasibility depends on the protection level of a cell, here: cell P2/A

Example: Checks for unclear instances

- Preprocess: Cells logically identical → Status identical (viz. „protected“)
- Supress all unprotected, non-zero cells

	BC	I	A	O
P2	1080	30 X	1000 X	50 X
C21	995	-	995	-
C22	85	30 X	X	50 X

Instance where feasibility depends on the protection level of a cell, here: cell P2/A

- Checks:
  - Either 0 or at least 2 suppressions in any row or column relation? ✓
  - At least 3 suppressions in relations with ≥ 2 „singletons“? ✓
  - Any primary suppression protected „sufficiently“ within its relations? ✓

    How about temporary primaries?
  - Subtable Audit: Feasibility intervals satisfy „relaxed“ protection levels?

Example: Checks for unclear instances

Recall: pl › 5

- Preprocess: Cells logically identical → Status identical (viz. „protected“)
- Supress all unprotected non-zero cells

	BC	I	A	O
P2	1080	X	X	X
C21	995	-	995	-
C22	85	X	X	X

Instance where feasibility depends on the protection level of a cell, here: cell P2/A

- Checks:
  - Either 0 or at least 2 suppressions in any row or column relation? ✓
  - At least 3 suppressions in relations with ≥ 2 „singletons“? ✓
  - Any primary suppression protected „sufficiently“ within its relations? ✓
    How about temporary primaries? ?
  - Subtable Audit: Feasibility intervals satisfy „relaxed“ protection levels?

Example: Relaxed Checks re. temporary primary‘s protection levels

Set protection level of temporary primaries to „almost 0“

- Either 0 or at least 2 suppressions in any row or column relation? ✓
- At least 3 suppressions in relations with ≥ 2 „singletons“? ✓
- **Any primary suppression protected „sufficiently“ within its relations?** ✓

  How about temporary primaries? ✓
- **Subtable Audit: Feasibility intervals satisfy „relaxed“ protection levels?** ✓

	BC	I	A	O
P2	1080	~~30~~ X	1080 X	~~50~~ X
C21	995	-	995	-
C22	85	~~30~~ X	X	~~50~~ X

protection level of a cell, here: cell P2/A

Example: Relaxed Checks re. temporary primary‘s protection levels

…Solves the problem in the particular example:

primary with pl :=4

primary with pl :=20

	BC	I	A	O
R	1645	90	1000	555
P1	550	50	-	500
P2	1080	30	1000	50
P3	15	10	-	5

Parent subtable (R,BC) to subtable (P2,BC).
Protection levels: (P1,O): 30 (R,A): 4

	BC	I	A	O
P2	1080	30 X	1000 X	50 X
C21	995	-	995	-
C22	85	30 X	X	50 X

protection level of a cell, here: cell P2/A

Example: Alternative approach

Step I. Assign penalty costs to „frozen“ cells

	BC	I	A	O
P2	1080	30	1000 X	50 X
C21	95 X	-	95 X	-
C22	5 X	30	5	50 X

Instance where feasibility depends on the protection level of a cell, here: cell P2/A

Example: Alternative approach

Step I. Assign penalty costs to „frozen“ cells

	BC	I	A	O
P2	1080	30	1000 X	90 X
C21	995 X	-	995 X	-
C22	85 X	30	5	50 X

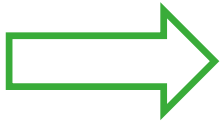
Instance where feasibility depends on the protection level of a cell, here: cell P2/A

Frozen cells get suppressed

Step II. Assign penalty costs to „frozen“ cells and set protection level of temporary primaries to „almost 0“

	BC	I	A	O
P2	1080	30	1000 X	90 X
C21	995	-	995	-
C22	85	30	X	50 X

Instance where feasibility depends on the protection level of a cell, here: cell P2/A



Tolerate underprotection risk and proceed

Last resort: Skip infeasible subtables

When would we skip an infeasible subtable?

- **Examples**
  - **I. If reducing protection levels is not a (selected) option, or**
  - **II. Cases of infeasibility corresponding to exact disclosure**

Region x NACE	- A	A1	A2
- R1	18	10	8
1A	4	-	4
1B	14	10	4

Fig. 1 Simple table, infeasible due to frozen cells

Recall: ***When protecting sets of linked tables*** $T_1,\ldots, T_N$ ***contained in a joint „cover table“, Modular processes the cover table skipping all subtables not contained in any of the*** $T_1,\ldots, T_N$ ***tables.***

Last resort: Skip infeasible subtables

What happens to skipped subtables?

- All interior cells (including zero cells) must be suppressed

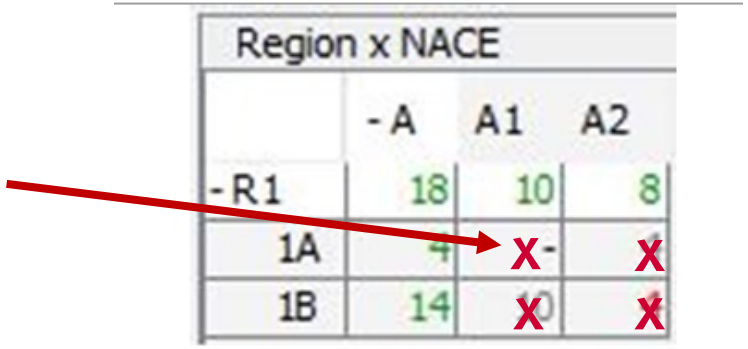


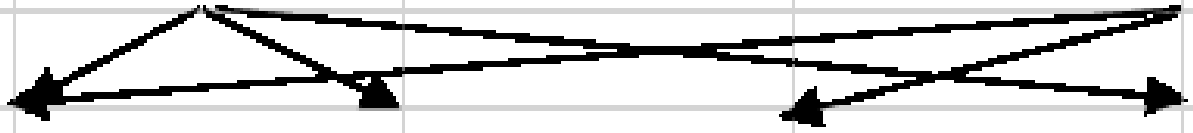
Fig. 1 Simple table, infeasible due to frozen cells

Which subtables to skip?

- Infeasible subtables, and…
  - any subtable *descending* directly or indirectly from an infeasible subtable

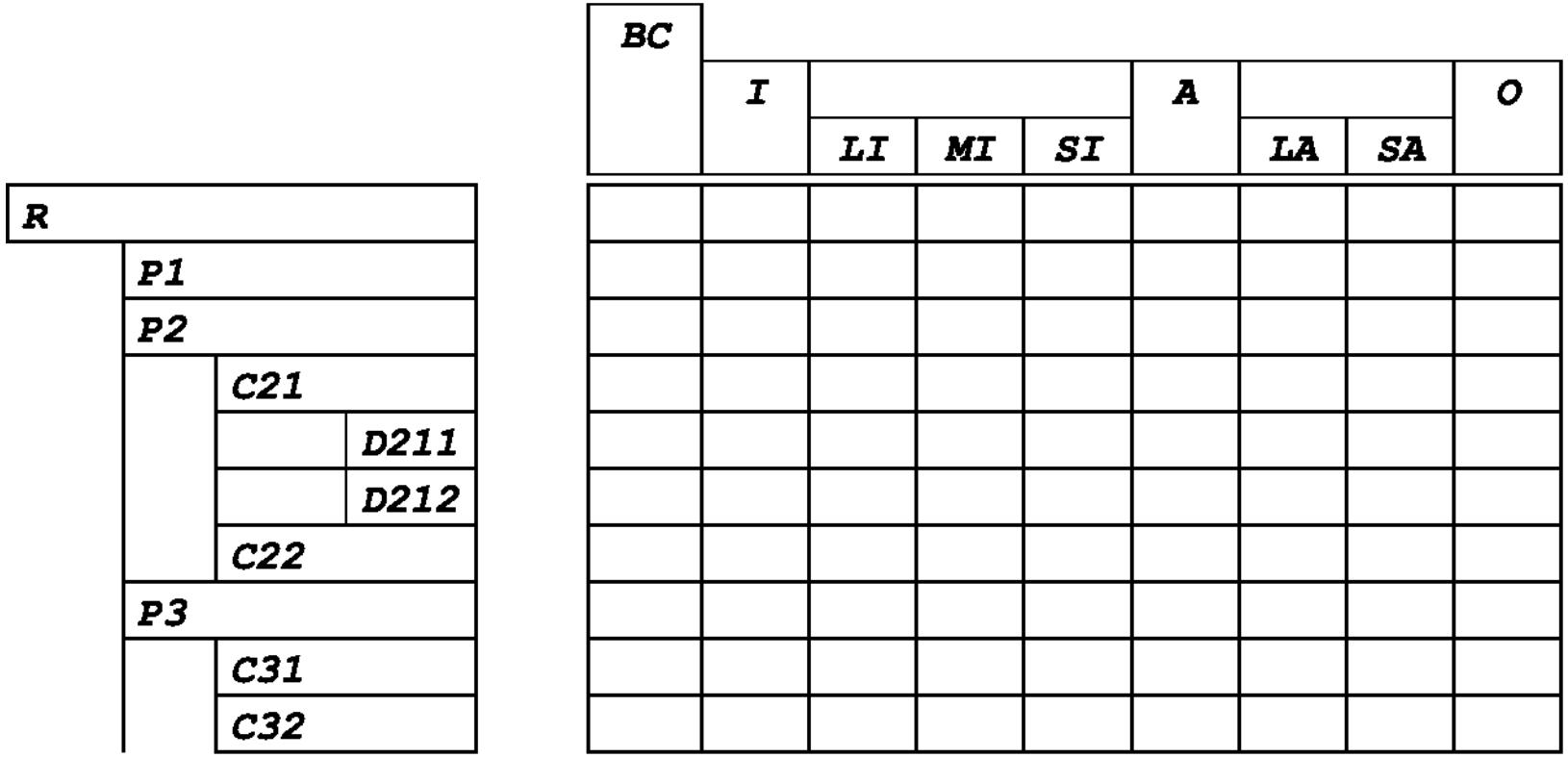
Subtable Descendancy

Generation					
n		Parent-ST1			Parent-ST2
n+1	Child-ST1		Child-ST2	...	Child-STn



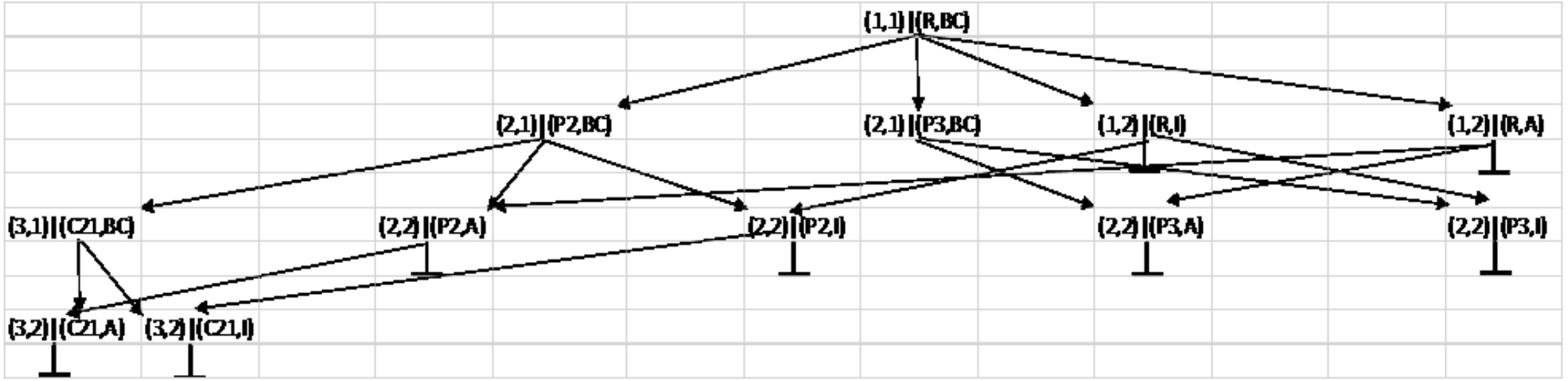
Subtable Descendancy

Recall hierarchical structure of the example (c.f. De Wolf, 2002)



Subtable Descendancy

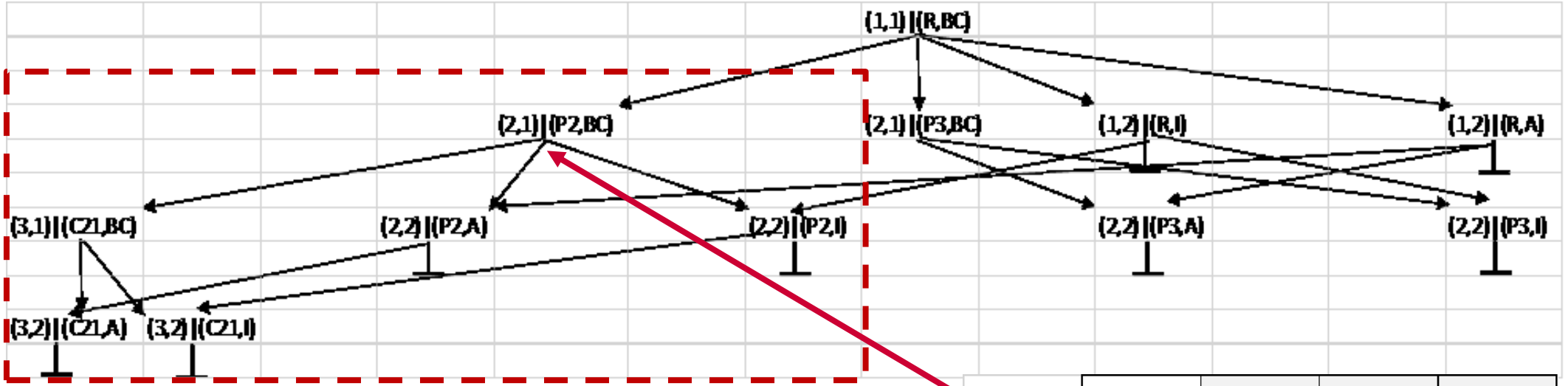
Descendancy tree of the example *)



*) Non-degenerate subtables, only

Subtable Descendancy

Example: Skipped descendant subtables of the infeasible subtable



	BC	I	A	O
P2	1080	30	1000	50
C21	995	-	995	-
C22	85	30	5	50

Instance where feasibility depends on the protection level of a cell, here: cell P2/A

Example with skipped subtables

Interior cells of the infeasible subtable

Interior cells of descendent subtables

R x BC									
	- BC	- I	LI	MI	SI	- A	LA	SA	O
- R	1645	90	80	5	5	1000	995	5	555
P1	550	50	50	-	-	-	-	-	500
- P2	1080	30	20	5	5	1000	995	5	50
- C21	995								
D211	105								
D212	890								
C22	85								
- P3	15	10	10	-	-	-	-	-	5
C31	5	-	-	-	-	-	-	-	5
C32	10	10	10	-	-	-	-	-	-

Fig. 5 Table *R* x *BC*, shaded bars covering interior cells of six skipped subtables

Summary / Conclusion

- τ-Argus Modular should offer a pragmatic solution for hierarchical tables with infeasible subtables due to the presence of protected cells.
- It should identify subtables that can be released safely
  - …or to somewhat relaxed protection standards (due to reduced protection levels of temporary primaries)
- It should skip infeasible subtables,
  - „hidden risks“ of previously released output to be accepted

Summary / Conclusion

Our Aim, after all:

- **Offer practical alternative to current „bad“ practice…**
  - **…tending to ignore risks of exact disclosure resulting from suppression of cells released in an earlier publication**

References

- **De Wolf, P.P. (2002), ‘HiTaS: A Heustic Approach to Cell Suppression in Hierarchical Tables’, In: ‘Inference Control in Statistical Databases’ Domingo-Ferrer (Ed.), Springer (Lecture notes in computer science; Vol. 2316)**