Mitigating Residual Seasonality While Preserving Accounting Relations in Hierarchical Time Series

Tucker McElroy[1]
U.S. Census Bureau[2]

December 14-15, 2023

[1] Joint work with Baoline Chen and Osbert Pang. Contact: tucker.s.mcelroy@census.gov

Outline

1. Background: removing Residual Seasonality (RS) from GDP
2. Data structure of GDP components
3. The window structure of accounting relations
   (a) Accounting relations
   (b) Windows, muntins, hubs, spokes, panes, terraces, and echelons
4. Methodology for removing RS while preserving accounting relations
5. Reconciliation for PCE Goods and Services
6. Conclusion

Background

- Components of GDP may have Residual Seasonality (RS), and also satisfy many accounting relations.

- Direct seasonal adjustment of (monthly or quarterly) components results in violations of accounting relations; this is undesirable to the Bureau of Economic Analysis (BEA).

- We seek to modify existing seasonal adjustments as little as possible, such that:
  - RS is removed
  - frequential (monthly to quarterly) and hierarchical (sub-aggregates to super-aggregates) accounting relations are preserved

Background

Primary Challenge: understanding the hierarchical lattice structure arising from the accounting relations.

- Self-intersecting lattice: modifications to one component can place constraints upon others that are involved in the same accounting relation

- Analysis of the lattice via topological notion of *terrace*: components that can be modified together, contingent only upon previously computed results from higher *echelons* (batches of terraces)

Background

Secondary Challenge: we propose a top-down algorithm for modifying monthly time series, removing RS as we move through echelons of terraces.

There is a nonlinear optimization for each series, whereby:

- monthly values are changed as little as possible;
- quarterly and hierarchical relations still hold;
- all resulting composites have no RS

Data Structure

We focus on the Goods and Services portion of Personal Consumption of Expenditure (PCE).

- PCE is published monthly and quarterly, and makes up about 2/3 of the components of GDP
- PCE has two kinds of accounting relations:
  - frequency aggregation: the average of the three months pertaining to a quarter transforms a monthly flow time series to a quarterly flow time series
  - hierarchical aggregation: each composite time series is expressed as an aggregate of component time series (at each time point, whether monthly or quarterly)

Data Structure

RS through aggregation: a monthly component time series with no RS, can manifest RS when aggregated by frequency or hierarchy. How?

- Frequency: suppose the monthly time series has strong latent monthly seasonality that is buried in noise, that is negatively correlated across months such that its quarterly aggregation is zero

- Hierarchical: suppose the various component time series have a common latent seasonality, each buried in independent noise, which averages down

Data Structure

PCE Hierarchical Structure: each composite is the aggregation of all components "below" it.

- Aggregation levels determined by major product type
- Published at annual, quarterly, and monthly flows
- Each component has 6-letter identification code

Data Structure

- We construct a directed graph for PCE, consisting of a node for each variable and a directed edge that connects it to any other nodes that are components of the composite (Figure 1)

- PCE breaks into Goods (DGDSRC) and Services (DSERRC).

- Goods breaks into Durable Goods (DDURRC) and Nondurable Goods (DNDGRC).

- Durable has four sub-aggregates: Motor Vehicles and Parts (DMOTRC), Furnishings and Durable Household Equipment (DFDHRC), Recreational Goods and Vehicles (DREQRC), and Other Durable Goods (DODGRC).

Data Structure

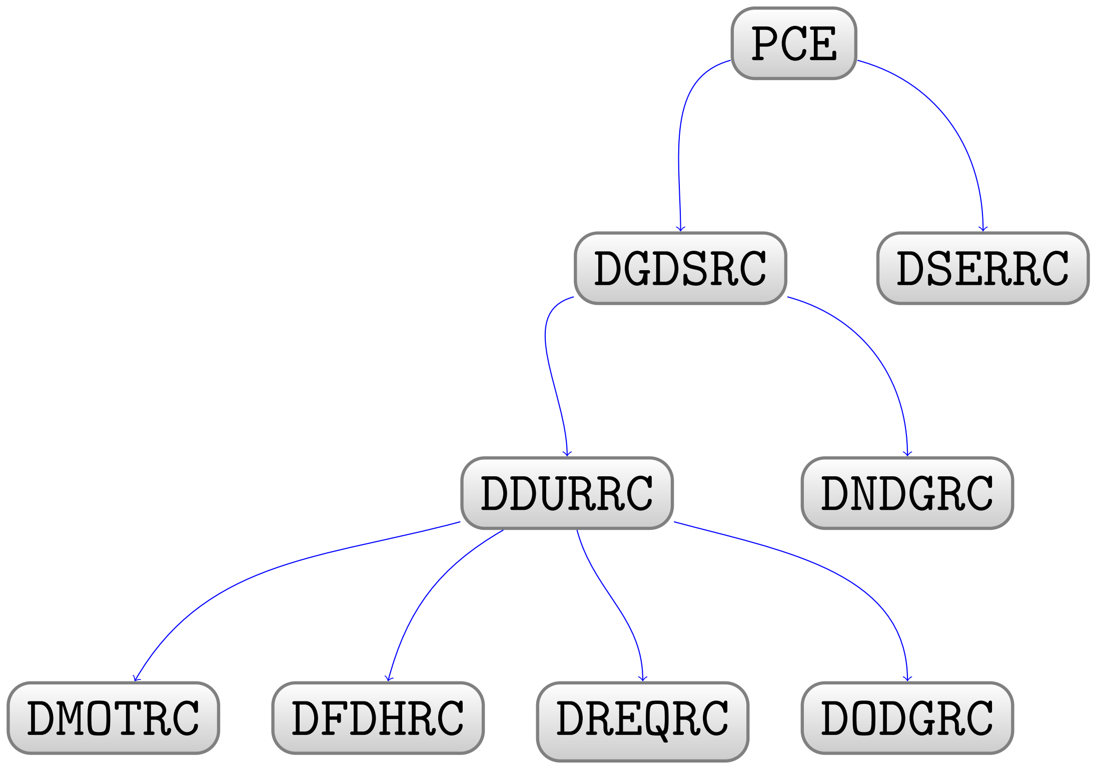

Figure 1: Top of the hierarchical lattice for the PCE sub-lattice of GDP.

Data Structure

Not a tree: components can have aggregation relations to more than one composite (which can be at different “levels”)!

It’s a window: self-intersecting lattice of a Gothic window

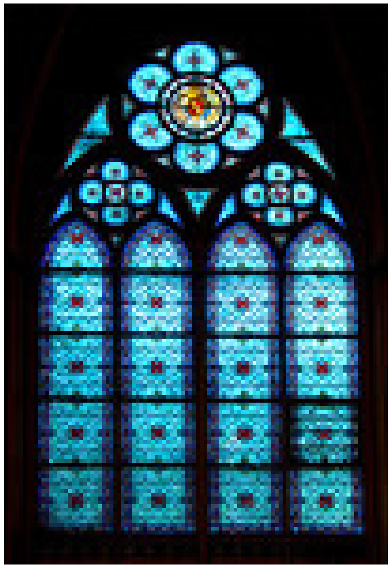

Figure 2: Notre Dame de Paris, stained glass window.

Window Structure

Goal: remove RS from all four variables, such that the accounting relations represented by the green (hierarchical) and blue (sampling frequency) arrows are preserved.

$$
\begin{tikzcd}
\boxed{\text{Quarterly High}} \arrow[r] \arrow[d] & \boxed{\text{Monthly High}} \arrow[d] \\
\boxed{\text{Quarterly Low}} \arrow[r] & \boxed{\text{Monthly Low}}
\end{tikzcd}
$$

Window Structure

Top-down Approach: proceed hierarchically as follows.

- Suppose Quarterly High and Monthly High are already adjusted
- Modify Monthly Low such that:
  1. It has no RS
  2. It aggregates to Monthly High
  3. Its Quarterly Low aggregate has no RS

Call such a procedure a **Frequency Aggregation** (FA) algorithm.

Window Structure: Fictitious Example

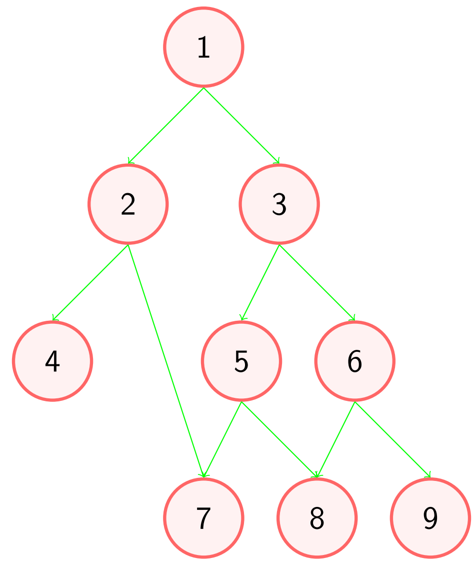

Figure 3: Example of a hierarchical lattice

Window Structure: Fictitious Example

Need to determine batches of nodes that can be adjusted jointly without fear of causing an incoherency due to other neglected accounting relations.

- Suppose node 2 in Figure 3 has already been adjusted; we want to adjust nodes 4 and 7 such that the accounting relation holds.
- But modifying node 7 affects accounting relation to node 5!
- Also node 8 is involved with node 7 via accounting relation to node 5.
- And so forth: need to jointly adjust nodes 4, 7, 8, and 9, subject to three accounting relations!

Windows, Muntins, Hubs, and Spokes

Window: a hierarchical lattice W is a called a "window". It has n nodes connected by directed edges, called **muntins**. Each muntin connects a **hub** node to a **spoke** node.

- Each hub has a single accounting relation
- A hub's spokes are the variables subject to that accounting relation
- Each hub has at least 2 spokes

Windows, Muntins, Hubs, and Spokes

Example: in Figure 3, the spokes of node 2 are $\{4, 7\}$. The hubs of node 8 are $\{5, 6\}$. The sill is $\{4, 7, 8, 9\}$.

Windows, Muntins, Hubs, and Spokes

An $N \times N$ *adjacency matrix* A encodes accounting relations:

- Let all variables be indexed from 1 through 357 (there are 147 series in PCE Goods, and 210 series in PCE Services), so $N = 357$
- $A_{jk} = 1$ if and only if variable j is a component (spoke) of variable (hub) k, equals 0 otherwise
- The non-zero entries of the kth column of A correspond to the spokes of node k.
- The non-zero entries of the jth row of A corresponds to the hubs of node j.

Tiers and Terraces

Tier: the collection of hubs for a given spoke in a window W.

We want to partition the window into batches of nodes to be jointly modified, without neglecting relevant accounting relations. We must find

- The tier of any spoke of the given node – this is called the node's **Pane**;
- Keep finding panes (iterate);
- When the set of such nodes is maximal, we stop – this is called the **Terrace**.

Terraces

Theorem 1. *The collection of terraces of a window W is a partition of W (excepting the "bottom" nodes).*

Terraces partition the (non-sill) window, and we can run the FA algorithm on each terrace, beginning with terraces on the top and proceeding down the window.

Terraces

Example: in Figure 3

- the terraces are $\{1\}$, $\{3\}$, and $\{2, 5, 6\}$, with $\{4, 7, 8, 9\}$ being the sill;
- to run FA on terrace $\{2, 5, 6\}$ we must already have adjusted those hubs in a previous step;
- this requires that we already ran FA on terraces $\{1\}$ (because this is the tier of node 2) and $\{3\}$ (which is the tier of nodes 5 and 6).

Echelons

Given two terraces B and C, we say that B is above C, and write $B > C$, if B has spokes in C (i.e., $B \cap \mathcal{T}(C) \neq \emptyset$). The **Echelon** of a terrace C, written $\mathcal{E}(C)$, is the collection of all terraces that are above C.

Ordering echelons: before we can run FA on a terrace C, we must run it on every terrace in its echelon $\mathcal{E}(C)$. The value ℓ such that $\mathcal{E}^{\ell}(C) = \{1\}$ for a given terrace C is its **Echelon rank**.

Proposition 1. *The tier of any terrace lies in an echelon of lower rank (higher up the lattice).*

Methodology for Removing RS

Proposition 1 supports the idea of using echelon rank to order our computations:

1. Compute terrace partition for the window.
2. Determine echelon rank of each terrace.
3. Run FA on each terrace (in any order) of a given rank; then increment the rank.

Methodology for Removing RS

Example: consider the terrace displayed in Figure 4.

- The terrace consists of DNRSRC, DHLRRC, and DHSPRC, denoted as k_1, k_2, and k_3 respectively.

- The spoke nodes are DFPNRC, DNPNRC, DOUSRC, DNPHRC, DFPHRC, and DGVHRC, denoted as $j_1, \ldots, j_6$, respectively.

Methodology for Removing RS

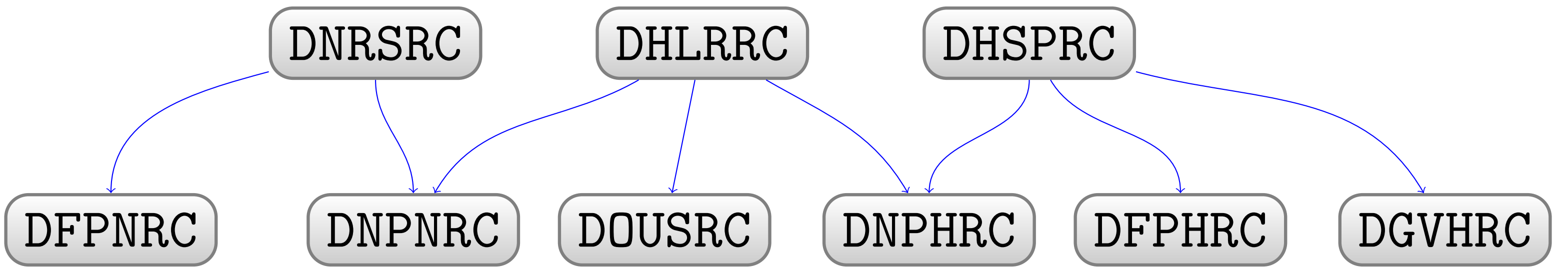

Figure 4: Example from PCE Services

Methodology for Removing RS

Hierarchical Relations: Consider a terrace C.

- Let $\{\underline{x}_t\}$ and $\{\underline{x}_n\}$ be monthly and quarterly spokes.
- Let $\{\underline{u}_t\}$ and $\{\underline{u}_n\}$ be monthly and quarterly hubs.
- Suppose there are $q = |C|$ constraints on $r = |\mathcal{S}(C)|$ spokes.
- Accounting relations of spokes to hub are summarized through a $q \times r$ matrix J such that

$$\underline{u}_n = J\underline{x}_n \qquad \underline{u}_t = J\underline{x}_t.$$

Methodology for Removing RS

Each row of J is a linear combination corresponding to a muntin. (Assume $q \leq r$.) In example (Figure 4),

$$J = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}.$$

Methodology for Removing RS

Frequency Relations: let Q_n denote the set of three months corresponding to quarter n, so that

$$\underline{x}_n = \sum_{t \in Q_n} \underline{x}_t / 3 \qquad \underline{u}_n = \sum_{t \in Q_n} \underline{u}_t / 3.$$

Methodology for Removing RS

Top-down Logic: When we come to a terrace C,

- Hubs $\underline{u}$ have already been modified to $\underline{w}$
- Want to modify $\underline{x}_t$ unto close (mean square relative difference) $\underline{y}_t$ such that $J\underline{y} = \underline{w}$
- Require

$$\underline{y}_n = \sum_{t \in Q_n} \underline{y}_t / 3 \tag{1}$$

to have no RS, and aggregates to $\underline{w}_n$.

Methodology for Removing RS

Let $\underline{a} \div \underline{b}$ denote the component-wise division of two vectors, and $\| \cdot \|$ denotes the Euclidean norm.

1. $\|\underline{y}_t \div \underline{x}_t - 1\|$ is small for some $\{y_t\}$

2. $\{\underline{y}_t\}$ has no monthly RS

3. $\{\underline{y}_n\}$ has no quarterly RS

4. $J\underline{y}_t = \underline{w}_t$

5. $J\underline{y}_n = \underline{w}_n$. (Given (1), implied by 4.)

United States™ Census Bureau

Methodology for Removing RS

Frequency Aggregation Algorithm: parse as constrained optimization problem.

- Assess RS with a seasonal adjustment diagnostic δ set at a p-value threshold τ.

- We use the Model-Based F (MBF) test (Lytras, Feldpausch, and Bell (2007)) with $\tau = .05$.

- Employ the inequality constraint $\delta \geq \tau$ for monthly and quarterly.

Reconciliation for PCE Goods and Services

Top-down Applied to PCE: we used testing threshold $\tau = .05$, and past 20 years for modeling (for MBF).

- 74 our of 357 nodes had monthly or quartely RS
- 22 from Goods, 52 from Services
- All seasonality removed while preserving accounting relations!
- Modifications were small in most cases.

Reconciliation for PCE Goods and Services

Aggregation Level	No. of Variables with RS (M or Q)	No. of Variables with RS (M, not Q)	No. of Variables with RS (Q, not M)	No. of Variables with RS (M and Q)
L2	1	1	0	0
L3	0	0	0	0
L4	13	9	3	1
L5	24	15	6	3
L6	27	17	6	4
L7	9	6	0	3
Sum	74	48	15	11

Table 1: Number of variables found to have RS by the MBF test statistic at 5% significance, sorted by aggregation level and sampling frequency.

Reconciliation for PCE Goods and Services

Meats and Poultry: all spokes and hubs have RS. Hub is DMAPRC; spokes are DBEERC, DPORRC, DMEARC, DPOURC.

Description	Level	PCE Code	p_M	p_Q	Max Abs Q	Max Abs M
Meats and poultry	5	DMAPRC	.0213	.9325	.56	.77
Beef and veal	6	DBEERC	.0120	.0000	.54	.79
Pork	6	DPORRC	.0251	.9075	.55	.61
Other meats	6	DMEARC	.0120	.5031	.76	1.30
Poultry	6	DPOURC	.0199	.9352	.48	.57

Table 2: Tier for Meats and poultry: RS test p-values and maximum absolute value of percent modification, at monthly and quarterly frequencies.

Reconciliation for PCE Goods and Services

Outpatient Services: hub DOUTRC (Outpatient services), and one spoke, DPHYRC, have quarterly RS; the other spoke, DDENRC, has monthly RS.

Description	Level	PCE Code	p_M	p_Q	Max Abs Q	Max Abs M
Outpatient services	3	DOUTRC	.9998	.0000	.21	.21
Physician services	4	DPHYRC	.9991	.0000	.66	.67
Dental services	4	DDENRC	.0000	.9733	.23	.27

Table 3: Tier for Outpatient services: RS test p-values and maximum absolute value of percent modification, at monthly and quarterly frequencies.

Reconciliation for PCE Goods and Services

Hospital Services with Health Services to Household: corresponds to Figure 4 terrace.

Description	Level	PCE Code	p_M	p_Q	Max Abs Q	Max Abs M
Hospital services	4	DHSPRC	.9610	.9143	.17	.18
Nonprofit hospital	4, 5	DNPHRC	.8551	.0000	.12	.18
Proprietary hospital	5	DFPHRC	.1122	.0000	.35	.37
Government hospital	5	DGVHRC	.8956	.2686	.68	.70
Health services to household	3	DHLRRC	.9152	.0000	.11	.17
Outpatient services	4	DOUSRC	.9909	.6286	.26	.28
Nonprofit nursing	4	DNPNRC	.3609	.7664	.16	.17

Table 4: Tiers for Hospital services and Health services to household: RS test p-values and maximum absolute value of percent modification, at monthly and quarterly frequencies.

Reconciliation for PCE Goods and Services

Receipts from Sales and Services: an example of interconnected tiers.

- DNPSRC is at level L2 of the PCE Lattice, and has 9 components (at level L3);
- two of these, DRELRC and DGIVRC, are also components of DSOCRC, which is at level L3 of the PCE Lattice;
- one of the spokes of DNPSRC is DHLRRC, which was a hub in previous example;
- the terrace of DHLRRC has echelon rank one greater than that of DNPSRC.

Reconciliation for PCE Goods and Services

Description	Level	PCE Code	p_M	p_Q	Max Abs Q	Max Abs M
Receipts from sales by nonprofit	2	DNPSRC	.9036	.7801	.09	.12
Health services to household	3	DHLRRC	.9152	.0000	.11	.17
Recreation services	3	DRCRRC	.9244	.6153	.01	.01
Education services	3	DEDRRC	.9661	.9511	.03	.04
Social services to household	3	DSSRRC	.3651	.7839	.02	.02
Religious organizations	3, 4	DRELRC	.0017	.9948	.82	.17
Foundations and	3, 4	DGIVRC	.0141	.7628	1.86	2.26
Services of social	3	DSASRC	.9062	.2576	.09	.00
Civic and social on	3	DCISRC	.5148	.2564	.02	.00
Professional advocate	3	DSNRRC	.8786	.5270	.01	.02
Social services and	3	DSOCRC	.0031	.2794	.52	.17
Social assistance	4	DSCWRC	.0385	.0000	.75	.17
Social advocacy	4	DSADRC	.8956	.1968	.09	.32

Table 5: Tiers for Hospital services and Health services to household: RS test p-values and maximum absolute value of percent modification, at monthly and quarterly frequencies.

Reconciliation for PCE Goods and Services

Motor Vehicle Leasing Services: in this example the early portions of the sample had larger amplitude in growth rates, resulting in somewhat larger modifications.

Description	Level	PCE Code	p_M	p_Q	Max Abs Q	Max Abs M
Motor vehicle leasing services	5	DMVLRC	.0003	.0253	1.74	2.93
Auto leasing	6	DALERC	.5793	.0657	1.70	3.11
Truck leasing	6	DTLERC	.0000	.0467	7.46	44.68

Table 6: Tiers for Motor vehicle leasing services: RS test p-values and maximum absolute value of percent modification, at monthly and quarterly frequencies.

Reconciliation for PCE Goods and Services

Professional Advocacy: DSNGRC has biennial periodicity (linked to U.S. election cycle).

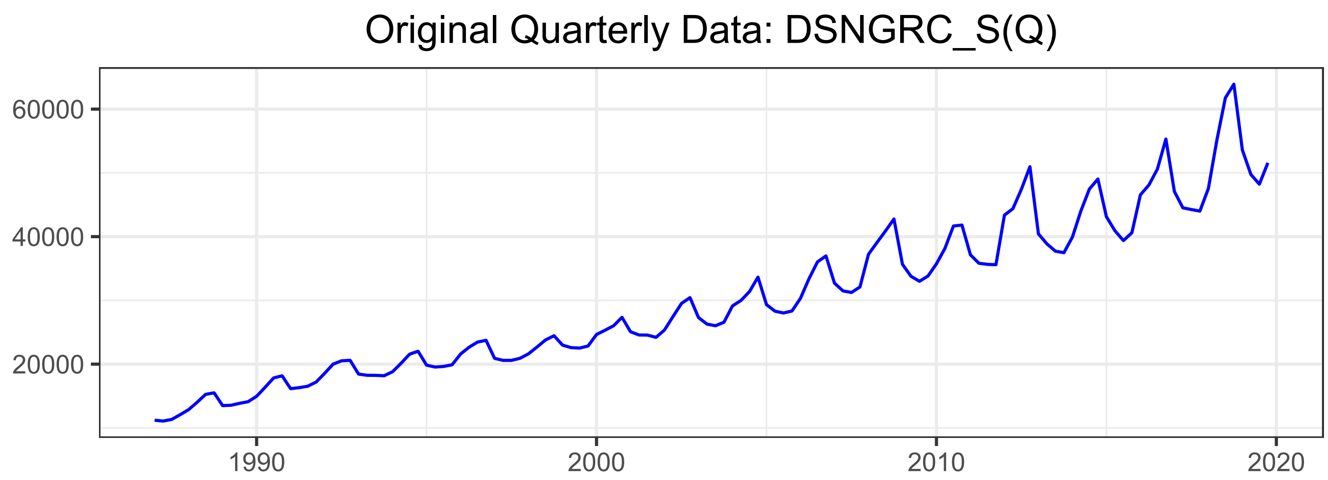

Figure 5: Plot of quarterly time series DSNGRC (Professional advocacy).

Reconciliation for PCE Goods and Services

Spoke (DSNGRC) has monthly and quarterly RS, but hub (DNPERC) has no RS. Our algorithm corrects RS but preserves biennial oscillation.

Description	Level	PCE Code	p_M	p_Q	Max Abs Q	Max Abs M
Gross output of nonprofit	2	DNPERC	.0807	.2901	.131	.162
Professional advocacy	3	DSNGRC	.0000	.0000	2.95	3.36

Table 7: Tiers for Gross output of nonprofit: RS test p-values and maximum absolute value of percent modification, at monthly and quarterly frequencies.

Summary

Problem Facets:

- A double (quarterly and monthly) hierarchical self-intersecting lattice data structure determined by accounting rules.
- Lattice nodes with multiple possible composites, yielding a tiered structure decomposed in terms of hubs, spokes, and muntins.
- The natural subsets for algorithmic application are terraces, which are collections of nodes invariant to the operation of determining all the hubs of all the spokes of a given node.

Summary

Problem Facets:

- A well-defined concept of echelon rank, facilitating top-down approach
- A nonlinear optimization problem whereby given monthly time series (which are spokes of a terrace) are modified as little as possible, subject to accounting relations given by the muntins unto hub variables of the given terrace, and such that the modified monthly spoke variables and their quarterly aggregates have no RS.

Merci!

United States™ Census Bureau