Hagtíðindi
Statistical Series

Working papers

10 December 2019

Effects of changes in working experience and education on the Icelandic Wage Index

Rannsókn á áhrifum starfsaldurs og menntunar á launavísitölu

Statistics Iceland

Icelandic summary

Í þessari greinargerð er greint frá helstu niðurstöðum rannsóknar á því hvort hækkandi starfsaldur og aukin menntun hafi áhrif á þróun launavísitölu Hagstofu Íslands. Rannsóknin er gerð í kjölfar úttektar á launavísitölu sem fór fram haustið 2018. Úttektin var á vegum nefndar um umbætur á úrvinnslu og nýtingu launatölfræðiupplýsinga sem skipuð var af forsætisráðherra og kom til vegna gagnrýni á útreikningum launavísitölu. Í því sambandi var bent á að launavísitalan ofmeti launabreytingar þar sem útreikningur hennar byggir á pöruðum samanburði. Hann felur í sér að mæl er breyting launa á milli samliggjandi tímapunkta hjá fastri einingu sem er algeng aðferð við gæðaleiðréttingar verðvísitalna.

Úttektin var framkvæmd af óháðum sérfræðingi, dr. Kim Zieschang, og var tilgangurinn að meta aðferðir og gögn sem liggja til grundvallar útreikningi á launavísitölu ásamt því að leggja til úrbætur eftir þörfum. Niðurstöður úttektar voru þær að launavísitalan er traust, byggir á aðferðafræðilega sterkum grunni og áreiðanlegum gögnum af nægilegri þekju. Í úttektinni var mælt með Hagstofan breyti hvorki gagnasöfnun né aðferðum sem liggja til grundvallar launavísitölu en kanni hvort launavísitalan feli í sér áhrif af auknum gæðum vinnuafls yfir tíma með hækkandi starfsaldri og aukinni menntun. Á grundvelli úttektar beindi nefndin þeim tilmælum til Hagstofunnar að rannsakað yrði hvort hækkandi starfsaldur og aukin menntun hefði áhrif á launavísitölu og yrði þar tekið mið af tillögum úttektaraðila. Rannsóknin tók því mið af tilmælum nefndar um umbætur á úrvinnslu og nýtingu launatölfræðiupplýsinga og samráð var haft við úttektaraðila, dr. Kim Zieschang um rannsóknina.

Menntun og starfsaldur eru dæmi um eiginleika vinnuaflsins sem breytast yfir tíma. Í tilfelli launavísitölu er mikilvægt að hafa í huga að breytingar á þessum eiginleikum vinnuaflsins tekur ekki til einstaklinga heldur ráðningasambands. Ráðningasamband er sú mælieining sem liggur til grundvallar launavísitölu, það er launagreiðandi, launamaður, starf og atvinnugrein. Ef einstaklingur fær hærri laun í kjölfar þess að hann skiptir um starf eða fer að vinna hjá öðrum launagreiðanda þá mælast þær breytingar ekki í launavísitölu þar sem ráðningarsamband hans hefur rofnað. Greining á gagnasafni launavísitölu leiðir í ljós að um helmingur ráðningasambanda hefur rofnað eftir þrjú ár. Hækkandi starfsaldur og aukin menntun geta skilað auknum gæðum vinnuafls og leitt til launahækkunar en þó er ekki hægt að ganga að því vísu, til dæmis þegar menntun er ótengd viðkomandi starfi.

Helstu aðferðir sem beitt er við gæðaleiðréttingar á verðvísitölum er pörun (e. matched sample models) og spár eða tilreikningar á grundvelli aðhvarfslíkana (e. hedonic models). Í launavísitölu er byggt á pöruðum breytingum reglulegs tímakaups milli tveggja samliggjandi mánaða að því tilskyldu að ráðningarsamband sé óbreytt. Áhrif breytinga menntunar og starfsaldurs á launavísitölu voru óþekkt en núverandi aðferðir mæla ekki þá þætti. Auk þess hafa áreiðanlegar upplýsingar um menntun og starfsaldur ekki verið aðgengilegar fyrr en síðustu ár. Í þessari rannsókn var byggt á aðhvarfslíkönum (e. hedonic models), sem þekkt eru úr fræðum verðvísitalna, til að meta áhrif starfsaldurs og menntunar á launavísitölu. Líkönin gera það kleift að bera saman mat á launabreytingum með og án áhrifaþátta og meta þannig áhrif menntunar.

Helstu niðurstöður rannsóknarinnar voru þær að frávik á milli vísitalna sem reiknaðar eru á grundvelli líkana með eða án menntunar og starfsaldurs eru ekki tölfræðilega marktæk á milli mánaða. Áhrifin eru ekki línuleg og geta ýmist verið til hækkunar eða lækkunar á mánaðarlegri launavísitölu. Niðurstöður, sem byggja á mánaðarlegum mælingum launavísitölu árin 2008 til 2018, benda hins vegar til þess áhrifin geti safnast upp til lengri tíma. Áhrifin eru að meðaltali 0,024% á mánuði (CI = [-0,2%, 0,3%]) og 2,9% að meðaltali eftir 10 ár. Hafa ber í huga að því lengra tímabil sem er skoðað því víðari verða öryggisbilin og þar sem þau innihalda alltaf núll er ekki hægt að útiloka að áhrifin séu engin.

Rannsóknin er einstök í sinni röð og eftir því sem Hagstofan veit best eru ekki til aðrar rannsóknir, hvorki innlendar né erlendar, um þetta efni. Ljóst er að fleiri rannsóknir þarf til að skýrari vísbendingar fáist. Þá þarf að hafa í huga að það á ekki alltaf við að aukinn starfsaldur og menntun séu gæðabreytingar. Einnig þarf að huga að því að tímabilið sem lagt er til grundvallar er sérstakt með tilliti til efnahagsþróunar og því er ekki hægt að útiloka að annað tímabil hefði skilað öðrum niðurstöðum.

Úttekt og rannsóknir á launavísitölunni gefa því ekki tilefni til að breyta útreikningum. En mikil gróska er í rannsóknum á aðferðum á sviði verðvísitalna um þessar mundir og telur Hagstofan áhugavert að rannsaka frekar fleiri aðferðir við útreikninga á verðbreytingum.

Auk umfjöllun um rannsóknina má finna í greinargerðinni lýsingu á aðferðum launavísitölu (sjá einnig lýsigögn um launavísitölu) og umfjöllun um mat á launaþróun ef notast er við niðurstöður launavísitölu annars vegar og meðallauna hins vegar.

Effects of changes in working experience and education on the Icelandic wage index

Working Paper

10.12.2019

In this paper we investigate the effect of changes in factors such as working experience (total and with employer) and education of employees on the Icelandic wage price index (IWPI) through time. This is due to the fact that the variables included in the calculation of the wage index, like employee-employer contract, occupation and economic activity are not always reflecting quality changes events at the desired level of detail. The data used for the present paper has been used by the current index calculation system for the past ten years and is enhanced with the corresponding employee attributes.

The main conclusions of the analysis are: (i) if the relative variation of the *monthly* index which is due to quality changes is defined by $\epsilon_0 = T/\tilde{T} - 1$, where $T, \tilde{T}$ are the wage index and its value when controlling for quality changes by wage modelling, then it has an average $\bar{\epsilon}_0 = 0.00024$, i.e. 0.024% and a 95% confidence interval of $[-0.002, 0.003]$ and thus not statistically significant; (ii) this average value grows linearly with the number of index chaining time steps, while the confidence interval expands as well. The *average* reaches a value of 0.0288 after about 10 years and its confidence interval still covers the value zero although much wider.

The difference between the IWPI and an index based on average earnings (AEI) is explained as well, in the second part of the paper. The difference is mainly due to the distinct sample structures used by the two approaches but is also generated by the differences between ratios of geometric and ratios of arithmetic means employed when calculating these two index numbers. The samples mirror structural and quality changes of the labour market such as employment levels or composition and amount of hours worked in the case of AEI but mainly market induced price changes in the case of IWPI.

1 Introduction

The goal of this paper is twofold:

- to evaluate the influence of factors such as working experience (total and with employer)[1]and education[2] of employees on the Icelandic wage price index (IWPI)
- to explain differences between the IWPI and indices based on other averaging formulae and sampling structures, like the average earnings formula and corresponding sampling

[1]Age is used as a proxy for total working experience

[2]ISCED levels grouped as low (ISCED less than 3) medium (ISCED between 3 and 5), high (ISCED=6), very high education (ISCED bigger than 7)

The first evaluation is suggested since factors such as experience, age and education of employees change in time while the data used when calculating the wage index, e.g. employee-employer contract, occupation and economic activity, do not always reflect these changes at the desired level of detail.

The scrutiny of multiple wage index definitions is due to the fact that different formulae/sample structures often generate divergent results while the causes are not always made explicit for the users. We build simple examples to illustrate and explain when each index is more appropriate, depending on the goal of the users' analysis. In the process, we define the index number based on average earnings in order to make the comparison meaningful.

The wage index is regarded as a price index for hourly wages. The general price index problem is to measure the changes in prices between two points in time at constant quality, changes solely determined by market forces, while: (i) the sample evolves with time (new and disappearing items), (ii) the quality of persistent items can change with time, (iii) long time series of index changes need to be defined, (iv) other errors can be present, including measurement errors.

There are two solutions for this problem:

1. based on observed wages, i.e. exclusively on data: this is the matched models approach for constant quality. If this is based on a direct index, as has been the CPI practice in the past, the sample becomes non-representative and residual quality effects accumulate. If a chained index is used, then this can lead to chaining drift, the drift size size depending on the index formula, on data characteristics and length of chaining. If not all quality changes are captured by the matching process, then they are captured by the index. However, there are cases when the drift can be kept to small values, i.e. by using superlative index formulae (Törnqvist or Fisher) and when employing data which is monotonous, as is the case for wages. Residual quality effects may also be kept at small or even non-significant levels, depending on the length of time series, on the matching process and the sample structures.

2. based on imputed wages, i.e. the hedonic models' approach for constant quality. A good example, in consumer price index (CPI) context is given in (1). Hedonic methods, like all modelling solutions generate in addition to the estimates for the index values, confidence intervals which need to be reported as well[3]. They can also be confronted with biases due to the prevalence of missing items, differences in wage variances over matched and unmatched items, or to the choice of functional form of the model. On the other hand, one may define transitive, drift free indices and may include all information about data at all time points when building some of the hedonic types of indices.

The literature has been growing recently due to research on hedonic modelling, we only mention a very good overview here given in (2) and references therein. The main research problems are related to: (i) the choice of prices (in our case wages) to be imputed: all or only the missing ones[4], (ii) the choice of model for fitting the data: for each time point, for two adjacent ones, for many successive time points and using rolling windows (since index revisions are not popular), (iii) the choice of the type of model: fixed effects model, regression based on sets of characteristics, functional form, (iv) the choice of index formula.

Differences between hedonic and matched models (for one chain link) are due to several factors: (i) how different the imputed wages of matched and non-matched items are, at the moment of

[3]The data based approach is also accompanied by sampling errors which can be kept low by using large samples

[4]I.e. for new/disappearing items at the moment before/after joining/separating from the sample of matched items

separation (leaving or joining the sample), (ii) how different the weights of the matched items are, between successive points in time. The ratio of imputed wages depends in turn on the regression coefficients and on differences in average characteristics over samples at successive points in time.

In this paper we compare the matched model index based on observed wages and a matched model index which controls for residual quality changes by using imputed values[5] based on hedonic models. This is done for bilateral indices in order to avoid combining drift and quality effects, while estimating the average long term effect by the product between the average monthly effect and the number of chaining steps. It is also done by comparing the corresponding multilateral, drift free indices.

The paper is structured as follows. In the next section we describe the main findings of our analysis, in section 3 we give a short review of the index and data. Section 4 is dedicated to the hedonic modelling of wages and its effect on the IWPI. We give several notes about shortcomings and advantages of such models in the next section and a detailed note on the differences in calculations between the wage price index and average earnings index in the last section of the paper. Details of theoretical calculations and illustrative examples are listed in the Appendices.

2 Summary of findings

The main results of data analysis, modelling and theoretical calculations are presented in this section.

1. *The dynamic nature of the sample, matching and chaining for the multilateral wage index*

   Any bilateral price index[6] aims at measuring the average market induced changes in prices, in our case hourly wages, between two points in time. If the population or sample used for such calculations is always the same (static universe), a bilateral, direct index could be calculated between any two moments, no matter how far from each other. In real life, the sets of hourly wages which can be measured are usually different from one moment to the next (dynamic universe). Moreover, the set of comparable wages between t_i and t_{i+1} (matched items) is usually different from the set of comparable wages at t_j and t_{j+1} when $i \neq j$.

   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.

   The events which create this continuous evolution of the samples and of the matched item sets are: appearance and disappearance of items, i.e old contracts are replaced by (different number of-) new ones. Chaining the bilateral index values is therefore a well known solution for building a multilateral index, although not without disadvantages, the most celebrated one being the chaining drift. One should note that, as previously confirmed for our wage data (3), based on (4), (5) and (6), the Törnqvist index used for the IWPI has the smallest drift of all superlative indices, especially for longer periods.

[5]All items change between two points in time since the employees' total experience for instance is always different between these points

[6]Measuring wage changes between two successive points in time

2. *The effects of quality changes*

   One may control for changes in quality as the ones defined by the time - evolving age, experience and education levels of employees by (hedonic regression) modelling the wage data as a function of these characteristics. The corresponding relative variation of the (bilateral, monthly) wage index depends then, as shown in section 4, on the products between: the regression coefficients and the corresponding variation in average sample characteristics between successive time points.

   Due to the dynamic nature of the hourly wage samples described above, the average values of the characteristics oscillate through time[7]. Figure 1 illustrates the average changes over all economic sectors[8], noting that such changes between successive points in time are small and have both negative and positive signs.

   The numerical values of the regression coefficients of the present data models are in agreement with studies on gender payed gap (GPG) of Statistics Iceland (8), (9). Details are found in the Appendix.

3. *The size of the relative variation of the IWPI induced by quality changes*

   The size of the monthly relative variation of the wage index T with respect to the index controlled for quality changes $\tilde{T}$ is defined by $T/\tilde{T}-1 = \epsilon_0$. The point estimate of $\bar{\epsilon}_0$ is 0.00024, with a 95% confidence interval of $[-0.002, 0.003]$ and a symmetric density distribution. This average value grows linearly with the number of chaining time steps (N), as also pointed out in (7) as $N\bar{\epsilon}_0$. The confidence interval expands accordingly. The empirical density distribution of the ϵ_0 term for the whole labour market is shown in the Appendix. The Appendix also includes the empirical distributions of ϵ_0 by occupation and economic activity (for the private sector) and by sub-sector (central and local) for the public one.

   The present analysis is concerned with the evaluation of the effect of quality changes on the wage index based on matched model Törnqvist - formula as used in practice. Our calculations are similar to the analysis in (2). In addition, we analysed the effect of aggregation and the influence of including fixed effects in the models with and without characteristics.

4. *The difference between the wage price index (IWPI) and average earnings based index* (AEI)

   As explained in section 6, where several illustrative examples are discussed, this is due to two main causes: (i) the sample structures used by the two different formulae are different, as previously explained in (3), (ii) ratios of geometric and arithmetic averages are not identical and their discrepancy has no preferred direction, (iii) the IWPI is a price index while the AEI involves quantity. The IWPI includes, at each point in time, matched items, i.e. items which have been in the sample at least at the previous moment in time. The purpose of IWPI is to reflect only price changes but not quality and structural changes. In contrast, the AEI includes all items that are sampled throughout the year, thus capturing exactly the effect of structure and quality changes, e.g. the ones driven by the variation in the number of employees or in the number/composition of hours worked for instance.

[7] This remark is in agreement with the calculations in (7), made in the context of evaluating the turnover effects.
[8] Analogue results stand for individual sectors

3 Short description of IWPI and data structures

The IWPI is a price index which measures the changes in hourly wages paid to an employee for fixed working hours within the same occupation by the same employer. These changes are combined in accordance with the price index theory by using a Törnqvist index formula for the lower aggregation levels and linear combination of the obtained index numbers for higher level of aggregation, as described in (3). The sector index numbers are chained for obtaining times series of index values.

The wages used in the calculation of IWPI are regular hourly wages, i.e. all basic wages paid for both day-time and shift-work hours as well as fixed wage contract hours, including additional payments like fixed overtime and bonuses settled regularly in each wage period. Overtime and other irregular payments are excluded as well as employers' social contributions and taxes[9].

The data used for calculating the index is mainly data collected through the Icelandic Survey on Wages, Earnings and Labour Costs (ISWEL; Launarannsókn Hagstofu Íslands) conducted by Statistics Iceland. The ISWEL data are aimed at adequately representing the population on the private market, municipalities and central government. Administrative data are also utilised in the calculations of the IWPI, primarily PAYE (Pay as your Earn) data. The monthly PAYE data reflects for example the sum of the taxable income of workers for the entire labour market, irrespective of sector or employer's size, but does not contain information on working hours. Henceforth, the use of PAYE data is primarily on constructing a sample frame for ISWEL and weights for various statistical products in the wage statistics such as the IWPI. For the present study, monthly IWPI data of all economic sectors for the period 2008-2018 were used, after additional characteristics of employees and employers, needed for hedonic modelling[10], have been added to the data set used in production.

4 Hedonic wage modelling effect on the IWPI

The goal of hedonic modelling is to control for the effects of variation in quality, as reflected by a set of characteristics, as originally developed for problems related to the CPI.

Two approaches were considered for such modelling of the enriched wage index data with further attributes of employees and companies: (i) a unique model (for each sector) where the regression coefficients of the characteristics are independent of time, data is pooled over many time periods and time dummy variables are introduced, and (ii) multiple independent hedonic models, for each point in time (and each or pooled cells in a sector). A good comparison of all types of hedonic models is shown in (1), where they all give very close results, although the distance between hedonic and matched models grows with time.

For the present analysis, a regression model[11] with time dummy variables and a complete set of characteristics of the wage index matched item data set was chosen since best fitting while minimizing omitted variable bias and small sample effects. The main time - incrementing factors were identified as : experience (of employees in a given company), education and age of the employees. Age is also a proxy for total experience in Iceland where population is rather homogeneous as illustrated by the age of starting to work which is very young in general. Higher powers of age and experience enter the model as well, $\beta_{age^2}(age_i^t)^2$ and $\beta_{experience^2}(experience_i^t)^2$.

[9] https://statice.is/publications/metadata?fileId=19576
[10] This new data structure has become available to Statistics Iceland in recent years.
[11] As also recommended in (7)

In this setting one defines new wage values where the effect of age, experience and education is controlled for, by:

$$\ln(\tilde{p})_i^t = \ln p_i^t - \sum_k \hat{\beta}_k \cdot z_{ik}^t \tag{1}$$

where $\hat{\beta}_k$ are model point estimates of the hedonic regression coefficients. Therefore, as shown in the Appendix, the relation between the corresponding wage indices is:

$$\ln T = ln\tilde{T} + \sum_k \hat{\beta}_k \left(\overset{(\Gamma_t)}{\overline{z}_k^t} - \overset{(\Gamma_{t_0})}{\overline{z}_k^{t_0}} \right) \tag{2}$$

The relative variation of the (occupation - economic activity cell or occupation - labour union cell) index with respect to its quality controlled value is:

$$T/\tilde{T} - 1 = \exp\left(\sum_k \hat{\beta}_k \left(\overset{(\Gamma_t)}{\overline{z}_k^t} - \overset{(\Gamma_{t_0})}{\overline{z}_k^{t_0}} \right) \right) - 1 \tag{3}$$

In the equations above, $\hat{\beta}_k$ are the estimated model parameters and the averages $\overset{(\Gamma_t)}{\overline{z}_k^t}$, $\overset{(\Gamma_{t_0})}{\overline{z}_k^{t_0}}$ of the characteristics are calculated *over the samples of (cell) items at time* t, t_0. Details, including the effect of multiple aggregation levels and chaining of the index, typical for price index numbers, are given in Lemma 1 - Lemma 4 of the Appendix.

If one fitted a model for each time point, the calculations would be very similar and the quality related factor would be:

$$\exp\left(\sum_k (\hat{\beta}_k^t \overset{(\Gamma_t)}{\overline{z}_k^t} - \hat{\beta}_k^{t_0} \overset{(\Gamma_{t_0})}{\overline{z}_k^{t_0}}) \right) - 1 \tag{4}$$

with $\hat{\beta}_k^t, \hat{\beta}_k^{t_0}$ the estimates obtained by fitting distinct for t and t_0 respectively.

Hedonic indices can be built in several ways: by using single imputations of wages which appear/disappear between two moments t, t_0 , by using double imputations, i.e. for all wages, present and not present in the matched samples , by using so-called exact/superlative hedonics as in (10) , (11) or (12) or full imputation time dummy indices as analysed in detail in (2).

The terms in the equations (3) or (4) depend on: (i) the coefficients of the hedonic models $\hat{\beta}_k$ or $\hat{\beta}_k^t$ (ii) the changes in average characteristics on the samples between two time points.

The differences in averages of characteristics between successive points in time are illustrated in Figure 1, for all the attributes. The figure shows how very little these quantities vary through time and that these variations can be either positive or negative. This is due to the fact that the sample of items entering the index at successive points in time changes continuously, which is the very reason why one needs to chain the index.

Two cases may be identified when analysing the quality related factor:

1. For a static sample, $\overline{z}_k^t - \overline{z}_k^{t_0}$ (where averages are calculated on the sample at time t and at time t_0 respectively, for unchanged, static, sample) would be the same as $\overline{(z_k^t - z_k^{t_0})}$, equal to one and always increasing between two successive time points for attributes like age and experience. This difference would grow linearly with time and a *direct*, i.e. non-chained index would suffer of sample degradation although not of chain drift.

2. For a dynamic sample, as in the case of the real life data, the changes in characteristic averages oscillate through time. In the case of the age-characteristic (proxy of total experience) for instance, this oscillations are around a mean of the order of 0.1 (months) as shown in Figure 1. One has the options of chaining a bilateral or calculating a multilateral, drift-free, time dummy hedonic index as described in (2), with equivalent results for the monthly quality effect.

The estimated monthly size of the quality effect defined as the relative variation of the index $T/\tilde{T} - 1 = \epsilon_0$ has an average $\bar{\epsilon}_0 = 0.00024$ or 0.024% and a density distribution showing that both negative and positive changes, albeit small, can be induced by changes in quality as the ones described previously. The effect is quite similar on the private and public sector: (i) the mean for the private sector is 0.00026 with a $95\% C.I. = [-0.006, 0.006]$ and (ii) the mean for the public sector is 0.00027 with a $95\% C.I. = [-0.005, 0.005]$. The average value of the quality effect does increase with the chaining number of steps to $N \cdot \bar{\epsilon}_0$ where N is the number of monthly time steps. It should be kept in mind that the width of the confidence interval of this value grows with time as well. The empirical density distribution[12] of the monthly quality factor is illustrated in the Appendix. Distributions are also shown for sub-aggregates of the main economic sectors, i.e. by occupation and economic activity for the private sector and by sub-sector (local or central) for the public one.

5 Advantages and disadvantages of using hedonic models

The use of hedonic regressions as described in the previous section is not the ideal solution to the problem of building an improved price index. Following are several difficulties which may arise during such an excercise and which concern both data and user requirements.

The estimated coefficients of such a models are biased, if one does not account for the exponential transformation involved, although this bias is in general small. Another type of bias of these indices may arise due to the effect of the prevalence of imputed wages of unmatched items.

If a time pooling approach is adopted, the model fitting is done for the whole data available at a given point in time and the coefficients of such model change when new data points are added. By contrast, revisions are not easily accepted by the users of a price index. A rolling window solution to this problem has been proposed in (13) and tested for a particular price index in (15).

If the opposite approach is attempted, i.e fitting a hedonic model for each point in time, some aggregates are not sufficiently large for the purpose and new aggregation structures have to be defined for each instance, based on yet to be defined homogeneity measures.

The quality controlled price indices obtained by following this type of procedure are characterised, in addition to modelling errors, by errors due to the structural form of the hedonic functions, by measurement errors for all characteristics, by errors in defining the set of characteristics, the fit of the models and the number of observations.

On the other hand, time dummy full imputation index numbers based on hedonic models offer a fast processing and a no drift, transitive multilateral type of index, once a model has been evaluated and defined as optimal for the data. In addition, such an index makes use of all sample data available, including both the separated and the matched items.

[12] An ergodic assumption is employed for this purpose.

6 When are the IWPI and AEI different

Since users may be interested in more than one measure of wage variations published by Statistic Iceland, we describe here what generates discrepancies between them. We conclude by explaining when one measure is more useful than the other for practical applications.

We start with four oversimplified examples for the purpose of illustrating the difference in the behaviour of the IWPI and an index based on average earnings (AEI), due to their very definitions as detailed in the Appendix 8.3. These are slight generalizations of the numerical examples in (17):

1. if only market based changes in wages take place between two successive points in time (T_0, T_1), without changes in the number of employees, hours worked and quality, then: the bilateral indices satisfy $IWPI = AEI$, at least approximately, the small difference being generated by technical reasons as explained in the paragraph below

2. if only the number and/or composition of hours worked changes between T_0, T_1, everything else being the same (hourly wages, number of employees, quality), then: IWPI is constant, while AEI increases if the number of hours increases or *decreases* if the number of hours does so

3. if only the number of employees increases between T_0, T_1, while wages, time worked and quality are constant, then IWPI is constant while AEI: (i) increases if the new employees have wages which are all higher than the average wage at T_0; (ii) *decreases* if the new employees have wages which are all lower than the average at T_0

4. if only wage changes due to factors unrelated to the market and not measured by any characteristic of the job contracts take place, while the number of employees and hours worked are constant between T_0 and T_1, then: AEI changes exactly like in the case 1 above, while IWPI: (i) does not change if the matched items mechanism can detect the causes (ii) changes in the same way as AEI

In general, the dynamics is determined by a mixture of all the cases above, in different proportions, and it is not easy to disentangle their individual contribution to the index movements.

Several technical differences make more subtle contributions as well:

1. the monthly IWPI is (as a Törnqvist index) the ratio of two geometric, weighted, means of hourly wages; the chained IWPI over a year is the ratio of two 'double geometric means' of hourly wages $p_i(\tau)$: over all items $i \in S$) and over time $\tau \in [T_0, T]$. Here S is an economic sector or a cell and T_0, T are successive years.

2. an index based on the average earnings over a year is the ratio of two 'double arithmetic means' of hourly wages $p_i(\tau)$: over time $\tau \in [T_0, T]$ and over items $i \in S$.

3. while an arithmetic mean is greater than a geometric mean (and equal only when all hourly wages would be identical), the difference between *the ratio of* two arithmetic means and *the ratio of* two geometric means does not have a fixed direction. The differences are greater for more diverse hourly wages.

4. the order of averaging, over time and over items in sample is different for the two procedures and

The procedures are based on different sample structures:

1. matched items, for the IWPI, with the goal of measuring only price changes and little else, versus
2. all items in sample throughout the year (weighted) with the goal of capturing changes in quality and in structure.

In conclusion, users may decide on AEI or IWPI depending on their goals. If they are trying to estimate market induced wage changes, then the IWPI should be the right tool. If their aim is to capture structural changes (e.g. employment) and to estimate current levels of average wages, then the AEI is the most appropriate measure.

7 Conclusions

In this paper we have analysed the effects, on the IWPI, of controlling for changes in quality as measured by several characteristics like age, experience with employer and education of employees. These can be regarded as quality changes as long as they are not reflected in the usual variables measured for the purpose of calculating the IWPI.

The monthly quality effect is small, its average grows linearly with time but its values show symmetric density distributions, with both positive and negative signs. This is due to the growth rates of the IWPI and of the wage index based on hedonic modelling which are different at each point in time. The difference does not have a fixed direction, although the IWPI is overall faster. Therefore applying potential quality controlling factors to the IWPI would decrease the current index values at some points in time but also increase it at others. The very long term chained indices would be accompanied by higher values of the total average quality effect although with wide confidence intervals as well. The latter are signaling high uncertainty about the value of the quality effect but also higher probabilities of observing both small or high actual values. Should there be any large oscillations in wages at the same time as large quantity changes, the risk of growing drift and quality effects would increase as well.

The modelling approach explored here could be considered as a candidate when designing a new type of index, after a careful testing along the lines mentioned in section 5. Practical solutions for avoiding to reach long term significant values of the quality effects while simultaneously avoiding growing drift values would include using shorter spans of index time series, building multilateral (time dummy type) indices or chaining with adaptive frequency. Employing any of these methods would require rigorous data driven evaluation. In addition, the effect of quality changes on the employer side (not only on the employee) should be explored and evaluated.

The comparison between IWPI and AEI was made explicit by using their definitions and specific limiting cases as examples in order to illustrate the factors which can and *should* create divergencies between them, since these two measures are related but not describing identical phenomena.

8 Appendices

8.1 Theoretical calculations of the hedonic modelling effects on the IWPI

Lemma 8.1. *: the quality factor of the cell Törnqvist index*

If for each item i, at each point in time t, the wages controlled for quality changes are defined by:

$$\ln \tilde{p}_i^t = \ln p_i^t - \sum_k \hat{\beta}_k \cdot z_{ik}^t \tag{5}$$

while the time dummy (D_t), with or without fixed effects (γ_i included or not included γ_i^*) wage models are:

$$\ln p_i^t = \sum_t \delta_t D_t + \gamma_i^{(*)} + \sum_k \beta_k \cdot z_{ik}^t + \epsilon_i^t \tag{6}$$

then the relative variation of the bilateral Törnqvist index of a cell aggregate Γ, between time t_0 and time t, due to the effect of quality changes, is:

$$\frac{T_\Gamma(t,t_0)}{\tilde{T}_\Gamma(t,t_0)} - 1 = B_\Gamma(t,t_0) + O(B^2) \tag{7}$$

with the factor :

$$B_\Gamma(t,t_0) = \sum_k \hat{\beta}_k \left(\overset{(\Gamma_t)}{\overline{z}_k^t} - \overset{(\Gamma_{t_0})}{\overline{z}_k^{t_0}} \right) \tag{8}$$

where $\overline{z}_k^{t_0}$ and $\overline{z}_k^t$ are weighted averages of the values of the characteristics of the items found in the cell Γ at time t_0 and at time t. We denoted $O(B^2)$ higher order terms in the series expansion of the exponential, which are much smaller than the first order term B_Γ.

Note that the index variation due to quality can be equivalently expressed, up to a small bias due to the structural form of the model, as:

$$\frac{T_\Gamma(t,t_0)}{\tilde{T}_\Gamma(t,t_0)} - 1 = exp(\delta_t - \tilde{\delta}_t) - 1 \tag{9}$$

when the same reference point t_0 is used for fitting two alternative models: one which includes and which does not include the quality reflecting characteristics.

Proof

Direct calculations, based on (1), series expansion of an exponential and Törnqvist index formula:

$$\ln T_\Gamma(t,t_0) = \sum_{i\in\Gamma} \omega_i(t,t_0) \ln \left(\frac{p_i^t}{p_i^{t_0}} \right) \tag{10}$$

where the weights $\omega_i(t,t_0) = \frac{1}{2}(\omega_i(t) + \omega_i(t_0))$.

Lemma 8.2. *: the quality related factor of the Törnqvist sector index*

If the variations of cell index numbers are given by (8), then the bilateral, sector Törnqvist index numbers satisfies:

$$\frac{T_S(t,t_0)}{\tilde{T}_S(t,t_0)} - 1 = B_S(t,t_0) + O(B^2) \tag{11}$$

where

$$B_S(t,t_0) = \sum_k \hat{\beta}_k \left(\overset{(S_t)}{\overline{z}_k^t} - \overset{(S_{t_0})}{\overline{z}_k^{t_0}} \right) \tag{12}$$

while $\overset{(S_t)}{\overline{z}_k^t}$ *and* $\overset{(S_{t_0})}{\overline{z}_k^{t_0}}$ *are the weighted averages of characteristics of the items in sector* S *at time* t *and time* t_0 *respectively. These sector weights are the product between the item weight in a cell and the cell weight in the sector.*

Proof

$$\ln T_S(t,t_0) = \sum \lambda_\Gamma T_\Gamma(t,t_0) \tag{13}$$

and

$$\ln \tilde{T}_S(t,t_0) = \sum \lambda_\Gamma T^a_\Gamma(t,t_0) \tag{14}$$

thus

$$\ln T_S(t,t_0) - \ln \tilde{T}_S(t,t_0) = \sum_{\Gamma\in S} \lambda_\Gamma \sum_k \hat{\beta}_k \left(\sum_{i\in\Gamma(t)} \omega_i z^t_{ik} - \sum_{i\in\Gamma(t_0)} \omega_i z^{t_0}_{ik} \right) \tag{15}$$

which is equivalent to (12) if we identify the sector averages with

$$\overset{(S_t)}{\overline{z}_k^t} = \sum_{\Gamma\in S} \sum_{i\in\Gamma(t)} \lambda_\Gamma \omega_i z^t_{ik} = \sum_{\Gamma\in S} \sum_{i\in\Gamma(t)} \Omega_{i\Gamma} z^t_{ik} \tag{16}$$

Lemma 8.3. *: the quality factor of the sector index used in practice*

If the relative variation of the sector Törnqvist index is given by (12) and the index used in practice is Laspeyres type as defined by:

$$L_S(t,t_0) = \sum_\Gamma \lambda_\Gamma T_\Gamma(t,t_0) \tag{17}$$

where λ_Γ *are cell weights, then:*

$$\frac{L_S(t,t_0)}{\tilde{L}_S(t,t_0)} = 1 + B^L_S(t,t_0) + O(B^2) \tag{18}$$

with $\tilde{L}_S(t,t_0)$ *the quality controlled index and with* $B^L_S(t,t_0)$ *a weighted average of Törnqvist cell factors, with "renormalised"weights depending on all cell weights* λ_Γ *and all cell index numbers* T_Γ*.*

Proof

Inserting the expression of $\tilde{T}_\Gamma(t,t_0) = T_\Gamma(t,t_0)e^{-B_\Gamma(t,t_0)}$ into

$$\tilde{L}_S(t,t_0) = \sum_\Gamma \lambda_\Gamma \tilde{T}_\Gamma(t,t_0) \tag{19}$$

and using (17), one obtains:

$$\frac{L_S(t,t_0)}{\tilde{L}_S(t,t_0)} = \left(\sum_\Gamma \frac{\lambda_\Gamma T_\Gamma(t,t_0)(1 - B_\Gamma + O(B^2_\Gamma))}{\sum_{\Gamma'} \lambda_{\Gamma'} T_{\Gamma'}(t,t_0)} \right)^{-1} \tag{20}$$

which is:

$$\frac{L_S(t,t_0)}{\tilde{L}_S(t,t_0)} = 1 + \sum_\Gamma (\chi_\Gamma) B_\Gamma + O(B^2_\Gamma)) = 1 + B^L_S(t,t_0) + O(B^2_\Gamma) \tag{21}$$

where

$$\chi_\Gamma = \frac{\lambda_\Gamma T_\Gamma(t,t_0)}{\sum_{\Gamma'} \lambda_{\Gamma'} T_{\Gamma'}(t,t_0)} \tag{22}$$

and

$$\sum_\Gamma \frac{\lambda_\Gamma T_\Gamma(t,t_0)}{\sum_{\Gamma'} \lambda_{\Gamma'} T_{\Gamma'}(t,t_0)} = 1 \tag{23}$$

which concludes the proof.

Lemma 8.4. *the quality factor of the chained sector index*

If the relative variation of the bilateral sector index is given by (12) then the relative variation of the chained index grows linearly with the number of chaining steps and satisfies:

$$\frac{L_S(t_N,t_0)}{\tilde{L}_S(t,t_0)} - 1 \approx \sum_{j=1}^{N} B_S^L(t_j,t_{j-1}) \approx N \cdot B_S^m \tag{24}$$

with B_S^m *the average monthly factor.*

Proof

By direct calculation, using chaining definition and the previous lemmas. A corollary of this lemma is that the average quality factor of a chained sector index is a multiple of the average monthly quality factor.

Notes: (i) the wage index based on time dummy models can be calculated directly from the coefficients of the time dummy variables for any two points in time, no matter how far from each other, if they are included while model fitting. Such an index is transitive and has a very small bias, but needs to be revised each time a new data point is measured. This also corresponds to the differences in average characteristics over the samples measured at these time points. In turn, these differences grow on average linearly with the number of time steps, giving the same result as obtained above by chaining two-point indices. (ii) the same type of relation holds when summing (weighted) sector terms.

8.2 Comparing regression coefficients across studies

The regression coefficients can be compared across studies if one takes into account the differences of scale due to time units, i.e. months in the present one versus years in the GPG studies:

- the estimate of the coefficient for experience of employee with a given company (linear term) is $3.5e-04$[13] when fitting a hedonic model to the monthly data of 2008-2018, or 0.0048 for a yearly fit, in agreement with 0.004 for both females and males in the GPG study (9). These values also agree with the earlier study where the coefficient for males was 0.004, although at that point in time the fixed effect model for females gave a smaller value for the coefficient of the experience linear term.
- the estimate of the age (linear term) coefficient is $1.6e-03$ on monthly data[14], i.e. 0.02 for a yearly fit, in agreement with 0.04 for females and 0.05 for males of the study (8). The values of (9) which only include the private sector, were slightly higher, for the years 2000-2007, i.e 0.06 and 0.07 respectively.

[13] standard error=$7.5e-06$, t-value 46.2 and p-value $< 2.2e-16$

[14] standard error=1.2e-05, t-value 124.7 and p-value $< 2.2e-16$

– the estimates of the coefficients of the second power of experience and age characteristics (negative values, of the order $e-06$ and $e-07$) are also in agreement with the previous studies.

– the coefficient of education level $ISCED = 6$ (university degree), when the reference level of the education factor is set at $ISCED < 3$ level, is estimated at 0.057 in our study when using fixed effects and characteristics and 0.02 when using only characteristics, and 0.06 (and 0.07) for males (and females) in (9). Although not directly comparable, since the education factor levels were somewhat differently defined there, the older study shows slightly smaller values[15].

8.3 Theoretical calculations: comparing IWPI and AEI

The IWPI calculations are based on Törnqvist index formula, which gives, when chaining through a whole year, a ratio of double weighted geometric averages, over time t_i and items k.

$$T_{yearly} = \prod_{t_i} T(t_i, t_{i-1}) = \prod_{t_i} \frac{\prod_k (p_k(t_i))^{\omega_k(t_i,t_{i-1})}}{\prod_k (p_k(t_{i-1}))^{\omega_k(t_i,t_{i-1})}} \tag{25}$$

such that the minimum $t_i = T$ and maximum is $t_i = T + 1$.

The AEI formula is a ratio of arithmetic averages, over items k for two successive values of the (yearly) time T .

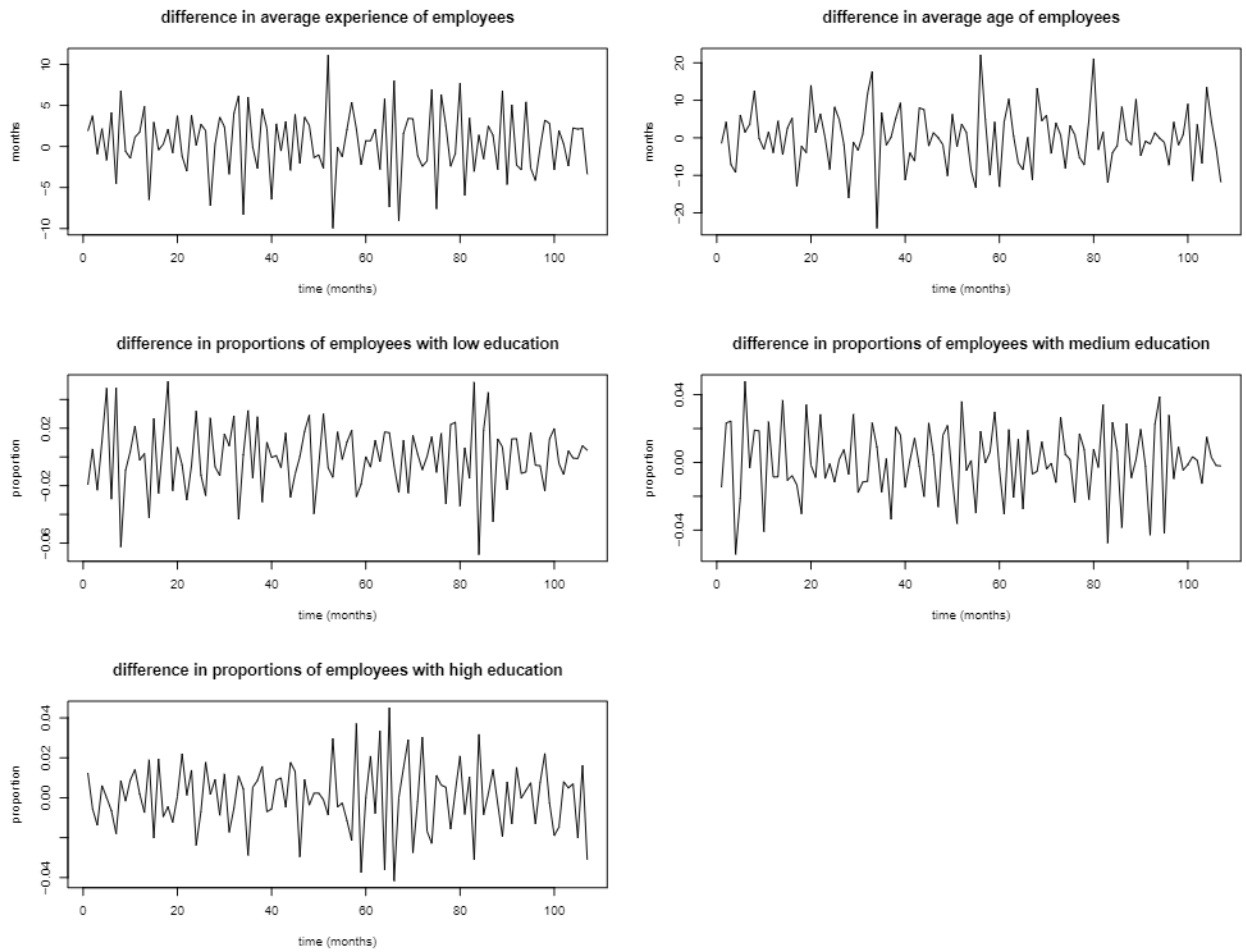
$$A_{yearly} = \frac{\sum_k p_k(T+1)\nu_k(T+1)}{\sum_k p_k(T)\nu_k(T)} \tag{26}$$

where $p_k(T)$ is the average wage over time (over a year T), for item k while $\nu_k(T)$) are the associated weights.

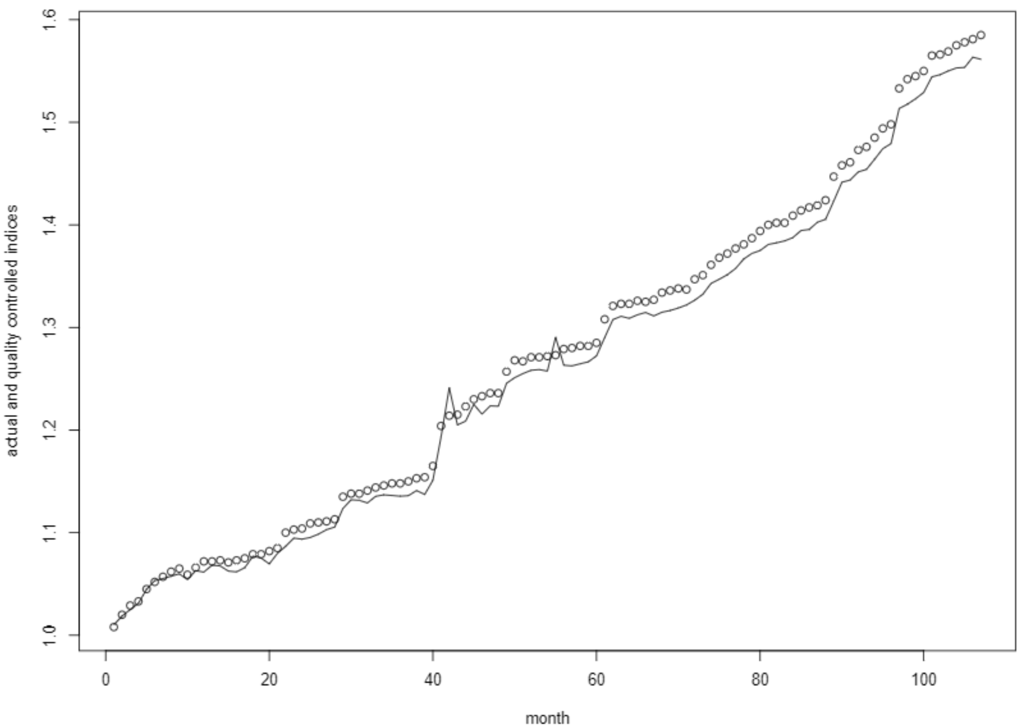
The difference between *the ratio of* two arithmetic means and *the ratio of* two geometric means does not have a fixed direction. The differences are greater for more diverse hourly wages.

8.4 Figures

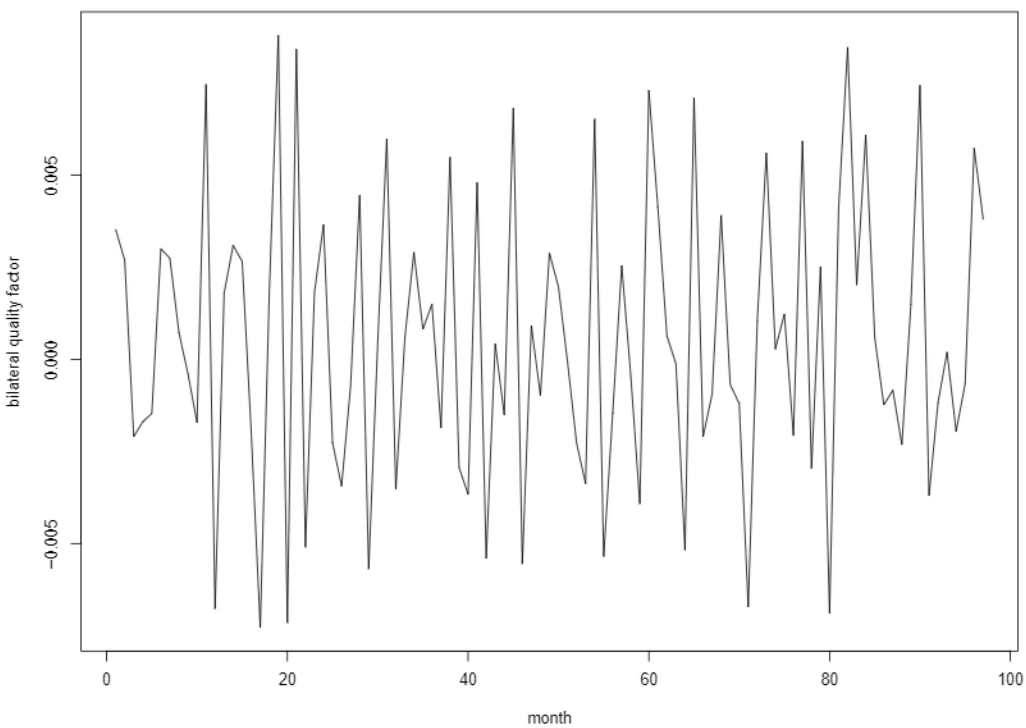
[15] 0.04 and 0.03 for males and females respectively



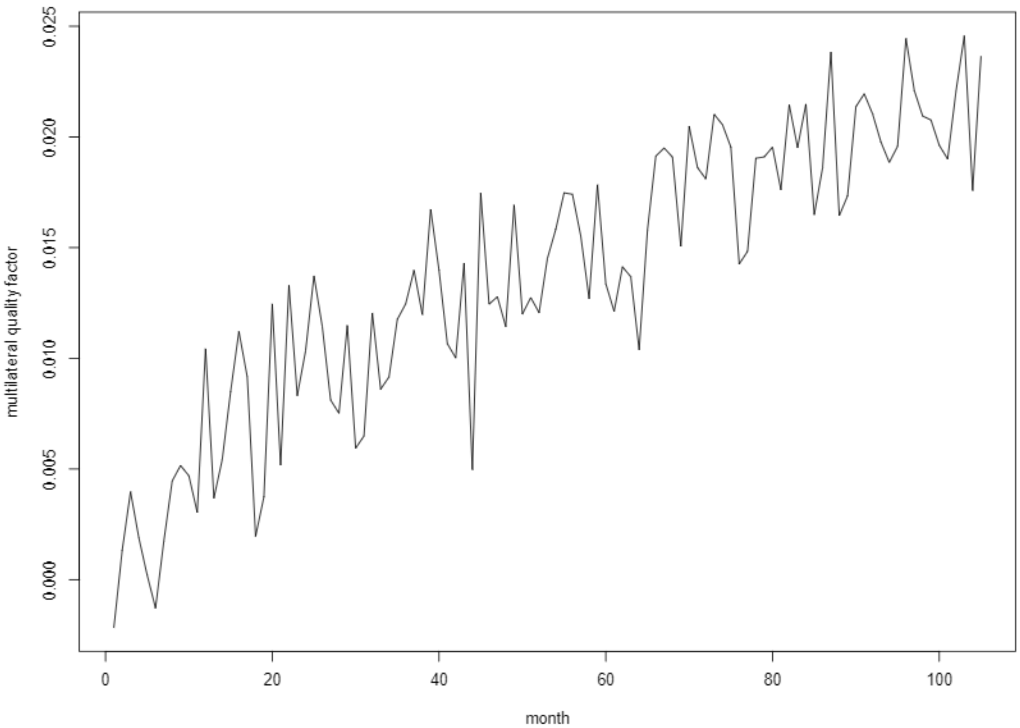
Mynd 1: Changes in average characteristics over the samples used by the IWPI, at successive points in time (appropriate statistical tests for these time series do not reject stationary hypotheses; statistical tests do not show seasonal variations either), all oscillating around values which are very 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.



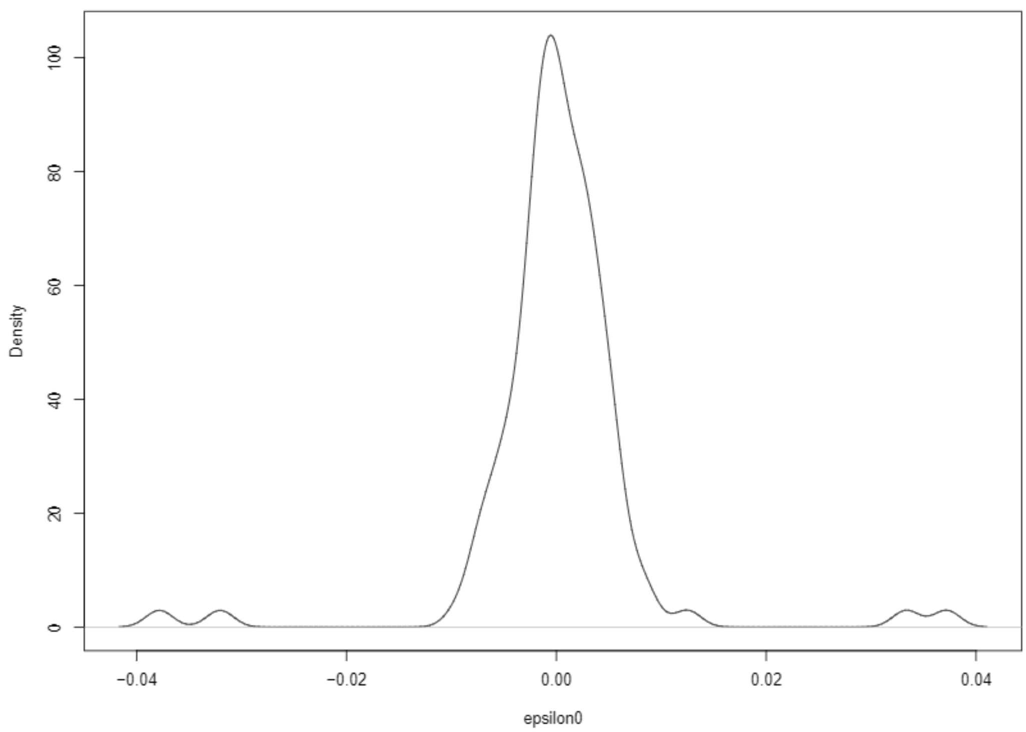
Mynd 2: The actual wage index, over ten years time, if the index reference point (index=1) is January 2008. The wage index values when quality controlled (lines).



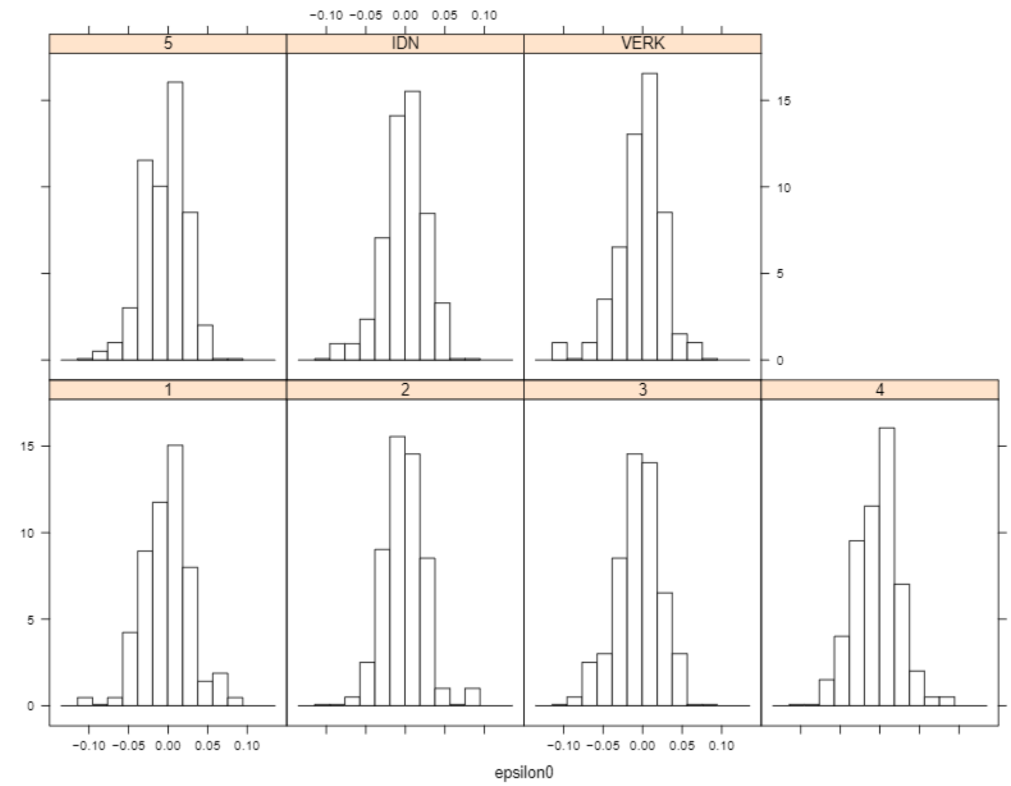
Mynd 3: The monthly quality factor, i.e. the relative difference between the monthly index and its quality controlled value.



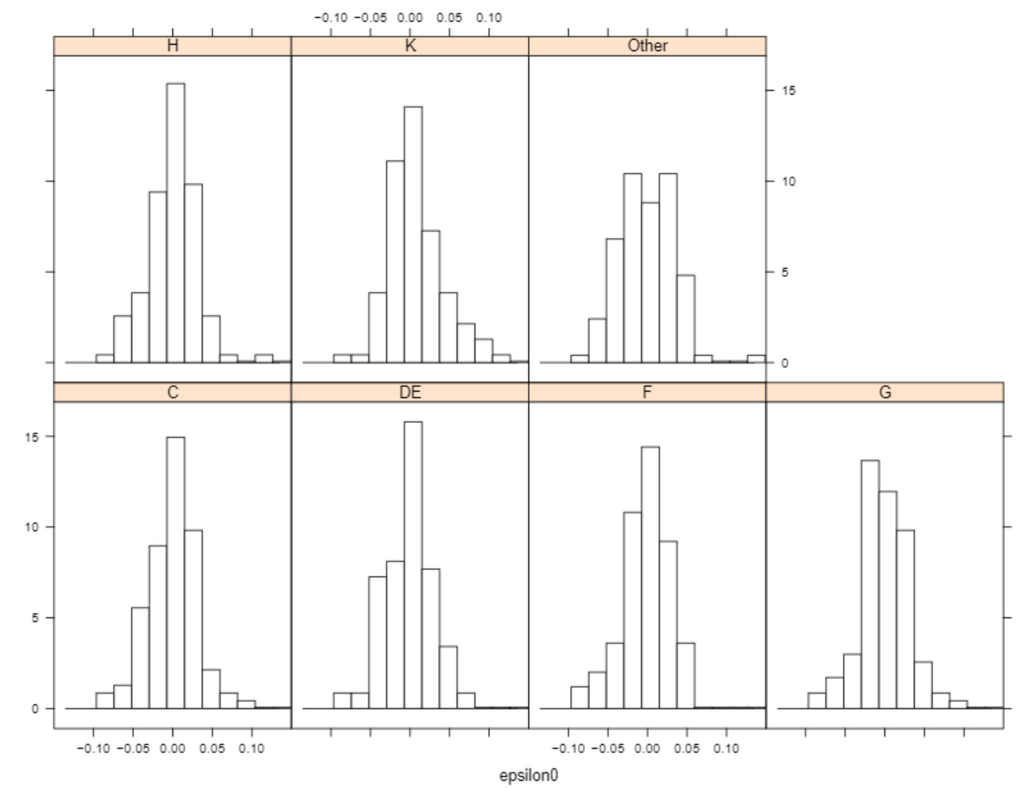
Mynd 4: The multilateral quality factor, i.e. the relative difference between the monthly index and its quality controlled value.



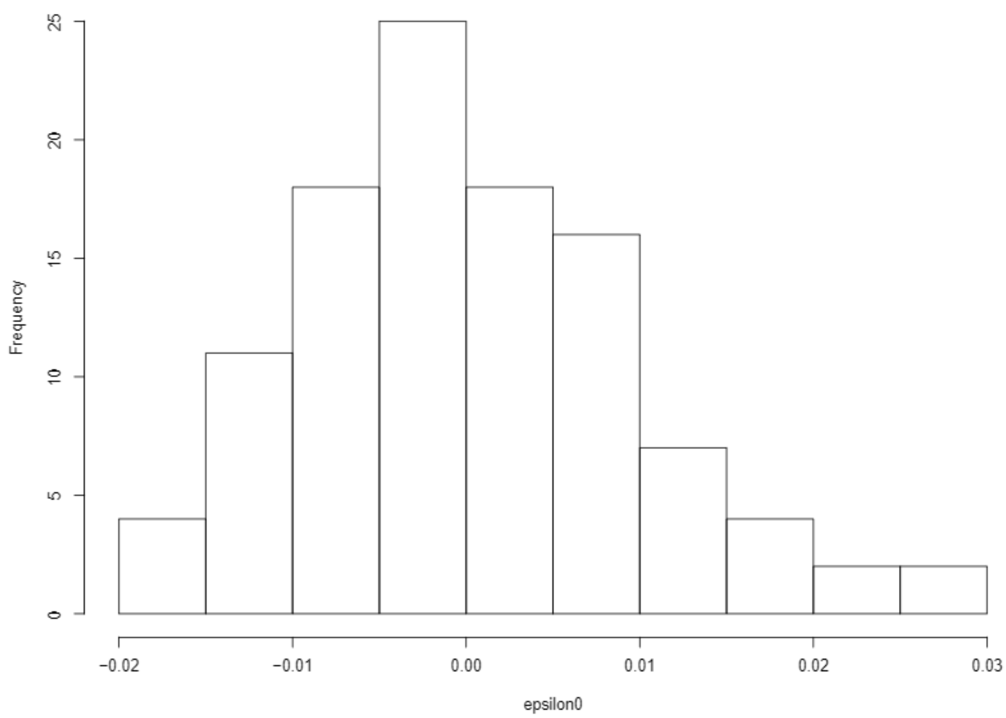
Mynd 5: Density distribution of the quality factor, i.e. the relative difference between the monthly index and its quality controlled value (epsilon0).



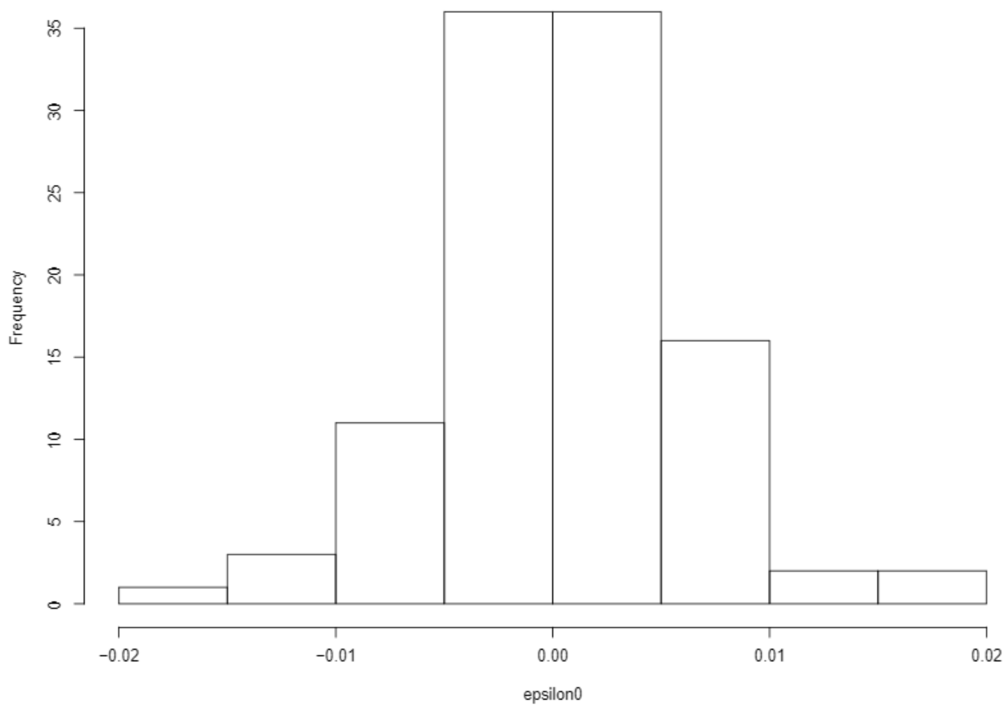
Mynd 6: Empirical density distribution of the quality factor (epsilon0) by occupation for the private sector: 1- managers, 2- professionals, 3- technicians and associate professionals, 4- clerks, 5- service workers and shop and market sales workers, IDN- craft workers, VERK- general, machine and specialized workers.



Mynd 7: Empirical density distribution of the quality factor (epsilon0) by economic activity for the private sector: C - manufacturing, DE - energy and water supply, waste treatment, F- construction, G- whole sale and retail trade, repair of motor vehicle, H- transportation and storage, K- financial and insurance activities, Other - information and communication (I), accommodation and food service (J), engineering activities (M).



Mynd 8: Empirical distribution of the monthly quality factor (epsilon0) for the public sector, local government.



Mynd 9: Empirical distribution of the monthly quality factor (epsilon0) for the public sector sector, central government.

Heimildir

[1] Manninen, K. *The Effects of the Quality Adjustment Method on Price Indices for Digital Cameras.* https://www.bea.gov/system/files/papers/WP2005-1.pdf . 2005.

[2] Diewert, E., de Haan, J. *Quality Change, Hedonic Regression and Price Index Construction.* Otawa group. 2017.

[3] Calian, V. and Experts from the Unit of wages, income and education *Methodology of the Icelandic Wage Index.* Statistical Series of Statistics Iceland, vol. 103, issue 20. http://hagstofan.s3.amazonaws.com/media/public/2018/e2c06929-b8a3-4866-8086-8f6291f943d9_SG5weZE.pdf. 2018.

[4] Hill, R. J. *Superlative index numbers: not all of them are super* Journal of Econometrics 130. 2006.

[5] Lent, J. *Chain Drift in Some Price Index Estimators.* Bureau of Transportation Statistics. http://www.asasrms.org/Proceedings/papers/2000_049.pdf. 2000.

[6] Ehemann, C. *Chain drift in leading superlative indices.* Working Paper no. 2005-09, Bureau of Economic Analysis, Washington, DC. 2005.

[7] Zieschang, K. *Review of the Icelandic WagePriceIndex* Report to the Wage Statistics Committee Government of Iceland. https://www.stjornarradid.is/lisalib/getfile.aspx?itemid=40395406-2616-11e9-942f-005056bc4d74 . 2018.

[8] Indridadottir, M. K. and Sigurdsson, E. and Gylfadottir, M. V. Skýrsla unnin vegna samstarfsverkefnis Hagstofu Íslands, Alþýðusambands Íslands og Samtaka atvinnulífsins um launamun kvenna og karla byggt á gagnasöfnum Hagstofunnar. https://hagstofa.is/media/49846/launamunur_kynjanna_lokaskyrsla_februar_2010.pdf . 2010.

[9] Gylfadottir, M. V. and Indridadottir, M. K. and Karlsson, A. *Analysis on Gender Pay Gap 2008–2016. Statistical Series, Statistics Iceland.* https://hagstofan.s3.amazonaws.com/media/public/4a70b304-09ff-4c6b-adcf-050ac2bef384/pub_doc_mqp71tr.pdf. 2018.

[10] Fixler, D. and Ziechang, K.D. *Incorporating ancillary measures of process and quality change into a superlative productivity index* The journal of Productivity Analysis, 2, 245-267. 1992.

[11] Feenstra, R.C. *Exact Hedonic Price Indices* Review of Economics and Statistics, vol. LXXVII, 934-654. 1995.

[12] Diewert, W. E., S. Heravi, and M. Silver *Hedonic Imputation versus Time Dummy Hedonic Indexes* Price Index Concepts and Measurement; Studies in Income and Wealth, 70, 87–116, University of Chicago Press, Chicago. 2009.

[13] Diewert, Erwin *Quality Adjustment and Hedonics: A Unified Approach.* Microeconomics.ca working papers , Vancouver School of Economics, revised 14 Mar 2019.

[14] de Haan, J. and F. Krsinich *Time Dummy Hedonic and Quality Adjusted Unit Value indices: Do They Really Differ?* Review of Income and Wealth, vol. 64:4. 2012.

[15] de Haan, J. and F. Krsinich *The Treatment of Unmatched Items in Rolling Year GEKS Price indices: Evidence from New Zealand Scanner Data.* paper presented at the Economic Measurement Group Workshop 2012, Australian School of Business, University of New South Wales, November 23. 2012.

[16] de Haan, J. and F. Krsinich *Scanner Data and the Treatment of Quality Change in Nonrevisable Price indices.* Journal of Business and Economic Statistics 32:3, 341-358.2014.

[17] ABS, Australian Bureau of Statistics *Average weekly earnings and wage price index - what do they measure.* 6302.0 Average Weekly earnings, Australia, May 2014.

Statistical Series *Working papers*
Vol. 104 • Issue 15 10 December 2019
ISSN 1670-4770
Supervision: Violeta Calian, chief methodologist, Unit of Administration and Cooperation
• violeta.calian@statice.is
Contributors: Experts from the Unit of Wages, income and education
• wages@statice.is
www.statice.is
Telephone (+354) 528 1000
Fax (+354) 528 1099