The Icelandic Wage Price Index

Violeta Calian, Oliver Snær Ægisson

violeta.calian@hagstofa.is
oliver.s.aegisson@hagstofa.is

Statistics Iceland

https://github.com/violetacln/iwpi_review

Goals

- Clarify: structure and performance of the IWPI
- Other index numbers, new data, details on analysis
- Collect: new ideas/suggestions

Main findings of the IWPI internal and external evaluation

(WP1, WP2, External Report)

- Attributes of IWPI:
  - good hierarchical structure
  - optimum index formula choice
  - calculated according to standard methods of price index theory
  - making use of comprehensive data
- Chaining drift effect: minimised by the choice of superlative index and monotonic data trends*

  (no fast oscillations as in scanner data!)
- Effect of quality changes: significant but small, C.I. of effect includes zero*
- AEI comparison:
  - AEI and IWPI reflect different phenomena
  - AEI not recommended as alternative of price index by the External Report
- Main challenges for the current index:
  - coverage and sample refresh
- **Methodology*:
  - analytical calculations /theoretical results (WP1, WP2, External Report, Literature)
  - empirical / data-based tests (WP1, WP2)

The *wage* price index: formulation of the problem

- To measure the changes in *hourly wages* between *two* points in time at constant quality, which are solely determined by market forces

while:

- the sample evolves with time: new & disappearing items
- the quality of persistent items can change with time
  (e.g. experience, work satisfaction, working conditions for fixed items)
- long time series of changes need to be defined
- *other errors* can be made when measuring/collecting data on wages

Challenges of index numbers

- Definition/choice of *elementary aggregate* and homogeneity condition
- Time series, *stochastic* process issues
- Estimated values based on *real world data* (i.e. sample data, discrete time, in/complete)
- Sources of *bias* and *errors*:
    - sampling:
        - coverage and representativeness (including rare patterns)
        - data collection errors
    - dynamic effects:
        - quality changes
        - new products
        - missing products
        - life cycle evolution / persistent products
        - seasonality
    - combining bilateral indices into time series of index values

Known solutions to known problems

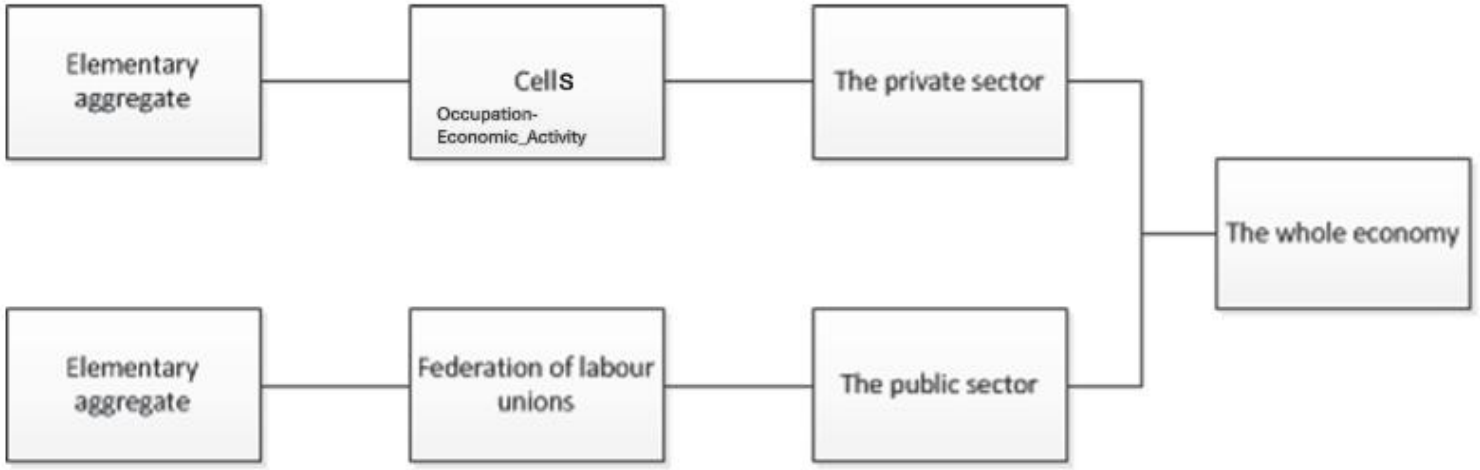
- Matched models
- Hedonic models / time-dummy models
- Sample refresh
- Weights updating
- Shadow/Imputed prices

- Chained index, fixed base index
- Multilateral index numbers

Data structure for the IWPI

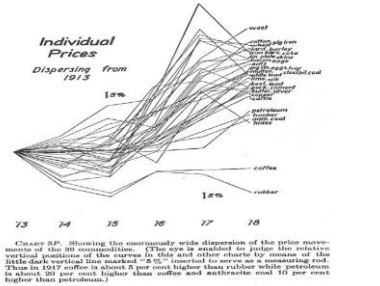
- Sampling (Icelandic Survey on Wages, Earnings and Labour Costs (ISWEL), Launarannsókn Hagstofu Íslands)

- Census type of data/ heildarsöfnun (ideal, future 100% coverage)



Formulation of the mathematical problem

Example 1 (source: Fisher, *The making of index numbers*, 1927)



- **Non-unique** decomposition of the change in a sum of products (value aggregates) into the product of two numbers, one due to price changes and one due to quantity changes $\frac{V^1}{V^0} = \frac{\sum_i p_i^1 q_i^1}{\sum_j p_j^0 q_j^0} = I_P I_Q$

- *IWPI* ***item/elementary aggregate****:*

  (employer, employee, occupation, economic activity) +*

  & Automatic and manual checks for the sudden changes in wages!

- *Note:*

an index is a generalized average of changes in prices or quantities

Main approaches and main index formula

- Axiomatic (set of conditions) *F*
- Stochastic (model) *T*
- Economic (optimization) *F, W, T*
- Fixed basket (constructive) *F, W*
- Divisia's theory ← a chain index as the first (*L*) or second (*F*) approximation of Divisia´s function

$$\frac{p(t)}{p(t_0)} = exp\left(\sum_i \int_0^t w_i(\tau)\frac{dp_i(\tau)}{d\tau}\frac{1}{p_i(\tau)}d\tau\right)$$

Used for: deriving definitions/weaknesses/strengths

- Laspeyres $L = \frac{\sum_i p_i^1 q_i^0}{\sum_j p_j^0 q_j^0} = \sum_i (p_i^1/p_i^0) w_i^0$, where $w_i^0 = \frac{p_i^0 q_i^0}{\sum_j p_j^0 q_j^0}$
- Paasche $P = \frac{\sum_i p_i^1 q_i^1}{\sum_j p_j^0 q_j^1} = \sum_i (p_i^1/p_i^0) w_i^1$, where $w_i^1 = \frac{p_i^0 q_i^1}{\sum_j p_j^0 q_j^1}$
- Fisher $F = (LP)^{1/2}$
- Törnqvist-Theil $T = \prod \left(\frac{p_i^1}{p_i^0}\right)^{w_i}$, where $w_i = (w_i{}^{t_0} + w_i{}^{t_f})/2$
- Walsh $W = \frac{\sum_i p_i^1 q_i}{\sum_j p_j^0 q_j}$, where $q_i = \left(q_i^0 q_i^1\right)^{1/2}$

Conclusion: $F \sim W \sim T$ when time-series data is *well behaved*

The **IWPI** index: Törnqvist (lower aggregation levels) and Laspeyres (highest levels)

Important **properties**

- consistency in aggregation (approximate/exact)
- time symmetry/circularity $I(t^0, t^1)I(t^1, t^2) = I(t^0, t^2)$

Property:

if log(prices) and weights have monotonic trends => circularity(Törnqvist) = exact

Tested on real data!

WPI based on *observed wages: the matched-models approach*

Use examples: WPI, CPI

- If index is direct, e.g. over (t1,t10)
  -> sample may become non-representative after long time, quality effects may accumulate
- If index is chained e.g. (t1, t2)(t2,t3)…(t9,t10)
  -> index may have chaining drift
- If quality changes are not eliminated by the matching process
  -> they will affect the index

- If index formula is superlative and data monotonous -> *drift is small* (theory and data tests)
- Quality effects may be small/non-significant (see proof)
- New/disappearing items: reduced influence due to weight and sample updates, chaining
- *Mitigating factors: good coverage*, **weight*** & sample updating*

The probability of measuring the hourly wage of the very same contract employee-employer (i.e. with same occupation, economic activity) for two successive points in time decreases with the length of the temporal interval between these events. Empirically, we found that less than a half of these contracts may still be found in the total sample after about three years (while almost five years for the public sector) from the moment they are included in it.

- Note: Törnqvist Index & Census data -> would give the ideal result!

WPI based on *Hedonic-models approach*

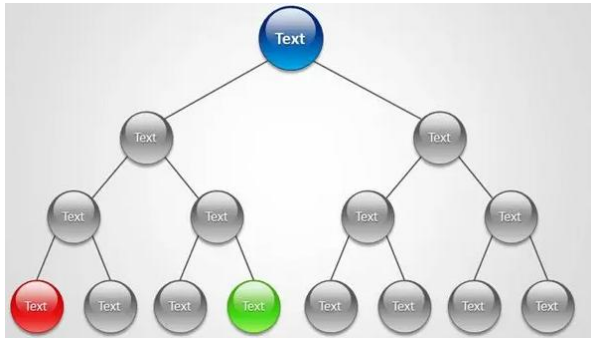
Why explore: recommended test by External Report for IWPI
Use examples: CPI (computers, phones, houses, etc)

- Multiple choices: {Functional form of model * Type of index * Type of imputation}
- *Confidence/prediction intervals* for the index: need to be reported since employing estimates
- May be biased due to:
  - proportion of missing items
  - different variances on observed versus estimated wages
  - functional form/model choice
  - omitted variables effects
- Errors due to measurement/choices of *characteristics sets*
- Regular revision are needed (model, characteristics, data)

It **should** include **other** "item" characteristics (e.g. of employer, working conditions, other subjective quality, …)

- May define transitive, drift free indices
- May include information about all data observed at all time points

Evaluating the contribution of the quality factor

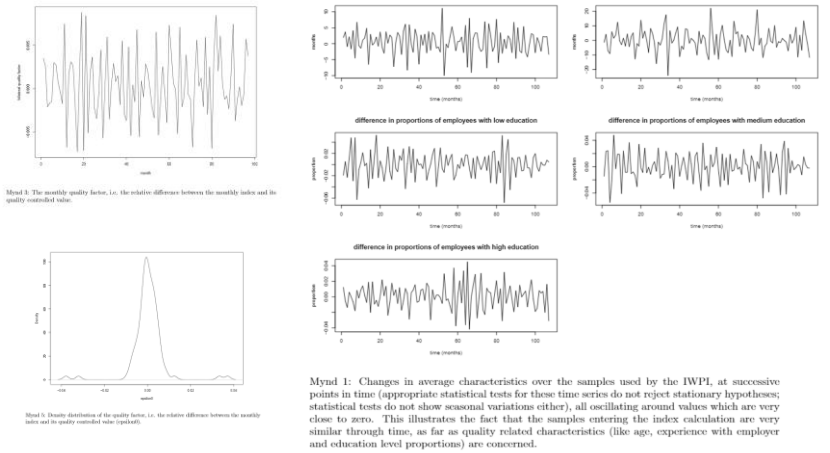


$$\ln p_i^t = \sum_t \delta_t D_t + \gamma_i^{(*)} + \sum_k \beta_k \cdot z_{ik}^t + \epsilon_i^t$$

$$T/\tilde{T} - 1 = \exp\left(\sum_k \hat{\beta}_k \left(\overline{z}_k^{(\Gamma_t)t} - \overline{z}_k^{(\Gamma_{t_0})t_0}\right)\right) - 1$$

$$\frac{T_\Gamma(t,t_0)}{\tilde{T}_\Gamma(t,t_0)} - 1 = B_\Gamma(t,t_0) + O(B^2)$$

$$B_\Gamma(t,t_0) = \sum_k \hat{\beta}_k \left(\overline{z}_k^{(\Gamma_t)t} - \overline{z}_k^{(\Gamma_{t_0})t_0}\right)$$



Mynd 1: Changes in average characteristics over the samples used by the IWPI, at successive points in time (appropriate statistical tests for these time series do not reject stationary hypotheses; statistical tests do not show seasonal variations either), all oscillating around values which are very close to zero. This illustrates the fact that the samples entering the index calculation are very similar through time, as far as quality related characteristics (like age, experience with employer and education level proportions) are concerned.

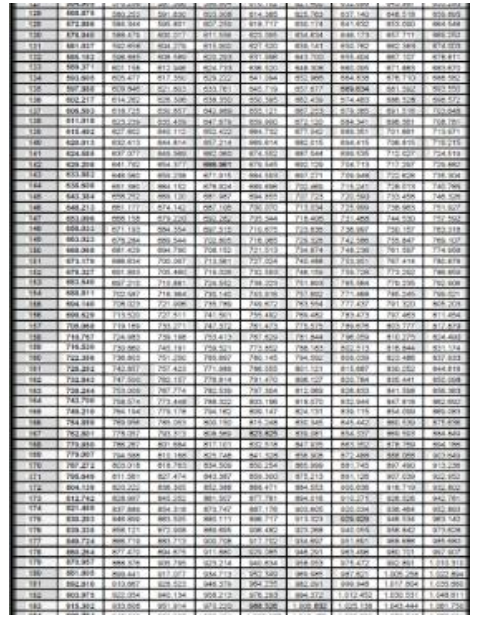
See calculations and data-based tests and plots in [WP2 and External report]

Conclusion

The quality factor for the cell index depends on the difference between the weighted averages (over the cell, at successive points in time) of the characteristics which do vary with time.

Average factor ~ 0.00024 i.e. 0.024%

CI [-0.002, 0.003]



But: *Launatafla*
Extra-effect!
(Input methods!?)

Note:

Hedonic models are not unique!

$$y_i^\tau = h_i^\tau(\beta, z) + \varepsilon_i^\tau = \alpha^\tau + g_i^\tau(\beta, z) + \varepsilon_i^\tau \quad \text{->}$$

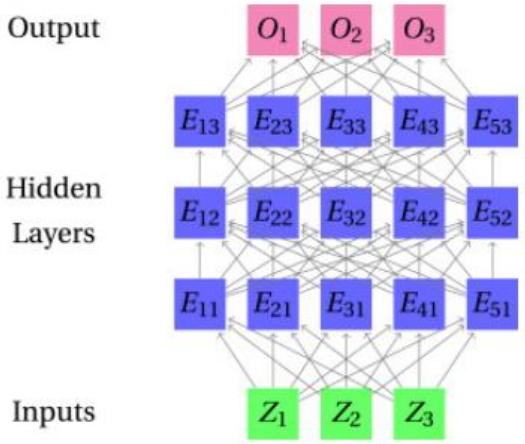
the quality adjustment ratio of the elementary aggregate is:

$$\frac{\hat{I}_\Gamma(t, t_0)}{\hat{I}_\Gamma^{adj}(t, t_0)} = \varphi^{-1}(\mathcal{A}_\Gamma(\hat{g}_i^t(\hat{\beta}, z)))/\varphi^{-1}(\mathcal{A}_\Gamma(\hat{g}_i^{t_0}(\hat{\beta}, z)))$$

- Which values are *estimated based on mode*
  - -*all* wages, at each point in time
  - -only *missing* wages (of new and disappearing items)
- What data is used to *fit* the model:
  - -data observed at *each* time point
  - -pooled data of *two adjacent* time points
  - -pooled data over *many successive* time points and a rolling window (since revisions not popular)
- What type of *model*:
  - - FE and RE effects, regression on which characteristics, which functional forms
- Which price *index* formula (elementary and higher)

Please, use up-to-date hedonic models, e.g. deep NN!

https://www.sciencedirect.com/science/article/pii/S030440762500106X#fig3



Download: Download high-res image (291KB)

Download: Download full-size image

Fig. 3. Standard architecture of a Deep Neural Network. In the hedonic price prediction network, the penultimate layer is interpreted as an embedding of the product's hedonic value and the output layer contains predicted hedonic prices in all time periods. In comparison, the networks used for text and image processing have very high-dimensional inputs and outputs, with intermediate hidden layers composed of neural sub-networks. The dense embeddings typically result from taking the last hidden layer of the network.

Note:

Confidence Intervals of hedonics *are large enough to include the matched model results*

Figure 2: Fixed Base Price Indices

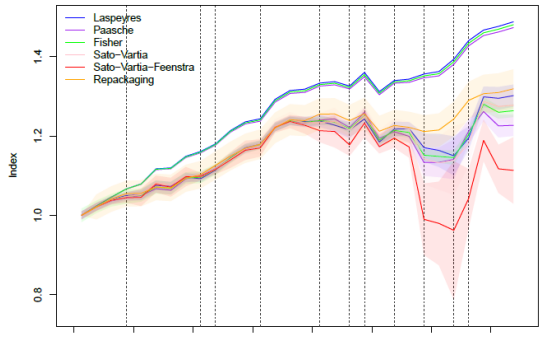
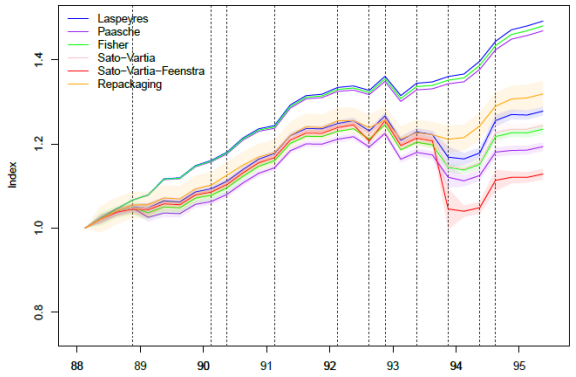


Figure 3: Chained Price Indices



Crawford, I., Neary, J.P., https://hdl.handle.net/10419/198889

Beers, M., DOI 10.1007/s10182-006-0015-9

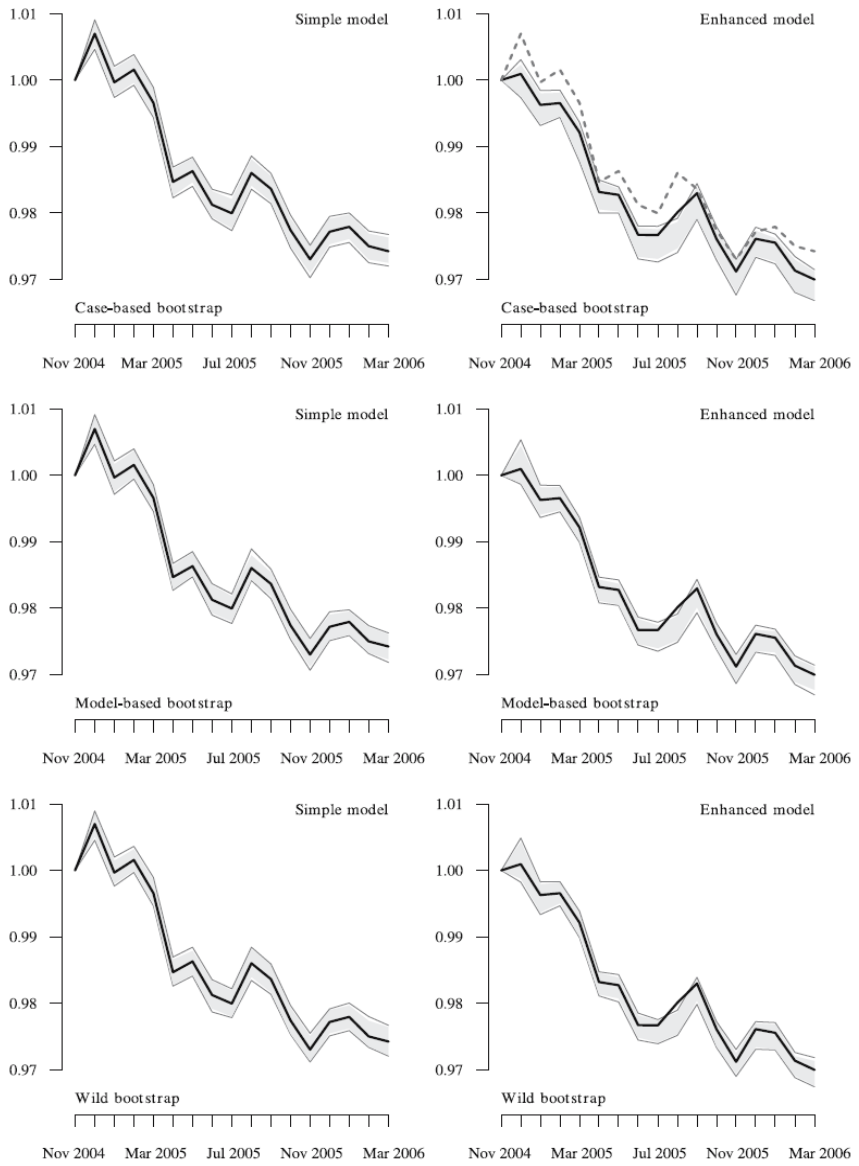


Fig. 1 Jevons-type hedonic elementary price index estimates for Nov. 2004–Mar. 2006 with confidence bands using the case-based, model-based and wild bootstrap approach (from *top* to *bottom*) based on $R = 199$ replications. The solid black line, the shaded area, and the solid grey lines represent the index estimates as well as the 90% and 95% confidence bands respectively. The dashed line in the upper-right frame equals the solid line in the upper-left frame; it is plotted for the sake of easier comparison.

Statistics Iceland: additional index/measures and a theoretical one

	Basic Wage Index	Wage Index	Total Wage index	Taxable wages	Average Earnings Index
Do other payments than basic salary affect the measure?	No	Yes - partly	Yes	Yes	Yes
Changes in regular payments such as shift payments and bonuses.	No	Yes	Yes	Yes	Yes
Changes in overtime payments	No	No	Yes	Yes	Yes
Changes in irregular payments	No	No	Yes	Yes	Yes
Do changes in working time arrangement affect the measure?	Yes - partly	Yes - partly	Yes - partly	Yes - partly	Yes - partly
Shorter workday according to collective agreements	Yes	Yes	Yes	No	No
Changes from full-time to part-time or vice versa	No	No	No	Yes	No
Changes in overtime hours	No	No	Yes	Yes	Yes
Do changes in the labourforce affect the measure?	No	No	Yes - partly	Yes	Yes
Changes in the proportion of low/high wage earners	No	No	Yes	Yes	Yes
Changes in the the number of employees	No	No	No	Yes	Yes
What is measured?	Basic hourly wages	Regular hourly wages	Total hourly wages	All taxable payments	Average wages, affected by composition and structural changes

EMBED DOWNLOAD IMAGE DOWNLOAD PDF

Statistics Iceland

Wage index numbers of other countries?

- Not too many WPIs, mainly due to difficulties of data collection
- The closest: the Australian WPI (matching jobs, see

https://www.abs.gov.au/methodologies/wage-price-index-australia-methodology/mar-2025#overview)

- LCI for European countries and New Zealand
- AWE for UK
- ECI for USA

Conclusions

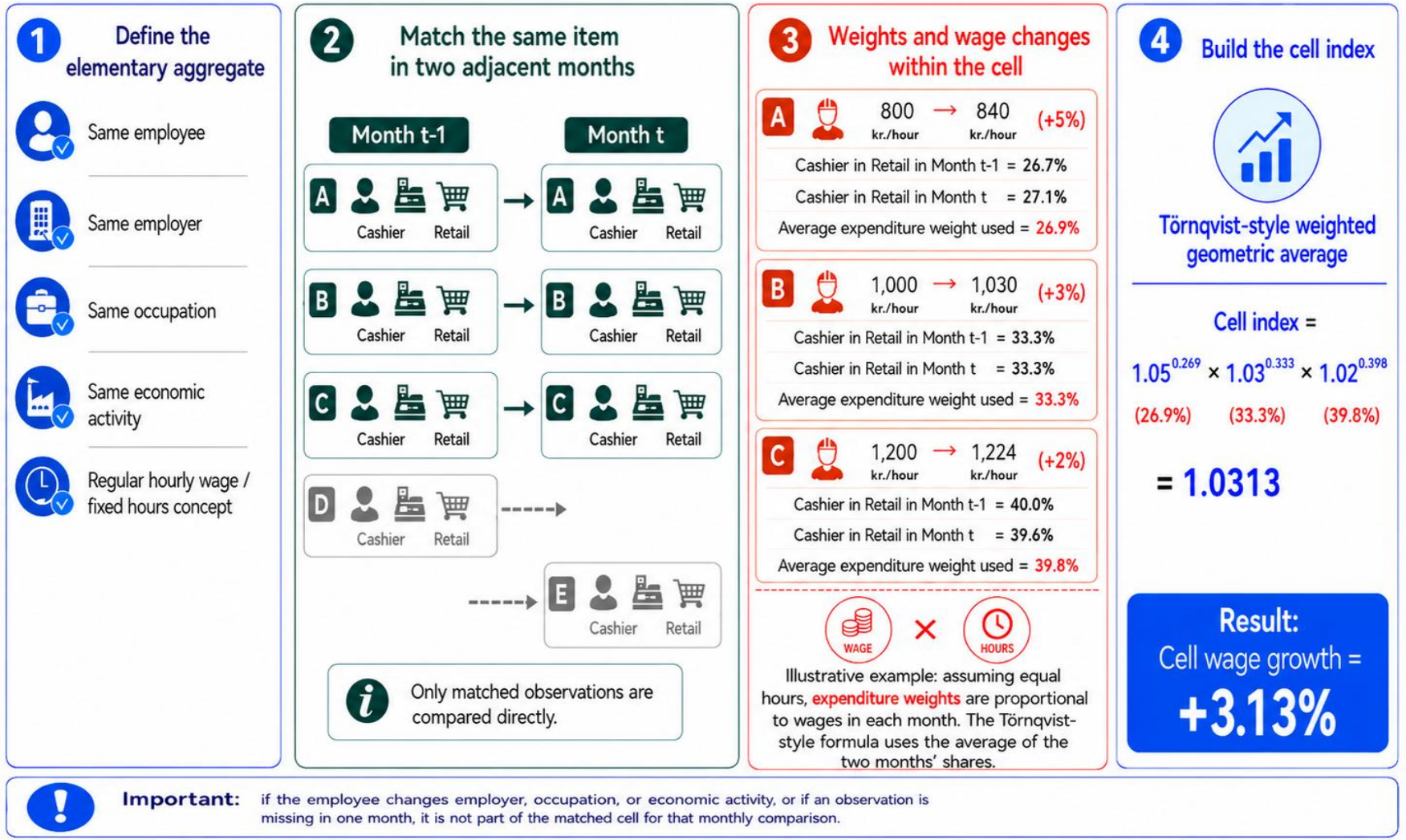
- Solid structure of the IWPI
- It minimises the unwanted effects of:
  - Chaining drift
  - Quality changes
- Challenges: coverage (increasing towards census) and sample refresh
- Publishing complementary index numbers, reflecting different phenomena related to wage dynamics

References*

- *Methodology of the Icelandic Wage Index*.
  Author Violeta Calian. Statistical series: Working paper, 103(20), 16 August 2018.
- *Effects of changes in working experience and education on the Icelandic Wage Index*.
  Author Violeta Calian. Statistical series: Working paper, 104(15), 10 December 2019.
- *Review of the Icelandic Wage Price Index Report to the Wage Statistics Committee Government of Iceland Kim Zieschang.*
- *Wage index Metadata*

Appendices

An Example of the Process Behind the Icelandic Wage Index



AI co-authored picture

Cell defined by:
Occupation & Economic Activity

Discussions / work in progress / future plans

- Data -> coverage, updates, refresh
- Report on the uncertainty bounds of any type of index
- Choices of higher-level index formula ->

    Törnqvist all/ other options

- Report on the effect of quality changes on index (and which ones)