

NOTE ON THE CONSTRUCTION OF THE IFLS CONSUMPTION EXPENDITURE AGGREGATES¹

Firman Witoelar
The World Bank

August 2009

1. Overview

The Indonesia Family Life Survey (IFLS) is an ongoing longitudinal household surveys that collect a vast amount of information about individuals, households, communities and facilities. The first wave of the survey (IFLS1) field in 1993 was representative of 83% population of Indonesia. Since then there have been three subsequent survey waves, IFLS2 (1997), IFLS3 (2000), and IFLS4 (2007). In 1998, a survey was conducted using a sub-sample of IFLS population.²

This note describes how the data files containing the consumption expenditure aggregates from five waves of the IFLS: IFLS1 (1993), IFLS2 (1997), IFLS2+ (1998), IFLS3 (2000), and IFLS4 (2007) were constructed. This note is written as the documentation for the construction of the following data files:

- pce93nom.dta
- pce97nom.dta
- pce97.dta
- pce98nom.dta
- pce98.dta
- pce00nom.dta
- pce00.dta
- pce07nom.dta

The files with ‘nom’ in the file names contain information about household expenditure aggregates and the various components including food expenditure categories, non-food expenditure categories and expenditure shares. All values are monthly figures and in nominal values.

The files pce97.dta, pce98.dta, pce00.dta, and pce07.dta contain only the total food expenditure, total non-food expenditures, and total household expenditure. All values are monthly figures and appear in both nominal and real terms. To obtain the real values, both temporal deflator and

¹ This work was partially funded by the World Bank Research Committee through the project “Equity and Growth” of the Research Group. The views expressed here do not necessarily reflect those of the World Bank or its member countries.

² The 1993, 1997, 2000, and 2007 datasets are available to registered users at no cost on the Rand website: www.rand.org/FLS/IFLS. The 1998 data are not in the public domain.

spatial deflator were used, using prices in December 2000 in Jakarta as the base. Since deflators were only available from 1997 onwards, there is no pce93.dta; for IFLS1 (1993), there is only a file for nominal values.

The note is organized as follows. Section 2 describes the components of household expenditure aggregates, how each variable and group of commodities were defined and constructed. Section 3 discusses how outliers and missing values were treated. Section 4 describes how the deflators were constructed. Appendix 1 discusses the comparison between data constructed in the new work with the data used in the writing of the book “Indonesian Living Standards: Before and After the Crisis” by Strauss et al (2004). Appendix 2 presents a comparison of the 1993 nominal consumption expenditure values in the data pce93nom.dta and in the data available from www.rand.org/FLS/IFLS (as part of the IFLS-1 RR Updates). Appendix 3 makes a comparison of the content and coverage of the consumption aggregate in 1993 compared to subsequent years (1997, 1998, 2000, and 2007).

2. Components of Household Consumption Aggregates

This section describes the various components of the household consumption aggregates, and the construction of the expenditure aggregate as well as expenditure share variables.

In the IFLS, household expenditures are categorized as follows:

1. Food expenditure
2. Non-food expenditures: frequently purchased goods/services
3. Non-food expenditures: less frequently purchased goods/services (including durables)
4. Education expenditures
5. Housing expenditures

Information about various transfers in and out of the households was also collected. In constructing household expenditures, transfers are treated differently and excluded from the totals.

2.1. Food Expenditures

Market-purchased and self-produced food

Data on food consumption were collected by asking households about the values of each of the 37 food items purchased within the past week (the *ks02* variable) and the values of the food items out of own production or from a gift (*ks03*) that were consumed within the past week. The values were then converted to monthly figures by multiplying them by 52 and dividing them by 12. New variables were then constructed by grouping several relevant food items, such as staple food, vegetables, dried food, etc.

For each food categories, three types of variables were created. Variables that begin with “m” are the monthly values of the food items that were purchased (from *ks02*). Variables beginning with

“i” are the monthly values food items out of own production or given to the household (from *ks03*). Variables that begin with “x” are the total (*ks02+ks03*) monthly expenditures.

Food transfer

A question is also asked about the value of the food given by the household to outside parties in the past week (*ks04b*). A variable *xfdtout* was created by converting *ks04b* to monthly values (by multiplying with a factor of 52/12).

Total food expenditure

The total monthly food consumption is created by adding all *ks02* and *k03*, or equivalently all of the food groups to form a new variable called *xfood*. A variable adding *xfood* and *xfdtout* was also created, called: *hhfood*.

2.2. Non-food Consumption: frequently purchased goods and services

Households were asked question about a number of non-food household expenses that were incurred during the past month. The expenditure categories (*ks06type*) are: electricity/water/phone, personal toiletries, household items, domestic services, recreation and entertainment, transportation, sweepstakes, *arisan*, and the values of non-food items given to other parties outside the household on a regular basis.

Total non-food

A variable containing the total of the above expenditure excluding arisan and transfer (*ks06 type F2 and G*) was constructed: *xnonfood2*. A variable containing all of the above expenditure including *arisan* and transfer was constructed: *totks06*.

Self-produced non-food

A question about the values of all of the non-food items above that were own-produced or given from outside the household was asked (*ks07a*). A variable called *inonfood* was constructed from *ks07a*.

2.3. Non-food expenditures: less frequently purchased goods/services including durables

Questions about annual expenditures (*ks08*) for various items for which purchases were relatively infrequent were asked. The categories include expenditure for clothing, furniture, medical, ceremonies, tax, and other. Questions about the same expenditure items but those that were self-produced were also asked (*ks09*).

The values were converted to monthly figures by dividing by 12. As in the case of food expenditure variables, three types of variables were created: m, i. and x, referring to those that were market-purchased, self-produced, and total, respectively.

A variable containing the total of the above expenditure transfer (*ks08 and ks09a type G*) was constructed: *xnonfood3*.

2.4. Housing

Households were asked how much money they had to pay for monthly rent (*kr04*). Households who own their own houses were asked how much money they would have to pay if they were renting their houses (*kr05*). Three variables were created: *xhouse*, *xhrent*, and *xhown*.

2.5. Education

Questions were asked about expenditure for education in the past year both for children living in the household for the following categories: tuition, uniform, and transportation (*ks10aa*, *ks11aa*, *ks12aa*) and tuition, uniform, transportation, and boarding for those living outside the households (*ks10ab*, *ks11ab*, *ks12ab*, *ks12bb*). The values were divided by 12 in order to create the monthly figures.

Education total

A variable adding all monthly schooling expenditures for children living in the household was constructed: *xeduc*. For children living outside, *xeducout*. Adding these two variables get the total monthly expenditure on education, *educall*.

2.7. Total household expenditures and per capita expenditure

Total household expenditure, *hhexp*, was constructed by adding the total food expenditure and the total non-food expenditure. The total food expenditure, *xfood*, is as described above. The total non-food expenditure, *xnonfood*, was constructed by adding non-food expenditures excluding *arisan* and transfer (*xnonfood2*), non-food expenditures excluding transfer (*xnonfood3*), education expenditures for children living in the household (*xeduc*), and non-food items that were self-produced (*inonfood*).

Note that this means that *arisan* and transfers were excluded, and so was education expenditures for children living outside the households.

$$hhexp = xfood + xnonfood$$

where

$$xnonfood = xnonfood2 + xnonfood3 + xeduc + inonfood$$

Households per capita expenditure variable, *pce*, was constructed by dividing household expenditure *hhexp* by the number of household members, *hhsiz*. The *hhsiz* variable was constructed by adding the number of household members.

2.8. Other Totals

A variable called *xtransfer* was created by adding transfers of food from the households, non-food transfers, and education expenditures on children living outside the household:

$$xtransfer = xfdtout + xtransf2 + xtransf3 + xeducout$$

A variable called *xhhx* was created, consisting of expenditures on utilities, household goods, domestic services and furniture. Variables called *xritax*, *xentn*, and *xdura* were created to stand for the expenditure for ritual and tax, entertainment, and durable goods, respectively.

2.9. Expenditure share variables

For several expenditure categories, share variables were created by dividing the monthly expenditure of the particular item by total household expenditure (*hhexp*). The categories are rice, staple, vegetable, meat and fish, alcohol and tobacco, food, oil, medical, clothing, dairy, all schooling expenditures, and housing

2.10. Outlier indicators

Variables were constructed to indicate whether a particular household contains any outlier in each of the expenditure sections. The variables are *_outlierkr04*, *_outlierks05*, *_outlierkr*, *_outlierks02*, *_outlierks03*, *_outlierks*, *_outlierks06*, *_outlierks08*, *_outlierks09*, *_outliersks0*.

**Table 1. Variable Descriptions for data sets
pce93nom.dta, pce97nom.dta, pce98nom.dta, pce00nom.dta, pce07nom.dta**

Variables	Description	Source data files	Source variables
<u>Identifier and location variables</u>			
hhidxx	Household ID, (xx=93, 97,98,00, 07)	htrack	hhidxx
origea	Original IFLS EA (<i>not available in pce98nom.dta</i>)	htrack	commid
Kecid	Kecamatan ID (<i>not available in pce98nom.dta</i>)	bk_sc	sc03
kabid	Kabupaten ID	bk_sc	sc02
provid	Province ID	bk_sc	sc01
<u>HH size</u>			
hhsz	Household size	bk_ar1	ar01a
<u>Monthly food consumption</u>			
mrice	Market-purchased (ks02): rice	b1_ks1	ks02+ks03 type A
mstaple	Market-purchased (ks02): staple (incl. rice)	b1_ks1	ks02+ks03 type A,B,C,D,E
mvege	Market-purchased (ks02): vegetable, fruit	b1_ks1	ks02+ks03 type F,G,H
mdried	Market-purchased (ks02): dried food	b1_ks1	ks02+ks03 type I, J
mmeat	Market purchased (ks03): meat	b1_ks1	ks02+ks03 type K,L,OA,OB
mfish	Market-purchased (ks02): fish	b1_ks1	ks02+ks03 type M,N
mdairy	Market-purchased (ks02): dairy	b1_ks1	ks02+ks03 type P,Q
mspices	Market-purchased (ks02): spices	b1_ks1	ks02+ks03 type R,S,T,U,V
msugar	Market-purchased (ks02): sugar	b1_ks1	ks02+ks03 type W,AA
moil	Market-purchased (ks02): oil	b1_ks1	ks02+ks03 type X,Y
mbeve	Market-purchased (ks02): beverages	b1_ks1	ks02+ks03 type Z,BA,CA,DA, EA
maltb	Market-purchased (ks02): alcohol/tobacco	b1_ks1	ks02+ks03 type FA,GA,HA
msnack	Market-purchased (ks02): snacks	b1_ks1	ks02+ks03 type IA
mfdout	Market-purchased (ks02): food out of home	b1_ks1	ks02+ks03 type IB
irice	Self-produced (ks03): rice	b1_ks1	ks02+ks03 type A
istaple	Self-produced (ks03): staple (incl. rice)	b1_ks1	ks02+ks03 type A,B,C,D,E
ivege	Self-produced (ks03): vegetable, fruit	b1_ks1	ks02+ks03 type F,G,H
idried	Self-produced (ks03): dried food	b1_ks1	ks02+ks03 type I, J
imeat	Self-produced (ks03): meat	b1_ks1	ks02+ks03 type K,L,OA,OB
ifish	Self-produced (ks03): fish	b1_ks1	ks02+ks03 type M,N
idairy	Self-produced (ks03): dairy	b1_ks1	ks02+ks03 type P,Q
ispices	Self-produced (ks03): spices	b1_ks1	ks02+ks03 type R,S,T,U,V
isugar	Self-produced (ks03): sugar	b1_ks1	ks02+ks03 type W,AA
ioil	Self-produced (ks03): oil	b1_ks1	ks02+ks03 type X,Y
ibeve	Self-produced (ks03): beverages	b1_ks1	ks02+ks03 type Z,BA,CA,DA, EA
ialtb	Self-produced (ks03): alcohol/tobacco	b1_ks1	ks02+ks03 type FA,GA,HA
isnack	Self-produced (ks03): snacks	b1_ks1	ks02+ks03 type IA
ifdout	Self-produced (ks03): food out of home	b1_ks1	ks02+ks03 type IB
xrice	Consumption, ks02+ks03: rice	b1_ks1	ks02+ks03 type A
xstaple	Consumption, ks02+ks03: staple (incl. rice)	b1_ks1	ks02+ks03 type A,B,C,D,E
xvege	Consumption, ks02+ks03: vegetable, fruit	b1_ks1	ks02+ks03 type F,G,H
xdried	Consumption, ks02+ks03: dried food	b1_ks1	ks02+ks03 type I, J

Variables	Description	Source data files	Source variables
xmeat	Consumption, ks02+ks03: meat	b1_ks1	ks02+ks03 type K,L,OA,OB
xfish	Consumption, ks02+ks03: fish	b1_ks1	ks02+ks03 type M,N
xdairy	Consumption, ks02+ks03: dairy	b1_ks1	ks02+ks03 type P,Q
xspices	Consumption, ks02+ks03: spices	b1_ks1	ks02+ks03 type R,S,T,U,V
xsugar	Consumption, ks02+ks03: sugar	b1_ks1	ks02+ks03 type W,AA
xoil	Consumption, ks02+ks03: oil	b1_ks1	ks02+ks03 type X,Y
xbeve	Consumption, ks02+ks03: beverages	b1_ks1	ks02+ks03 type Z,BA,CA,DA, EA
xaltb	Consumption, ks02+ks03: alcohol/tobacco	b1_ks1	ks02+ks03 type FA,GA,HA
xsnack	Consumption, ks02+ks03: snacks	b1_ks1	ks02+ks03 type IA
xfdtout	Consumption, ks02+ks03: food out of home	b1_ks1	ks02+ks03 type IB
xfdtout	Food transfer, ks04b	b1_ks0	ks04b
xfood	Food consumption, ks02 and ks03	b1_ks1	ks02+ks03 A to IB
hhfood	Food consumption and transfer (ks02,ks03, ks04b)	constructed	xfood + xfdtout
<u>Non-food: frequently purchased goods/services</u>			
xutility	Electricity/water/phone ks06A	b1_ks2	ks06 type A
xpersonal	Personal goods ks06B	b1_ks2	ks06 type B
xhhgood	Household goods ks06C	b1_ks2	ks06 type C
xdomest	Domestic goods ks06C1	b1_ks2	ks06 type C1
xrecreat	Recreation ks06D	b1_ks2	ks06 type D
xtransp	Transport. ks06E	b1_ks2	ks06 type E
xlottery	Lottery ks06F1	b1_ks2	ks06 type F1
xarisan	Arisan ks06F2	b1_ks2	ks06 type F2
xtransf2	Transfer ks06G	b1_ks2	ks06 type G
xnonfood2	Non-food, transfer & arisan excluded.	b1_ks2	ks06 type A,B,C,C1,D,E,F1
totks06	Non-food, transfer & arisan included	b1_ks2	ks06 type A,B,C,C1,D,E,F1,F2,G
inonfood	Monthly non-food own-produce (ks07a)	b1_ks0	ks07a
<u>Non-food: less frequently purchased goods/services including durable goods</u>			
xcloth	Clothing,ks08A+ks09aA	b1_ks3	ks08+ks09 type A
xfurn	Furniture,ks08B+ks09aB	b1_ks3	ks08+ks09 type B
xmedical	Medical ,ks08C+ks09aC	b1_ks3	ks08+ks09 type C
xcerem	Ceremony, ,ks08D+ks09aD	b1_ks3	ks08+ks09 type D
xtax	Tax, ks08E	b1_ks3	ks08 type E
xother	Other,ks08F+ks09aF	b1_ks3	ks08+ks09 type F
xtransf3	Transfer,ks08G	b1_ks3	ks08 type G
xnonfood3	Non-food , all ks08,ks09a excl. ks08G	b1_ks3	ks08+ks09 type A,B,C,D,E,F
<u>Non-food: Housing</u>			
xhrent	Housing: rent (kr04)	b2_kr1	kr04
xhown	Housing: own (kr05)	b2_kr1	kr05
xhouse	Housing: rent or own (kr04/kr05)	b2_kr1	kr04 or kr05

Variables	Description	Source data files	Source variables
Non-food: Education			
<u>Children living in the household</u>			
xedutuit	Tuition, ks10aa	b1_ks0	ks10aa
xeduunif	Uniform, ks11aa	b1_ks0	ks11aa
xedutran	Transport, ks12aa	b1_ks0	ks12aa
<u>Children living outside the household</u>			
xedutuitout	Tuition, ks10ab	b1_ks0	ks10ab
xeduunifout	Uniform, ks11ab	b1_ks0	ks11ab
xedutranout	Transport, ks12ab	b1_ks0	ks12ab
xedubordout	Boarding, ks12bb	b1_ks0	ks12bb
<u>Totals</u>			
xeduc	Education, children in hh, ks10aa-ks12aa	constructed	xedutuit+xeduunif+xedutran
xeducout	Education, children outside hh, ks10ab-ks12bb	constructed	xedutuitout+xeduunifout+ xedutranout
xeducall	Education, all (ks10aa-ks12bb)	constructed	xeduc+xeducout
<u>Total Non-food</u>			
xnonfood	All non-food expenditure excluding transfers	constructed	xnonfood2 + xnonfood3 + xhouse + xeduc + inonfood
<u>Household expenditure and Per capita Expenditure</u>			
hhexp	Monthly household expenditure	constructed	xfood+xnonfood
pce	Monthly per capita expenditure	constructed	hhexp/hhsize
<u>Other variables</u>			
xtransfer	Monthly transfer	constructed	xfdout+transf2+transf3 + educout
xritax	Monthly expend. on ritual and tax	constructed	xcerem + xtax
xentn	Monthly expend. on entertainment	constructed	xcerem + xrecreat + xlottery
xdura	Monthly expend. on durables	constructed	xfurn + xother + transf3
<u>Expenditure share variables</u>			
wrice	Expend. share: staple food	constructed	xrice/hhexp
wstaple	Expend. share: staple food	constructed	xstaple/hhexp
wvege	Expend. share: vegetable	constructed	xvege/hhexp
wfood	Expend. share: food	constructed	xfood/hhexp
woil	Expend. share: oil	constructed	xoil/hhexp
wmedical	Expend. share: Medical cost	constructed	xmedical/hhexp
wcloth	Expend. share: clothing	constructed	xcloth/hhexp
wdairy	Expend. share: dairy	constructed	xdairy/hhexp
weducall	Expend. share: Education (all)	constructed	xeducall/hhexp
whous	Expend. share: housing	constructed	xhouse/hhexp
wmtfs	Expend. share: Meat+fish	constructed	(xmeat+xfish)/hhexp
<u>Outlier indicators</u>			
_outlierks02	=1 if hh has any outlier in ks02	constructed	
_outlierks03	=1 if hh has any outlier in ks03	constructed	
_outlierks	=1 if hh has any outlier in ks	constructed	
_outlierkr04	=1 if hh has any outlier in kr04	constructed	
_outlierkr05	=1 if hh has any outlier in kr05	constructed	

Variables	Description	Source data files	Source variables
_outlierkr _outlierks06 _outlierks08 _outlierks09 _outlierks0	=1 if hh has any outlier in kr =1 if hh has any outlier in ks06 =1 if hh has any outlier in ks08 =1 if hh has any outlier in ks09 =1 if hh has any outlier in ks04a, ks07b, ks10aa, ks11aa, ks11ab, ks10ab, ks11ab, ks12ab, ks12bb	constructed constructed constructed constructed constructed	

3. Outliers and Missing Values

3.1 Outliers

An ad hoc approach was taken to identify gross outliers in the data set. Each variable was tabulated and observations at the higher end of distribution were labeled as outliers when the difference between the values of those observations and the next lower one was 'large'. This turns out to be a very conservative approach. Typically, an observation that was identified as an outlier would be more than three times the standard deviation more than the mean values.

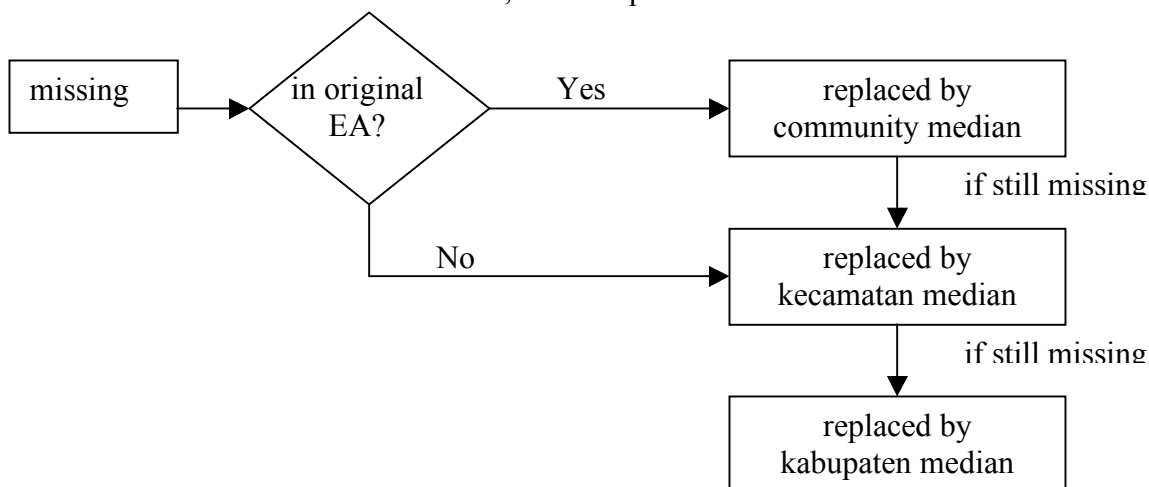
These outliers were replaced by missing values and indicator variables were created to identify which households contain outliers in each of the consumption components.

No imputations were done to replace the outliers, therefore all households containing any number of outliers will also have missing values for their expenditure aggregate variables.

3.2. Missing values and Imputation

When values were missing in the original data sets due to "Don't Know" or "Missing", they were replaced, when possible, by imputed values.

The values used for imputation are the median values of the variable taken at the community level, *kecamatan* level, or *kabupaten* level. Missing values for households living in an original IFLS EAs would be replaced by the community median values provide the median values existed. Otherwise, the kecamatan medians were used. If kecamatan medians were not available, the kabupaten medians were used.



The method is very similar to that employed in constructing the pce variables used in the book “Indonesian Living Standards”.

The difference is that in constructing the data sets for the book, to determine at what level the imputation was done, instead of following the rule above, an ad hoc judgment were made by looking at the number of each cells (consumption variable x community, kecamatan, or kabupaten). In practice, the data constructed using this ad hoc approach and the ones using the community for median for IFLS original EAs and kecamatan or kabupaten for the movers are similar.

However, in order to make the newly constructed data sets for 1997 and 2000 consistent with the ones used for the book, the code lines for the imputation from the do-files constructing the data for the book were used here.

As is discussed in the next section, because the versions of the source data are different, there are some differences between the new data sets with the book version.

Implementation

In the do files to create pce97nom.dta and pce00.dta, the approach to mimic what was done for the book was implemented by adding the code lines between the heading:

```
***** START OF THE 'MANUAL' IMPUTATION OF <VARIABLE NAME>*****;
```

<ad hoc imputation>

```
***** END OF THE 'MANUAL' IMPUTATION OF <VARIABLE NAME>*****;
```

Users who wish not to use this approach may delete all the lines between the headings.

Imputation in the 1998 data set

For the 1998 data set, the approach that was taken was to use community median for those non-movers, and use kecamatan, kabupaten median if they are missing, and use kecamatan or kabupaten median for the movers.

4. Deflators: data files deflate_hh97.dta, deflate_hh98.dta, deflate_hh00.dta

The data files pce97.dta, pce98.dta, and pce00.dta contain the nominal and real values of household food expenditures, non-food expenditures, and total expenditures. These data files also contain the per capita expenditure, nominal and real, derived by dividing household expenditure by the size of the households.

Two sets of deflators were used. The first set of deflators is the temporal deflators (*cpi_tornq*), using December 2000 as the base. The second set of deflators is the spatial deflators (*cpi_sp00*) using Jakarta as the base.

To construct the real household expenditure (*rtotal*) the following formula was used:

$$rpce = (hhexp \times cpi_tornq) / cpi_sp00$$

so the household expenditure was first deflated by dividing it with the temporal deflator and it was then deflated again by dividing it with the spatial deflator.

4.1. Background

In order to convert nominal Rupiah values into real values based on Jakarta December 2000 price levels, you must both temporally and spatially adjust values using the temporal and spatial indexes assigned to each household:

(1) You must inflate locale Rupiah values to December 2000 prices levels by multiplying the nominal value with the Tornquist inflator. Since Indonesia has experience price inflation from 1997 to 2000, the real value in local prices will be higher than the nominal value after applying the Tornquist inflator.

(2) You must then convert this into Jakarta-area prices by dividing with the spatial deflator. The real value will be higher (lower) than the nominal value in areas/locations with prices which are lower (higher) than Jakarta after applying the spatial deflator.

However, if computing poverty rates using province/urban/rural specific poverty lines, then per capita consumption should only be adjusted using the Tornquist deflator.

4.2. Temporal deflators (cpi_tornq)³

Tornquist indices were created for each household based on the month and year of the household interview and the location (province, urban/rural status) of the household). The Tornquist indicator adjusts for price changes within the household's location from the date of interview to December 2000, creating real local Rupiah amounts (base to December 2000). Urban households are assigned the Tornquist deflator based on price data for the nearest city from the BPS list of 43 (only 34 of those are actually needed to match to the IFLS sample) and rural households are assigned the Tornquist deflator based on price data for their province of residence.⁴

The Tornquist indices were constructed separately for urban and rural prices, using consumption shares from the 1996 and 1999 SUSENAS consumption modules as weights for the price increases from the consumer price index (cpi) data.^{5 6} By considering consumption shares from both years, the Tornquist index allows for the fact that households will substitute away from expensive items, such as rice, towards cheaper ones as relative prices change. This substitution will mitigate the welfare impact of price changes that should in principle be accounted for in a cost of living index. Other indices such as Laspeyres do not account for such substitution.

Using SUSENAS share weights has an advantage over BPS procedures, at least for their urban price indices, because in calculating mean urban shares, BPS weights household shares using weights formed from total household expenditure and are not adjusted for

³ Section 4.2 is taken from Strauss, John, Kathleen Beegle, Agus Dwiyanto, Yulia Herawati, Daan Pattinasarany, Elan Satriawan, Bondan Sikoki, Sukamdi, Firman Witoelar. 2004. *Indonesian Living Standards: Before and After the Financial Crisis*. Rand Corporation, USA and Institute of Southeast Asian Studies. See, in particular [Appendix 3A. Calculation of Deflators and Poverty Lines](#).

⁴ Price series for the urban are from the *Urban CPI* series from the *Cost of Living Surveys*, that collect prices from 40+ major cities (i.e. provincial capital + other major cities). The respondents for this survey are urban households. Price series for the rural are from the "*Survey Harga Konsumen Pedesaan*" (Rural Consumer Price Survey). The respondents of this survey are farmers/farm workers. This series is a province-level series.

⁵ The Tornquist formula applied to our case is:

$$\log cpi_T = \sum_i 0.5(w_{i,1999} + w_{i,1996}) * \log(p_{i,1} / p_{i,0})$$

where $w_{i,1999}$ is the budget share of commodity i in 1999, taken from SUSENAS; $w_{i,1996}$ is the budget share in the base period, 1996; $p_{i,1}$ and $p_{i,0}$ are the prices of commodity i in periods 1 and 0 (in our case period 1 will correspond to Dec 2000 and period 0 to the month and year of interview of the household).

⁶ This required that we match a list of commodities from the urban price indices to separate lists from both the 1996 and 1999 SUSENAS (the two SUSENAS' have different commodity code numbers) and conduct an analogous procedure for the rural price indices. Correspondences worked out by Kai Kaiser, Tubagus Choesni and Jack Molyneux (Kaiser et al. 2001) proved very valuable in helping us do this, although we re-did the exercise and made a number of changes. Other studies have used the quantities in the SUSENAS' to form unit prices (see, for example, Deaton and Tarozzi, 2000). For us this is not appropriate since we need prices deflators for months and years not covered by the SUSENAS. IFLS is not a very good source for prices for the purpose of constructing cpi's. Unit prices are not available in the household expenditure module because quantities are not collected. While some price information is collected in the household questionnaire and separately in the community questionnaire, from local markets; there are only a limited number of commodities available, and so we do not use them

household size.⁷ This results in rich households getting a very high weight compared to poor households, which would not be the case if household size was used instead (Deaton and Grosh, 2000, note that this is a common problem in many countries). The particular problem this causes in Indonesia over this time period is that the food share BPS uses is very low, 38% on average over all urban areas, compared to a share of 55% found in the 1996 SUSENAS module (both shares being for the same year) or 53% in IFLS. In addition, food price inflation was higher over the period 1997-2000 than non-food inflation, so that a lower food share will understate inflation, and thus overstate real income growth over this period. Obviously this will overstate any recovery in pce levels.⁸

4.2. Spatial deflators (cpi_sp00)

In addition to needing to deflating nominal Rupiah amounts temporally, it is also necessary to adjust for spatial price differences. The spatial deflator variable is the ratio of the location (province, urban/area) poverty line (in December 2000 prices) to the Jakarta poverty line. Thus, it converts the local December 2000 values into Jakarta December 2000 values.

⁷ For rural shares it is not clear whether expenditure-based or population-based weights were used by BPS.

⁸ On the other hand, the BPS consumer expenditure survey collects expenditures for a far more disaggregated commodity list than does the SUSENAS module (which is the longer form of the two SUSENAS consumption surveys). It is especially more detailed on the non-food side. Having less detail on non-foods is thought to lead to serious underestimates of non-food consumption and thus an overstatement of food shares (Deaton and Grosh, 2000).

Appendix 1

IFLS2 and IFLS3 Consumption Expenditure Aggregates:

Comparison of "Indonesian Living Standards Before and After the Crisis" version and the current version

This note shows the comparison of consumption expenditure aggregates constructed for the book "Indonesian Living Standards Before and After the Crisis" with the aggregates constructed in this project.⁹ There are differences in the two versions because the data sources that were used were different.

For the 2000 data sets, the book version used the February 8 2002 data set (a beta, pre-public release version). The new data set is constructed using the public release data sets (dated April 7, 2004). The differences between the data sets include differences in location of the households (which imply differences in the median values used for imputation), and differences in the values of the variables.

The book version of the 1997 data set use the data files dated August 20, 2001, while the new data set is constructed using the public release data files (dated May 28, 2004). The differences between the two data sets are few, but they do exist.

2000 consumption expenditure aggregates comparison

The comparison between the two resulting data sets shows that the new version of the consumption expenditure aggregate data file (pce2000.dta) has 10,222 observations while the book version has 10,220 observations. Out of those households about 10,210 are the same households. Among the same households, 10,115 households (99.07) percent have exactly the same total household expenditures in the two data files. Around 0.93 percent (95 households) have different household expenditures.

Among all the 10,210 households, the mean difference of the hh expenditure is around Rp 158, with a standard deviation of around Rp 13,000. Among the 95 households whose expenditures are different, the mean difference is around Rp 17,000 (standard deviation around Rp 137,000). The largest difference is around Rp 800,000.

⁹ Strauss, John, Kathleen Beegle, Agus Dwiyanto, Yulia Herawati, Daan Pattinasarany, Elan Satriawan, Bondan Sikoki, Sukamdi, Firman Witoelar. 2004. Indonesian Living Standards: Before and After the Financial Crisis. Rand Corporation, USA and Institute of Southeast Asian Studies.

1997 consumption expenditure aggregates comparison

The new version of the 1997 consumption expenditure aggregate data file (pce1997.dta) has 7,536 observations while the book version has 7,518 observations. Out of those households about 7,517 are the same households. Among the same households, 7,510 households have exactly the same total household expenditures in the two data files. 7 households have different household expenditures.

Among all the 7,510 households, the mean difference of the household expenditure is around Rp 0.13, with a standard deviation of around Rp 48. Among the 7 households whose expenditures are different, the mean difference is around Rp 142 (standard deviation around Rp 1,689). The largest difference is around Rp 3,250.

Appendix Table 1. Summary statistics of the book version and the new version:

	2000		1997	
	Book version	New version	Book version	New version
Total number of households	10,220	10,222	7,518	7,536
<i>Among households appearing in both files:</i>				
Value at 5% of distribution	206,433.3	206,316.7	95,750	95,750
Value at 25% of distribution	420,166.7	419,983.3	204,216.7	204,216.7
Value at 50% of distribution (median)	671,675	671,731.3	338,983.3	338,983.3
Value at 75% of distribution	1,116,650	1,115,600	590,758.4	590,758.4
Value at 95% of distribution	2,770,767	2,770,767	1,615,083	1,615,083
Mean	992,469	992,311.3	607,608.5	607,608.4
Number of observations	10210		7,510	
Number of cases with different values	95 (0.93%)		7 (0.09%)	
Median difference	0		0	
Mean difference	157.695		.132	

Appendix 2

IFLS1 Consumption Expenditure Aggregate:

Comparison of previous version (expend2.dta) and new version (pce93nom.dta)

Data set	Variable	Obs	Mean	Variable label
pce93nom.dta				
	xfoodtot	7149	157,868.9	Monthly food expend: ks02+ks03+ks04
	xnonfoodtot	7191	140,226.3	Monthly non-food expend: ks06+ks08+educ
	xhhexp	7136	297,892.1	Monthly hh expenditure
	pce	7136	71,924.11	Monthly HH per capita expenditure
expend2 (available with the IFLS-1 RR data at www.rand/FLS/IFLS)				
	foodtot	7206	158,360.0	monthy food item+transfer
	nfdamt	7206	130,355.1	monthy nonfood expenses
	expend	7205	288,742.6	nominal monthy hhld expenditures
	expendpc	7204	68,007.02	nominal monthy per capita expenditures

Appendix 3

IFLS Consumption Expenditure Aggregate:

Comparison of IFLS1 and IFLS2/2+/3/4

The IFLS questionnaire is not identical across waves, and this is also true for the questionnaire portions used to construct the consumption aggregate. Specifically, the questionnaire was changed slightly from the first round (IFLS1 in 1993) to the second round (IFLS2 in 1997 and onwards), whereas the questions are identical for IFLS2, IFLS2+, IFLS3, and IFLS4. These differences include:

- Wording of questions KS05, KS07, and KS09 in regards to whether the expenditure amount is for the entire household or not.
- Questions ks02 and ks03 item O in 1993 was separated into two categories (OA and OB in 1997, 1998, 2000, and 2007)
- Questions ks02 and ks03 item IA in 1993 was separated into two categories (IA and IB in 1997, 1998, 2000, and 2007)
- Question ks04a in 1993 does not exist in 1997, 1998, 2000, 2007)
- Question ks06 item C in 1993 was separated into two categories (C and C1 in 1997, 1998, 2000, and 2007)
- Question ks06 item F in 1993 was separated into two categories (F1 and F2 in 1997, 1998, 2000, and 2007)
- Question ks12a did not exist in 1993 and was a new category added in 1997, 1998, 2000, and 2007