Price Index for Dwelling Rents in Finland

Vartia, Yrjö, University of Helsinki
Suoperä, Antti, Statistics Finland
Vuorio, Johanna, Statistics Finland
2019

Abstract

The price index for dwelling rents in Finland have been measured from data that is collected in connection with the Labour Force Survey (LFS). The construction of index series is based on the base strategy with log-Laspeyres as the index number formula. Some problems have emerged: First, the data did not allow detailed regional classification. Second, the purpose of data collection was mainly for labour markets survey, not for dwelling rents (i.e. weighting into target population). Third, the selection of proper index number formula had not been adequately analyzed.

In this study we analyze these three questions by new monthly collected data having over 400 000 observations from whole country quite representatively. The data includes dwelling rents from the Social Insurance Institution's register of housing allowances and from two major private rental housing companies. It includes government-subsidized and non-subsidized dwellings.

1 Introduction

During years 2018-2019, Statistics Finland/Housing statistics renewed the data collection of the rents of dwellings. The data includes dwelling rents from the Social Insurance Institution's register of housing allowances and from two major private rental housing companies including more than 400 000 rented dwellings per month. New data source is more comprehensive than ever. It contains more than hundredfold amount of observations compared to the old data of rents. We make a new partition for these 400 000 rented dwellings including about 400 strata. The partition is based on finance of dwelling (private or government subsided), detailed regional classification and dwelling-type (one-room, two-rooms and three or more). The estimation of weights for calibration of observations into target population according to partition is carefully studied.

Each dwelling has unique identifier (ID) such that we can make price-link for each observation. Our price-links are defined to be $0 \rightarrow t$, where 0 is a base and t an observation period. We define the base period as average quarter of the previous year and the observation period as the quarter of current year. Practically this means that we construct index series by the base strategy that is free of chain error (drift).

We analyze two sets of index number formulas. The first set is based on formulas using old or new weights and are called as *basic formula* (i.e. old weights: Laspeyres (L), Log-Laspeyres (Ll) and Harmonic-Laspeyres (Lh) and new weights: Palgrave (Pl), Log-Paasche (Lp) and Paasche (P)). The second set of index numbers include six formula: Stuvel (S), Montgomery-Vartia (MV), Törnqvist (T), Fisher (F), Walsh (W) and Sato-Vartia (SV). We call these index number formulas as *excellent* (Vartia & Suoperä, 2018).

We show, that index number formulas can be classified to formulas having upward or downward bias and index formulas, that must be classified as 'unbiased' according to the index number theory, yet for our data the quantity index equal unity. We show that most of analyzed index number formulas together with the base strategy may be selected for official production of price index for dwelling rents. Practically this means that only Log-Laspeyres (Ll), Harmonic-Laspeyres (Lh), Palgrave (Pl) and Log-Paasche (Lp) are contingently biased and should not be used. We also show that Jevons should not be used.

We formulate 10 questions, which together specify the index number problem for rented dwellings (Vartia, 1976, pp. 92-95). We group these questions into three categories: 1. Intended use of the price index, 2. preliminary specification of the relevant information and 3. technical choices. The questions for each category are presented in chapter two. The empirical analysis in this study follows these ten questions simply by answering them.

2 Notations, Basic Concepts and the Index Number Problem

In this chapter we define all that is needed for the index number calculations for dwelling rents. First, we define commonly used notations and second, the basic concepts and the index number problems with 10 questions. This is done very shortly but such, that all relevant issues for constructing index numbers are presented.

2.1 Notations

Our notation for index number calculations is the following:

Commodities:	$a_1, a_2, \ldots, a_{n_t}$ are rented dwellings in period t. The number n_t is over 400 000.
Time periods:	$t = 0, 1, 2, \ldots$ are the compared situations and are quarters.
Quantity:	q_i^t is the quantity of a_i in period t (square meters).
Unit value:	$p_i^t = v_i^t / q_i^t$ is the unit price of a_i in period t (square meter price).
Value:	$v_i^t = p_i^t q_i^t$ is the value of a_i in period t (dwelling rent).
Total value:	$V^t = \sum_i v_i^t$ is the total value in period t (total sum of dwelling rents)
Total quantity:	$Q^t = \sum_i q_i^t$ is the total quantity in period t (total sum of square meters).
Price relatives:	$p_i^{t/0} = p_i^t / p_i^0$ is the price relative of a_i from period 0 to t.
Quantity relatives:	$q_i^{t/0} = q_i^t / q_i^0$ is the quantity relative of a_i from period 0 to t.
Value relatives:	$v_i^{t/0} = v_i^t / v_i^0$ is the value relative of a_i from period 0 to t.
Value shares:	$w_{it} = v_i^t / \sum_i v_i^t$ is the value share of a_i in period t.
Index numbers:	$P^{t/0}$ and $Q^{t/0}$ are some price and quantity indices for price- and quantity-links $0 \rightarrow t$.

We assume that all prices and quantities are strictly positive (contain no zeros). This implies that all values, price, quantity and value relatives and value shares are also well-defined and strictly positive.

2.2 Basic Concepts and the Index Number Problem

We formulate 10 questions, which together forms all that is necessary for construction of price index for dwelling rents. We do it in this chapter very shortly and more detailed description and statistical analysis is presented in chapter three. These questions show all the relevant problems within index numbers and in their production. This list of 10 problems was originally presented by Vartia (1976, pp. 92-95). In this study we are analyzing dwelling rents and the questions are presented by following three categories.

Intended use of the price index

1. What is the general characterization of the set *A* of commodities? Rented dwellings in Finland.
2. What is definition of the economic agents? Household/consumer.
3. What is definition of the time periods? For official production of price index for dwelling rents time periods are quarters and for the CPI time periods are months.

Preliminary specification of the relevant information

4. What is definition of the partition of commodities? We divide target population into government-subsidized and non-subsidized category. Both categories include similar classification which is based on cartesian product of detailed regions and dwelling-type (one-room, two-rooms and three or more).
5. How the data is collected? The data is obtained from the Social Insurance Institution's register of housing allowances and from two major private rental housing companies.
6. What kind of weights are used? Weights for calibration of observations into target population according to sample theory and weights that index number formula dictate.

Technical choices

7. What are the index number formulas? Binary comparisons of periods based on suitable axioms. We category index number formulas into basic and excellent ones. We analyze several index numbers.
8. What is the strategy for construction of index numbers? Bilateral (or base) strategy that uses previous year normalized to average quarter as its base period.
9. Treatment of quality changes in bilateral compilations? Quality changes of dwellings, like renovation works, are not registered in data and are ignored (i.e. all dwellings in compilations are comparable in quality).
10. Treatment of the new or vanishing commodities or *one-sided nulls*? If binary price-links cannot be formed, one-sided nulls are ignored.

2.2.1 Index Number Formulas

In Table 1, we collect together all index number formulas that are analyzed. We present them mostly (exception *Lh* and *Pl* additive form) in multiplicative form, including Laspeyres and Paasche, which are derived from their logarithmic representations (see Vartia, 1976, p.128). Practically this means, that aggregation of price changes in logarithmic form is much simpler. We analyze here also the basic index numbers to get information about their bias. We show by the 'Fisher´s Five-Tined Fork' how much the basic index number formulas deviate from index numbers that are 'unbiased' according to the index number theory (Vartia & Suoperä, 2018).

These index number formulas are used when price relatives are aggregated into crude aggregates like 'dwelling-type' in some region according to the 'finance-type' etc. Aggregation is always done in logarithmic form. The index number formulas are presented for the time periods 0 and 1. Replacing one by time period *t*, we get our strategy: In this study we use the base strategy, where the base period is previous year normalized as average quarter and the observation period *t* is a quarter of current year. This means that at the first quarter of a current year we update our base period as previous year. This is most natural choice, because it is free of chain error (or drift).

Table 1: Analyzed index number formulas (see Vartia & Suoperä, 2017, 2018)

Basic formula		
Symbol and name of formula	$P^{1/0}$	Weights of the formula, w_i
L Laspeyres	$\prod (p_i^1/p_i^0)^{w_i^0}$	$w_i^0 = \frac{L(p_i^1 q_i^0, p_i^0 q_i^0)}{L(p^1 q^0, p^0 q^0)}$,
l log- Laspeyres	$\prod (p_i^1/p_i^0)^{w_i^0}$	$w_i^0 = \frac{v_i^0}{V^0}$
Lh Harmonic-Laspeyres	$1/\sum w_i^0\,(p_i^0/p_i^1)$	$w_i^0 = \frac{v_i^0}{V^0}$
Pl Palgrave	$\sum w_i^1\,(p_i^1/p_i^0)$	$w_i^1 = \frac{v_i^1}{V^1}$
p log-Paasche	$\prod (p_i^1/p_i^0)^{w_i^1}$	$w_i^1 = \frac{v_i^1}{V^1}$
P Paasche	$\prod (p_i^1/p_i^0)^{w_i^1}$	$w_i^1 = \frac{L(p_i^0 q_i^1, p_i^1 q_i^1)}{L(p^0 q^1, p^1 q^1)}$,

Excellent formula		
T Törnqvist	$\prod \left(p_i^1/p_i^0\right)^{\bar{w}_i}$	$\bar{w}_i = 0.5 * (w_i^0 + w_i^1)$
SV Sato-Vartia	$\prod \left(p_i^1/p_i^0\right)^{\bar{w}_i}$	$\bar{w}_i = \frac{L(w_i^1, w_i^0)}{\sum L(w_i^1, w_i^0)}$
MV Montgomery-Vartia	$\prod \left(p_i^1/p_i^0\right)^{\bar{w}_i}$	$\bar{w}_i = \frac{L(v_i^1, v_i^0)}{L(V^1, V^0)}$
W Walsh	$\prod \left(p_i^1/p_i^0\right)^{\bar{w}_i}$	$\bar{w}_i = \sqrt{w_i^0 \cdot w_i^1} \Big/ \sum \sqrt{w_i^0 \cdot w_i^1}$
S Stuvel	A+$\sqrt{A^2 + V^1/V^0}$, where A=0.5· $\left(L^{1/0} - \frac{V^1/V^0}{P^{1/0}}\right)$	$L^{1/0}$ = Laspeyres and $P^{1/0}$ = Paasche are price indices
F Fisher	$\left(L^{1/0} \cdot P^{1/0}\right)^{1/2}$	

3 Empirical Results

We analyze here two set of index numbers presented in Table 1. The data is exceptional case in index compilation because the quantity index equals unity for each quantity-links 0 → t, where 0 is a base and t an observation period. This forms a special case for basic index numbers, because the 'Fisher's five-tined fork' comes true exactly. This is clearly one from few cases, where Laspeyres (L) and Paasche (P) are 'unbiased' index numbers in the sense of index number theory (see Vartia & Suoperä, 2018). This extraordinary data reveals also that there is something wrong with Törnqvist weights, because the factor antithesis $V^{t/0}/\prod\left(q_i^t/q_i^0\right)^{\bar{w}_i}$ (i.e. FA) do not equals $\prod\left(p_i^t/p_i^0\right)^{\bar{w}_i}$.
Practically this means that the FA of Törnqvist equals precisely to 'unbiased' index number, like Fisher, Stuvel, Montgomery-Vartia,… , but Törnqvist price index does not.

3.1 Fisher's Five-Tined Fork – The Basic Index Numbers

According to Vartia (1978), Vartia & Vartia (1984) and Vartia & Suoperä (2018, pp. 7-8), Fisher (1922) visualized the basic index numbers as forks containing a certain number of tines. This is an important visualization of differences between various basic indices and other choices of calculation. In his data, the basic index numbers happened to fall exactly into five separate groups or "tines", which forms "the five-tined fork". This occurs, because very special case $Q^{t/0} = 1$ is realized. Practically this means that L and P equals, which is not generally or even usually true. For this *special data* we may say like Fisher that these formulas are "very good", but for some other data these formulas are almost surely contingently up- or downward biased with doses of bias (or quanta) = $0, \pm 1, \pm 2$ (see Vartia & Suoperä, 2018). In Figures one to four we show some examples of Fisher's five-tined fork.

Figure 1: Fisher's 'Five-Tined Fork' for the basic index number formulas for non-subsidized dwellings in Helsinki, region one for one-room', 2015=1.

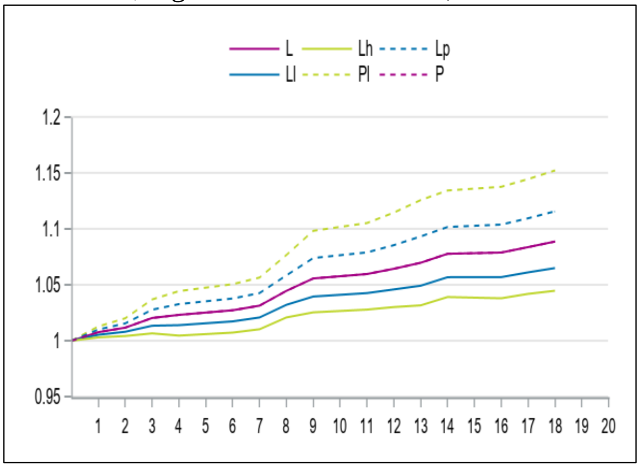
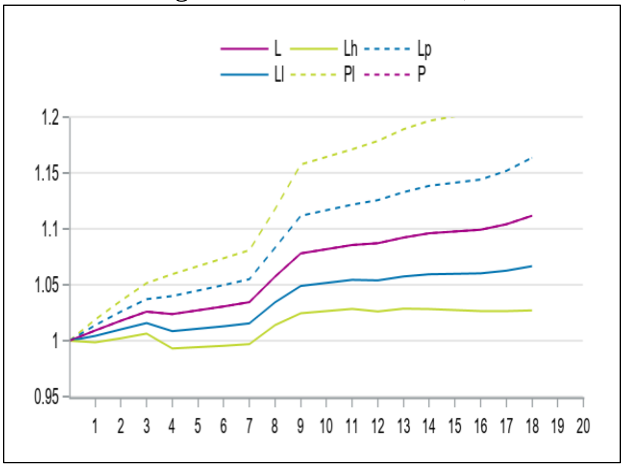


Figure 2: Fisher's 'Five-Tined Fork' for the basic index number formulas for non-subsidized dwellings in Helsinki, region one for two-rooms', 2015=1.



Figures one to four tell us that Log-Laspeyres (*Lh*), Harmonic-Laspeyres (*Lh*), Palgrave (*Pl*) and Log-Paache (*Lp*) are seriously biased and should never be used as official statistic. For our very exceptional data Laspeyres, Paasche and naturally Fisher (i.e. $L = P = F$) are unbiased having 0 doses of bias (see Vartia & Suoperä, 2018).

Figure 3: Fisher's 'Five-Tined Fork' for the basic index number formulas for government-subsidized dwellings in Helsinki, region one for one-room', 2015=1.

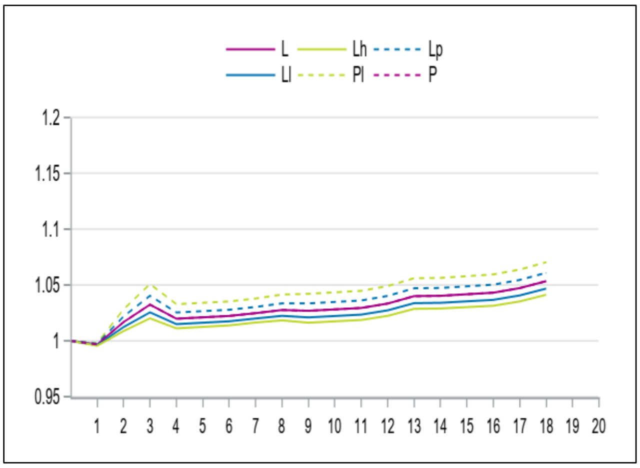
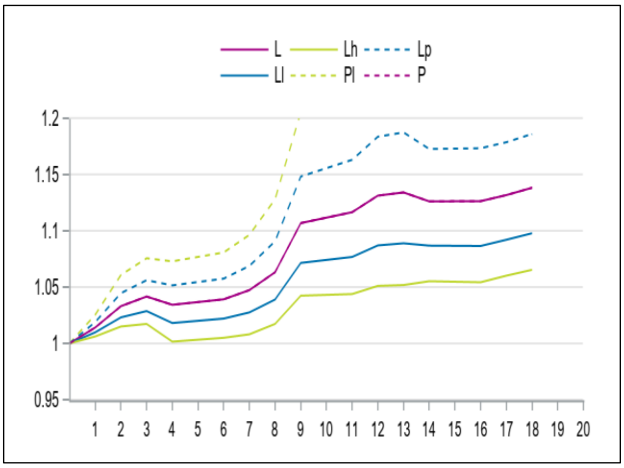


Figure 4: Fisher's 'Five-Tined Fork' for the basic index number formulas for government-subsidized dwellings in Helsinki, region one for two-rooms', 2015=1.



3.1.1 Laspeyres, Paasche and Jevons

Jevons is the unweighted (or equally weighted) geometric average of observed price ratios. We estimate Jevons for our partition and compare it with Laspeyres (L), Paasche (P) and Fisher (F). In Figures 5 and 6 we see that

Figure 5: Laspeyres, Paasche, Fisher and Jevons for non-subsidized dwellings in Espoo, region one for one-room', 2015=1.

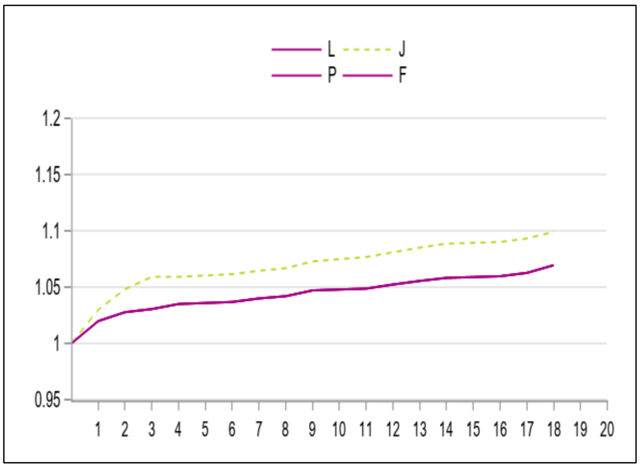
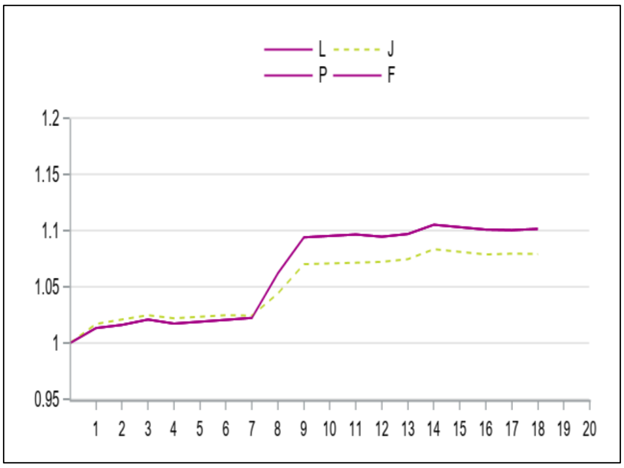


Figure 6: Laspeyres, Paasche, Fisher and Jevons for non-subsidized dwellings in Espoo, region one for two-rooms', 2015=1.



Jevons may have serious up- or downward bias compared to the 'unbiased' Fisher. So, Jevons should never be used as official statistics.

3.2 Excellent Index Numbers

The set of excellent index numbers have been analyzed carefully in Vartia & Suoperä (2018) and this study follow it. Our data of dwelling rents is quite extraordinary compared to ordinary complete micro datasets because the quantity index equal unity for each category of aggregation. This holds for most of index number formulas. Practically this means that $V^{t/0} = P^{t/0}$ for the best ones.

We show here a new result for the Törnqvist formula by deriving Törnqvist price index by factor antithesis (FA), that is, first using Törnqvist formula for quantities, and then returning back to the price space by calculating its co-factor $FA(P) = V^{t/0}/\prod\left(q_i^t/q_i^0\right)^{\overline{w}_i}$. The Törnqvist price index derived by the FA equals precisely with price indices of Fisher, Stuvel, Sato-Vartia and Montgomery -Vartia, but not with the price index of Törnqvist, that is, $FA(P) \neq \prod\left(p_i^t/p_i^0\right)^{\overline{w}_i}$. This is seen in all Figures below. Two dashed lines, Törnqvist and Walsh, deviate slightly from best index numbers (Montgomery-Vartia, Sato-Vartia, Stuvel and Fisher). Deviations are 'harmless' such that any of them may be selected for official statistic for dwelling rents.

Figure 7: Excellent index numbers for non-subsided dwellings in Helsinki, region one for one-room', 2015=1.

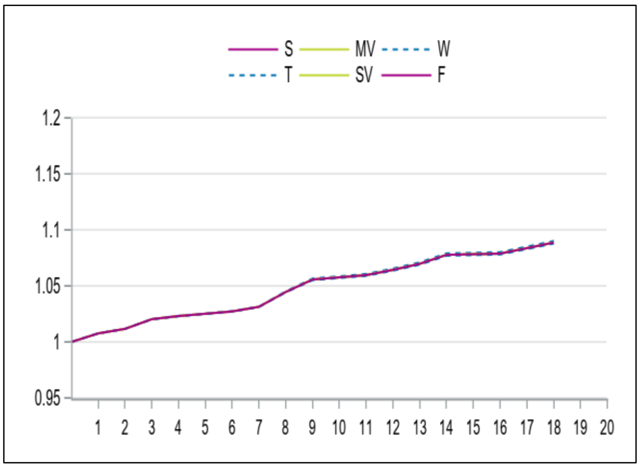


Figure 8: Excellent index numbers for non-subsidized dwellings in Helsinki, region one for two-rooms', 2015=1.

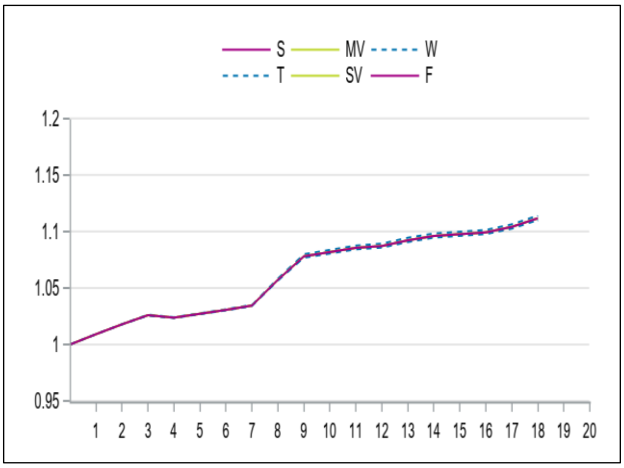


Figure 9: Excellent index numbers for government-subsidized dwellings in Helsinki, region one for one-room', 2015=1.

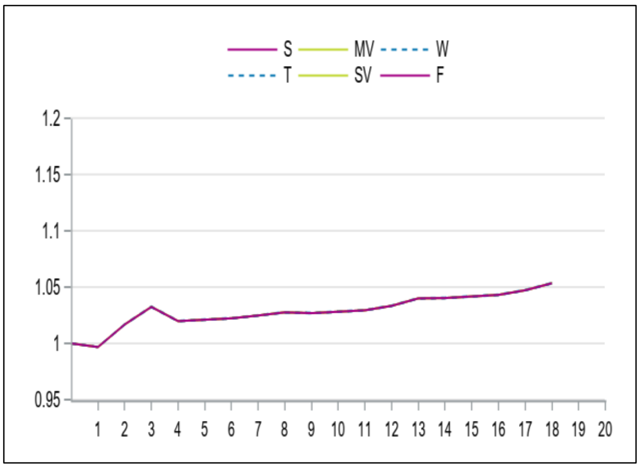
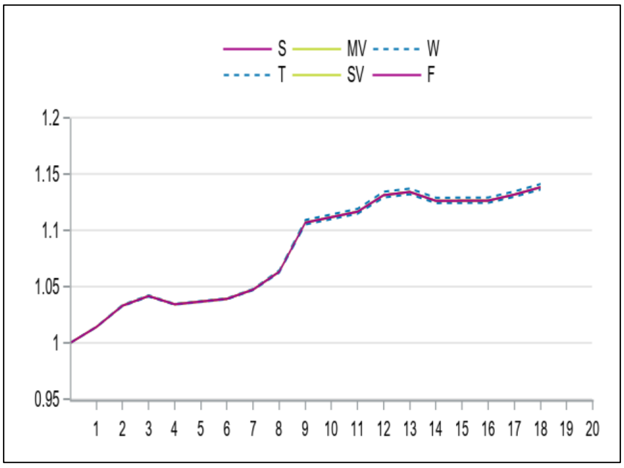


Figure 10: Excellent index numbers for government-subsidized dwellings in Helsinki, region one for two-rooms', 2015=1.



Figures 7 -10 present changes of dwelling rents according to finance-type for 'one-room' and 'two-rooms' in Helsinki region one. Our partition consisting of 400 stratas give quite similar graphs - first, government-subsided rents increases more compared to non-subsided, second, rents for 'two-rooms' increases faster compared to 'one-room' and 'three-rooms or more' that behave quit similarly.

4 Conclusion

Statistics Finland renewed dwelling rent data collection by introducing new monthly collected data having over 400 000 observations from whole country. The data was first carefully edited. Then the estimation of weights for calibration of observations into target population according to partition was carefully studied. The partition includes about 400 strata covering whole Finland. The new data makes it possible to renew official statistics for average prices and price index for rents. This study gives following results:

1. We define price-links as $0 \rightarrow t$, where the base period (i.e. 0) is a previous year normalized to average quarter and the observation period (i.e. t) is a quarter of current year. We construct our index series by base strategy.

2. We show that the basic index number formulas form the ‘Fisher’s five-tined fork’ precisely. Practically this means that Log-Laspeyres (*Ll*), Harmonic-Laspeyres (*Lh*), Palgrave (*Pl*) and Log-Paasche (*Lp*) are contingently biased and should never be used as official statistic. Mostly bias is serious.
3. In our test data quantity change equals unity for all observations and for all aggregation category. Practically this means that the price index number equals value change for the best index number formulas that are Stuvel, Montgomery-Vartia, Sato-Vartia and Fisher, also Laspeyres and Paasche. We show quite surprising issue for the Törnqvist formula – the Törnqvist formula do not decompose value change precisely into price and quantity changes (similarly as most Diewert superlative index numbers (Diewert,1976) except Fisher), but using the factor antithesis of Törnqvist as price index then also Törnqvist decompose the value change exactly into the price and quantity changes. This is probably a new result in index theory, and we may ask ‘what is wrong with Törnqvist weights’?
4. ‘Bias’ of Törnqvist and Walsh compared to the best index numbers (i.e. Stuvel, Montgomery-Vartia, Sato-Vartia and Fisher are all equal) is so harmless that any excellent formula may be selected as official production of rent statistic.

We suggest selecting Törnqvist that satisfies certain economic axioms, even though does not decompose the value change exactly into the price and quantity changes.

References

Diewert, Erwin (1976): Exact and superlative index numbers, *Journal of econometrics*, 4, 115-145.

Fisher, Irving (1922): **The making of index numbers,** Houghton Mifflin Company, Boston.

Vartia, Yrjö (1976): **Relative changes and index numbers,** Research Institute of the Finnish Economy ETLA, series A4 (dissertation in statistics).

Vartia, Yrjö (1978): Fisher´s five-tined fork and other quantum theories of index numbers, *In Theory and applications of economic indices* (Ed. W. Eichhorn, R. Henn, O. Opitz, R. W. Shephard), Physica-Verlag, Wurzburg.

Vartia, Y. & Suoperä, A. (2017)**:** "Index number theory and construction of CPI for complete micro data". (http://www.stat.fi/meta/menetelmakehitystyo/index_en.html).

Vartia, Y. & Suoperä, A. (2018)**:** "Contingently biased, permanently biased and excellent index numbers for complete micro data", 2018. (http://www.stat.fi/static/media/uploads/meta_en/menetelmakehitystyo/contingently_biased_vartia_suopera_updated.pdf)

Vartia, Yrjö and Vartia, Pentti (1984): Descriptive index number theory and the Bank of Finland currency index, *Scandinavian Journal of Economics*, 86, 3, 352-364.