Data Guide

Data Structures, Variable Naming and Derived Variables

International Tobacco Control Policy Evaluation Project
University of Waterloo, Ontario, Canada

ITC Data Management Core
January 23, 2019

Table of Contents

List of Tables

1. Introduction

In order that all researchers participating in the International Tobacco Control Policy Evaluation Project can report research findings in a consistent manner, derived variables and administrative variables are included in all ITC data releases. *Derived variables* are computed from survey data, and include variables such as Heaviness of Smoking Index (HSI). *Administrative variables* are neither survey data nor derived, but provide vital information to data users. For example, the variable *cohort* indicates the wave at which a respondent was recruited. There is variation between country projects in which derived and administrative variable are included, though a key set of variables remain constant across projects and waves.

The first section of this document provides the reader with a general overview of how ITC data is structured and named. The second section of this document describes the administrative variables in the core datasets. Similarly, the third section of this document describes the derived and administrative variables in the wave datasets. The fourth section discusses strata and clusters and other geographic variables, as well as survey weight variables. The final section is an index of all derived and administrative variables.

1.1 Common Data Structure

All ITC projects follow a common data structure. The data for every project are organized into several datasets, a *core* dataset and *wave* datasets, and for some projects, sample subsets. The core datasets contain basic information for every respondent ever recruited into the project. This dataset contains few variables but many records (one for each respondent ever recruited for that project). Variables stored in the core dataset are generally do not change from wave to wave, for example, wave of recruitment, age at recruitment, sex, and so on.

Wave datasets contain survey questions, derived variables, and administrative variables for every individual project wave. These datasets contain hundreds of variables and records for only those who participated in that particular survey wave. While all country projects sample adult smokers, some projects also sample other groups such as adult non-smokers or youth. In this case, there are core and wave datasets for each sub-group. For example, there would be a core dataset for smokers, and a core dataset for non-smokers, and a smoker wave 1 dataset, a non-smoker wave 1 dataset, and so on.

If you are a SAS user, you will also notice an additional *formats* dataset. This dataset simply contains variable and response option label formatting information. These formats are unique to each country project and must be applied to the datasets prior to analysis. Stata and SPSS formats are already embedded in the core and wave datasets and therefore this extra step is not necessary for users of these statistical software packages.

The basic structure of ITC data is illustrated in Table 1.

Table 1. Example of ITC common data structure: Four Country and Southeast Asia.

Project	Dataset	Number of Records	Number of Variables
Four Country	Core	29,821	57
	Wave 1	10,290	509
	Wave 2	8,818	754
	Wave 3	9,030	815
	Wave 4	8,701	774
	Wave 5	9,320	736
	Wave 6	9,104	727
	Wave 7	7,393	772
	Wave 7.5	678	529
	Wave 8	6,153	820
	Wave 8.5	1,507	515
	Wave 9	9,706	1,577
	Wave 10	2,946	684
Southeast Asia	Smokers Core	7,694	27
	Smokers Wave 1	4,006	410
	Smokers Wave 2	3,721	356
	Smokers Wave 3	4,443	364
	Smokers Wave 4	4,328	400
	Smokers Wave 5	4,182	431
	Smokers Wave 6	4,162	482
	Non-smokers Core	2,663	21
	Non-smokers Wave 1	1,560	210
	Non-Smokers Wave 2	1,581	93
	Non-Smokers Wave 3	1,492	106
	Youth Core	3,561	26
	Youth Wave 1	2,011	211
	Youth Wave 2	1,774	157
	Youth Wave 3	1,858	178
	Youth Wave 4	1,837	164
	Youth Wave 5	1,889	194
	Youth Wave 6	1,774	206

Note that this is not a comprehensive list of all projects nor all waves currently available.

Table 2 summarizes all datasets for all ITC country projects that are currently available for use by internal researchers.

Most researchers will require data from multiple datasets to conduct their analysis. All datasets share a common identifier variable called *uniqid*. This variable identifies each individual respondent with a unique number, allowing researchers to link the core and wave datasets by respondent within country projects. *Uniqid* is not linkable across projects.

Table 2. Summary of ITC projects and currently available data.

Project Code	Project Countries	Sample Composition	Available Data
4C / ITC4	Four Country – Canada, United States, United Kingdom & Australia	Smokers	Waves 1 to 9 or 10
BD	Bangladesh	Smokers & Non-smokers	Waves 1 to 4
BR	Brazil	Smokers	Wave 1 & 2
BT	Bhutan	Smokers & Non-smokers	Wave 1
CN	China	Smokers & Non-smokers	Waves 1 to 5
DE	Germany	Smokers & Non-smokers	Waves 1 to 3
FR	France	Smokers & Non-smokers	Waves 1 to 3
IES	Ireland/Scotland/United Kingdom	Smokers & Non-smokers	Waves 1 to 4
IN	India	Smokers & Non-smokers	Waves 1 & 2
KE	Kenya	Smokers & Non-smokers	Wave 1
KR	South Korea	Smokers	Waves 1 to 3
MU	Mauritius	Smokers & Non-smokers	Waves 1 to 3
MX	Mexico	Smokers	Waves 1 to 7
NL	Netherlands	Smokers	Waves 1 to 9
NZ	New Zealand	Smokers	Waves 1 & 2
SEA	South East Asia – Thailand & Malaysia	Smokers, Non-smokers & Youth	Waves 1 to 6
UY	Uruguay	Smokers	Waves 1 to 5
ZM	Zambia	Users & Non-users	Waves 1 to 2

1.2 Common Variable Names

ITC participants are asked many questions about their smoking behaviors, opinions and attitudes toward tobacco control policies, which results in a large number of variables in any one dataset. Survey question variables and variables derived from survey questions follow a common naming convention in all ITC datasets, so that survey questions can be readily identified across datasets and country projects. The exception to this naming convention is the administrative variables, which instead have names that are intended to intuitively describe the variable. For example, the sex of the respondent is indicated by *sex*, and the country in which the respondent resides is indicated by *country*, and so on.

The standard ITC variable name format is such that each part of the variable name has a specific meaning. For example, for the variable aDE31212v:

a = wave identifier
DE = domain code
31 = country code
212 = domain number
v = optional suffix

Wave identifier. The wave identifier 'a' indicates that this variable is part of a Wave 1 dataset. Wave 2 data are indicated by a 'b', and Wave 3 by a 'c', and so on.

Domain code. The 'DE' refers to the variable category demographics. Table 4 lists the domain codes used in the ITC surveys. Note that NR refers to the same content domain as SM and CH combined. For the first 3 waves of the Four Country study, NR was used for all variables pertaining to stop-smoking medications and cessation help. The domain was then split into SM (for stop-smoking medications) and CH (cessation), and these two domain codes were used for all datasets thereafter.

Country code. The '31' indicates that this variable comes from an ITC China dataset. Note that data from English-speaking countries do not have a country code (i.e., aDE212v vs. aDE31212v); they have an implicit country code of '00'. **This variable name format is the default used to describe derived variables throughout this document.** Table 3 lists the country codes used in the ITC surveys.

Domain number. The same variable numbers usually refer to the same survey questions across waves and across country projects. In this case, 212 refers to an income variable. Where exactly the same question and response options are used in different countries, the same variable number is used, but sometimes the same number will also be used for a country's nearest equivalent in terms of question content. Also, some countries use multiple questions for the same content that have only one question in other countries (e.g., frequency of use of tobacco products is sometimes measured by a frequency scale that has "not at all" as one of the options, but in other countries the frequency question is preceded by a separate yes/no question that filters out those who do not use the product at all), so exact matching of variable names is not possible.

Optional suffix. The most common suffixes are 'v' and 'o', where v refers to a derived or computed variable that has been added to the dataset. In this case, the 'v' suffix on aDE31212v indicates that this China wave 1 income variable has been derived. An 'o' suffix refers to an 'other' response where the respondent can provide an open-ended response to a survey question (e.g., aBR31663o for a second cigarette brand that a respondent might usually smoke).

Other possible suffixes include an alphabetic series of 'a', 'b', 'c' and so on. This type of suffix series is only used in instances where the same survey question may be answered in more than one way, most often by using different units of measurement. For example, for survey question BR501, "About how long have you been smoking [your current brand of cigarettes]?" Respondents may provide their answer in days (BR501a), weeks (BR501b), months (BR501c), or years (BR501d).

Occasionally the suffixes w, x, y, and z are used to refer to the countries in the Four Country study, where the response categories are different – as for income (e.g., DE211wx is income for Canada and the United States; DE211y is income for the United Kingdom, and DE211z for Australia) and education.

Table 3. ITC country codes.

Country	Country Project	Country Code
Australia	Four Country	00
Canada	Four Country	00
United Kingdom	Four Country	00
United States	Four Country	00
Ireland	Ireland/Scotland/United Kingdom	00
Scotland	Ireland/Scotland/United Kingdom	00
New Zealand	New Zealand	00
Malaysia Adults	Southeast Asia	11
Thailand Adults	Southeast Asia	11
Malaysia Youth	Southeast Asia	12
Thailand Youth	Southeast Asia	12
China	China	31
Kenya	Kenya	72
Korea	Korea	21
France	France	41
Netherlands	Netherlands	42
Germany	Germany	44
Mexico	Mexico	51
Uruguay	Uruguay	52
Brazil	Brazil	57
India	India	61
Bangladesh	Bangladesh	62
Bhutan	Bhutan	64
Mauritius	Mauritius	73
Sudan	Sudan	77
Zambia	Zambia	74

Table 4. ITC domain codes.

Prefix	Domain	Description	Product(s)
AC	Anti-smoking Campaigns	Anti-smoking info and campaigns	NCT
AD	Ads & Anti-smoking campaigns	Combined for and against cigarettes	Cig
AI	After Interview	Contact info, compensation, future studies	--
AQ	Quit Attempts	Past attempts to quit – timing, strategies	NCT
BI	Before Interview	Household rostering, screening, conversion	--
BQ	Beliefs about Quitting		Cig
BR	Brands (cigarette brands)	Usual, last-purchased	Cig
CA	Anti-smoking Campaigns	Anti-smoking info and campaigns	Cig
CH	Cessation Help	Other than meds	Cig
DE	Demographics	Income, education, ethnicity, etc	--
DI	Traits & States (individual differences)		--
EA	Ads	AD201-699 – ads for cigs AD100-199, 701-999 – anti-tobacco campaigns	E-cig
EB	Brands (e-cigarette brands)	Usual, last-purchased, used for LQA	E-cig
EC	Characteristics, Types (characteristics of e-cigs	Not including brands	E-cig
ED	Dependence (e-cig dependence)	Addiction, susceptibility	E-cig
EE	Experiences with E-cigs	Perceptions	E-cig
EF	Frequency & Usage	Includes non-freq patterns of use	E-cig
EH	History of Use	Usage in the past	E-cig
EI	Industry & Regulation		E-cig
EK	Knowledge, Info, WL	Knowledge, sources of info, health warnings	E-cig
EL	(Last) Purchase (e-cig last purchase)	Last-purchase, other purchases	E-cig
EP	Use in Public (e-cigs in public)		E-cig
EQ	Cessation Help (quitting with e-cigs)	Cessation of cigs using e-cigs	Cig
ER	Reasons for Use & Non-use (e-cig reasons)	When & why used; negative effects	E-cig
ET	Use in Public		Cig

Prefix	Domain	Description	Product(s)
	(environmental tobacco smoke)		
ES	Psychosocial (e-cigs and social influence)	Beliefs and attitudes	E-cig
FR	Frequency & Usage	Includes reasons for use, non-freq patterns of use	Cig
HE	Health		--
HI	History of Use	Usage in the past	Cig, NCT
HK	Hookah		Hookah
IN	Industry & Regulation		Cig
KN	Knowledge	Diseases, harm mechanisms	Cig
KT	Knowledge	Diseases, harm mechanisms	NCT
LM	Light/ Mild	Light/ mild, menthol	Cig
LP	(Last) Purchase	Last-purchase, other purchases	NCT
ME	Media	Mass media usage, internet	--
MK	Midwakh/ Dokha	Only in UAE	Midwakh
NC	Non-cigarette Tobacco, by product	NT (main prefix)	NCT
NR	Cessation Meds & Assistance (old) (nicotine replacement)	Meds & other forms of help	Cig
NT	Non-cigarette Tobacco, by product	NC (spill-over from NC)	Cigars, cigarillos, filtered cigars
OT	Psychosocial (opinions about tobacco)	Beliefs and attitudes	NCT
PR	Perceived Risk		Cig, NCT
PS	Psychosocial	Beliefs and attitudes	Cig
PU	(Last) Purchase	Last-purchase, other purchases	Cig
QA	Quit Attempts	Past attempts to quit – timing, strategies	Cig
RE	Perceived risk (risks from e-cigs)		E-cigs
RH	Relative Harm	Harm compared over products	NCT
SB	Dependence (smoking behaviour)	Addiction, changes noticed (plain packaging), thoughts about	Cig, NCT
SL	Smokeless Tobacco (old)		SL
SM	Cessation Meds (stop-smoking meds)		Cig
SO	Sources of Tobacco	Stores and other sources, for last and other purchases	Cig
SS	Use in Public (second-hand smoke)		Cig, NCT
ST	Smoked Tobacco (old)		ST

Prefix	Domain	Description	Product(s)
TA	Ads		NCT
TB	Brands (tobacco brands)	Usual, last-purchased	NCT
TC	Tobacco Change	Adding, quitting, switching tobacco products	Cig, NCT
TF	Frequency & Usage	Includes non-freq patterns of use	NCT
TI	Industry & Regulation		NCT
TP	Use in Public (tobacco in public)		NCT
TQ	Beliefs about Quitting (beliefs about quitting tobacco)		NCT
TS	Availability & Cultivation (tobacco sales)		NCT
TU	Reasons for Use & Non-use (tobacco use)	When & why used; negative effects	NCT
WL	Warning Labels		Cig
WH	Wealth	Household expenditures, income	--
WT	Warning labels		NCT

1.3 Non-Response & Missing Value Coding

Non-responses are consistently coded as 7, 8, and 9. The number of digits used in non-response codes is dependent upon on the highest number used in response codes. Non-responses to open-ended questions are coded with 2 or more -digit NR codes (such as 77, 88, 99 or 777, 888, 999 etc.).

7,77,777, … = Not Applicable (NA)
8,88,888,… = Refused to answer (Ref)
9,99,999, … = Don't know (DK)

Any question that a respondent should not have been asked due to skip patterns is coded as NA. Responses are coded as system-missing ('.') where respondents were not present for a given wave, and also for questions that should have been asked but do not have a code-able response or non-response.

1.4 Analytic Data Sets

For many of the projects and waves, so-called "analytic data sets" have been created, for common or straightforward analyses, particularly analyses carried out by new users. Table A2 in Appendix 3 shows the countries and waves for which analytic data sets exist. The plan is to

construct them for all of the data sets eventually. The analytic data sets cover the same domains as the wave datasets as described in Section 3, but have been simplified in the following ways:

- Unneeded components of derived variables have been excluded.
- An effort has been made to minimize the number of variables that do not have responses for everyone, or everyone of a certain smoking status, by constructing derived variables which have responses for all. For example, SB226 (frequency of butting out before finishing) has a response only from those who have said "Yes" to SB221 ("In the last month, have you butted out a cigarette before you finished it because you thought about the harm of smoking?"). In the analytic data sets, SB226 is replaced by SB226v, which is given the value 0 = Never if the answer to SB221 is "No".
- Where possible, open ended responses have been coded.
- In the analytic data sets, instead of using 7, 8, 9 or 77, 88, 99… to indicate "not applicable", "refused to answer" and "do not know", ".n", ".r" and ".d" are used in SAS datasets.
- Consistent derived variables have been provided for levels of education and annual income.
- Simplified derived variables for ethnicity have been provided.
- Prices are given on a common basis, and in one currency (i) using exchange rates and (ii) using purchasing power parity.

In this document, unless otherwise indicated, the derived variables are present in the wave datasets specified in the tables. Not all of them will be present in the analytic data sets. It will be indicated where a variable has been introduced specifically for purposes of an analytic data set.

2. Core Dataset Variables

Core datasets contain basic information on every respondent ever recruited into a given country project. The core datasets comprise administrative variables which do not follow the standard ITC variable naming convention, but rather have names meant to intuitively describe the variable. It is important to note that this document contains a comprehensive listing of all variables in all core datasets, but many of these variables are not available in every core dataset. Also note that core variables associated with geography and sampling procedures are discussed in *4.1 Strata, Clusters, & Other Geographic Variables.*

2.1 Key Identifier & Demographics

Table 5 summarizes the key identifier and demographic variables, and Table 6 lists the availability of these variables in the various ITC core datasets.

Table 5. Key identifier and demographic variables in core datasets.

Variable	Label
uniqid	Unique respondent ID number
cohort	Wave of recruitment
sex	Sex
age	Age at recruitment (continuous)
ageGrp	Age at recruitment (categorical)
country	Country
lang	Language spoken other than French
ethnic	Ethnicity
race	Malaysian/Thai race
status	Smoking status
houseCmp	Household composition

uniqid – unique respondent ID number. The variable *uniqid* uniquely identifies respondents across all waves within a given country project. In order to link the same respondent's data between datasets, the identifier remains the same in all datasets. Note that *uniqid* is not linkable between country projects, as respondents only participate in one country project. The numeric value of *uniqid* is computed differently for each country project, but ultimately is irrelevant except as a unique numeric string used to link data. The *uniqid* variable is present in all datasets.

cohort – wave of recruitment. The variable *cohort* indicates at which wave of the country project each respondent was recruited. The response '1' indicates that the respondent was recruited in the first wave, '2' in the second wave, and so on.

Table 6. Variable listing of key identifier and demographic variables in core datasets.

Variable	4C	BD smokers & non-smokers	BR	BT	CN smokers & non-smokers	DE	FR smokers	FR nonsmokers	IES	IN smokers & non-smokers	KE	KR	MU smokers & non-smokers	MX	NL	NZ	SEA smokers & non-smokers	SEA youth	UY	ZM users & non-users
uniqid	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
cohort	×	×	×	×	×	×	×	×	×	×		×	×	×	×	×	×	×	×	×
sex	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
age	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
ageGrp	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
country	×	×	×		×	×	×	×	×	×		×		×	×	×	×	×	×	×
lang							×	×												
ethnic	×				×				×							×	×	×		
race																	×	×		
status					×															
houseCmp	×					×	×		×			×								

sex. The variable *sex* indicates the sex of the respondent as discerned by the survey interviewer or provided by the respondent at the time of recruitment. Sex is coded '1' as female and '2' as male.

age – age at recruitment (continuous). The variable *age* measures respondent's age in years and is computed by subtracting the date of birth from the year of recruitment, in *rDate*. This variable is present on all core datasets.

ageGrp – age at recruitment (categorical). The variable *ageGrp* categorizes respondents' ages in years. The age values are taken from *age* to compute the *ageGrp* categories. For most country projects, the computed age categories are:

1 = 18 to 24 years
2 = 25 to 39 years
3 = 40 to 54 years
4 = 55 years and older

In the Southeast Asia Youth datasets, the age categories are:

1 = 13 to 15 years
2 = 16 to 17 years

country – country of residence. The variable *country* indicates the country in which the respondent resides. The country variable may seem redundant for many projects, as the country name is frequently reflected in the project name, such as for the ITC Germany project. However, in the case of multinational projects such as the ITC Four Country project or the ITC Southeast Asia project, this variable is useful. As well, the *country* variable is useful when conducting inter-country comparisons.

lang – Language spoken other than French. The variable *lang* is present only on the France core datasets. It identifies whether respondents speak any languages in addition to French, where the response options are:

→ 1 = French only
→ 2 = French and another language
→ 3 = Another language other than French only
→ 88 = Refused
→ 99 = Don't know

ethnic – ethnicity. The variable *ethnic* is a dichotomous variable that indicates the ethnicity of the respondent as either part of the majority group in the resident's country, or other. Ethnicity was measured differently in every country and was generalized differently in every country project. Refer to **Table 7** for the ethnicity response options and coded responses for each country.

race – Malaysian/Thai race. The variable *race* indicates the ethnicity (or race) of Southeast Asian respondents, as provided at the time of recruitment. The *race* variable was used to compute the more generalized *ethnic* variable. Response options are

101 = Malay
102 = Chinese
103 = Indian
104 = Iban
105 = Kadazan
106 = Murut
107 = Bajau
108 = Other
201 = Thai
202 = Chinese
203 = Lao
204 = Yavee
205 = Eko
206 = Yow
207 = Kaling
208 = Other tribesman language
209 = Kamre
210 = Sui
211 = Other
777 = NA
888 = Refused
999 = Don't know

status – Smoking status. The variable *status* indicates the respondent's frequency of smoking at the time of recruitment as daily, weekly or monthly. The variable *rSmoke* is also listed in some datasets; *rSmoke* is interchangeable with *status*. *rSmoke* and *status* are similar to *FR309v* of the wave datasets, which indicates the respondent's frequency of smoking at each wave. Response options are:

1 = Daily smoker
2 = Weekly smoker
3 = Nonsmoker

houseCmp – smoking status of adults in household. The variable *houseCmp* indicates the smoking status of all adults who reside within the respondent's household. The total number of adults in the household and the number of adult smokers is ascertained at the time of recruitment.

1 = Single adult smoker
2 = All adult smokers
3 = Mixed adult household
4 = Number of smokers not stated (France only)
7 = NA
8 = Refused
9 = Don't know

Table 7. Ethnicity Response Options and Coded Responses

Country	Interview Response Options	Project	Responses in Datasets
China	Han	CN	1 = Han
	Zhuang, Man, Hui, Miao, Uygur, Yi, Tujia, Mongolian, Tibetan, Other		2 = Others
Ireland	Irish	IES	1 = White
	Other		2 = Other
Scotland	White	IES	1 = White
	Asian/Asian-British, Black/Black-British, Chinese, Mixed, Other		2 = Other
Canada	White	ITC4	1 = Majority group
	Chinese, South Asian, Black, Filipino, Latin American, South East Asian, Arab, West Asian, Korean, Aboriginal, Other.		2 = Minority group
US	White	ITC4	1 = Majority group
	Black, Hispanic, Asian, Native, Other		2 = Minority group
Australia	Speaks only English at home	ITC4	1 = Majority group
	English & one of: Italian, Greek, Cantonese, Mandarin, Arabic, Vietnamese, Other		2 = Minority group
UK	White	ITC4	1 = Majority group
	Asian/Asian-British, Black/Black-British, Chinese, Mixed, Other		2 = Minority group
New Zealand	Maori	NZ	1 = Maori
	Pacific		2 = Pacific
	Asian		3 = Asian
	Other		5 = Other
Malaysia	Malay	SEA	1 = Majority Group
	Chinese, Indian, Other		2 = Minority Group
Thailand	Thai	SEA	1 = Majority Group
	Other		2 = Minority Group

2.2 Recruitment

Table 8 summarizes the variables describing recruitment and Table 9 lists the availability of these variables in the various ITC core datasets.

Table 8. Recruitment variables in core dataset

Variable	Label
Mode	Survey mode
rCentre	Recruitment call centre
rDate	Recruitment survey date
rDisp	Recruitment call disposition
rIntvwr	Recruitment interviewer ID
rLength	Length of recruitment survey
rSmoke	Smoking status at recruitment
rStatus	Smoking status at recruitment (binary)
statusRecr	Tobacco use status at the time of recruitment
tobaccoStatus	Type of tobacco user at recruitment
caFrench	French interview
Aflag_hh	Smoking type in enumeration survey

Table 9. Variable listing of recruitment variables in core datasets

Variable	4C	BD smokers & non-smokers	BR	BT	CN smokers & non-smokers	DE	FR smokers	FR nonsmokers	IES	IN smokers and non-smokers	KE	KR	MU smokers	MU nonsmokers	MX	NL	NZ	SEA smokers & non-smokers	SEA youth	UY	ZM users & non-users
Mode																×					
rCentre	×																				
rDate	×						×	×					×	×	×		×			×	×
rDisp	×																				
rIntvwr	×																				
rLength	×																				
rSmoke	×				×	×	×					×	×		×					×	
rStatus							×	×									×				
statusRecr																					×
tobaccoStatus										×											

Variable	4C	BD smokers & non-smokers	BR	BT	CN smokers & non-smokers	DE	FR smokers	FR nonsmokers	IES	IN smokers and non-smokers	KE	KR	MU smokers	MU nonsmokers	MX	NL	NZ	SEA smokers & non-smokers	SEA youth	UY	ZM users & non-users
caFrench	×																				
aflag_hh				×																	

Mode – Survey mode. The variable *mode* indicates whether the combined recruitment & wave 1 interview was web-based (CASI) or conducted by telephone (CATI). The response options are:

1 = CATI
2 = CASI

rCentre – recruitment call centre. The variable *rCentre* indicates the call centre which made the recruitment calls. The response options are:

1 = Toronto
2 = Montreal
3 = Calgary
4 = Melbourne
5 = Auckland

rDate – recruitment survey date. The variable *rDate* indicates on which date the call centre made the recruitment call. The dates are presented in the format dd-mm-yyyy.

rDisp – recruitment call disposition. The variable *rDisp* indicates the call disposition for the final phone call attempt for recruitment. Dispositions indicate whether the respondent completed the entire recruitment survey or whether any questions were skipped. "Next BD" means "next birthday". The response options are:

0 = Not in service
1 = Fax, modem
2 = TCI fault
3 = Non-residential
4 = Cell phone, mobile
8 = Resp completes to Q19 but hangs up before end
9 = Resp completes recruitment script with no skips
10 = Appt 0: with initial resp for next BD smoker
11 = Appt 1: with initial resp to complete introduction
12 = Appt 2: with next BD resp to complete screener
13 = Appt 3: initial resp makes appt for next BD smoker
17 = Household refusal to get resp
18 = Initial resp refuses to give household composition
19 = No one 18 or older
20 = No adult smoker, household size given
21 = No adult smoker, household size refused

22 = Resp refuses but answers Q1d,e, >0 smokers
27 = Next BD smoker refuses to participate
28 = Initial resp refuses to get next BD smoker
29 = Resp unavailable this wave
30 = Next BD smoker refuses to give age
32 = Next BD smoker too young, no others
33 = Next BD smoker <100 cigarettes, no others
34 = Next BD smoker smokes < monthly, no others
35 = Resp refuses
40 = Rings only
41 = Busy, no ans machine
42 = Ans mach, not clear if residential
43 = Ans mach, residential
77 = NA
80 = Initial resp hangs up in intro
81 = Next BD smoker hangs up during screener
82 = Resp hangs up during Q5-Q19
83 = Resp skips question –s in Q5-Q19
84 = Resp completes to Q19, skips later ques, reaches end
85 = Complete, no income
88 = Refused
90 = Term: initial resp has language problem
91 = Term: intial resp is incompetent
92 = Term: other problem with initial resp
93 = Term: Next BD smoker has language problem
94 = Term: Next BD smoker is incompetent
95 = Term: Other problem with Next BD smoker
99 = Don't know

rIntvwr – recruitment interviewer ID. The variable *rIntvwr* indicates the interviewer who conducted the recruitment call with a numeric identifier.

rLength – length of recruitment survey. The variable *rLength* indicates the length of time in minutes it took to complete the recruitment survey.

rSmoke – smoking status at recruitment. The variable *rSmoke* indicates the respondent's smoking status at recruitment. The response options for France and Four Country are:

1 = Daily
2 = Weekly
3 = Monthly
4 = Quit less than 1 month ago
5 = Quit 1 to 6 months ago
6 = Quit more than 6 months ago

The response options for Mexico and Uruguay are:

1 = Daily
2 = Weekly
3 = Monthly

The response options for Korea are:

1 = Daily
2 = Weekly

rStatus – smoking status at recruitment (binary). The variable *rStatus* indicates the respondent's smoking status at recruitment as being either smoking or non-smoking. The response options are:

1 = Smoker

2 = Nonsmoker

statusRecr – tobacco use status at the time of recruitment. The variable *statusRecr* indicates the respondent's smoking status at recruitment. This variable is only available in Zambia dataset.

tobaccoStatus –type of tobacco user at recruitment. The variable *tobaccoStatus* indicates the respondent's tobacco-using status at recruitment. This variable is only available in India dataset. The response options are:

1 = Smoked tobacco
2 = Mixed use
3 = Smokeless
4 = Non-user

caFrench – interview conducted in French. The binary variable *caFrench* indicates whether the interviews in Canada were conducted in French. The response options are:

1 = Yes
2 = No

aflag_hh – smoking type in wave 1 enumeration survey. The variable *aflag_hh* is present only in Bhutan core dataset.

2.3 Participation

Table 10 summarizes the variables describing survey participation and Table 11 lists the availability of these variables in the various ITC core datasets.

Table 10. Survey participation variables in core

Variable	Label
inM*x*	Respondent present in wave *x* main survey
inP*x*	Respondent present in wave *x* replenishment survey
inR*x*	Respondent present in wave *x* recruitment survey
lost	Wave at which respondent was lost to follow up

Table 11. Variable listing for survey participation variables in core dataset.

Variable	4C	BD smokers & non-smokers	BR	BT	CN smokers & non-smokers	DE	FR smokers & non-smokers	IES	IN smokers and non-smokers	KE	KR	MU smokers & non-smokers	MX	NL	NZ	SEA smokers & non-smokers	SEA youth	UY	ZM users & non-users
inM*x*	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
inP*x*	×	×	×		×	×	×	×	×		×	×	×	×		×	×	×	×
inR*x*	×										×								
lost	×							×	×			×				×	×		×

*inM*x – *respondent present in wave* x *main survey.* The *inMx* series of variables indicates whether a respondent completed the main/cohort survey at wave x, where *x* represents a wave number, 1, 2, 3, and so on. At wave 1, all participants are part of the main survey. At wave 2 and beyond, all respondents who are successfully recontacted and complete the wave 2 survey are part of the main sample. The response options are:

0 = Absent from the wave
1 = Present in the wave
7 = Not applicable

inPx – respondent present in wave x replenishment survey. The *inPx* series of variables indicates whether a respondent was present as a new recruit for a given survey (where *x* represents a wave number, 1, 2, 3, and so on). ITC replaces those respondents who are lost to follow-up, and new respondents are thus part of the replenishment survey. As such there is no variable *inP1* since there is no replenishment survey at wave 1. Respondents become part of the main sample of every wave in which they participate subsequent to the replenishment wave. The response options are identical to those in the *inMx* series.

inRx – respondent present in wave x recruitment survey. The *inRx* series of variables indicates those who were recruited at any given wave (where *x* represents a wave number, 1, 2, 3, and so on). All of those who are present in the replenishment sample are necessarily part of the recruitment sample, since they first completed the recruitment screening survey and then participated in the full ITC survey. There is a differentiation between recruitment and replenishment because individuals who participate in the recruitment survey do not necessarily complete the replenishment survey. It is important to note that in most cases, the recruitment survey and the main (or replenishment) survey are combined into a single interview, and as such, *inRx* is only applicable where recruitment is conducted in a separate interview as in the Four Country Survey. The response options are identical to those in the *inMx* series.

lost – wave at which respondent was lost to follow up. The variable *lost* identifies those respondents who were lost to follow-up, and at which wave they were lost, where '2' indicates lost at wave 2, '3' at wave 3, and so on. For participants who continued to participate in the most recent survey wave, this variable is entered as '0'.

2.4 Other Core Variables

Table 12 summarizes all other variables and Table 13 lists the availability of these variables in the various ITC core datasets.

Table 12. Other core variables.

Variable	Label
*y*flag_hh	smoking type in wave *y* enumeration survey
*y*flag_main	smoking type in wave *y* main survey
*y*flag_sc	smoking type in wave *y* screener survey
biomrkrID	Link to biomarker/cotinine study
chiIdx	socio-economic status based on CASHPOR Housing Index tertiles
dqFlag2	Data quality flag at wave 2
dqFlag3	data quality flag at Wave 3
hhid	household id (combined)
NOFEMALENE	# of female non/ex-smokers
NOFEMALEUS	# of female smokers
NOMALENEVE	# of male non/ex-smokers
NOMALEUSER	# of male smokers
TOTALHOUSE	population of household
personid	Person id
remark	Remark in wave 2
sample	Sample type (based on division/district/upazila)-numeric
sampleType	Sample type (based on division/district/upazila)-character
sformno	Census form number
wave	ITC wave

yflag_hh – smoking type in wave y enumeration survey. The variable *yflag_hh* is present only in Bhutan. As there is only one wave to date, it is called *aflag_hh* in the data set.

yflag_main – smoking type in wave y main survey. The variable *yflag_main* is present only in Bhutan.

yflag_sc – smoking type in wave y screener survey. The variable *yflag_sc* is present only in Bhutan.

biomrkrID – link to biomarker/cotinine study. The variable *biomrkrID* provides a numeric value that links respondents in the Mexico datasets to another research study, which collected biological samples for a pilot study.

childx – socio-economic status based on CASHPOR Housing Index tertiles. The variable is present only in Bangladesh. 'Low', 'medium' and 'high' socio-economic status are assigned by tertiles of CASHPOR housing index.

householdID – household identifier. This variable is present only in the Zambia dataset.

hhid – household id (combined). The variable *hhid* is present only in the Bhutan dataset. The household id is combined by value of variables *dzongkhag, urbanrural, gewog, ea,* and *household*.

hhIdId –unique household ID. The variable *hhIdId* is present only in the India dataset.

NOFEMALENE – # of female non/ex-smokers. The variable *NOFEMALENE* is present only in the Bhutan dataset and identifies the population of female non-smokers or ex-smokers in the household.

NOFEMALEUS – # of female smokers. The variable *NOFEMALEUS* is present only in the Bhutan dataset and identifies the population of current female smokers in the household.

NOMALENEVE – # of male non/ex-smokers. The variable *NOMALENEVE* is present only in the Bhutan dataset and identifies the population of male non-smokers or ex-smokers in the household.

NOMALEUSER – # of male smokers. The variable *NOMALEUSER* is present only in the Bhutan dataset and identifies the population of current male smokers in the household.

TOTALHOUSE – population of household. The variable *TOTALHOUSE* is present only in the Bhutan dataset and identifies the population of the household, which is the sum of variables *NOFEMALENE, NOMALENEVE, NOMALEUSER,* and *NOMALEUSER.*

personid. The variable *personid* numbers the respondents based on the order in which they were surveyed within the household.

sample/sampleType – These two variables *sample* and *sampleType* are present only in Bangladesh data. *Sample* is in numeric data type and *sampleType* is in character data type.

sformno. The variable *sformno* is present only in Bangladesh data and identifies the census form number for respondents.

wave. The variable *wave* is present only in Bangladesh data and identifies the ITC wave.

remark. The variable *remark* is present only in China data.

dqFlag2. The variable *dqFlag2* is present only in Germany data and is for internal usage only.

dqFlag3. The variable *dqFlag3* is present in the 4-country data and is for internal usage only.

Table 13. Variable listing for other core variables.

Variable	4C	BD smokers & non-smokers	BR	BT	CN smokers & non-smokers	FR smokers & non-smokers	DE	IES	IN smokers and non-smokers	KE	KR	MU smokers & non-smokers	MX	NL	NZ	SEA smokers & non-smokers	SEA youth	UY	ZM users & non-users
yflag_hh				×															
yflag_main				×															
yflag_sc				×															
biomrkrID													×						
chilIdx		×																	
dqFlag2							×												
dqFlag3	×																		
householdID																			×
hhid				×															
hhIdId									×										
NOFEMALENE				×															
NOFEMALEUS				×															
NOMALENEVE				×															
NOMALEUSER				×															
TOTALHOUSE				×															
personid													×						
remark					×														
sample		×																	
sampleType		×																	
sformno		×																	
wave		×																	

3. Wave Dataset Variables

A wave dataset contains all survey questions, derived variables and administrative variables for a given survey wave, and only contains records for those who participated in that particular wave. It is important to note that this document contains a comprehensive listing of derived variables in all wave datasets, but not all of these variables are found in every wave dataset.

Also, it is important to note that the variables that follow the standard ITC variable naming convention (e.g., aDE31212v) are referenced in this document without wave identifiers or country codes (e.g., DE212v). This alteration was made so that variables may be easily cross-referenced in this document with ITC datasets from all waves and all countries.

3.1 Wave Administrative Variables

Table 14 summarizes the key identifier and demographic variables found in wave datasets and Table 15 lists the availability of these variables in the various ITC wave datasets.

Table 14. Wave administrative.

Variable	Label
uniqid	Unique respondent ID number
cohort	Wave of recruitment
country	Country
Sex	Sex of respondent
BIRTHYR	Year of birth
status	Smoking status
linked	Linked to wave 1
Mode	Survey mode
yage	Age for wave y
yCentre	Survey call centre for wave *y*
yDate	Survey date for wave *y*
yDisp	Survey call disposition for wave *y*
yflag_main	Smoking type in main survey for wave y
yflag_sc	Smoking type in screener survey for wave y
yIntvwr	Interviewer ID for wave y
yLang	Language of survey for wave *y*
yLength	Length of survey for wave *y*
yMode	Mode of interview for wave *y*
ySrvyd	Wave participation total at current interview
Language	Language interview condu
eDateFlag	Flag for data quality issues
qDateFlag	Flag for quit date: questionable data for quit history questions
smkFlag	Flag for smoking status
biomrkrID	Link to biomarker/cotinine study

Variable	Label
hhOutcome	Household outcome code
casenum	Case number (Malaysia)
rrkey	Respondent identifier

Table 15. Variable listing for wave administrative variables.

Variable	4C	BD smokers & non-smokers	BR	BT	CN smokers & non-smokers	DE smokers & non-smokers	FR smokers & non-smokers	IES	IN smokers and non-smokers	KE	KR	MU smokers & non-smokers	MX	NL	NZ	SEA smokers	SEA non-smokers	SEA youth	UY	ZM users & non-users
uniqid	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
cohort	×	×	1		1-3		×	×	×		1-2	×		×	×	×	×	×		×
country	×		1		×		×	×			1-2				×	×	×	×		
SEX						3		4						×						
BIRTHYR														1-7						
status					×															
linked																2		2		
Mode														×						
*y*age																		×		
*y*Centre								×												
*y*Date	×	×	×	×	×	×	×	×	×	×	1-2	×	×			×	×	×	×	×
*y*Disp	1-7							1-3									2			
*y*flag_main				×																
*y*flag_sc				×																
*y*Intvwr	1-4							1-3								2	2	2		
*y*Lang																2		2		
*y*Length	1,2,4-10	×			1-3			×	×	×	1-2	×	6			×	×	×		×
*y*Mode	×															2		2		
*y*Srvyd	1-7							1-3												
Language				×																

Variable	4C	BD smokers & non-smokers	BR	BT	CN smokers & non-smokers	DE smokers & non-smokers	FR smokers & non-smokers	IES	IN smokers and non-smokers	KE	KR	MU smokers & non-smokers	MX	NL	NZ	SEA smokers	SEA non-smokers	SEA youth	UY	ZM users & non-users
eDateFlag	5																			
qDateFlag																2	2			
smkFlag	4																			
biomrkrID													1 -5							
hhOutcome																		1		
casenum																	2	2	2	
rrkey																	2		2	
survey				×																
releaseDate									×											

In cases where a variable is not present on all waves of a project, only the wave numbers where it is present are listed.

For *uniqid, cohort and country,* see Section 2.4 *Other Core Variables*. *yCentre – call centre for wave y*. The *yCentre* variable indicates the call centre from which the phone interview was conducted at a given wave. The wave is indicated with a wave identifier preceding the variable name (where *y* represents a wave identifying letter, such as 'a' for Wave 1, or 'b' for Wave 2, and so on). The response options are:

1 = Toronto
2 = Montreal
3 = Calgary
4 = Melbourne
5 = Auckland

yage – age for wave y. The *yage* variable indicates the age when the interview was conducted.

yDate – survey date for wave y. The *yDate* variable indicates on what date the interview was conducted. The dates are presented in the format dd-mmm-yyyy.

yDisp – call disposition for wave y. The *yDisp* variable indicates the call disposition for the final phone call attempt for recruitment. The response options for Ireland/Scotland are:

9 = Complete

81 = Refuses 1 or more ques, complete 2
83 = Partial (skip 1+ in q11-q22)
84 = Partial (skip 1+ after q22)

The response options for Four Country are:

0 = Non-contact, not in service, line problem
9 = Contact, all questions answered – complete 1
15 = Contact, rescheduled appt not kept
27 = Resp calls to withdraw
28 = Contact, household refusal to get resp
29 = Contact, resp unavailable this wave
30 = Contact, resp has died
31 = Emerg # sought, no contact
32 = Emerg # sought, contact, unknown
33 = Emerg # sought, contact, refused
34 = Emerg # sought, contact, obtained
40 = Non-contact, rings only
41 = Non-contact, busy
42 = Non-contact, ans machine, fax or modem
77 = NA
80 = Contact, resp refuses before Qxx is finished – refusal
81 = Contact, all ques asked, refuses 1+ quest – complete 2
83 = Interruption, never completed – incomplete
84 = NBD smoker completes screener, skips later questions
85 = Resp completes all but income – complete 1
86 = Contact, age, sex, smoking status unknown
88 = Refused
95 = Interviewer termination
96 = Non-contact, unsuccessful attempt to trace
99 = Don't know

Length – length of survey for wave y. The *yLength* variable indicates the length of time in minutes it took to complete the interview/survey for a given wave *y*.

yMode – mode of interview for wave y. The *yMode* variable indicates the mode of interview, where '1' is face-to-face, and '2' is telephone interivew, for a given wave *y.*

ySrvyd – total number of interviews (including current one) a respondent has participated . This variable indicates the cumulative number of waves in which the respondents have participated as of the completion of a given wave *y*. *ySrvyd* is derived from the *cohort* variable.

eDateFlag — Flag for data quality issues. This variable is present only on the Four Country wave 5 dataset. It indicates where survey dates were not provided by the survey firm, and as such,

approximate survey dates were imputed based on the median survey date for that particular wave. Where *eDateFlag* equals '1', data quality issues exist, and `0' otherwise.

qDateFlag – Flag for quit date: questionable data for quit history questions. This variable is present only on the Southeast Asia adult wave 2 datasets. It flags observations where the time since self-reported quit is longer than the time elapsed between waves 1 and 2 (31 months). As a result, the data for *QA441a* to *QA441e* were problematic even after attempts were made to verify the data. As such, this variable flags these quit history question data quality issues, when *qDateFlag* is equal to '1', and `0' otherwise.

smkFlag — Flag for smoking status. This variable is only present in the Four Country wave 4 dataset. It flags observations where the variables used to compute smoking status did not seem to provide enough data to actually compute smoking status (*FR309v*). Note that *FR309v,* as computed by the survey firm, does not appear to have any data quality issues, but rather this variable flags where the 'paper trail' seems to have been lost. The lost paper trail is indicated when *smkFlag* is equal to '1'.

Survey – complete survey. This variable is only present in Bhutan wave 1 datasets and indicates if the respondents completed main (individual) survey.

releaseDate - date dataset was last updated/released. This variable is only present in India dataset and indicates the last updated/released date of the data.

3.2 Advertising/Promotion/Monitoring of Anti-Tobacco Campaigns

Table 16 summarizes the advertising derived variables and Table 17 lists the availability of these variables in the various ITC wave datasets.

Table 16. Advertising of anti-tobacco campaigns derived variables.

Variable	Label
AD105v	Heard about Tak Nak/Participated in Quit & Win
AD107v	On posters & billboards
AD108v	Media score
AD109v	Media source score
AD111v	Quit & Win made you think about quitting, overall
AD135v	Main message 1 coded
AD136v	Media of message 1 -poster & billboard
AD137v	Main message 2 coded
AD138v	Media of message 2 - poster & billboard
AD139v	Main message 3 coded
AD140v	Media of message 3 - posters & billboards

Variable	Label
AD141v	Most effective medium of Tak Nak ad - poster & billboard
AD146v	Discussion amongst family & friends
AD151v	Discourage smoking in children -- coded
AD152v	Discourage smoking in young adults --coded
AD153v	Discourage smoking in older adults -- coded
AD171v	Picture: finance meets parents - recode
AD172v	Picture: audition for model - recode
AD173v	Picture: tar in lungs - recode
AD174v	Picture: rotting lungs - recode
AD175v	Picture: blood clot in brain - recode
AD176v	Picture: lung cancer - recode
AD177v	Advertising recall - score
AD181v	Finance meet parents -recode
AD182v	Audition to be model - recode
AD183v	Tar in lungs - recode
AD184v	Rotting lungs - recode
AD185v	Blood clot in brain - recode
AD186v	Lung cancer - recode
AD187v	Advertising recall - score
AD240v	Noticed any advertising
AD759v	Noticed any anti-smoking info
AD901v	Teeth decay
AD902v	Medium mentioned for teeth decay
AD903v	Rotting lungs
AD904v	Medium mentioned for rotting lungs
AD905v	Cause cancer
AD906v	Medium mentioned for cause cancer
AD907v	Blood clots
AD908v	Medium mentioned for blood clots
AD909v	Smoking kills
AD910v	Medium mentioned for smoking kills
AD911v	Every puff damages
AD912v	Medium mentioned for every puff
AD913v	Quit smoking
AD914v	Medium mentioned for quit smoking
AD915v	Smoking is addictive

Variable	Label
AD916v	Medium mentioned for smoking addictive
AD919v	Premature ageing
AD920v	Medium mentioned for premature ageing
AD921v	Cause impotence
AD922v	Medium mentioned for causes impotence
AD925v	Endangers health
AD926v	Medium mentioned for endanger health
AD927v	Say no to cigs
AD928v	Medium mentioned for say no to cigs
AD929v	Don't smoke
AD930v	Medium mentioned for don't smoke
AD931v	Smoking forbidden
AD932v	Medium mentioned for smoking forbidden
AD933v	Less smoking habits
AD934v	Medium mentioned for less smoking habits
AD950v	Tak Nak message recall

Table 17. Variable Listing for advertising derived variables.

Variable	4C	BD smokers & non-smokers	BR	BT	CN smokers	CN nonsmokers	DE smokers & non-smokers	FR smokers & non-smokers	IES	IN smokers and non-smokers	KE	KR	MU smokers & non-smokers	MX	NL	NZ	SEA smokers & non-smokers	SEA youth	UY	ZM users & non-users
AD105v					1												1			
AD107v																		1		
AD108v																		1		
AD109v																	1			
AD111v					1															
AD135v																		1		
AD136v																		1		
AD137v																		1		
AD138v																		1		
AD139v																		1		

Variable	4C	BD smokers & non-smokers	BR	BT	CN smokers	CN nonsmokers	DE smokers & non-smokers	FR smokers & non-smokers	IES	IN smokers and non-smokers	KE	KR	MU smokers & non-smokers	MX	NL	NZ	SEA smokers & non-smokers	SEA youth	UY	ZM users & non-users
AD140v																		1		
AD141v																	1	1		
AD146v																	1			
AD151v																	1			
AD152v																	1			
AD153v																	1			
AD177v																	1			
AD187v																		1		
AD240v	6-8.5																			
AD759v	6-9																			
AD901v																		1		
AD902v																		1		
AD903v																		1		
AD904v																		1		
AD905v																		1		
AD906v																		1		
AD907v																		1		
AD908v																		1		
AD909v																		1		
AD910v																		1		
AD911v																		1		
AD912v																		1		
AD913v																		1		
AD914v																		1		
AD915v																		1		
AD916v																		1		
AD919v																		1		
AD920v																		1		
AD921v																		1		
AD922v																		1		

Variable	4C	BD smokers & non-smokers	BR	BT	CN smokers	CN nonsmokers	DE smokers & non-smokers	FR smokers & non-smokers	IES	IN smokers and non-smokers	KE	KR	MU smokers & non-smokers	MX	NL	NZ	SEA smokers & non-smokers	SEA youth	UY	ZM users & non-users
AD925v																		1		
AD926v																		1		
AD927v																		1		
AD928v																		1		
AD929v																		1		
AD930v																		1		
AD931v																		1		
AD932v																		1		
AD933v																		1		
AD934v																		1		
AD950v																		1		

In cases where a variable is not present on all waves of a project, only the wave numbers where it is present are listed.

AD105v – heard about Tak Nak on billboards & posters/Participate in Quit & Win, overall. This variable is present on both the Southeast Asia adult smoker & non-smokers wave 1 datasets and the China smokers wave 1 dataset, but in each case the meaning is unique. In the Southeast Asia datasets, this variable identifies whether respondents had ever heard of the Tak Nak anti-tobacco campaign, where '1' is mentioned and '2' is not mentioned. In the China dataset, it identifies whether respondents participated in the Quit & Win program, where the response options are:

1 = Yes/mentioned
2 = No/not mentioned

AD107v – on posters and billboards. This variable combines the posters (*AD104*) and billboards (*AD105*) responses, for the survey questions asking on which media had respondents noticed Tak Nak messages. This combination was done to make these responses more easily comparable to other surveys where poster and billboards are always combined. This variable is only present in the Southeast Asia wave 1 youth survey.

AD108v and AD109v – media score. These variables sum *AD101* to *AD107,* that is, the number of media sources the respondents reported noticing Tak Nak messages; AD108v is the derived variable found in the Southeast Asia youth datasets and AD109v in the adult datasets.

AD111v – Quit & Win made you think about quitting, overall. In the China smokers dataset, *AD111* is only asked of those who said in *AD100* that they had heard of Quit & Win at AD100. The *AD111v* variable also includes those who had not heard of Quit & Win, coding their response for 'Quit & Win made me think about quitting' as 'not at all'. The response options are:

1 = Very much
2 = A little
3 = Not at all

AD135v, AD137v, AD139v – main message 1, 2 & 3 coded. These variables code the corresponding open-ended variables on Tak Nak messages (e.g., *AD135o*) into categorized summary variables. The categories for these variables are:

1 = Teeth decay
2 = Rotting lungs
3 = Cause cancer
4 = Blood clots
5 = Smoking kills
6 = Every puff
7 = Quit smoking
8 = Smoking is addictive
9 = Hard drugs
10 = Premature aging
11 = Causes impotence
12 = Children's IQ
13 = Endangers health
14 = Say no to cigarettes
15 = Don't smoke
16 = Smoking is forbidden
17 = Less smoking habits

AD136v, AD138v, AD140v – media of messages 1, 2 & 3 – posters and billboards. These variables combine the count of the corresponding variables on posters (e.g., *AD136c*) and billboards (e.g., *AD136d*). The categories for these variables are:

1 = TV
2 = Radio
3 = Posters & billboards
5 = Newspapers

AD141v – most effective medium of Tak Nak ad – poster and billboard. This variable combines the count of the corresponding variables on posters (*AD141c*) and billboards (*AD141d*). The categories for this variable are:

1 = TV
2 = Radio
3 = Posters & billboards
5 = Newspapers
6 = None of the above
7 = NA/Don't know anything about Tak Nak
9 = Not sure

AD146v – discussion among family and friends. This variable combines the responses from the variables that ask whether Tak Nak messages prompted discussion among family (*AD145*) and discussion among friends (*AD146*). Response options are:

1 = Yes
2 = No

AD151v, AD152v, AD153v – discourage smoking in children, young adults or older adults -- coded. These variables recode the responses from the corresponding variables that ask whether respondents think Tak Nak messages will be effective in discouraging smoking in different age groups. Originally there were 5 responses, 'Definitely yes', 'probably yes', 'probably not', 'definitely not', 'can't say'. The two positive and two negative responses are collapsed into one 'yes' and one 'no' category, leaving a total of 3 response categories in the derived variables:

1 = Yes
2 = No
3 = Can't say

AD177v & AD187v –advertising recall - score. These variables sum the number of advertisements that respondents recalled in variables *AD171* to *AD176* and in *AD181* to *AD186*, respectively.

AD240v – noticed any advertising. This variable summarizes the responses from variables *AD211*, *AD221*, *AD251* and *AD231*, where if a respondent gave a 'yes' response to any of the 4 variables then *AD240v* was also coded as 'yes'. In this binary variable, '0' is no and '1' is yes.

AD759v – noticed any anti-smoking info. This variable summarizes the responses from variables *AD711, AD716, AD721, AD726, AD731, AD736, AD741, AD746* and *AD751*, where if a respondent gave a 'yes' response to any of the 9 variables then *AD759v* was also coded as a 'yes'. In this binary variable, '0' is no and '1' is yes.

AD901v to AD933v (odd numbers) – Tak Nak advertising messages. These variables recode the main Tak Nak messages that respondents identified in *AD136o, AD138o* and *AD140o*. Each variable from *AD901v* to *AD933v* codes a different main Tak Nak message, with responses coded as either '0' for not mentioned or '1' for not mentioned. There is no *AD917v* or *AD923v*.

AD902v to AD934v (even numbers) – medium mentioned for Tak Nak advertising messages. These variables recode the media sources mentioned in *AD137, AD139* and *AD141*. Note that there is no *AD918v* or *AD924v*. The categories for these variables are:

→ 1 = TV
→ 2 = Radio
→ 3 = Posters
→ 4 = Billboards
→ 5 = Newspapers

AD950v – Tak Nak message recall. This variable counts how many Tak Nak messages out of the three variables *AD136o, AD138o* and *AD140o* respondents could recall.

3.3 Beliefs about Quitting

Table 18 summarizes the beliefs about quitting derived variables and Table 19 lists the availability of these variables in the various ITC wave datasets.

Table 18. Beliefs about quitting derived variables.

Variable	Label
BQ146v	Firm date to quit smoking

Table 19. Variable listing for beliefs about quitting derived variables

Variable	4C	BD smokers & non-smokers	BR	BT	CN smokers	CN non-smokers	DE smokers & non-smokers	FR smokers	FR non-smokers	IES	IN smokers and non-smokers	KE	KR	MU smokers	MU non-smokers	MX	NL	NZ	SEA smokers & non-smokers	SEA youth	UY	ZM users	ZM non-users
BQ150v	×				1			×		×			1-2	×									×	

Table 20. Variable listing for beliefs about quitting derived variables in analytic datasets

Variable	4C	BD smokers & non-smokers	BR	BT	CN smokers	CN non-smokers	DE smokers	DE non-smokers	FR smokers	FR non-smokers	IES	IN smokers and non-smokers	KE	KR	MU smokers	MU non-smokers	MX	NL	NZ	SEA smokers & non-smokers	SEA youth	UY	ZM users	ZM non-users
BQ146v	1-8						×							×	×			×						

BQ146v – firm date to quit smoking. This variable is derived for the analytic datasets. *BQ146v* was derived using *BQ141* (intention to quit) and *BQ146* (firm date set for those planning to quit within the next month) in order to include all respondents. If the respondent reported that they plan to quit "in the next 6 months" or "beyond 6 months" (*BQ141*=2 or 3), then *BQ146v* is coded as "no immediate plans to quit". If the respondent is "not planning to quit" (*BQ141*=4), then *BQ146v* is coded as "not planning to quit". If the respondent plans to quit "in the next month" (*BQ141*=1), then the respondent was asked *BQ146*. If the respondent planning to quit within the next month has set a firm quit date (*BQ146*=1; coded as "yes"), then *BQ146v* is coded as "yes". Similarly, if *BQ146*=2 (coded as "no"), then *BQ146v* is coded as "no".

BQ150v – Intention to quit. The variable *BQ150v* is a binary code for respondents' intentions to quit. If the response to *BQ141* indicates any intention to quit (i.e., responses 1 to 3) then *BQ150v* is coded as '1' for 'Yes, intending to quit'; and '0' for 'No, not intending to quit'. Non-response options of NA, DK and Refused are also carried forward.

3.4 Cigarette Brands

BR310v – name of cigarette brand coded. The variable BR310v categorizes the cigarette brands that respondents reported as their usual brand in BR310. Note that BR310v only codes the cigarette brand names; other responses are coded in BR311v. The categories are:

→ 1 = Dunhill
→ 2 = Marlboro
→ 3 = Pall Mall
→ 4 = Sempoema
→ 5 = Winston
→ 6 = Others

Table 20 summarizes the cigarette brand derived variables and Table 21 lists the availability of these variables in each of the ITC wave datasets.

Table 21. Cigarette brand derived variables

Variable	Label
BR302tv	Type of cigarette
BR309v	Usual brand FM or RYO
BR310v	Name of cigarette brand
BR311v	Brand of cigarette usually smoked
BR365v	Number of packs smoked
BR405v	(Derived) tar level of 1st brand
BR502v	How long smoked this brand
BR505v	How long smoked current brand

Table 22. Variable listing for cigarette brand derived variables.

Variable	4C	BD smokers & non-smokers	BR	BT	CN smokers	CN non-smokers	DE smokers	DE non-smokers	FR smokers	FR non-smokers	IES	IN smokers and non-smokers	KE	KR	MU smokers & non-smokers	MX	NL	NZ	SEA smokers & non-smokers	SEA youth	UY	ZM users & non-users
BR302tv																			1			
BR309v	8-10																					
BR310v																				1		
BR311v																				1		
BR365v																						
BR405v					×																	
BR502v	3,4								1													
BR505v	4,5						×		1		3,4			1				2				

In cases where a variable is not present on all waves of a project, only the wave numbers where it is present are listed.

BR302tv – type of cigarette. This variable summarizes the type of cigarettes used into three categories based on *BR301t*– brand smoked most. This variable is applicable only to Thai respondents; hence respondents from Malaysia are always coded as '.' (i.e., system missing). The categories for this variable are:

→ 1 = Thai factory
→ 2 = Import
→ 3 = Hand made

BR309v – usual brand FM or RYO. The variable *BR309v* categorizes whether the usual brand that respondents reported is factory made cigarettes or roll-your-own. It is derived from BR312 (if Country = 1), BR322 (if Country = 2), BR331 (if Country = 3) or BR341 (if Country = 4). The categories are:

→ 1 = Factory-made / packet cigarettes
→ 2 = Roll-your-own cigarettes
→ 3 = Other

BR310v – name of cigarette brand coded. The variable *BR310v* categorizes the cigarette brands that respondents reported as their usual brand in *BR310*. Note that *BR310v* only codes the cigarette brand names; other responses are coded in *BR311v*. The categories are:

→ 1 = Dunhill
→ 2 = Marlboro
→ 3 = Pall Mall

→ 4 = Sempoema
→ 5 = Winston
→ 6 = Others

BR311v – brand of cigarette usually smoked recode. The variable *BR311v* categorizes the responses about usual cigarette brands from *BR311*. Response options are:

1 = factory
2 = no usual brand
3 = smokes hand rolled cigarettes
4 = don't know/can't remember

BR365v – Number of packs smoked per week. The variable *BR365v* is derived from BR363, BR364a and BR364b to represent the packs respondent smoked per week (on average). If respondent reports less than 1 pack per week (BR363 = 0), BR365v is equal to 0. If a range is reported (BR363=76, and BR364a/BR364b are not missing), BR365v is set to be the mean of BR364a and BR364b.

BR405v – Derived tar level of 1st brand. The variable *BR405v* is derived from NR405. If NR405 is in (1, 15), NR405v = NR405. Otherwise, BR405v = 97.

BR502v – how long smoking this brand. The *BR502v* variable recalculates responses to *BR501* into one common measurement of time. For the France smoker wave 1 dataset, *BR502v* measures the length of time respondents have been smoking their current brand of cigarettes in years; in the Four Country datasets it is measured in months. In Four Country, any value greater than 10 years is coded as '997' for '> 10 years'

BR505v – current brand smoked how long. The *BR505v* variable codes responses to *BR501* into a binary variable, where responses are either coded as '1' for 'current brand smoked less than one year' or '2' for 'current brand smoked longer than one year'.

3.5 Cessation Help

Table 22 summarizes the cessation help derived variables and Table 23 lists the availability of these variables in the various ITC wave datasets.

Table 23. Cessation help derived variables.

Variable	Label
CH809v	Doctor's advice substantial, overall
CH811v	Received advice from doctor to quit, overall
CH813v	Referral from doctor to quit, overall
CH815v	Quitting Rx from doctor, overall
CH817v	Quitting pamphlet from doctor, overall
CH819v	Test to assess nicotine dependence, overall
CH821v	Support for having quit, overall
CH823v	Referral from doctor to help stay quit, overall
CH827v	Pamphlet from doctor on how to stay quit, overall
CH872v	Doctor's advice on harm reduction, overall

Table 24. Variable listing for cessation help derived variables.

Variable	4C	BD smokers & non-smokers	BR	BT	CN smokers & non-smokers	DE smokers & non-smokers	FR smokers	FR non-smokers	IES	IN smokers and non-smokers	KE	KR	MU smokers & non-smokers	MX	NL	NZ	SEA smokers & non-smokers	SEA youth	UY	ZM users & non-users
CH809v	9																			
CH811v	5-9						×					2,3				×				
CH813v	5-9						×					2,3				×				
CH815v	5-9						×					2,3				×				
CH817v	5-9						×					2,3				×				
CH819v							×									1				
CH821v	5-9											2,3				×				
CH823v	5-7.5											2,3				×				
CH825v												2,3				×				
CH827v	5-7.5											2,3				×				
CH872v	9																			

In cases where a variable is not present on all waves of a project, only the wave numbers where it is present are listed.

CH809v – Doctor's advice substantial, overall. CH809v is the same as CH809.

CH811v to CH927v (odd numbers) – cessation help, overall. The *CH811* to *CH927* series of variables were only asked of those who reported having visited a doctor or other health professional since the last survey date; hence, those who had no visit were skipped and responses for those variables were set to 'NA'. These variables, *CH811v* to *CH927v*, re-include those who had no doctor visits by changing the 'NA' responses to '2' for 'No'; positive responses were kept as '1' for 'Yes'. Note that the domain NR was the predecessor to CH, and as a result *NR811v* to *NR827v* are essentially the same variables. See also Section *3.10 Nicotine Replacement Therapy*.

CH872v – Doctor's advice substantial, overall. CH872v is the same as CH872.

3.6 Demographics

Table 24 summarizes the demographic derived variables and Table 25 lists the availability of these variables in the various ITC wave datasets. Note that all weight variables, *DE911v* to *DE959v*, are discussed in Section *4.2 Cross-sectional & Longitudinal Weights*.

Table 25. Demographic derived variables.

Variable	Label
DE211sv	Annual household income - coded (MY)
DE211tv	Derived income (TH)
DE212v	Income categories
DE235sv	Occupation recoded (MY)
DE312v	Highest level of education (derived)
DE512v	Ethnicity recoded (CA)
DE572v	Ethnicity recoded (US)
DE612v	Ethnicity recoded (UK)
DE791v	Number of adults living in the household
DE861v	Age of oldest child in home
HHSIZE_GFK	Household Size - GFK
HOUSE_GfK	Housing Type - GfK
ETHNIC_GfK	Race/Ethnicity - GfK
INCOME_GfK	Household income - GfK
INTERNET_GfK	Household internet Acess - GfK
MARITAL_GfK	Marital Status – GfK
T01_GfK	Presence of Household Members - Children 0-1 - GfK
T1317_GfK	Presence of Household Members - Children 13-17 - GfK

Variable	Label
T18OV_GfK	Presence of Household Members - Adults 18+ - GfK
T25_GfK	Presence of Household Members - Children 2-5 - GfK
T612_GfK	Presence of Household Members - Children 6-12 - GfK
WORK_GfK	Current Employment Status - GfK
EDUC_GfK	education (Highest Degree Received) - GfK

Table 26. Variable listing for demographic derived variables.

Variable	4C	BD smokers & non-smokers	BR	BT	CN smokers & non-smokers	DE smokers & non-smokers	FR smokers & non-smokers	IES	IN smokers and non-smokers	KE	KR	MU smokers & non-smokers	MX	NL	NZ	SEA smokers	SEA non-smokers	SEA youth	UY	ZM users & non-users
DE211v																				×
DE211sv																1	1			
DE211tv																1				
DE212v	×	×			×		×	×	×	×		2-3				×				×
DE213v																				×
DE214v																				×
DE235sv																1	1			
DE312v	×	×			×		×	×	×	×		2-3				×				×
DE791v																				×
DE861v	6-9																			
HHSIZE_GFK	9																			
HOUSE_GfK	9																			
ETHNIC_GfK	9																			
INCOME_GfK	9																			
INTERNET_GfK	9																			
MARITAL_GfK	9																			
T01_GfK	9																			

Variable	4C	BD smokers & non-smokers	BR	BT	CN smokers & non-smokers	DE smokers & non-smokers	FR smokers & non-smokers	IES	IN smokers and non-smokers	KE	KR	MU smokers & non-smokers	MX	NL	NZ	SEA smokers	SEA non-smokers	SEA youth	UY	ZM users & non-users
T1317_GfK	9																			
T18OV_GfK	9																			
T25_GfK	9																			
T612_GfK	9																			
WORK_GfK	9																			
EDUC_GfK	9																			

In cases where a variable is not present on all waves of a project, only the wave numbers where it is present are listed.

DE211v – monthly household income per adult household member. DE211v is calculated based on DE211 (household income month) and DE791v (number of adults living in the household).

DE211sv – annual household income – coded (MY). DE211sv categorizes income from *DE211s* from Malaysian respondents. Income is measured in Ringgit Malaysia (RM) and divided into 6 categories:

→ 1 = RM 10,000 or less
→ 2 = RM 10,001 – 20,000
→ 3 = RM 20,001 – 30,000
→ 4 = RM 30,001 – 40,000
→ 5 = RM 40,001 – 50,000
→ 6 = Greater than RM 50,000

DE212tv – derived income (TH). This variable categorizes the income values in *DE211t* from Thai respondents. Values are measured in baht and are divided into 3 categories:

→ 1 = 0 to 70,000 Baht
→ 2 = 70,0001 – 195,749 Baht
→ 3 = Greater than 195,749 Baht

DE212v – income categories. This variable categorizes income into three broad categories: low, moderate, and high. Note that a country-specific income variable is present on all datasets; however, this income standardization variable, *DE212v* is currently only available for some project datasets. Note that re-contacts are asked if their income status has changed since last survey date. If no, then their current income is set to the same as their previous income.

Table 26 summarizes the annual income response options available to respondents from various countries, and the re-coded responses found in *DE212v*.

Table 27. Income Response Options and Coded Responses.

Country	Original Responses	Project	Recoded Response
France	Less than € 18,000 € 18,000 - 35,999 € 36,000 or greater	France	1=Low 2=Moderate 3=High 9=Not stated
Ireland, Scotland & UK	Less than £30,000 £30,000 - 44,999 £45,000 or greater	Ireland/Scotland	1=Low 2=Moderate 3=High 9=Not answered
UK	Less than £15,000 £15,001 - 30,000 £30,001 or greater	Four Country	1=Low 2=Moderate 3=High 9=No answer
Canada, US & Australia	Less than $30,000 $30,000 - 59,999 $60,000 or greater	Four Country	1=Low 2=Moderate 3=High 9=No answer
India*	Less than INR 5,000 INR 5,000 – 15,000 INR 15,000 or greater	India	1=Low 2=Moderate 3=High 9=Not stated
Kenya*	Less than KES 23,670 KES 23,670 – 120,000 KES 120,000 or greater	Kenya	1=Low 2=Moderate 3=High
Mauritius*	Less than MUR 14,999 MUR 15,000 – 29,999 MUR 30,000 or greater	Mauritius	1=Low 2=Moderate 3=High 9=Not stated
Mexico*	Less than $3,000 $3,001 – 8,000 $8,001 or greater	Mexico	1=Low 2=Moderate 3=High 9=Not stated
Uruguay*	Less than $7,000 $7,001 – 30,000	Uruguay	1=Low 2=Moderate

Country	Original Responses	Project	Recoded Response
	$30,001 or greater		3=High
			9=Not stated
Zambia*	Less than $165,000	Zambia	0=Low
	$165,001 – 265,000		1=Moderate
	$265,001 or greater		2=High
			9=Not stated

*DE212v was derived specifically for the analytic dataset and the original response is MONTHLY income rather than ANNUAL income

DE235sv – occupation recoded (MY). This variable recodes the open-ended responses about occupation in *DE235s* from Malaysian respondents. The responses were categorized into 7 groups:

→ 1 = Professional
→ 2 = Administrative
→ 3 = Service
→ 4 = Skilled
→ 5 = Unskilled
→ 6 = Pensioners, Students, Unemployed
→ 7 = Others

DE312v – highest level of education (derived). This variable categorizes education into three broad categories: low, moderate, and high. Note that a country-specific education variable is available on all datasets; however, this education standardization variable *DE312v* is currently only available for some country project wave datasets. Note that re-contacts are asked if their education status has changed since last survey date. If no, then their current education is set to the same as their previous education. Table 27 summarizes the education response options available to respondents from various countries, and the re-coded responses found in *DE312v*.

Table 28. Education Response Options and Coded Responses.

Country	Original Responses	Project	Recoded Response
France	Less than completed high school	France	1=Low
	Completed high school/some university		2=Moderate
	Completed university or post-graduate		3=High
Ireland, Scotland & UK	Secondary/vocational 3 or less	Ireland/ Scotland	1=Low
	College/university (no degree)		2=Moderate
	Completed university or post-graduate		3=High
UK	Secondary/vocational 3 or less	Four Country	1=Low
	College/university (no degree)		2=Moderate
	Completed university or post-graduate		3=High
Canada & US	Completed high school or less	Four Country	1=Low
	Community college/trade/technical school/some university (no degree)		2=Moderate
	Completed university or post-graduate		3=High

Australia	Completed high school or less Technical/trade/some university (no degree) Completed university or post-graduate	Four Country	1=Low 2=Moderate 3=High
Kenya	Illiterate or Nursery/Kindergarten Primary or Post-Primary/Vocational Secondary School and above	Kenya	1= Low 2=Moderate 3=High
India	Illiterate to Middle School class V to VII Secondary School (ITI course, class XII/X or intermediate) Graduate (BA/ BSc/ Diploma etc.) and above	India	1= Low 2=Moderate 3=High
Mauritius*	Less than primary school/ Primary school Form 1-4/ SC completed/ HSC completed/Vocational training University (no degree)/(degree completed)	Mauritius	1=Low 2=Moderate 3=High
Mexico*	Completed middle school or less Some technical/commercial/highschool or some university (incomplete) University (completed) or post graduate	Mexico	1=Low 2=Moderate 3=High
Uruguay*	Completed middle school or less Highschool or technical school Army/police & others, teachers/trainers, or university	Uruguay	1=Low 2=Moderate 3=High
Zambia*	Less than primary school Primary school/post-primary/wocational school Completed secondary school or higher	Zambia	1=Low 2=Moderate 3=High

*DE312v was derived specifically for the analytic dataset

DE512v – ethnicity recoded (CA). The variable is derived for the analytic datasets. Responses from *DE511* to *DE541* were collapsed into a single measure, stored in *DE512v*. If respondents provided more than 1 ethnicity, they were classified as "mixed ethnicity". Otherwise, they were grouped into a single ethnicity. Note that some ethnic groups were collapsed. Chinese, Japanese, Korean, Southeast Asian, and Filipino are classified as "Asian". Any mention of "First Nations" ethnicity (*DE533*) are coded as "First Nations", even if other ethnicities were also reported (under the rationale that in Canada, even if a person is part First Nations, they still have aboriginal benefits). The other categories are White, South Asian, Black, Latin American, Arab, Mixed, and Unknown.

DE572v – ethnicity recoded (US). The variable is derived for the analytic datasets. Responses from *DE551* to *DE571* were collapsed into a single variable, *DE572v*. If more than one ethnicity was provided, the respondent was coded as "mixed ethnicity". Classifications were made to resemble those needed for the NIH reporting requirements. The classifications include White, Black, Hispanic, Asian, Native, Mixed, and Unknown.

DE612v – ethnicity recoded (UK). The variable is derived for the analytic datasets. Similar to the approach used for the US, this variable is derived by collapse *DE611* and *DE616* into one variable, *DE612v*. Asian reflected a broad Asian category (including Chinese). South Asian included

Bangladeshi, Indian, and Pakistani only. Open-ended responses were combined into specific groups, based on the response, where ethnicities that did not fit were coded as "Other". Other categories are White, Black, and Mixed.

DE791v – number of adults living in the household. This variable is derived from the enumeration data and it is based on the total number of adults that was enumerated in each household.

HHSIZE_GFK – Household Size – GfK. This variable is provided by GfK and applied for respondents who were recruited from GfK panel.

*HOUSE_GfK – Housing Type – GfK***.** This variable is provided by GfK and applied for respondents who were recruited from GfK panel. The responses were categorized into 5 groups:

→ 1 = A one-family house detached from any other house
→ 2 = A one-family house attached to one or more houses
→ 3 = A building with 2 or more apartments
→ 4 = A mobile home
→ 5 = Boat, RV, van, etc.

ETHNIC_GfK - Race/Ethnicity – GfK. This variable is provided by GfK and applied for respondents who were recruited from GfK panel. The responses were categorized into 5 groups:

→ 1 = White, Non-Hispanic
→ 2 = Black, Non-Hispanic
→ 3 = Other, Non-Hispanic
→ 4 = Hispanic
→ 5= 2+ Races, Non-Hispanic

INCOME_GfK - Household income – GfK. This variable is provided by GfK and applied for respondents who were recruited from GfK panel. The responses were categorized into 19 groups:

→ 1 = Less than $5,000
→ 2 = $5,000 to $7,499
→ 3 = $7,500 to $9,999
→ 4 = $10,000 to $12,499
→ 5= $12,500 to $14,999
→ 6 = $15,000 to $19,999
→ 7 = $20,000 to $24,999
→ 8 = $25,000 to $29,999
→ 9 = $30,000 to $34,999
→ 10= $35,000 to $39,999
→ 11= $40,000 to $49,999
→ 12= $50,000 to $59,999
→ 13= $60,000 to $74,999
→ 14= $75,000 to $84,999
→ 15= $85,000 to $99,999
→ 16= $100,000 to $124,999
→ 17= $125,000 to $149,999

→ 18= $150,000 to $174,999
→ 19= $175,000 or more

INTERNET_GfK- Household internet Acess – GfK. This variable is provided by GfK and applied for respondents who were recruited from GfK panel.

MARITAL_GfK - Marital Status – GfK. . This variable is provided by GfK and applied for respondents who were recruited from GfK panel. The responses were categorized into 6 groups:

→ 1 = Married
→ 2 = Widowed
→ 3 = Divorced
→ 4 = Separated
→ 5 = Never married
→ 6 = Living with partner

T01_GfK - Presence of Household Members - Children 0-1 – GfK. This variable is provided by GfK and applied for respondents who were recruited from GfK panel.

T1317_GfK - Presence of Household Members - Children 13-17 – GfK. This variable is provided by GfK and applied for respondents who were recruited from GfK panel.

T18OV_GfK - Presence of Household Members – Adults 18+ – GfK. This variable is provided by GfK and applied for respondents who were recruited from GfK panel.

T25_GfK - Presence of Household Members - Children 2-5 - GfK. This variable is provided by GfK and applied for respondents who were recruited from GfK panel.

T612_GfK - Presence of Household Members - Children 6-12 - GfK. This variable is provided by GfK and applied for respondents who were recruited from GfK panel.

WORK_GfK - Current Employment Status - GfK. This variable is provided by GfK and applied for respondents who were recruited from GfK panel. The responses were categorized into 6 groups:

→ 1 = Working - as a paid employee
→ 2 = Working - self-employed
→ 3 = Not working - on temporary layoff from a job
→ 4 = Not working - looking for work
→ 5 = Not working – retired
→ 6 = Not working – disabled
→ 7 = Not working - other

EDUC_GfK - education (Highest Degree Received) - GfK. This variable is provided by GfK and applied for respondents who were recruited from GfK panel. The responses were categorized into 6 groups:

→ 1 = No formal education
→ 2 = 1st, 2nd, 3rd, or 4th grade
→ 3 = 5th or 6th grade

→ 4 = 7th or 8th grade
→ 5 = 9th grade
→ 6 = 10th grade
→ 7 = 11th grade
→ 8 = 12th grade NO DIPLOMA
→ 9 = HIGH SCHOOL GRADUATE - high school DIPLOMA or the equivalent (GED)\
→ 10= Some college, no degree
→ 11= Associate degree
→ 12= Bachelors degree
→ 13= Masters degree
→ 14= Professional or Doctorate degree

DE861v – age of oldest child in home. This variable is derived from the variables that code for the number of children in the home in different age categories (i.e., *DE831, DE836, DE841, DE846* & *DE851*). To create *DE861v*, the oldest age group, 13 to 17 years, is reviewed first. If there are no children in the oldest age group, the next age group is reviewed, and so on. Table 28 summarizes the categories for *DE861v* for the age of oldest child in the home.

Table 29. Age categories for the oldest child in the home.

Category	Age Group	Age
0	No children < 18	n/a
1	Infant	<1
2	Primary	1 to 5
3	Pre-teen	6 to 12
4	Teen	13 to 17

3.7 Environmental Tobacco Smoke/Smoking Restrictions

Table 29 summarizes the environmental tobacco smoke/smoking restrictions derived variables and Table 30 lists the availability of these variables in the various ITC wave datasets.

Table 30. Environmental tobacco smoke/smoking restrictions derived variables.

Variable	Label
ET601v	Work outside the home (all respondents)

Table 31. Variable listing for environmental tobacco smoke derived variables.

Variable	4C	BD smokers & non-smokers	BR	BT	CN smokers & non-smokers	DE smokers & non-smokers	FR smokers & non-smokers	IES	IN smokers & non-smokers	KE	KR	MU smokers & non-smokers	MX	NL	NZ	SEA smokers	SEA non-smokers	SEA youth	UY	ZM users & non-users
ET601v	6-10																			

In cases where a variable is not present on all waves of a project, only the wave numbers where it is present are listed.

ET601v – work outside the home (all respondents). The variable *ET601v* combines the responses for the questions on working outside the home, from *ET601a*, which was asked of the main sample, and *ET601b*, which was asked of replenishment. The response options are:

1 = yes
2 = no
9 = not stated

3.8 Frequency of Smoking

Table 31 summarizes the frequency of smoking derived variables and Table 32 lists the availability of these variables in the various ITC wave datasets.

Table 32. Frequency of smoking derived variables.

Variable	Label
FR112v	Age when first smoked whole cig (derived)
FR230v	How often smoke in LM - coded
FR245v	Cigarettes per day, continuous
FR250v	Cigarettes per day, categorical
FR260v	Heaviness of Smoking Index (HSI)
FR309v	Smoking status
FR310v	Smoking Status
FR333v	Use of factory-made Vs roll-your-own cigarettes

Table 33. Variable listing for frequency of smoking derived variables.

Variable	4C	BD smokers	BD nonsmokers	BR	BT	CN smokers	CN nonsmokers	DE smokers	DE nonsmokers	FR smokers	FR nonsmokers	IES	IN smokers	IN non-smokers	KE	KR	MU smokers	MU	MX	NL	NZ	SEA smokers	SEA	SEA youth	UY	ZM smokers	ZM nonsmokers
FR112v																								1			
FR230v																								1			
FR245v	×			×		×		×		×		×	×		×	×	×			×	×					×	
FR250v	×			×		×		×		×		×	×		×	×	×			×	×					×	
FR260v	×			×		×		×		×		2-4				×	×			×	×					×	
FR309v	×	×		×				×		×		×	×		×	×	×		×	×	×	2			×	×	
FR310v												3-4									2						
FR333v	×					×				×		×				×					×						

In cases where a variable is not present on all waves of a project, only the wave numbers where it is present are listed.

FR112v – age when first smoked whole cigarette (derived). This variable recodes the age when first smoked a whole cigarette from *FR113* into 5 categories:

→ 1 = Less than 13 years
→ 2 = 13 years
→ 3 = 14 years
→ 4 = 15 years
→ 5 = 16 years or older

FR230v – how often smoke in last month - coded. This variable recodes the frequency of smoking in the last month as provided in *FR230* into a binary variable:

1 = not at all
2 = smoked at least once in last month

FR245v – cigarettes per day, continuous. The variable *FR254v* is a continuous measure of the number of cigarettes respondents smoke per day. If the respondent is a daily smoker (from *FR309v*), then cigarettes per day (CPD) is equal to the stated number of cigarettes consumed per day. If the respondent is a weekly smoker, then CPD is equal to stated number of cigarettes consumed per week divided by 7. If the respondent is a monthly smoker, then CPD is equal to stated number of cigarettes consumed per month divided by 30.4. If the respondent has quit smoking or the smoking status is NA, then CPD is classified as '777', (Not Applicable). If smoking status is Refused then CPD is classified as '888' (Refused). If smoking status is DK then CPD is classified as '999' (Don't Know).

FR250v – cigarettes per day, categorical. The variable *FR250v* derives the CPD categories from *FR245v*, continuous CPD. Continuous CPD is rounded to nearest integer and categorized as:

→ 0 = 1-10 cigarettes per day

→ 1 = 11 – 20 cigarettes per day
→ 2 = 21-30 cigarettes per day
→ 3 = 31+ cigarettes per day

FR260v – heaviness of smoking index (HSI). The Heaviness of Smoking Index in *FR260v* is calculated by summing the value of the categorical CPD (*FR250v*) and categorical time to first cigarette (*SB013v*). Both *FR250v* and *SB013v* have category values from 0 to 3, which translates to *FR260v* having values ranging from 0 to 6. If either value is missing or coded as a non-response, then HSI is also classified as missing or non-response.

FR309v – smoking status. For new recruits, smoking status is ascertained in the first interview. First respondents are asked if they smoke daily (*FR211)*. If yes, they are coded as daily smokers for *FR309v*; if not, they are asked if they smoke at least once per week (*FR221)*. If yes, they are coded as weekly smokers; if not, they are asked if they smoke at least once per month (*FR231)*. If yes, they are coded as monthly smokers; if not, they cannot be included in the survey. For recontacted respondents, smoking status is reviewed at every wave. If respondents were smokers at the last wave (i.e., they had not quit), then they are asked if they still smoke as often as last wave (*FR301)*. If they answer yes, then their smoking status is carried forward from the previous wave. If they answer no, then their new smoking status is ascertained (*FR306)*. If respondents had quit smoking last wave, they are asked if they are still quit (*QA337)*, and if not then their new smoking status is ascertained (*FR307)*. The smoking frequency categories for *FR309v* are:

→ 1 = Daily smoker
→ 2 = Weekly smoker
→ 3 = Monthly smoker
→ 4 = Quit for less than 1 month
→ 5 = Quit for 1 to 6 months
→ 6 = Quit for more than 6 months
→ 7 = Quit at LSD and ever since

FR310v - smoking status. This variable categorizes all respondents as either:

1 = smokers
2 = quitters
3 = non-smokers

FR333v – use of factory-made versus roll-your-own cigarettes. This variable recodes the responses to *FR326*, which asks if they smoke factory-made, roll-your-own (RYO) or both types of cigarettes, and *FR331*, which asks for every ten cigarettes smoked, how many are factory-made and RYO. If respondents reported in *FR326* that they only smoke factory-made, then *FR333v* is coded as '1', only factory-made, and similarly, if they reported only smoking RYO then *FR333v* is coded as '5', only RYO. If respondents reported in *FR326* that they smoke both, then *FR333v* is calculated using responses to *FR331*. If respondents reported smoking 0 to 3 RYO out of ten cigarettes then *FR333v* was coded as '2', mainly factory-made. If respondents reported smoking 4 to 6 RYO out of ten cigarettes then *FR333v* was coded as '3', factory-made and RYO are similar. If respondents reported smoking 7 to 10 RYO out of ten cigarettes, then *FR333v* was coded as '4', mainly RYO. Note that extreme values in *FR331* (i.e., 0 and 10) are taken at face value, given that

it is plausible to smoke one type so rarely that out of 10 cigarettes, it would average out to 0, or 10.

3.9 Knowledge of Health Effects/Tobacco Constituents

Table 33 summarizes the knowledge of health effects/tobacco constituents derived variables and Table 34 lists the availability of these variables in the various ITC wave datasets.

Table 34. Knowledge of health effects/tobacco constituents derived variables.

Variable	Label
KN101v	Smoking causes lung decay (binary recode)
KN102v	Smoking causes stained teeth (binary recode)
KN103v	Smoking causes premature aging (binary recode)
KN221v	Smoking causes stroke (binary recode)
KN231v	Smoking causes impotence (binary recode)
KN241v	Smoking causes lung cancer in smokers (binary recode)
KN251v	Smoking causes lung cancer in nonsmokers (binary recode)
KN256v	Smoking causes mouth cancer (binary recode)
KN261v	Knowledge of health effects (score)

Table 35. Variable listing for knowledge derived variables.

Variable	4C	BD smokers & non-smokers	BR	BT	CN smokers & non-smokers	DE smokers & non-smokers	FR smokers & non-smokers	IES	IN smokers and non-smokers	KE	KR	MU smoker & non-smokers	MX	NL	NZ	SEA smokers & non-smokers	SEA youth	UY	ZM smokers & non-smokers
KN101v																1-2			
KN102v																1-2	1		
KN103v																1-2	1		
KN221v																1-2			
KN231v																1-2			

Variable	4C	BD smokers & non-smokers	BR	BT	CN smokers & non-smokers	DE smokers & non-smokers	FR smokers & non-smokers	IES	IN smokers and non-smokers	KE	KR	MU smoker & non-smokers	MX	NL	NZ	SEA smokers & non-smokers	SEA youth	UY	ZM smokers & non-smokers
KN241v																1-2	1		
KN251v																1-2	1		
KN256v																2			
KN261v																1-2	1		

In cases where a variable is not present on all waves of a project,
only the wave numbers where it is present are listed.

KN101v, KN102v, KN103v, KN221v, KN231v, KN241v, KN251v, KN256v – binary recode of knowledge of health effects. These variables recode the corresponding variables from *KN101* to *KN256*, which were originally coded as 'yes', 'no', and don't know. These derived variables dichotomize responses to 'yes' and 'no/don't know' by collapsing 'no' and 'don't know' responses from the original variables.

KN261v – knowledge of health effects (score). This variable sums the number of 'yes' responses from *KN101* to *KN256.*

3.10 Nicotine Replacement Therapy

Table 35 summarizes the nicotine replacement therapy derived variables and Table 36 lists the availability of these variables in the various ITC wave datasets.

Table 36. Nicotine replacement therapy derived variables.

Variable	Label
NR001v	Number of NRT medications used over 1 year ago
NR002v	Number of NRT medications used in last year
NR112v	Number of medications used since last survey date/in last 6 months (counter)
NR202v	Number of medications used at the same time (counter)
NR302v	Medication used recently in last 6 months
NR307v	Duration using NRT products (days)
NR811v	Received advice from doctor to quit, overall

Variable	Label
NR813v	Referral from doctor to quit, overall
NR815v	Quitting prescription from doctor, overall
NR817v	Quitting pamphlet from doctor to quit, overall
NR821v	Support for having quit, overall
NR823v	Referral from doctor to help stay quit, overall
NR825v	Stop-smoking medication prescription, overall
NR827v	Pamphlet from doctor on how to stay quit, overall

Table 37. Variable listing for nicotine replacement therapy derived variables.

Variable	4C	BD smokers & non-smokers	BR	BT	CN smokers	CN non-smokers	DE smokers & non-smokers	FR smokers & non-smokers	IES	IN smokers and non-smokers	KE	KR	MU smokers	MU non-smokers	MX	NL	NZ	SEA smokers	SEA non-smokers	SEA youth	UY	ZM smokers	ZM non-smokers
NR001v					×																		
NR002v					×																		
NR112v	1-4								×			×											
NR202v	1																						
NR302v									2														
NR307v									3,4														
NR811v	1-3				×				×			×											
NR813v	1-4				×				×			×											
NR815v	1,2,4								×			×											
NR817v	1-3				×				×			×											
NR821v	3,4											×											
NR823v	3,4								3,4			×											
NR825v									3,4			×											
NR827v	3,4								3,4			×											

In cases where a variable is not present on all waves of a project, only the wave numbers where it is present are listed.

NR001v – number of NRT meds used over 1 year ago. This variable sums the number of NRT medications cited in *NR111, NR113, NR116, NR108, NR109, NR119* and *NR119o* that were used more than one year ago.

NR002v – number of NRT meds used in last year. This variable sums the number of NRT medications cited in *NR111, NR113, NR116, NR108, NR109, NR119* and *NR119o* that were used in the past year

NR112v – number of medications used since last survey date/in last 6 months (counter). This variable sums the number of NRT medications cited in *NR121* to *NR141* as having been used in the last 6 months or since last survey date.

NR202v – number of medications used at the same time (counter). This variable counts the number of NRT medications cited in *NR221* to *NR241* as having been used at the same time.

NR302v – medication used recently in last 6 months. This variable categorizes most recently used NRT medications as cited in N*R121* to *NR141* if only one medication was mentioned in these variables; otherwise, it is the NRT medications cited in *NR301(which med. Used most recently).* The categories are:

→ 1 = Nicotine water
→ 2 = Nicotine gum
→ 3 = Nicotine patch
→ 4 = Nicotine lozenge
→ 5 = Nicotine tablets
→ 6 = Nicotine inhaler
→ 7 = Nicotine nasal spray
→ 8 = Zyban
→ 9 = Wellbutrin
→ 10 = Other
→ 77 = NA
→ 88 = Refused
→ 99 = Don't know

NR307v – duration using NRT products (days). In the variable *NR307*, respondents could provide a time value measured in days, weeks, or months. This variable recalculates these responses provided in *NR307* into time measured in days.

NR811v to NR827v (odd numbers) – cessation help usage, overall. Recall that the domain NR is the predecessor to domains CH (cessation help) and SM (stop-smoking medications). For the variables *NR811v* to *NR827v*, see Section *3.5 Cessation Help*. In the analytic datasets, these variables are derived slightly differently than described in Section *3.5 Cessation Help*. In the analytic datasets, *NR801* (visited a doctor or health professional) was combined with *NR811*, *NR813*, *NR815*, and *NR817*, respectively, such that *NR811v* to *NR817v* represent all respondents, taking on the following values: 1 = yes; 2 = no; 3 = did not visit doctor (i.e., *NR801*="no"). (Note: In the regular derivation, those who did not visit a doctor were coded as 'no' rather than 'did not visit doctor').

3.11 Purchases of Tobacco

Table 37 summarizes the demographic derived variables and Table 38 lists the availability of these variables in the various ITC wave datasets.

Table 38. Purchases of tobacco derived variables.

Variable	Label
PU201v	Reported price for cigarettes is for...
PU221v	Packs per carton (derived)
PU222v	Number of packs per carton (categorical, derived)
PU321v	Cigarettes per pack (continuous, derived)
PU322v	Number of cigarettes per pack (categorical, derived)
PU555v	Price per unit of tobacco purchase
CPDIR	Affordability index (or cigarette price to daily income ratio)
pricePerCig	Price paid per cigarette/stick
PU601v	Total number of cartons purchased (derived)
PU603v	Total number of packs purchased (derived)
PU605v	Total number of cigarettes per pack (derived)
PU607v	Total number of cigarettes purchased (derived)
PU621v	Price paid for total purchase in local currency units (derived)
PU623v	Price paid per cigarette/stick in local currency units (derived)
PU625v	Price paid for total purchase in US dollars, purchasing power parity (derived)
PU627v	Price paid per cigarette/stick in US dollars, purchasing power parity (derived)
PU701v	Total number of pouches purchased (RYO only; derived)
PU721v	Price paid for total pouch purchase in local currency units (RYO only; derived)
PU723v	Price paid per pouch in local currency units (RYO only; derived)
PU725v	Price paid for total pouch purchase in US dollars, purchasing power parity (RYO only; derived)
PU727v	Price paid per pouch in US dollars, purchasing power parity (RYO only; derived)

Table 39. Variable listing for purchases of tobacco derived variables in analytic datasets.

Variable	4C	BD smokers & non-smokers	BR	BT	CN smokers & non-smokers	DE smokers & non-smokers	FR smokers & non-smokers	IES	IN smokers and non-smokers	KE	KR	MU smokers	MU non-smokers	MX	NL	NZ	SEA smokers	SEA non-smokers	SEA youth	UY	ZM smokers	ZM non-smokers
PU201v	6-10																					
PU221v	1,2							1														
PU222v	3																					
PU321v								1														
PU321v	1,2																					
PU322v	3																					
PU555v	5-10										×											

In cases where a variable is not present on all waves of a project, only the wave numbers where it is present are listed.

Table 40a. Variable listing for purchases of tobacco derived variables.

Variable	4C	BD smokers & non-smokers	BR	BT	CN smokers & non-smokers	DE smokers	DE non-smokers	FR smokers & non-smokers	IES	IN smokers and non-smokers	KE	KR	MU smokers	MU non-smokers	MX	NL	NZ	SEA smokers	SEA non-smokers	SEA youth	UY	ZM smokers	ZM non-smokers
CPDIR																							
pricePerCig																							
PU601v	1-8		2			×					×	×	×			×							
PU603v	1-8		2			×					×	×	×			×							
PU605v	1-8		2			×					×	×	×			×							
PU607v	1-8		2			×					×	×	×			×							
PU621v	1-8		2			×					×	×	×			×							
PU623v	1-8		2			×					×	×	×			×							
PU625v	1-8		2			×					×	×	×			×							
PU627v	1-8		2			×					×	×	×			×							
PU701v	1-8					×					×					×							

Variable	4C	BD smokers & non-smokers	BR	BT	CN smokers & non-smokers	DE smokers	DE non-smokers	FR smokers & non-smokers	IES	IN smokers and non-smokers	KE	KR	MU smokers	MU non-smokers	MX	NL	NZ	SEA smokers	SEA non-smokers	SEA youth	UY	ZM smokers	ZM non-smokers
PU721v	1-8					×					×					×							
PU723v	1-8					×					×					×							
PU725v	1-8					×					×					×							
PU727v	1-8					×					×					×							

In cases where a variable is not present on all waves of a project, only the wave numbers where it is present are listed.

PU201v – reported price for cigarettes is for… This variable provides the form in which the cigarettes were purchased (e.g., loose, pack, carton, RYO tobacco, etc.). It complements *PU555v* by providing the units for the price per unit of tobacco purchase. The categories are:

→ 1 = Carton
→ 2 = Pack
→ 3 = Loose
→ 4 = RYO tobacco
→ 5 = Respondent has quit
→ 9 = Price not reported

PU221v – number of packs per carton (derived). This continuous variable combines responses to *PU221a* ('How many packs of cigarettes were in the carton?') and *PU222b* ('How many packs of cigarettes were in each carton?').

PU222v – number of packs per carton (categorical, derived). The categorical responses from *PU222a* ('How many packs of cigarettes were in the carton?') and *PU222b* ('How many packs of cigarettes were in each carton?') were combined in *PU222v*. The categories are:

→ 1 = Four
→ 2 = Five
→ 3 = Six
→ 4 = Seven
→ 5 = Eight
→ 6 = Ten
→ 7 = Twenty
→ 8 = Other
→ 77 = NA
→ 88 = Refused
→ 99 = Don't know

PU321v – cigarettes per pack (continuous, derived). The continuous responses from *PU321a* ('How many cigarettes were in the pack?') and *PU321b* ('How many cigarettes were in each pack?') were combined in *PU321v.*

PU322v – number of cigarettes per pack (categorical, derived). The categorical responses from *PU322a* ('How many cigarettes were in the pack?') and *PU322b* ('How many cigarettes were in each pack?') were combined in *PU322v*.

→ 1 = Ten
→ 2 = Twenty
→ 3 = Twenty-five
→ 4 = Thirty
→ 5 = Thirty-five
→ 6 = Forty
→ 7 = Fifty
→ 8 = Other, specify
→ 9 = Don't know
→ 77 = NA
→ 88 = Refused
→ 99 = DK

PU555v - price per unit of tobacco purchase. This variable derives the price per unit of tobacco purchase regardless of the packaging. *PU555v* equals the price for 1 unit, if provided (in *PU231, PU331, PU431,* or *PU531*). Or, if the respondent gave total price for all units, *PU555v* is the total price divided by the number of units.

CPDIR – affordability index (or cigarette price to daily income ratio). The variable *CPDIR* was derived for the *Tobacco Price and Taxation: ITC Cross-Country Comparison Report*[1]. This variable is derived using three measures: self-reported price paid for cigarettes (measured as price paid per cigarette/stick; see the derived variable *pricePerCig*); daily income, based on the annual or monthly household income reported by respondents; and average daily dose of cigarettes smoked per day, using the derived variable FR245v (see section *3.8 Frequency of Smoking*). In most countries, income is collected as a categorical variable. In these cases, the imputed income was taken to be the midpoint of each range with the exception of the upper endpoint, where the low end of the range was used. The affordability index is then computed as:

CPDIR = price per dose / daily income

where:

price per dose = cigarettes smoked per day * price per cigarette
daily income = annual income/365.24 OR = monthly income/30.4

Note that RYO smokers are not included in the computation of CPDIR due to the difficulty in assessing price per dose, and that unreasonable values are set to missing (i.e. CPDIR > 1 which implies that respondents pay more for their daily dose of cigarettes than the typical amount of money they would earn in one day).

[1] ITC Project (March 2012). *Tobacco Price and Taxation: ITC Cross-Country Comparison Report*. University of Waterloo, Waterloo, Ontario, Canada.

pricePerCig – price paid per cigarette/stick. This variable is used in the calculation of CPDIR and derives the price per cigarette/stick regardless of whether the price reported was given as per total purchase (e.g. price for all cartons, packs etc.) or per unit purchased (e.g. price per carton, pack etc.). Where the price reported was given as per total purchase, carton, pack etc., the number of cartons purchased, packs per carton and cigarettes per pack etc. was also reported in order to determine the total number of cigarettes purchased. Using this information, the price paid per cigarette can be derived by dividing the reported price per total/unit purchase by the total number of cigarettes purchased.

PU601v – total number of cartons purchased (derived). This variable is derived for the analytic datasets. It is derived by combining *PU201* (unit purchased, i.e., carton, pack, single etc.) and *PU211* (number of cartons purchased). If the respondent reports that their last unit of purchase (*PU201*) was a carton, then *PU601v = PU601v*. Otherwise, *PU601v* is coded as 'not applicable'.

PU603v – total number of packs purchased (derived). This variable is derived for the analytic datasets. *PU603v* is derived using *PU201* (unit purchased, i.e., carton, pack, single), *PU222* (number of packs per carton) and *PU311* (number of packs purchased). If the respondent reports that their last unit of purchase (*PU201*) was a carton, then *PU603v = PU222*. If the respondent reports that their last unit of purchase (*PU201*) was in packs, *PU603v = PU311*. Otherwise, *PU603v* is coded as 'not applicable'.

PU605v – total number of cigarettes per pack (derived). This variable is derived for the analytic datasets. For the ITC-4 Country analytic datasets, *PU226* (number of cigarettes per pack) is an open-ended variable, thus *PU605v* is equal to *PU226*. For the MU analytic datasets, *PU226* and *PU322* (number of cigarettes per pack) had the response options, 1=10 cigarettes per pack and 2= 20 cigarettes per pack. So if *PU201* (unit purchased, i.e., carton, pack, single) was reported as 'carton', *PU605v* is equal to 10 multiplied by *PU226*. Similarly, if *PU201* (unit purchased) was reported as 'pack', *PU605v* is equal to 10 multiplied by *PU322*. If *PU201* (unit purchased) is reported as 'single, loose', then *PU605v* is equal to *PU411* (number of loose cigarettes purchased).

PU607v – total number of cigarettes purchased (derived). This variable is derived for the analytic datasets. If *PU201* (unit purchased, i.e., carton, pack, single) is reported as 'carton', then *PU607v* is equal to the product of *PU601v* (number of cartons purchased), *PU603v* (number of packs per carton) and *PU605v* (number of cigarettes per pack). If *PU201* (unit purchased) is reported as 'pack', then *PU607v* is equal to the product of *PU603v* (number of packs per carton) and *PU605v* (number of cigarettes per pack). If *PU201* (unit purchased) is reported as 'single, loose', then *PU607v* is equal to *PU411* (number of loose cigarettes purchased).

PU621v – price paid for total purchase in local currency units (derived). This variable is derived for the analytic datasets. For the ITC-4 Country analytic dataset, if *PU201* (unit purchased, i.e., carton, pack, single) is reported as 'carton', *PU621v* is equal to *PU231a* (cost per carton) if *PU211* (number of cartons purchased) is equal to 1. *PU621v* is equal to *PU211*PU231b* if the number of cartons purchased is greater than one. Otherwise, *PU621v=PU241* (amount paid). This variable is derived in the same way if *PU201* is reported as 'packs' or 'singles, loose'. The variables *PU311* (number of packs purchased), *PU331a*, *PU331b* (cost per pack), and *PU341* (amount paid) are used if respondent bought packs, and *PU411* (number of cigarettes purchased), *PU431a*, *PU431b* (cost per cigarette), and *PU441* (amount paid) are used if the respondent bought single/loose

cigarettes. For the MU analytic dataset, if *PU201* is reported as 'carton', then *PU621v* is equal to the product of *PU231* (cost of each carton) and *PU601v* (total number of cartons purchased). If *PU201* is reported as 'pack', then *PU621v* is equal to the product of *PU331* (cost of each pack) and *PU311* (number of packs purchased). In the MU Wave 1 survey, the price was not asked for those respondents who reported purchasing single/loose cigarettes, so *PU621v* was set to missing.

PU623v – price paid per cigarette/stick in local currency units (derived). This variable is derived for the analytic datasets by dividing *PU621v* (total price paid in local currency units) by *PU607v* (total number of cigarettes purchased).

PU625v – price paid for total purchase in US dollars, purchasing power parity (derived). This variable is derived for the analytic datasets. *PU625v* is derived by converting the reported local price paid for last total purchase (regardless whether purchase was cartons, packs, single cigarettes, etc.) into US dollars. This is done by dividing *PU621v* by the purchasing power parities conversion factor.

PU627v – price paid per cigarette/stick in US dollars, purchasing power parity (derived). This variable is derived for the analytic datasets and derives the price per cigarette/stick in US dollars regardless of whether the price reported was given as per total purchase (e.g. price for all cartons, packs etc.) or per unit purchased (e.g. price per carton, pack etc.). This is done by dividing *PU623v* by the purchasing power parities conversion factor.

PU701v – total number of pouches purchased (RYO only; derived). This variable is derived for the analytic datasets. *PU701v* is equal to *PU511* (number of pouches purchased). If PU511 is coded as 'not applicable', 'refused', or 'don't know', *PU701v* was coded likewise.

PU721v – price paid for total pouch purchase in local currency units (RYO only; derived). This variable is derived for the analytic datasets. For the ITC-4 Country analytic dataset, *PU721v* is equal to *PU521a* (price for one pouch) if *PU511* (number of pouches purchased) is equal to 1. *PU721v* is equal to *PU701v*PU521b* if the number of pouches purchased is greater than one. Otherwise, *PU721v=PU541* (amount paid for all pouches).

PU723v – price paid per pouch in local currency units (RYO only; derived). This variable is derived for the analytic datasets by dividing *PU721v* (total price paid in local currency units) by *PU701v* (total number of pouches purchased).

PU725v – price paid for total pouch purchase in US dollars, purchasing power parity (RYO only; derived). This variable is derived for the analytic datasets. *PU725v* is derived by dividing *PU721v* by the purchasing power parities conversion factor.

PU727v – price paid per pouch in US dollars, purchasing power parity (RYO only; derived). This variable is derived for the analytic datasets and derives the price per pouch in US dollars by dividing *PU723v* by the purchasing power parities conversion factor.

3.12 Quit Attempts

Table 39 summarizes the quit attempts derived variables and Table 40 lists the availability of these variables in the various ITC wave datasets.

Table 41. Quit attempts derived variables.

Variable	Label
QA106v	Number of previous quit attempts
QA108v	Any attempts to quit in past year (derived)
QA111v	Longest time off smoking (derived)
QA112v	Longest time off smoking, measured in days (derived)
QA231v	Days since end of most recent quit attempt (derived)
QA232v	Time since end of most recent quit attempt (composite, months)
QA235v	Time smokefree on most recent quit attempt (composite, months)
QA236v	Duration smoke-free (in days)
QA239v	Past year quit attempts
QA242v	Longest time smoke free (derived, in days)
QA442v	Days since start of most recent quit attempt
QA446v	Derived – most recent QA(number of days)
QA447v	Derived – most recent QA(number of years)
QA514v	Number of days smoke-free covering last survey date (derived)
QA553v	Longest time smoke-free since last survey date (days)
QA662v	Number of days since quit attempt ended

QA106v – number of previous quit attempts. This variable is derived for the analytic datasets. This variable combines *QA101* (ever tried to quit) and *QA106* (number of times ever tried to quit). If the respondent answers "no" to ever having tried to quit smoking (*QA101*) then *QA106v* is equal to 0. Otherwise, *QA106v* equals the number of times given in *QA106*.

QA108v – any attempts to quit in past year (derived). This variable is derived from *QA232v*, which is a composite variable that measures the number of months since the most recent quit attempt ended. If the number of months in *QA232v* is 12 or less, then *QA108v* is coded as '1' for 'any attempt'; otherwise, *QA108v* is coded as '0' for 'no attempt'.

QA111v – longest time off smoking (derived). In the *QA111* variables, respondents could provide a time value measured in hours, days, weeks, months or years. This variable recalculates these values from *QA111a* to *QA111e* and converts the responses into time measured in days. Responses in hours are divided by 24, responses in weeks are multiplied by 7, months are multiplied by 30.4 and years are multiplied by 365.24; the recalculated values are rounded to the nearest integer.

QA112v – longest time off smoking, measured in days (derived). This variable is derived for the analytic datasets. *QA112v* recalculates the responses from *QA111a* to *QA111e*, where respondents can choose the time units used for their report of longest time smoke-free (hours, days, weeks, months, and years, respectively) into days such that there is one common measurement of time.

QA231v – days since end of most recent quit attempt (derived). In the *QA231* variables, respondents could provide a time value measured in days, months or years. This variable recalculates these responses provided in *QA231a* to *QA231c* into time measured in days. Responses in months are multiplied by 30.4 and responses in years are multiplied by 365.24; the recalculated values are rounded to the nearest integer.

QA232v – time since end of most recent quit attempt (composite, months). Like *QA231v*, this variable also recalculates the responses provided in *QA231a* to *QA231c*, but it recalculates values into months rather than days.

Table 42. Variable listing for quit attempts derived variables.

Variable	4C	BD smokers and non-smokers	BR	BT	CN smokers and non-smokers	DE smokers	DE non-smokers	FR smokers	FR non-smokers	IES	IN smokers and non-smokers	KE	KR	MU smokers	MU non-smokers	MX	NL	NZ	SEA smokers	SEA non--smokers	SEA youth	UY	ZM smokers	ZM non-smokers
QA108v			×					×																
QA111v	5-10																							
QA231v	5-10		×			1							×	×			×						×	
QA232v								×																
QA235v			×			1		×				×					×	×						
QA236v	5-10													×					1	1				
QA241v												×												
QA242v														×										
QA358v											2													
QA658v											2													
QA442v	4-10					2				3-4														
QA446v																		×						
QA447v																		2						
QA514v	5-10																							
QA553v	4-8.5					2				3-4														

Variable	4C	BD smokers and non-smokers	BR	BT	CN smokers and non-smokers	DE smokers	DE non-smokers	FR smokers	FR non-smokers	IES	IN smokers and non-smokers	KE	KR	MU smokers	MU non-smokers	MX	NL	NZ	SEA smokers	SEA non--smokers	SEA youth	UY	ZM smokers	ZM non-smokers
QA662v	5-10					2				3-4								×						

In cases where a variable is not present on all waves of a project, only the wave numbers where it is present are listed.

QA235v – time smoke-free on most recent quit attempt (composite, months). In the *QA235b* to *QA235e* variables, respondents could provide a time value for how long they remained smoke-free on their most recent quit attempt, measured in days, weeks, months or years. *QA235v* recalculates these responses into months.

QA236v – duration smoke-free (in days). This variable is similar to *QA235v*. In *QA235a* to *QA235d*, respondents could provide a time value for how long they remained smoke-free on their most recent quit attempt, measured in hours, days, weeks or months. *QA236v* recalculates these responses into days.

QA239v – past year quit attempts. This variable was derived because it has often been used in the analysis of ITC data (both academic papers and national reports). In Wave 1 of the Four Country Survey, *QA239v* is based on *QA232v* (number of days since last quit attempt). If *QA232v* is between 0 and 365 (days) or if respondents had quit between the recruitment survey and the main survey in Wave 1, *QA239v* is set equal to 1.

QA241v, QA242v – Longest time smoke free (derived, in days). In the *QA241* variables, respondents could provide a time value measured in hours, days, weeks and months. This variable recalculates these values from *QA241a* to *QA241d* and converts the responses into time measured in days. Responses in hours are divided by 24, responses in weeks are multiplied by 7 and months are multiplied by 30.4; the recalculated values are rounded to the nearest integer.

QA358v - any attempts to quit smoked tobacco in the last year (derived).

QA658v - any attempts to quit smokeless tobacco in the last year (derived).

QA442v – days since start of most recent quit attempt. In variables *QA441a* to *QA441c* respondents can provide a time value for the length of time since their most recent quit attempt began, measured in days, weeks or months. Respondents can also provide an exact date that they started their most recent quit attempt in variables *QA441d* to *QA441f*. *QA442v* recalculates any of these values into

a number of days since start of most recent quit attempt. If a date is provided in *QA441d* to *QA441f,* then the survey date is subtracted from this date to obtain *QA442v.* If no date is provided, then the time values from *QA441a* to *QA441c* are used to calculate *QA442v* in days.

QA446v – most recent quit attempt (in days). The variable *QA446v* is similar to *QA442v*. In variables *QA445a* to *QA451c* respondents can provide a time value for the length of time since their most recent quit attempt began, measured in days, weeks or months. Respondents can also provide an exact date that they started their most recent quit attempt in variables *QA445d* to *QA445f*. *QA446v and QA447v* recalculates any of these values into a number of days since start of most recent quit attempt. If a date is provided in *QA445d* to *QA445f,* then the survey date is subtracted from this date to obtain *QA446v and QA447v.* If no date is provided, then the time values from *QA445a* to *QA445c* are used to calculate *QA446v* in days and *QA447v* in years.

QA514v – number of days smoke-free covering last survey date (derived). In variables *QA513a* to *QA513d,* respondents can provide a time value for the length of time they spent smoke-free spanning the last survey date, measured in hours, days, weeks or months. *QA514v* recalculates these responses into days.

QA553v – longest time smoke-free since last survey date (days). Respondents are only asked *QA553* if they respond 'yes' to *QA448,* which asks whether they have been smoke-free longer since the last survey date or than either values in *QA236v* or *QA514v*. In variables *QA553a* to *QA553d,* respondents provide the longest period of time they spent smoke-free since last survey date, in hours, days, weeks, or months. *QA553v* recalculates the responses in *QA551a* to *QA551d* into days; note that if the time value was less than 24 hours, then the value was not rounded up - in that case, *QA553v* was set to 0 days. If the respondent was not asked *QA551,* then *QA553v* selects the longest smoke-free time period of either *QA236v* or *QA514v.*

QA662v – number of days since quit attempt ended. In variables *QA661a* to *QA661c* respondents can provide a time value for the length of time since their most recent quit attempt ended, measured in days, weeks or months. Respondents can also provide an exact date that the quit attempt ended in variables *QA661d* to *QA661f*. *QA662v* recalculates any of these values into number of days since end of most recent quit attempt. If a date is provided in *QA661d* to *QA661f,* then the survey date is subtracted from this date to obtain *QA662v*. If no date is provided, then the time values from *QA661a* to *QA661c* are used to calculate *QA662v* in days.

3.13 Smoking Behaviour

Table 41 summarizes the smoking behaviour derived variables and Table 42 lists the availability of these variables in the various ITC wave datasets.

Table 43. Smoking behaviour derived variables.

Variable	Label
SB012v	Time to first cigarette in minutes (derived, continuous)
SB013v	Time to first cigarette in minutes (derived, categorical)

Variable	Label
SB021v	Total minutes to first cigarette after waking -- Daily Smoker
SB026v	Total minutes to first cigarette after waking -- Non-daily Smoker
SB226v	Overall frequency of butting out before finishing

Table 44. Variable listing for smoking behaviour derived variables.

Variable	4C	BD smokers and non-smokers	BR	BT	CN smokers	CN non-smokers	DE smokers	DE non-smokers	FR smokers	FR non-smokers	IES	IN smokers and non-smokers	KE	KR	MU smokers	MU non-smokers	MX	NL	NZ	SEA smokers non-smokers	SEA youth	UY	ZM smokers	ZM non-smokers
SB012v	×		×				×		×		×			×	×			×	×					
SB013v	×		×		×		×		×		×			×				×	×				×	
SB226v	×								×		×			×					×					

SB012v – time to first cigarette in minutes (derived, continuous). This variable provides a continuous value in minutes for the length of time between waking and smoking a first cigarette. For daily smokers (FR309v=1), this value is yielded from variables *SB021a* and *SB021b* which provide time values in minutes or hours. For weekly or monthly smokers, (FR309v=2 or 3) then this values is yielded from variables *SB026a* and *SB026b*, which have a slightly differently worded survey question, and also provide time values in minutes or hours. *SB012v* recalculates the values from either *SB021* or *SB026* into minutes. If values were provided in hours, a value in minutes was acquired by multiplying the hour value by 60 and rounding to the nearest minute.

In some cases, it was difficult to determine time to first cigarette because values were entered for *both* time in minutes and time in hours. In those cases, time to first cigarette was determined as follows:

- If respondents did not know how long they wait to their first cigarette after waking, then time to first cigarette was set to 999 (“Don’t Know”).
- If both time in minutes and time in hours had non-zero, non-NA values, time to first cigarette was set to missing. This was done to prevent the possibility of misclassification based on whether minutes or hours were used. (For example, if a respondent reported time in minutes as 15 and time in hours as 0.5, different values for the categorical measure of time to first cigarette would result).

- If both time in minutes and time in hours had non-zero, non-NA values but time in hours was greater than 24, then time in minutes was accepted as the value for time to first cigarette.
- If both time in minutes and time in hours had non-zero, non-NA values *and* there was no possibility for misclassification using either value, then time in minutes was accepted as the value, provided it was consistent with the number of cigarettes smoked per day.
- If one of time in minutes or time in hours was 0, the other value was accepted as the time to first cigarette. Note that a value of 0 was accepted if one value of time to first cigarette was 0 and the other was NA. That is, 0 was accepted in cases where the skip pattern was followed correctly. In cases where it wasn't, then the above decision rule was applied.

SB013v – time to first cigarette in minutes (derived, categorical). This variable categorized the values for time to first cigarette as calculated in *SB012v*. Categorical time to first cigarette is simplified into 4 categories:

→ 0 = > 60 min
→ 1 = 31-60 min
→ 2 = 6-30 min
→ 3 = < 5 min

SB226v – overall frequency of butting out before finishing. This variable combines responses from *SB221* ("In the last month have you butted out a cigarette before you finished it because you thought about the harm of smoking?") and from *SB226* ("Was that once, a few times, or lots of times?"). If respondents said 'no' to *SB221* then their response in *SB226v* was coded as 'Never' (0). If respondents said 'yes' to *SB221* then their response from *SB226* was used as their response to *SB226v*. Hence, SB226v provides a measure, for all respondents, on whether respondents ever butt out their cigarette before finishing. The response options are:

→ 0 = Never
→ 1 = Once
→ 2 = A few times
→ 3 = Lots of times
→ 7 = NA
→ 8 = Refused
→ 9 = Don't know

3.14 Stop-smoking Medications

Table 43 summarizes the stop-smoking medications derived variables and Table 44 lists the availability of these variables in the various ITC wave datasets.

Table 45. Stop-smoking medications derived variables.

Variable	Label
SM112v	Ever heard of/used stop smoking medications
SM141v	Number of other smoking medications used in last 12 months (counter)
SM161v	Use of stop smoking medications in a quit attempt, overall
SM270v	Time used, using Rx med -- composite (days)
SM330v	Time used, using last NRT -- composite (days)
SM331v	Current length of NRT use (days)
SM382v	Number of smoking medications ever used (counter)
SM388v	Number of reasons listed for stopping medication use (counter)
SM550v	Length of non-quit NRT use in days

Table 46. Variable listing for stop-smoking medications derived variables.

Variable	4C	BD smokers and non-smokers	BR	BT	CN smokers and non-smokers	DE smokers and non-smokers	FR smokers	FR non-smokers	IES	IN smokers and non-smokers	KE	KR	MU smokers and non-smokers	MX	NL	NZ	SEA smokers and non-smokers	SEA youth	UY	ZM smokers and non-smokers
SM112v	6-8																			
SM141v							×													
SM161v							3													
SM270v							×	×												
SM330v							×	×												
SM331v	6-8,9																			
SM382v							×													
SM388v							×													
SM550v	6-8,9																			

In cases where a variable is not present on all waves of a project, only the wave numbers where it is present are listed.

SM112v – ever heard of/used stop smoking medications. SM112v combines responses from *SM101, SM106, SM111, SM269* and *SM369*. New recruits are asked whether they have heard of stop-smoking medications (*SM101)* and if they have ever used SSM (*SM106)*. If no, to the first or both questions, then *SM112v* is classified as '0'. If yes to both questions, then respondents are

asked whether they have used any medications in the last year (*SM111).* If no, then *SM112v* is classified as '1'. Next respondents are asked whether they are still currently using them (*SM269* & *SM329).* If not currently using them then *SM112v* is classified as '2', otherwise if they report currently using them they are classified as '3'. For recontacts, respondents are asked *SM111,* and if they report having used medications since the last survey date, then the pattern continues as for new recruits. If recontacts report not having used SSM since the last survey date, then their previous *SM112v* is carried forward from the previous wave. The categories are:

→ 0 = Never heard of/never used
→ 1 = Used in the past, sometime before last survey date
→ 2 = Used since last survey date
→ 3 = Currently used
→ 7 = NA
→ 8 = Refused
→ 9 = Don't know

SM141v – number of other smoking medications used in last 12 months (counter). This variable sums the number of other stop-smoking medications used in the last 12 months, as indicated in the responses to *SM121, SM122, SM124, SM125, SM130, SM131, SM 132, SM135* and *SM140*.

SM161v – use of stop smoking medications in a quit attempt overall. This variable is derived in the France Wave 3 Survey and is computed using *SM112* (use of SSM on last quit attempt) and *SM161* (use of SSMs to stop smoking completely). If respondent answers 'yes' to *SM112* or *SM161* then *SM161v* = 1 ('yes'). Otherwise, *SM161v* = 2 ('no').

SM270v – time used, using prescription medication – composite (days). This variable combines time values from *SM270a, SM270b* and *SM270c* on how long respondents used/have been using a stop-smoking prescription medication. It recalculates the responses into a value measured in days.

SM330v – time used, using last NRT – composite (days). This variable combines time values from *SM330a, SM330b* and *SM330c* on how long respondents used/have been using a nicotine replacement therapy medication. It recalculates the responses into a value measured in days.

SM331v – current length of NRT use (days). This variable combines time values from *SM330a, SM330b* and *SM330c* on how long respondents have currently been using a nicotine replacement therapy medication. If respondents answered 'yes' to *SM329* ("Are you still using [referent NRT medication(s)]?") then their responses to the *SM330* variables are included, otherwise the value for *SM331v* is set to missing. When *SM330* values are combined in *SM331v,* the responses are recalculated into a value measured in days.

SM382v – number of smoking medications ever used (counter). This variable counts the number of smoking medications ever reported being used from *SM201* to *SM220.*

SM388v – number of reasons listed for stopping medication use (counter). This variable counts the number of reasons mentioned by the respondent for ceasing to use stop-smoking medication, as provided in variables *SM332* to *SM339*.

SM550v – length of non-quit NRT use in days. This variable combines the time values from *SM550a* to *SM550c* and converts the units provided into days.

3.15 Sources of Tobacco Products

Table 45 summarizes the sources of tobacco products derived variables and Table 46 lists the availability of these variables in the various ITC wave datasets.

Table 47. Sources of tobacco products derived variables.

Variable	Label
SO401v	Number of locations where purchased cigarettes (counter)
SO502v	Number of sources of cheap cigarettes (counter)

Table 48. Variable listing for sources of tobacco products derived variables.

Variable	4C	BD smokers and non-smokers	BR	BT	CN smokers and non-smokers	DE smokers and non-smokers	FR smokers	FR non-smokers	IES	IN smokers and non-smokers	KE	KR	MU smokers and non-smokers	MX	NL	NZ	SEA smokers and non-smokers	SEA youth	UY	ZM smokers and non-smokers
SO401v	4-10						×													
SO502v	1-3																			

In cases where a variable is not present on all waves of a project, only the wave numbers where it is present are listed.

SO401v – number of locations where purchased cigarettes (counter). This variable counts the number of locations where respondents mentioned purchasing cigarettes, in *SO411, SO421, SO437* and *SO439*.

SO502v - number of sources of cheap cigarettes (counter). This variable counts the number of locations where respondents mentioned purchasing less expensive cigarettes, in *SO511* to *SO551* (odd numbers).

3.16 Smoked Tobacco Products

Table 47 summarizes the smoked tobacco product derived variables and Table 48 lists the availability of these variables in the various ITC wave datasets.

Table 49. Smoked tobacco products derived variables.

Variable	Label
ST201v	Number of other tobacco products used daily in last month
ST202v	Number of other tobacco products used less than daily in last month
ST302v	Number of non-cigarette tobacco products used in the last month (counter)
ST507v	Number of less harmful cigarettes named (counter)
ST552v	Number of less harmful cigarettes ever tried (counter)
ST602v	Number of less harmful cigarettes tried since last survey date (counter)
ST652v	Number of less harmful cigarettes still using (counter)

Table 50. Variable listing for smoked tobacco products derived variables.

Variable	4C	BD smokers	BD non-smokers	BR	BT	CN smokers	CN non-smokers	DE smokers	DE non-smokers	FR smokers	FR non-smokers	IES	IN smoker and non-smokers	KE	KR	MU smokers	MU non-smokers	MX	NL	NZ	SEA smokers	SEA non-smokers	SEA youth	UY	ZM smokers	ZM non-smokers
ST201v						×																				
ST202v						×																				
ST302v	×							2		×		×			×					×						
ST507v	1-8																									
ST552v	1-8														×											
ST602v	1-8																									
ST652v	2-5																									

In cases where a variable is not present on all waves of a project, only the wave numbers where it is present are listed.

ST201v – number of other tobacco products used daily in last month. This variable counts the number of other tobacco products besides cigarettes that respondents report in *ST352, ST358, ST360, ST362* and *ST388* as using daily (response value of 1).

ST202v – number of other tobacco products used less than daily in last month. This variable counts the number of other tobacco products besides cigarettes that respondents report in *ST352, ST358, ST360, ST362* and *ST388* as using less than daily (response value of 2).

ST302v - number of non-cigarette tobacco products used in the last month (counter). This variable counts the number of non-cigarette tobacco products used in the last month, as mentioned by respondents in *ST311, ST313, ST315, ST317* and *ST331* (response value of 1, 'mentioned').

ST507v – number of less harmful cigarettes named (counter). This variable counts the number of less harmful cigarettes that respondents could name, as mentioned in *ST511, ST513, ST515, ST519, ST523* and *ST531* (response value of 1, 'mentioned').

ST552v – number of less harmful cigarettes ever tried (counter). This variable counts the number of less harmful cigarettes the respondent reported ever trying, as mentioned in *ST561* to *ST581*.

ST602v – number of less harmful cigarettes tried since last survey date (counter). This variable counts the number of less harmful cigarettes the respondent reported trying, as mentioned in *ST611* to *ST631*.

ST652v – number of less harmful cigarettes still using (counter). This variable counts the number of less harmful cigarettes the respondent is currently using, as mentioned in variables *ST661* to *ST681*. These variables were dropped in Four Country wave 6, and as such *ST652v* is not included beyond wave 5.

3.17 Cigarette Warning Labels

Table 49 summarizes the cigarette warning labels derived variables and Table 50 lists the availability of these variables in the various ITC wave datasets.

Table 51. Cigarette warning labels derived variables.

Variable	Label
WL211v	Read or looked closely at the warning labels on cigarette packages
WL313v	Used any strategy to avoid warning labels

Table 52. Variable listing for cigarette warning labels derived variables.

Variable	4C	BD smokers	BD non-smokers	BR	BT	CN smokers	CN non-smokers	DE smokers	DE non-smokers	FR smokers	FR non-smokers	IES	IN smoker and non-smokers	KE	KR	MU smokers	MU non-smokers	MX	NL	NZ	SEA smokers	SEA non-smokers	SEA youth	UY	ZM smokers	ZM non-smokers
WL313v	6-10																									

In cases where a variable is not present on all waves of a project, only the wave numbers where it is present are listed.

WL211v – Read or looked closely at the warning labels on cigarette packages. This variable is derived for the analytic datasets and combines *WL201* (noticed health warning in the last month) and *WL211* (read or looked closely at warning labels). If respondent answers "never" to having noticed health warning (*WL201*) then *WL211*v is also coded as "never". Otherwise, *WL211v=WL211*.

WL313v – used any strategy to avoid warning labels. This variable takes responses from *WL313*, which asks smokers if they had avoided the warning labels on cigarette packs in the past month, and includes an additional response option 'does not smoke' = 0 to include the respondents who have quit since the last survey date.

1 = no
2 =yes

3.18 Health Related Variables

Table 51 summarizes the health related derived variables and Table 52 lists the availability of these variables in the various ITC wave datasets.

Table 53. Health related derived variables.

Variable	Label
PR109v	Body mass index (BMI); derived

Table 54.Variable listing for cigarette warning labels derived variables.

Variable	4C	BD smokers	BD non-smokers	BR	BT	CN smokers	CN non-smokers	DE smokers	DE non-smokers	FR smokers	FR non-smokers	IN smoker and non-smokers	IES	KE	KR	MU smokers	MU non-smokers	MX	NL	NZ	SEA smokers	SEA non-smokers	SEA youth	UY	ZM smokers	ZM non-smokers
PR109v										3	3															

In cases where a variable is not present on all waves of a project, only the wave numbers where it is present are listed.

PR109v – Body mass index (BMI), in kg/m^2. This variable is derived from *PR105* (weight) and *PR107* (height) as *PR109v = PR105*/(*PR107*)^2. If the units for *PR105* and *PR107* are not in kg and m, respectively, these are converted to kg and m during the computation of BMI. Where there are missing values or non-responses (i.e., refused, don't know) in either of *PR105* or *PR107*, *PR109v* is also set as missing or refused/don't know.

3.19 Method of time duration variables conversion.

There are some time duration derived variables in the data. They are usually computed from several time values measured in different units, such as hours, day, weeks etc. Although the data only provide these variables in one unit, these variables could be converted into any units according to the requirement. The primary principle is

1) 1 day =24 hours.
2) 1 week = 7 days.
3) 1 month=30.4 days.
4) 1 year=365.24 days.

The following table shows useful formulae for conversion.

Unit	hours	days	weeks	months	years
hours	*Original*	*hours/24*	*(hours/24)/7*	*(hours/24)/30.4*	*(hours/24)/365.24*
days	*days*24*	*original*	*days/7*	*days/30.4*	*days/365.24*
weeks	*weeks*7*24*	*weeks*7*	*original*	*weeks*7/30.4*	*weeks*7/365.24*
months	*months*30.4*24*	*months*30.4*	*months*30.4/7*	*original*	*months/12*

years	*years*365.24*24*	*years*365.24*	*years*365.24/7*	*years*12*	*original*

Example:

QA242v – Longest time smoke free (derived, in days). The original unit of *QA242v* in the data is days. If *QA242v=199 (in days), then*

QA242v =199 days
*=199*24 hours=4776 hours*
=199/7 weeks=28.429 weeks
=199/30.4 months =6.546 months
=199/365.24 years =0.545 years.

4. Strata, Clusters & Weights

ITC country projects employ sampling procedures which usually involve stratification and sometimes clustering. Strata, clusters, and other geography variables from both core and wave datasets are discussed in this section of the document. All data also have been weighted for cross-sectional analysis, and where multiple waves exist, for longitudinal analysis. The weight variables are briefly discussed in this section. For more details on the sampling procedures and how weights were calculated, refer to the country project Technical Reports.

The weight variables follow a standard ITC variable naming convention. As noted earlier, these types of variables (e.g., aDE31211v) are referenced in this document without wave identifiers or country codes (i.e., DE211v). This alteration was made so variables may be easily cross-referenced in this document with the various ITC datasets from all waves and all countries.

4.1 Strata, Clusters, & Other Geographic Variables

Table 53 summarizes the strata and cluster variables for the ITC country projects, where applicable.

Table 55. Strata and geographical variables.

Project	Strata	Strata variable	Clusters	Cluster variables
4C / ITC4	Province/State/Region	aStrata, bStrata, etc.	None	n/a
BD	Division	strata	Upazila	psu
BR	City	aStrata,bStrata,etc.	None	n/a
BT	District	District	gewog	gewog
CN	City	Strata / City	Street Districts & City Blocks	psu1 psu2
DE	Region	Strata	None	n/a
FR	None	n/a	None	n/a
IN	State & Urban	Strata	Villages(rural) / Wards(urban)	psu
IES	Region	Strata	None	n/a
KE	Province & District	Strata	Enumeration area	psu
KR	City/Province/Seoul	Strata	None	n/a
MU	District	Strata	Enumeration area	psu
MX	City	Strata	AGEB	psu
NL	Province	Province	None	n/a
NZ	District Health Board (DHB)	Strata	Meshgblock	psu
SEA	Zone(MY)//Region(TH)	Strata / STATEPROV	District(MY)/District.or Enumeration block(TH)	psu
UY	City Regions (secciones)	Strata	District (segmentos)	psu
ZM	Province	strata	District	psu

At present, for some country projects the strata and geographical variables are kept in the wave datasets. In most cases, these variables are kept in the core datasets.

Table 54 summarizes the strata and geographical variables that are found in the core datasets and Table 55 lists the availability of these variables in the various core datasets.

Table 56. Strata and geographical variables in core datasets.

Variable	Label
strata	Sampling region
*y*Strata	Wave *y* Strata information
psu	Primary sampling unit (cluster)
psu1	First stage cluster (Jie Dao)
psu2	Second stage cluster (Ju Wei Hui)
city	City
region	Region
regionUda	Region uda
ieRegion	IE region
scRegion	Scotland Health Authority/Board Regions
ukRegion	UK government region
STATEPROV	Malaysian state/Thai province
urban	urban/rural
moved	If and when respondents moved between sampling strata
hhldperson	Household id - person id
houseid	Household identifier (folio)
street	CODE2.Street code
block	CODE3.Block code
household	CODE4.Household code
AD	ID - AD
BU	ID - BU
CD	ID - CD
EB	ID - EB
LQ	ID - LQ
district	district
division	division
upazila	upazila
mno	Household member number
village	village

Table 54. Strata and geographical variables in core datasets(Continued).

Variable	Label
yState	State you live in
dwelling	Dwelling ID
enumArea	Enumeration area ID
yProvince	Province in wave y
dzongkhag	district
gewog	gewog

Table 57. Variable listing for strata and geographic variables in core datasets.

Variable	4C	BD smokers and non-smokers	BR	BT	CN smokers and non-smokers	DE	FR smokers and non-smokers	IN smoker and non-smokers	IES	KE	KR	MU smokers and non-smokers	MX	NL	NZ	SEA smokers and non-smokers	SEA youth	UY	ZM smokers and non-smokers
strata		×			×			×	×	×		×			×	×	×		×
*y*Strata	×		×			×					×			×					
psu		×						×		×		×			×	×	×		×
psu1					×														
psu2					×														
city					×														
region							×												
regionUda							×												
ieRegion									×										
scRegion									×										
ukRegion									×										
STATEPROV																×	×		
urban		×						×								×	×		×
moved	×																		
hhldperson																		×	

Variable	4C	BD smokers and non-smokers	BR	BT	CN smokers and non-smokers	DE	FR smokers and non-smokers	IN smoker and non-smokers	IES	KE	KR	MU smokers and non-smokers	MX	NL	NZ	SEA smokers and non-smokers	SEA youth	UY	ZM smokers and non-smokers
houseid													×						×
street					×														
block					×														
household					×													×	
AD															×	×			
BU															×	×			
CD															×	×			
EB															×	×			
LQ															×	×			
district		×						×			×	×							
division		×																	
upazila		×																	
mno		×																	
village		×						×											
yState						×													
dwelling											×	×							
enumArea											×	×							
yProvince													×						
province																		×	×
dzongkhag				×															
gewog				×															
state								×											
subdistirct								×											

strata – sampling region. ITC project samples are stratified by geographical regions. The variable s*trata* indicates the stratum in which the respondent resided at the time of interview for wave 1.

yStrata – strata information for a given wave. ITC project samples are stratified by geographical regions. The *yStrata* series of variables indicates the stratum in which the participant resided at a given wave (where *y* represents a wave identifying letter, such as 'a' for Wave 1, or 'b' for Wave 2, and so on). See also Section *1.2 Common Variable Names* for more information on wave identifiers.

psu – primary sampling unit (cluster). The sampling plan for the Southeast Asia project involved clustering. The variable *psu* identifies these clusters using 3-digit numbers.

psu1 – first stage cluster (Jie Dao). For the China project, the sampling plan involved clustering. The first level of clusters were *Jie Dao*, or street districts within cities. The variable *psu1* numbers the street districts in which respondents reside.

psu2 – second stage cluster (Ju Wei Hui). The second level of clustering within the China datsets were *Ju Wei Hui*, or neighbourhood blocks within the street districts. The variable *psu2* numerically labels the neighbourhood blocks in which respondents reside.

region. The France datasets identify the geographical regions in which respondents reside, using the *region* variable.

RegionUda. The France datasets identify a second level of geographical regions in which respondents reside, using the *regionUda* variable.

moved – respondent moved between sampling strata. The variable *moved* indicates if and when a respondent moved between sampling strata.

urban – urban/rural. The variable *urban* is binary and identifies a respondent as residing in either an urban or rural setting.

city. In the China datasets, the variable *city* names the city in which a respondent resides. This variable is used as the *strata* variable in the China datasets.

hhldperson – household id – person id. The variable *hhldperson* identifies individuals with a numeric indicator - numbers indicate both the household and the person, in the order in which they were surveyed in the household.

houseid – household identifier (folio). The variable *houseid* identifies the household with a unique numeric indicator.

street – CODE2.Street code. The variable *street* uses a numeric code to anonymously identify the streets on which respondents reside.

block – CODE3.Block code. The variable *block* uses a numeric code to anonymously identify the blocks on which the respondents reside.

household – CODE4.Household code. The variable *household* uses a numeric code to anonymously identify the households in which respondents reside.

ieRegion – IE region. The variable *ieRegion* is present only in the Ireland/Scotland data and identifies the geographical regions of Ireland in which respondents reside.

scRegion – Scotland Health Authority/Board Regions. The variable *scRegion* is present only in the Ireland/Scotland data and identifies the Scotland Health Authority regions in which respondents reside.

ukRegion – UK government region. The variable *ukRegion* is present only in the Ireland/Scotland data and identifies the UK government regions in which respondents reside.

STATEPROV –Malaysian state/Thai province. The variable *STATEPROV* is present only in the Southeast Asia data and identifies the Malaysian states or Thai provinces in which respondents reside.

AD, BU, CD, EB LQ. These are 'census-like' variables that numerically identify the administrative district (*AD*), building unit (*BU*), census division (*CD*), enumeration block (*EB*), and living quarters (*LQ*). They are used in combination with *IDR* and *IDSTRATA* to create *uniqid*. This variable is present only on the Southeast Asia core datasets.

district–district. The variable *district* is present in the Bangladesh and Mauritius data and identifies the district in which respondents reside.

division–division. The variable *division* is present only in the Bangladesh data.

village–village. The variable *village* is present only in the Bangladesh data.

yState– State you live in. The variable *yState* is present only in the Germany data and identifies the Germany state in which respondents reside.

dwelling– Dwelling ID. The variable *dwelling* is present only in the Mauritius data.

enumArea– Enumeration area ID. The variable *enumArea* is present only in the Mauritius data and identifies the enumeration area which respondents reside.

yProvince– Province in wave y. The variable *yProvince* is present only in Netherlands data and identifies the province which respondents reside in corresponding waves.

province– Province (sampling strata). The variable *province* is present only in Zambia data and identifies the province which respondents are sampled from.

Dzongkhag – The variable *dzongkhag* is present only in Bhutan data and identifies the district which respondents reside.

Gewog – The variable *gewog* is present only in Bhutan data and identifies the gewog which respondents reside.

State – state. The variable s*tate* is present only in India data.

Subdistrict – sub- district. The variable *subdistrict* is present only in India data.

Table 56 summarizes the strata and geographic variables currently kept in the wave datasets, and Table 57 lists the availability of these variables on the wave datasets.

Table 58. Strata and geographic variables in wave datasets.

Variable	Label
strata	Strata (city/seccion)
psu	Primary sampling unit (AGEB/segmento)
AGEB	AGEB
manzana	Manzana (block)
seccion	Seccion
segmento	Segmento
Province	Province
hhkey	Household identifier
GEW1	Weighting factor (sex, age and region)
GEW2	Weighting internet
city	State you live in
scregion	Scotland region
ukregion	UK government region
suburb	Suburb

Table 59. Variable listing for strata and geographic variables in wave datasets.

Variable	4C	BD smokers	BD non-smokers	BR	BT	CN smokers	CN non-smokers	FR smokers	FR non-smokers	IES	IN smokers and non-smokers	KE	KR	MU smokers	MU non-smokers	MX	NL	NZ	SEA smokers	SEA non-smokers	SEA youth	UY	ZM smokers	ZM non-smokers
strata																×						×		
psu																×						×		

Variable	4C	BD smokers	BD non-smokers	BR	BT	CN smokers	CN non-smokers	FR smokers	FR non-smokers	IES	IN smokers and non-smokers	KE	KR	MU smokers	MU non-smokers	MX	NL	NZ	SEA smokers	SEA non-smokers	SEA youth	UY	ZM smokers	ZM non-smokers
AGEB																×								
manzana																×						×		
seccion																						×		
segmento																						×		
Province																	×							
hhkey																			×	×	×			
city				×																				
scregion										4														
ukregion										4														
suburb										4														

For *mode*, see Section *2.2 Recruitment*. For *hhkey*, see Section *2.4 Other Core Variables*

psu – primary sampling unit (AGEB/segmento). The sampling plans for Mexico and Uruguay both involve clustering. The variable *psu* identifies these clusters using 1 and 2-digit numbers for Uruguay based on *segmento*, and 4-digit numbers for Mexico based on *AGEB*.

AGEB. This variable indicates the primary sampling unit, a level above the *manzana* blocks. Note that *AGEB* is a string variable, consisting of both letters and numbers. The letters were converted to zeros to create the numeric *psu* variable.

manzana - block. The *manzana* variable identifies the city blocks on which respondents reside. This variable is present only on the Mexico and Uruguay wave datasets.

seccion. The *seccion* variable identifies census regions, which were used as strata in Uruguay.

segmento. The variable *segmento* is the cluster variable for Uruguay, interchangeable with *psu.*

Province. The *Province* variable is present on the Netherlands dataset only and numerically labels the provinces in which respondents reside.

GEW1 & *GEW2 – weighting factors.* These two variables are present only on the Netherlands wave dataset.

city. The *city* variable is present in Brazil wave1 data and identifies the city in which respondents reside.

scregion. The *scregion* variable is present only in Ireland wave 4 data and is available for only Scotland residents.

ukregion. The *ukregion* variable is present only in Ireland wave 4 data and is available for only UK residents.

suburb. The *suburb* variable is present only in Ireland wave 4 data.

4.2 Cross-sectional & Longitudinal Weights

This section briefly describes the cross-sectional and longitudinal weights present on the ITC wave datasets. For more detailed information on how weights are constructed, refer to the country project Technical Reports.

It is important to note that the weight variables follow the standard ITC variable naming convention (e.g., aDE31911v) and are referenced in this document without wave identifiers or country codes (i.e., DE911v). This alteration was made so variables may be easily cross-referenced in this document with ITC datasets from all waves and all countries.

Table 58 summarizes the availability of the weight variables in the ITC wave datasets.

Table 60. Variable listing for weight variables.

Variables	4C	BD	BR	BT	CN	DE	FR	IN	IES	KE	KR	MU	MX	NL	NZ	SEA	UY	ZM
DE902v	9																	
DE903v	9																	
DE904v	9																	
DE905v	9																	
DE906v	9																	
DE909v	9																	
DE910v	9			×														
DE911v	1-7, 7.5		2				2-3				1							
DE913v																1-2		
DE914v																1-2		
DE915v	×	×	×	×	1-4		×	×	×		×	×	×		1	×	3-5	×
DE916v									3-4							1		
DE917v	2-10										2-3	2-3				3		
DE919v	×	×	2	×	×	×	×	×	2-4	×	×	×	×	×	2	×	×	×
DE920v					3-5		3											
DE921v	2-5	2-4	2		2-5	2-3	2-3	2	2-4		2-3	2-3	2-7	2-9		2-6	2-5	2
DE922v					3-5		3											
DE923v	3-5	3-4			4-5	3	3		3		3	3		3-9		3-6	3-5	
DE925v	4-6	4			5								4-7	4-9			4-5	
DE927v	5-6												5-7	5-9			5	
DE929v	6												6-7	6-9				
DE931v													7	7-9				
DE933v														8-9				

Variables	4C	BD	BR	BT	CN	DE	FR	IN	IES	KE	KR	MU	MX	NL	NZ	SEA	UY	ZM
DE935v														9				
DE951v		2-4						2					4-7			2-6	3-5	2
DE953v		3-4														3-6	3-5	
DE955v	6-10	4											4-7				4-5	
DE957v	6-10												5-7				5	
DE959v	6-10												6-7					
DE961v	7-10												7					
DE963v	7.510-																	
DE965v	8-9																	
DE967v	8.5-10																	
DE969v	9-10																	
DE971v	10																	

The weights account for the disproportionate probability of respondents being selected based on a variety of factors, including age, sex, ethnicity, number of phone lines in the household (in telephone interviews), geographical regions, countries, etc. Variables *DE911v* to *DE919v* (no *DE912v* or *DE918v*) are assigned as cross-sectional weights in ITC wave datasets, with *DE913v, DE917v,* and *DE919v* representing rescaled cross-sectional weights. Variables *DE921v* to *DE929v* (odd numbers) represent actual longitudinal weights. Variables *DE951v, DE955v, DE957v* and *DE959v* are rescaled longitudinal weights. The rescaled weights are normalized so that they have a mean of 1 by country (or city, in China, Mexico and Uruguay) to avoid an issue in pooled analyses where a country or region with the largest sample size dominates the analysis.

Table 59 lists the weight variables in the ITC wave datasets from various countries.

Table 61. Weight variables in wave datasets.

Country	Wave	Variable	Label
4-Country	Wave 1	aDE911v	cross-sectional weight for R1
		aDE915v	cross-sectional weight for M1
		aDE919v	rescaled cross-sectional weight for M1
	Wave 2	bDE911v	cross-sectional weight for R2

Country	Wave	Variable	Label
		bDE915v	cross-sectional weight for P2
		bDE917v	cross-sectional weight for M2
		bDE919v	combined cross-sectional weight for M2-P2
		bDE921v	longitudinal weight for M1-M2 continuers
	Wave 3	cDE911v	cross-sectional weight for R3
		cDE915v	cross-sectional weight for P3
		cDE917v	cross-sectional weight for M2/P2 continuers
		cDE919v	combined cross-sectional weight for M2/P2-P3
		cDE921v	longitudinal weight for M1-M2-M3 continuers
		cDE923v	longitudinal weight for M2/P2-M3 continuers
	Wave 4	dDE911v	cross-sectional weight for R4
		dDE915v	cross-sectional weight for P4
		dDE917v	cross-sectional weight for M3/P3 continuers
		dDE919v	combined cross-sectional weight for M3/P3-P4
		dDE921v	longitudinal weight for M1-M2-M3-M4 continuers
		dDE923v	longitudinal weight for M2/P2-M3-M4 continuers
		dDE925v	longitudinal weight for M3/P3-M4 continuers
	Wave 5	eDE911v	recruitment cross-sectional weight for P5
		eDE915v	replenishment cross-sectional weight for P5
		eDE917v	main survey cross-sectional weight for M5
		eDE919v	combined Wave 5 cross-sectional weight
		eDE921v	longitudinal weight for M1-M2-M3-M4-M5 continuers
		eDE923v	longitudinal weight for M2/P2-M3-M4-M5 continuers
		eDE925v	longitudinal weight for M3/P3-M4-M5 continuers

Country	Wave	Variable	Label
		eDE927v	longitudinal weight for M4/P4-M5 continuers
	Wave 6	fDE911v	recruitment cross-sectional weight for P6
		fDE915v	replenishment cross-sectional weight for P6
		fDE917v	main survey cross-sectional weight for M6
		fDE919v	combined, rescaled Wave 6 cross-sectional weight
		fDE925v	longitudinal weight for M3/P3-M4-M5-M6 continuers
		fDE927v	longitudinal weight for M4/P4-M5-M6 continuers
		fDE929v	longitudinal weight for M5/P5-M6 continuers
		fDE955v	rescaled longitudinal weight for M3/P3-M4-M5-M6 continuers
		fDE957v	rescaled longitudinal weight for M4/P4-M5-M6 continuers
		fDE959v	rescaled longitudinal weight for M5/P5-M6 continuers
	Wave 7	gDE911v	recruitment cross-sectional weight for P7
		gDE915v	replenishment cross-sectional weight for P7
		gDE917v	main survey cross-sectional weight for M7
		gDE919v	combined, rescaled Wave 7 cross-sectional weight
		gDE955v	rescaled longitudinal weight for M3/P3-M4-M5-M6-M7 continuers
		gDE957v	rescaled longitudinal weight for M4/P4-M5-M6-M7 continuers
		gDE959v	rescaled longitudinal weight for M5/P5-M6-M7 continuers
		gDE961v	rescaled longitudinal weight for M6/P6-M7 continuers
	Wave 7.5	gDE917v	main survey cross-sectional weight for M7.5
		gDE919v	combined, rescaled Wave 7.5 cross-sectional weight
		gDE955v	

Country	Wave	Variable	Label
			rescaled longitudinal weight for M3/P3-M4-M5-M6-M7-M7.5 continuers,
		gDE957v	rescaled longitudinal weight for M4/P4-M5-M6-M7-M7.5 continuers,
		gDE959v	rescaled longitudinal weight for M5/P5-M6-M7-M7.5 continuers
		gDE961v	rescaled longitudinal weight for M6/P6-M7-M7.5 continuers
		gDE963v	rescaled longitudinal weight for M7/P7-M7.5 continuers
	Wave 8	hDE911v	recruitment cross-sectional weight for P8
		hDE915v	replenishment cross-sectional weight for P8
		hDE917v	main survey cross-sectional weight for M8
		hDE919v	combined, rescaled Wave 8 cross-sectional weight
		hDE955v	rescaled longitudinal weight for M3/P3-M4-M5-M6-M7-M8 continuers
		hDE957v	rescaled longitudinal weight for M4/P4-M5-M6-M7-M8 continuers
		hDE959v	rescaled longitudinal weight for M5/P5-M6-M7-M8 continuers
		hDE961v	rescaled longitudinal weight for M6/P6-M7-M8 continuers
		hDE963v	rescaled longitudinal weight for M7/P7-M8 continuers
		hDE965v	rescaled longitudinal weight for M7/P7-M75-M8 continuers
	Wave 8.5	hhDE915v	replenishment cross sectional weight for P8.5
		hhDE917v	main survey cross-sectional weight for M8.5
		hhDE919v	combined, rescaled Wave 8.5 cross-sectional weight
		hhDE955v	rescaled longitudinal weight for M3/P3-M4-M5-M6-M7-M8-M8.5 continuers
		hhDE957v	rescaled longitudinal weight for M4/P4-M5-M6-M7-M8-M8.5 continuers

Country	Wave	Variable	Label
		hhDE959v	rescaled longitudinal weight for M5/P5-M6-M7-M8-M8.5 continuers
		hhDE961v	rescaled longitudinal weight for M6/P6-M7-M8-M8.5 continuers
		hhDE963v	rescaled longitudinal weight for M7/P7-M8-M8.5 continuers
		hhDE967v	rescaled longitudinal weight for M8/P8-M8.5 continuers
	Wave 9	iDE902v	replenishment weight for 25+ yrs old smokers(including current quitters) recruited in Wave 9
		iDE903v	replenishment weight for 18-24 yrs old smokers(including current quitters) recruited in Wave 9
		iDE904v	replenishment weight for 18-24 yrs old tobacco users (including current quitters) recruited in Wave 9
		iDE905v	replenishment weight for 18+ yrs old tobacco users (including current quitters) recruited in Wave 9
		iDE906v	replenishment weight for 25+ yrs old tobacco users (including current quitters) recruited in Wave 9
		iDE909v	replenishment weight for 18-24 yrs old non tobacco users recruited in Wave 9
		iDE915v	rescaled replenishment cross sectional weight for P9
		iDE917v	rescaled main survey cross sectional weight for M9
		iDE919v	combined, rescaled Wave 9 cross-sectional weight
		iDE955v	rescaled longitudinal weight for M3/P3-M4-M5-M6-M7-M8-M9 continuers
		iDE957v	rescaled longitudinal weight for M4/P4-M5-M6_M7-M8-M9 continuers
		iDE959v	rescaled longitudinal weight for M5/P5-M6-M7-M8-M9 continuers
		iDE961v	rescaled longitudinal weight for M6/P6-M7-M8-M9 continuers

Country	Wave	Variable	Label
		iDE963v	rescaled longitudinal weight for M7/P7-M8-M9 continuers
		iDE965v	rescaled longitudinal weight for M7.5-M9 continuers (US only)
		iDE967v	rescaled longitudinal weight for M8/P8-M9 continuers
		iDE969v	rescaled longitudinal weight for M8.5/P8.5-M9 continuers
	Wave 10	jDE915v	rescaled replenishment cross sectional weight for P10
		jDE917v	rescaled main survey cross sectional weight for M10
		jDE919v	combined, rescaled Wave 9 cross-sectional weight
		jDE955v	rescaled longitudinal weight for M3/P3-M4-M5-M6-M7-M8-M8.5-M9-M10 continuers
		jDE957v	rescaled longitudinal weight for M4/P4-M5-M6_M7-M8-M8.5-M9-M10 continuers
		jDE959v	rescaled longitudinal weight for M5/P5-M6-M7-M8-M8.5-M9-M10 continuers
		jDE961v	rescaled longitudinal weight for M6/P6-M7-M8-M8.5-M9-M10 continuers
		jDE963	rescaled longitudinal weight for M7/P7-M8-M8.5-M9-M10 continuers
		jDE967v	rescaled longitudinal weight for M8/P8-M8.5-M9-M10 continuers
		jDE969v	rescaled longitudinal weight for M8.5/P8.5-M9-M10 continuers
		jDE971v	rescaled longitudinal weight for M9/P9-M10 continuers
Bangladesh	Wave 1	aDE62915v	Wave 1 cross-sectional weight
		aDE62919v	Wave 1 rescaled cross-sectional weight
	Wave 2	bDE62915v	Wave 2 cross-sectional weight
		bDE62919v	Wave 2 rescaled cross-sectional weight

Country	Wave	Variable	Label
		bDE62921v	Longitudinal weight for M1-M2 continuers
		bDE62951v	Rescaled longitudinal weight for M1-M2 continuers
	Wave 3	cDE62915v	Wave 3 cross-sectional weight
		cDE62919v	Wave 3 rescaled cross-sectional weight
		cDE62921v	Longitudinal weight for M1-M2-M3 continuers
		cDE62923v	Rescaled longitudinal weight for M2/P2-M3 continuers
		cDE62951v	Rescaled longitudinal weight for M1-M2-M3 continuers
		cDE62953v	Rescaled longitudinal weight for M2/P2-M3 continuers
	Wave 4	dDE62915v	Wave 4 cross-sectional inflation weight
		dDE62919v	Wave 4 rescaled cross-sectional inflation weight
		dDE62921v	Longitudinal weight for M1-M2-M3-M4 continuers
		dDE62923v	Longitudinal weight for M2/P2-M3-M4 continuers
		dDE62925v	Longitudinal weight for M3/P3-M4 continuers
		dDE62951v	Rescaled longitudinal weight for M1-M2-M3-M4 continuers
		dDE62953v	Rescaled longitudinal weight for M2/P2-M3-M4 continuers
		dDE62955v	Rescaled longitudinal weight for M3/P3-M4 continuers
Brazil	Wave 1	aDE57915v	Combined rescaled cross-sectional weight
	Wave 2	bDE57911v	Rescaled cross-sectional weight for M2
		bDE57915v	Rescaled cross-sectional weight for P2
		bDE57919v	Combined cross-sectional weight for M2-P2 -rescaled
		bDE57921v	Rescaled waves 1-2 longitudinal weights
Bhutan	Wave 1	aDE64910v	Wave 1 screener weight
		aDE64915v	Wave 1 cross-sectional weight
		aDE64919v	Wave 1 rescaled cross-sectional weight

Country	Wave	Variable	Label
China	Wave 1	aDE31915v	Wave 1 cross-sectional weight
		aDE31919v	Wave 1 rescaled cross-sectional weight
	Wave 2	bDE31915v	Wave 2 cross-sectional weight for M2-P2
		bDE31919v	Wave 2 rescaled cross-sectional weight
		bDE31921v	longitudinal weights for M1-M2 continuers
	Wave 3	cDE31915v	Wave 3 cross-sectional weight for M3-P3
		cDE31919v	Wave 3 rescaled cross-sectional weight
		cDE31920v	Waves 123 longitudinal weights
		cDE31921v	Waves 123 rescaled longitudinal weights
		cDE31922v	Waves 23 longitudinal weights
	Wave 4	dDE31915v	Wave 4 cross-sectional weights for M4-P4
		dDE31919v	Wave 4 rescaled cross-sectional weights for M4-P4
		dDE31920v	Waves 1234 longitudinal weights
		dDE31921v	Waves 1234 rescaled longitudinal weights
		dDE31922v	Waves 34 longitudinal weights
		dDE31923v	Waves 34 rescaled longitudinal weights
	Wave 5	eDE31919v	Wave 5 rescaled cross-sectional weights for M5-P5
		eDE31920v	Waves 1-2-3-4-5 longitudinal weights
		eDE31921v	Waves 1-2-3-4-5 rescaled longitudinal weights
		eDE31922v	Waves 3-4-5 longitudinal weights
		eDE31923v	Waves 3-4-5 rescaled longitudinal weights
		eDE31924v	Waves 4-5 longitudinal weights
		eDE31925v	Waves 4-5 rescaled longitudinal weights
Germany	Wave 1	aDE44919v	Wave 1 rescaled cross-sectional weights

Country	Wave	Variable	Label
	Wave 2	bDE44919v	Wave 2 rescaled cross-sectional weights
		bDE44921v	Waves 1-2 rescaled longitudinal weights
	Wave 3	cDE44919v	Wave 3 rescaled cross-sectional weights
		cDE44921v	Waves 1-2-3 rescaled longitudinal weights
		cDE44923v	Waves 2-3 rescaled longitudinal weights
France	Wave 1	aDE41915v	Wave 1 cross-sectional weight
		aDE41919v	Wave 1 rescaled cross-sectional weight
	Wave 2	bDE41911v	Cross-sectional weight for M2
		bDE41915v	Cross-sectional weight for P2
		bDE41919v	Cross-sectional weight for P2/M2
		bDE41921v	Longitudinal weight for M1-M2 continuers
	Wave 3	cDE41911v	Cross-sectional weight for M3
		cDE41915v	Cross-sectional weight for P3
		cDE41919v	Combined rescaled cross-sectional weight for P3/M3
		cDE41920v	Waves 1-3 longitudinal weight for M1-M2-M3 continuers
		cDE41921v	Waves 1-3 rescaled longitudinal weight for M1-M2-M3 continuers
		cDE41922v	Waves 2-3 longitudinal weight for M2-M3 continuers
		cDE41923v	Waves 2-3 rescaled longitudinal weight for M2-M3 continuers
India	Wave 1	aDE915v	Wave 1 city/area level cross-sectional weight
		aDE919v	Rescaled Wave 1 cross-sectional weight
	Wave 2	bDE61915v	Wave 2 cross-sectional weight
		bDE61919v	Re-scaled Wave 2 cross-sectional weight
		bDE61921v	Wave 1 – Wave 2 longitudinal weight
		bDE61951v	Re-scaled Wave 1 – Wave 2 longitudinal weight

Country	Wave	Variable	Label
Ireland	Wave 1	aDE915v	Wave 1 Survey Weight
	Wave 2	bDE915v	cross-sectional weight for P2
		bDE919v	combined, rescaled cross-sectional weight for M2/P2
		bDE921v	longitudinal weight for M1 continuers
	Wave3	cDE915v	cross-sectional weight for P3 recruits
		cDE916v	rescaled cross-sectional weight for P3 recruits
		cDE919v	combined cross-sectional weight for M2-M3/P3
		cDE921v	longitudinal weight for M1-M2-M3 continuers
		cDE923v	longitudinal weight for M2-M3 continuers
	Wave 4	dDE915v	cross-sectional weight for P4 recruits
		dDE916v	rescaled cross-sectional weight for P4 recruits
		dDE919v	combined cross-sectional weight
		dDE921v	longitudinal weight for M1-M2-M3-M4 continuers
Kenya	Wave 1	aDE72919v	rescaled cross-sectional weight for M1
Korea	Wave1	aDE21911v	recruitment weight for M1 (cross-sectional weight)
		aDE21915v	main weight for M1 (cross-sectional weight)
	Wave2	bDE21915v	rescaled cross-sectional weights for P2
		bDE21917v	rescaled cross-sectional weights for M2
		bDE21919v	Combined cross-sectional weights for M2/P2
		bDE21921v	rescaled longitudinal weights for M1-M2 continuers
	Wave 3	cDE21915v	rescaled cross-sectional weights for P3
		cDE21917v	rescaled cross-sectional weights for M3
		cDE21919v	Combined cross-sectional weights for M3/P3
		cDE21921v	rescaled longitudinal weights for M1-M2-M3 continuers

Country	Wave	Variable	Label
		cDE21923v	rescaled longitudinal weights for M2-M3 continuers
Mauritius	Wave 1	aDE73915v	wave 1 cross-sectional weight
		aDE73919v	wave 1 rescaled cross-sectional weight
	Wave 2	bDE73915v	Wave 2 cross-sectional weights for new recruits (P2)
		bDE73917v	Wave 2 cross-sectional weights for M1-M2 continuers
		bDE73919v	rescaled Wave 2 cross-sectional weight
		bDE73921v	rescaled longitudinal weight for M1-M2 continuers
	Wave 3	cDE73915v	Wave 3 cross-sectional weights for new recruits (P3)
		cDE73917v	Wave 3 cross-sectional weights for M2/P2-M3 continuers
		cDE73919v	Rescaled Wave 3 cross-sectional weight
		cDE73921v	Rescaled longitudinal weight for M1-M2-M3 continuers
		cDE73923v	Rescaled longitudinal weight for M2/P2-M3 continuers
Mexico	Wave 1	aDE51915v	wave 1 cross-sectional weight
		aDE51919v	wave 1 rescaled cross-sectional weight
	Wave 2	bDE51915v	Wave 2 final city level weight
		bDE51919v	Wave 2 rescaled city level weight
		bDE51921v	Waves 12 longitudinal weights rescaled to age sex group
		bDE51951v	Waves 12 longitudinal weights rescaled to city sample size
	Wave 3	cDE51915v	Wave 3 cross-sectional city level weight
		cDE51919v	rescaled Wave 3 cross-sectional weights
		cDE51921v	Waves 123 longitudinal weights rescaled to age sex group
		cDE51951v	Waves 123 longitudinal weights rescaled to city sample size
	Wave 4	dDE51915v	Wave 4 cross-sectional city level weight
		dDE51919v	rescaled Wave 4 cross-sectional weights

Country	Wave	Variable	Label
		dDE51921v	Waves 1234 longitudinal weights rescaled to age sex group
		dDE51925v	Waves 34 longitudinal weights rescaled to age sex group
		dDE51951v	Waves 1234 longitudinal weights rescaled to city sample size
		dDE51955v	Waves 34 longitudinal weights rescaled to city sample size
	Wave 5	eDE51915v	Wave5 cross-sectional weights
		eDE51919v	rescaled Wave5 cross-sectional weights
		eDE51921v	Waves 12345 longitudinal weights rescaled to age sex group
		eDE51925v	Waves 345 longitudinal weights rescaled to age sex group
		eDE51927v	Waves 45 longitudinal weights rescaled to age sex group
		eDE51951v	Waves 12345 longitudinal weights rescaled to city sample size
		eDE51955v	Waves 345 longitudinal weights rescaled to city sample size
		eDE61957v	Waves 45 longitudinal weights rescaled to city sample size
	Wave 6	fDE51915v	Wave 6 cross-sectional city level weight
		fDE51919v	Rescaled wave 6 cross-sectional weight
		fDE51921v	Wave 123456 longitudinal weights rescaled to age sex group
		fDE51925v	Wave 3456 longitudinal weights rescaled to age sex group
		fDE51927v	Wave 456 longitudinal weights rescaled to age sex group
		fDE51929v	Wave 56 longitudinal weights rescaled to age sex group
		fDE51951v	Wave 123456 longitudinal weights rescaled to city sample size
		fDE51955v	Wave 3456 longitudinal weights rescaled to city sample size
		fDE51957v	Wave 456 longitudinal weights rescaled to city sample size
		fDE51959v	Wave 56 longitudinal weights rescaled to city sample size

Country	Wave	Variable	Label
	Wave 7	gDE51915v	Wave 7 cross-sectional city level weight
		gDE51919v	Rescaled wave 7 cross-sectional weight
		gDE51921v	wave 1234567 longitudinal weights rescaled to age sex group
		gDE51925v	wave 34567 longitudinal weights rescaled to age sex group
		gDE51927v	wave 4567 longitudinal weights rescaled to age sex group
		gDE51929v	wave 567 longitudinal weights rescaled to age sex group
		gDE51931v	wave 67 longitudinal weights rescaled to age sex group
		gDE51951v	wave 1234567 longitudinal weights rescaled to city sample size
		gDE51955v	wave 34567 longitudinal weights rescaled to city sample size
		gDE51957v	wave 4567 longitudinal weights rescaled to city sample size
		gDE51959v	wave 567 longitudinal weights rescaled to city sample size
		gDE51961v	wave 67 longitudinal weights rescaled to city sample size
Netherlands	Wave 1	aDE42919v	Wave 1 rescaled cross-sectional weights (CASI+CATI)
	Wave 2	bDE42919v	Wave 2 rescaled cross-sectional weight
		bDE42921v	Waves 1-2 rescaled longitudinal weights (CASI only)
	Wave 3	cDE42919v	Wave 3 rescaled cross-sectional weights (CASI+CATI)
		cDE42921v	Waves 1-3 rescaled longitudinal weights (CASI+CATI)
		cDE42923v	Waves 2-3 rescaled longitudinal weights (CASI only)
	Wave 4	dDE42919v	Wave4 rescaled cross-sectional weights(CASI only)
		dDE42921v	Waves1-4 rescaled longitudinal weights(CASI only)
		dDE42923v	Waves2-4 rescaled longitudinal weights(CASI only)
		dDE42925v	Wave3-4 rescaled longitudinal weights(CASI only)
	Wave 5	eDE42919v	Wave 5 rescaled cross-sectional weights (CASI only)

Country	Wave	Variable	Label
		eDE42921v	Waves 1-5 rescaled longitudinal weights (CASI only)
		eDE42923v	Waves 2-5 rescaled longitudinal weights (CASI only)
		eDE42925v	Waves 3-5 rescaled longitudinal weights (CASI only)
		eDE42927v	Waves 4-5 rescaled longitudinal weights (CASI only)
	Wave 6	fDE42919v	Wave 6 rescaled cross-sectional weights (CASI only)
		fDE42921v	Waves 1-6 rescaled longitudinal weights (CASI only)
		fDE42923v	Waves 2-6 rescaled longitudinal weights (CASI only)
		fDE42925v	Waves 3-6 rescaled longitudinal weights (CASI only)
		fDE42927v	Waves 4-6 rescaled longitudinal weights (CASI only)
		fDE42929v	Waves 5-6 rescaled longitudinal weights (CASI only)
	Wave 7	gDE42919v	Wave 7 rescaled cross-sectional weights (CASI only)
		gDE42921v	Waves 1-7 rescaled longitudinal weights (CASI only)
		gDE42923v	Waves 2-7 rescaled longitudinal weights (CASI only)
		gDE42925v	Waves 3-7 rescaled longitudinal weights (CASI only)
		gDE42927v	Waves 4-7 rescaled longitudinal weights (CASI only)
		gDE42929v	Waves 5-7 rescaled longitudinal weights (CASI only)
		gDE42931v	Waves 6-7 rescaled longitudinal weights (CASI only)
	Wave 8	fDE42919v	Wave 8 rescaled cross-sectional weights (CASI only)
		fDE42921v	Waves 1-8 rescaled longitudinal weights (CASI only)
		fDE42923v	Waves 2-8 rescaled longitudinal weights (CASI only)
		fDE42925v	Waves 3-8 rescaled longitudinal weights (CASI only)
		fDE42927v	Waves 4-8 rescaled longitudinal weights (CASI only)
		fDE42929v	Waves 5-8 rescaled longitudinal weights (CASI only)
		fDE42931v	Waves 6-8 rescaled longitudinal weights (CASI only)

Country	Wave	Variable	Label
		fDE42933v	Waves 7-8 rescaled longitudinal weights (CASI only)
	Wave 9	iDE42919v	Wave 9 rescaled cross-sectional weights (CASI only)
		iDE42921v	Waves 1-9 rescaled longitudinal weights (CASI only)
		iDE42923v	Waves 2-9 rescaled longitudinal weights (CASI only)
		iDE42925v	Waves 3-9 rescaled longitudinal weights (CASI only)
		iDE42927v	Waves 4-9 rescaled longitudinal weights (CASI only)
		iDE42929v	Waves 5-9 rescaled longitudinal weights (CASI only)
		iDE42931v	Waves 6-9 rescaled longitudinal weights (CASI only)
		iDE42933v	Waves 7-9 rescaled longitudinal weights (CASI only)
		iDE42935v	Waves 8-9 rescaled longitudinal weights (CASI only)
New Zealand	Wave 1	aDE915v	Cross-sectional weight for Wave 1
	Wave 2	bDE919v	Cross-sectional weight for Wave2
SEAsia	Wave 1	aDE11913v	Wave 1 state-level cross-sectional for Malaysia ONLY
		aDE11914v	Wave 1 rescaled state-level cross-sectional weight for Malaysia ONLY
		aDE11915v	Wave 1 cross-sectional survey weight
		aDE11916v	Wave 1 pooled rescaled cross-sectional survey weight (MY = aDE11914v, TH = aDE11919v)
		aDE11919v	Wave 1 rescaled cross-sectional survey weight
	Wave 2	bDE11913v	Wave 2 state-level cross-sectional weight for new recruits, Malaysia ONLY
		bDE11914v	Wave 2 rescaled state-level cross-sectional weight for new recruits, Malaysia ONLY
		bDE11915v	Wave 2 cross-sectional weight for new recruits, Thailand ONLY
		bDE11919v	Wave 2 rescaled cross-sectional weight

Country	Wave	Variable	Label
		bDE11921v	Wave 2 longitudinal weight for respondents present in Waves 1 and 2
		bDE11951v	Wave 2 rescaled longitudinal weight for respondents present in Waves 1& 2
	Wave 3	cDE11915v	Wave 3 cross-sectional weight for new recruits
		cDE11917v	Wave 3 rescaled cross-sectional weight for new recruits
		cDE11919v	Wave 3 rescaled combined cross-sectional weight for everyone present at Wave 3
		cDE11921v	Wave 3 longitudinal weight for respondents present in Waves 1, 2 and 3
		cDE11923v	Wave 3 longitudinal weight for respondents present in Waves 1 and 3
		cDE11951v	Wave 3 rescaled longitudinal weight for respondents present in Waves 1, 2&3
		cDE11953v	Wave 3 rescaled longitudinal weight for respondents present in Waves 1 & 3
	Wave 4	dDE11915v	Wave 4 cross-sectional weight
		dDE11919v	Wave 4 combined cross-sectional weight for everyone present at Wave 4
		dDE11921v	Wave 4 longitudinal weight for everyone present in waves 1, 2, 3 and 4
		dDE11923v	Wave 4 longitudinal weight for everyone present in waves 1 and 4
	Wave 5	eDE11915v	Wave 5 cross-sectional weight
		eDE11919v	Wave 5 rescaled cross-sectional weight for everyone present at Wave 5
		eDE11921v	Wave 5 longitudinal weight for everyone present in waves 1, 2, 3, 4, and 5
		eDE11923v	Wave 5 longitudinal weight for everyone present in waves 1 and 5

Country	Wave	Variable	Label
	Wave 6	fDE11915v	wave 6 cross-sectional weight
		fDE11919v	wave 6 rescaled cross-sectional weight for everyone present at wave 6
		fDE11921v	wave 6 longitudinal weight for everyone present in waves 1, 2, 3, 4, 5, and 6
		fDE11923v	wave 6 longitudinal weight for everyone present in waves 1 and 6
Uruguay	**Wave 1**	aDE52919v	wave 1 rescaled cross-sectional weight
	Wave 2	bDE52919v	wave2 rescaled cross-sectional weight
		bDE52921v	Waves 1-2 rescaled longitudinal weight
	Wave 3	cDE52915v	Wave 3 cross-sectional weight
		cDE52919v	Wave 3 rescaled cross-sectional weight
		cDE52921v	Waves 1-2-3 rescaled to gender/age group longitudinal weight
		cDE52923v	Waves 2-3 rescaled to gender/age group longitudinal weight
		cDE52951v	Waves 1-2-3 rescaled to sample size longitudinal weight
		cDE52953v	Waves 2-3 rescaled to sample size group longitudinal weight
	Wave 4	dDE52915v	Wave 4 inflated cross-sectional weight
		dDE52919v	Wave 4 rescaled cross-sectional weight
		dDE52921v	Wave 1-2-3-4 inflated longitudinal weight
		dDE52951v	Wave 1-2-3-4 rescaled longitudinal weight
		dDE52923v	Wave 2-3-4 inflated longitudinal weight
		dDE52953v	Wave 2-3-4 rescaled longitudinal weight
		dDE52925v	Wave 3-4 inflated longitudinal weight
		dDE52955v	Wave 3-4 rescaled longitudinal weight
	Wave 5	eDE52915v	Wave 1 cross-sectional weight
		eDE52919v	Rescaled cross-sectional weight

Country	Wave	Variable	Label
		eDE52921v	Wave 1-2-3-4-5 inflated longitudinal weight
		eDE52951v	Wave 1-2-3-4-5 rescaled longitudinal weight
		eDE52923v	Wave 2-3-4-5 inflated longitudinal weight
		eDE52953v	Wave 2-3-4-5 rescaled longitudinal weight
		eDE52925v	Wave 3-4-5 inflated longitudinal weight
		eDE52955v	Wave 3-4-5 rescaled longitudinal weight
		eDE52927v	Wave 4-5 inflated longitudinal weight
		eDE52957v	Wave 4-5 rescaled longitudinal weight
Zambia	**Wave 1**	aDE74915v	National level cross-sectional inflation weight for Wave 1
		aDE74919v	Rescaled Wave 1 cross-sectional weight
	Wave 2	bDE74915v	Wave 2 cross-sectional sampling weight (inflation weight)
		bDE74919v	Wave 2 rescaled cross-sectional sampling weight
		bDE74921v	W1-W2 longitudinal sampling weight
		bDE74951v	W1-W2 rescaled longitudinal sampling weight

5. Index of Derived Variables

Appendix 1: New Derived Variables in ITC 4C Wave 8 Dataset

Data Management Core - ITC Project
University of Waterloo
October, 2012

The ITC 4-country Wave 7.5 survey was conducted by telephone and only in the United States between November 2, 2009 and January 10, 2010. From the United States recontact cohort, 915 eligible individuals who were daily smokers and smoked at least 10 manufactured cigarettes per day in Wave 7 were invited to participate the survey. Subsequently 678 out of 915 respondents completed the Wave 7.5 survey successfully.

In ITC 4-country Wave 8 survey, all respondents who completed the Wave 7 survey, whether they were in the Wave 7.5 survey or not, were invited. As a result, the "LSD" in a question such as

"Have you made any attempts to stop smoking since we last talked with you, that is, since [LSD, if smoking at LSD/you went back to smoking] ?"

sometimes has a double meaning: it refers to Wave 7 for participants who were not in the Wave 7.5 survey; and to Wave 7.5 for participants who successfully completed the Wave 7.5 survey. Hence, the measurements of number of quit attempts, and similar variables, were not measured on a uniform time period. Respondents who completed the Wave 7.5 survey were asked about number of quit attempts between Wave 7.5 and Wave 8, while other respondents were asked about quit attempts between Wave 7 and Wave 8. To solve this issue, it was decided to provide derived variables which "splice" the Wave 7 to Wave 7.5 and Wave 7.5 to Wave 8 information for Wave 7.5 respondents, for comparison across countries.

Note: For questions in Wave 8 referring to "LSD" where there was no corresponding question in Wave 7.5, "LSD" was the Wave 7 date for all respondents. Also, for questions in the Cessation Help (CH) section, "LSD" was Wave 7 for all respondents.

Table 1 summarizes the new derived variables. Please note the records for which additional variables are to be newly computed must satisfy the condition: *inM7_5=1* and *inM8=1*.

Table A1. New derived variables in ITC 4C Wave 8 dataset

Variable	Label
hQA331gh	any attempts to stop smoking- derived (Wave 7 to Wave 8)
hQA561gh	number of new quit attempts after LSD(Wave 7)-derived
hQA442vgh	number of days since start of most recent quit attempt -- composite, # days (Wave 7 to Wave 8)
hQA553vgh	longest time smoke-free since LSD -- composite in days (Wave 7 to Wave 8)
hCH801gh	visited doctor since LSD (Wave 7)-derived
hCH811gh	advice from doctor to quit-derived
hCH811vgh	received advice from doctor to quit,overall-derived
hCH812gh	advice made you think about quitting-derived
hCH813gh	referral from doctor-derived
hCH813vgh	referral from doctor to quit,overall-derived
hCH815gh	quitting Rx from doctor (smokers)-derived
hCH815vgh	quitting Rx from doctor, overall-derived
hCH817gh	pamphlet from doctor-derived
hCH817vgh	quitting pamphlet from doctor, overall-derived
hCH821gh	support from doctor for having quit-derived
hCH821vgh	support for having quit,overall-derived
hCH861gh	received quitting info from quit line-derived
hCH865gh	received quitting info from internet-derived
hCH869gh	received quitting info from clinic-derived
hSM111gh	used any quitting meds-derived (Wave 7 to Wave 8)
hSM161gh	SSM used to quit-derived
hSM161vgh	SSM used to quit-derived
hSM160gh	SSM used to stay quit-derived
hSM162gh	SSM used to reduce amount-derived

Table A1. New derived variables in ITC 4C Wave 8 dataset (Continued)

Variable	Label
hSM163gh	SSM used to cope with non-smoke situations-derived
hSM164gh	SSM used for other reason-derived
hSM164ogh	SSM used for what other reason-derived
hSM201gh	used gum for last QA-derived
hSM202gh	used patch for last QA-derived
hSM203gh	used lozenges for last QA-derived
hSM204gh	used tablets for last QA-derived
hSM210gh	used other NRT for last QA-derived
hSM211gh	used Zyban/ Bupropion/ Wellb for last QA-derived
hSM212gh	used Chantix/ Varenicline for last QA-derived
hSM215gh	used other Rx med for last QA-derived
hSM220gh	used other SSM (gen) for last QA-derived
hSM210ogh	used what other NRT for last QA-derived
hSM215ogh	used what other Rx for last QA-derived
hSM220ogh	used what other med for last QA
hSM259gh	last Rx SSM left over from earlier QA-derived
hSM260gh	got last Rx SSM by prescription-derived
hSM262gh	got last Rx SSM from a friend-derived
hSM263gh	got last Rx SSM free from a doctor-derived
hSM265gh	paid full price for last Rx SSM-derived
hSM266gh	got last Rx SSM at a discount-derived
hSM267gh	got last Rx SSM free-derived
hSM268gh	got last Rx SSM partially covered-derived
hSM269gh	still using Rx SSM-derived
hSM270vgh	used, using Rx SSM in (# days)-derived
hSM310gh	got last NRT by prescription-derived
hSM311gh	got last NRT over the counter-derived
hSM312gh	got last NRT from a friend-derived

Table A1. New derived variables in ITC 4C Wave 8 dataset (Continued)

Variable	Label
hSM313gh	got last NRT free, from a doctor-derived
hSM314gh	got last NRT at non-pharmacy store-derived
hSM315gh	paid full price for last NRT-derived
hSM316gh	got last NRT at a discount-derived
hSM317gh	got last NRT free-derived
hSM318gh	got last NRT partially covered-derived
hSM319gh	last NRT left over from earlier QA-derived
hSM329gh	still using last NRT-derived
hSM330vgh	used, using last NRT in (#days)-derived
hSM345gh	smoked while using last NRT-derived
hSM458gh	used to-quit NRT daily or less often-derived
hSM459gh	# to-quit NRT pieces per day-derived
hSM380gh	same SSMs to stay quit as to quit-derived
hSM401gh	used gum to stay quit-derived
hSM402gh	used patch to stay quit-derived
hSM403gh	used lozenges to stay quit-derived
hSM404gh	used tablets to stay quit-derived
hSM410gh	used other NRT to stay quit-derived
hSM411gh	used other Rx to stay quit-derived
hSM412gh	used other unknown/ generic SSM to stay quit-derived
hSM410ogh	which other NRT used to stay quit-derived
hSM411ogh	which other Rx used to stay quit-derived
hSM412ogh	which other SM used to stay quit-derived
hSM420gh	got stay-quit NRT by prescription-derived
hSM421gh	got stay-quit NRT over the counter-derived
hSM422gh	got stay-quit NRT from a friend-derived
hSM423gh	got stay-quit NRT free, from a doctor-derived
hSM424gh	got stay-quit NRT at non-pharmacy store-derived

Table A1. New derived variables in ITC 4C Wave 8 dataset (Continued)

Variable	Label
hSM429gh	stay-quit NRT left over from QA-derived
hSM435gh	paid full price for stay-quit NRT-derived
hSM436gh	got stay-quit NRT at a discount-derived
hSM437gh	got stay-quit NRT free-derived
hSM438gh	got stay-quit NRT partly covered-derived
hSM462gh	used stay-quit NRT daily or less often-derived
hSM464gh	# stay-quit NRT pieces per day-derived
hSM119gh	since LSD, used other SSMs to quit-derived
hSM121gh	since LSD, used gum-derived
hSM122gh	since LSD, used patch-derived
hSM123gh	since LSD, used lozenges-derived
hSM124gh	since LSD, used tablets-derived
hSM130gh	since LSD, used other NRT-derived
hSM135gh	since LSD, used other Rx med-derived
hSM140gh	since LSD, used other SSM (gen)-derived
hSM124gh	since LSD, used tablets-derived
hSM130gh	since LSD, used other NRT-derived
hSM135gh	since LSD, used other Rx med-derived
hSM140gh	since LSD, used other SSM (gen)-derived
hSM130ogh	since LSD, which other NRT used-derived
hSM135ogh	since LSD, which other Rx med used-derived
hSM140ogh	since LSD, which other Rx med used-derived
hSM498gh	used SSMs to cut down separately from to-quit-derived
hSM499gh	used SSMs for non-quit separately from to-quit-derived
hSM500gh	cut down or cope, which use most recent-derived
hSM501gh	used gum for non-quit-derived
hSM502gh	used patch for non-quit-derived
hSM503gh	used lozenges for non-quit-derived

Table A1. New derived variables in ITC 4C Wave 8 dataset (Continued)

Variable	Label
hSM504gh	used tablets for non-quit-derived
hSM510gh	used other NRT for non-quit-derived
hSM511gh	used prescription med for non-quit-derived
hSM512gh	used other med for non-quit-derived
hSM510ogh	which other NRT for non-quit-derived
hSM511ogh	which prescription med for non-quit-derived
hSM512ogh	used which other med for non-quit-derived
hSM520gh	got non-quit NRT by prescription-derived
hSM521gh	got non-quit NRT over the counter-derived
hSM522gh	got non-quit NRT from a friend-derived
hSM523gh	got non-quit NRT free, from a doctor-derived
hSM524gh	got non-quit NRT at non-pharmacy store-derived
hSM529gh	non-quit NRT leftover from QA-derived
hSM535gh	paid full price for non-quit NRT-derived
hSM536gh	got non-quit NRT at a discount-derived
hSM537gh	got non-quit NRT free-derived
hSM538gh	got non-quit NRT partly covered-derived
hSM544gh	used NRT to cut down after failed QA-derived
hSM550vgh	used, using non-quit NRT in (# days)-derived
hSM552gh	used non-quit NRT daily or less often-derived
hSM554gh	# non-quit NRT pieces per day-derived
hSM350gh	consider yourself addicted to NRT-derived
hSM351gh	enjoy NRT or only find useful-derived
hSM352gh	which non-quit NRT preferred-derived
hSM361gh	meds make quitting easier-derived
hSM362gh	can quit without meds-derived
hSM363gh	meds too expensive-derived
hSM364gh	Don’t know how to use meds-derived

Table A1. New derived variables in ITC 4C Wave 8 dataset (Continued)

Variable	Label
hSM365gh	meds too hard to get-derived
hSM366gh	meds might harm health-derived
hSM370gh	meds less harmful than reg cig-derived
hSM373gh	meds how much less harmful-derived
hSM375gh	meds more harmful or same as reg cig-derived
hCH811vgh	received advice from doctor to quit,overall-derived (Wave 7 to Wave 8)
hCH813vgh	referral from doctor to quit,overall-derived (Wave 7 to Wave 8)
hCH815vgh	quitting Rx from doctor, overall-derived (Wave 7 to Wave 8)
hCH817vgh	quitting pamphlet from doctor, overall-derived (Wave 7 to Wave 8)
hCH821vgh	support for having quit,overall-derived (Wave 7 to Wave 8)

Variables in QUITTING Section

hQA331gh - any attempts to stop smoking- derived (Wave 7 to Wave 8). The variable hQA331gh combines responses from *ggQA331* and *hQA331*. If respondents have any quit attempts either between Wave 7 and Wave 7.5 (*ggQA331=1*) or between Wave 7.5 and Wave 8 (*hQA331=1*), then *hQA331gh* is coded as 'yes' (1). Otherwise, *hQA331gh* is coded as their response from *hQA331*.

hQA561gh – number of new quit attempts after LSD (Wave 7)-derived. hQA561gh combines responses from *ggQA561* and *hQA561, which both count number of new quit attempts since the previous interview*.

- If both *hQA561* and *ggQA561* have valid value (not missing, "NA", "refused" or "don't know"), then *hQA561gh=hQA561+ggQA561*;
- If *hQA561* has a valid value, but *ggQA561* does not have a valid value, then *hQA561gh=hQA561*;
- If *ggQA561* has a valid value, but *hQA561* does not have a valid value, then *hQA561gh=ggQA561*;
- If both *hQA561* and *ggQA561* do not have valid values, then *hQA561gh="NA"*.

hQA442vgh – number of days since start of most recent quit attempt-- composite, # days (Wave 7 to Wave 8). The variable *hQA442vgh* combines information from *hQA442v* and *ggQA442v*. Respondents who were quitters in both the Wave 7.5 & Wave 8 survey and had been quit the entire time since Wave 7.5 (*hFR309v >3* and *ggFR309v>3* and *hQA337=2* and *hQA341=1*) were not eligible to answer *QA442v* in wave8. For them, *hQA442vgh* is computed by the sum of *ggQA442v* and the number of days between the Wave 7.5 survey date and the Wave 8 survey date. For other respondents, if *hQA442v* does not have a valid value, but *ggQA442v* does, then *hQA442vgh* is set equal to *ggQA442v*. Otherwise, *hQA442vgh* is set equal to *hQA442v*.

hQA553vgh - longest time smoke-free since LSD -- composite in days (Wave 7 to Wave8). The variable *hQA553vgh* is recalculated from *hQA551a-d, hQA442vgh* and *hQA514v* following the algorithm in the derived variable document.

Variables in STOP-SMOKING MEDICATIONS Section

hSM111gh - used any quitting meds-derived (Wave 7 to Wave 8). The variable *hSM111gh* combines information from *hSM111* and *ggSM111*. If respondents said 'Yes' to *ggSM111* and 'No' to *hSM111*, then *hSM111gh* was coded as 'Yes' (1). If *hSM111* is 'NA', then *hSM111gh* is coded as their response in *ggSM111*. Otherwise, *hSM111gh* is coded as their response in *hSM111*.
Note: The variable *SM111* is a filter for some subsequent questions which ask about use in the previous time period.

hSM161gh - SSM used to quit-derived. The variable *hSM161gh* combines information from *hSM161* and *ggSM161*. If respondents said 'Yes' to either *hSM161* or *ggSM161*, then *hSM161gh* was coded as 'Yes'. If *hSM161*. If *hSM161* has invalid value, then *hSM161gh* is coded as their response in *ggSM161*. Otherwise, *hSM161gh* is coded as their response in *hSM161*.

hSM161vgh - SSM used to quit-derived. The variable *hSM161vgh* is recalculated from *hSM111vgh* and *hSM161gh* following the algorithm in the derived variable document.

hSM201gh, hSM202gh, hSM203gh, hSM204gh, hSM210gh, hSM211gh, hSM212gh, hSM215gh, SM220gh, hSM500gh, hSM550vgh, hSM552gh, hSM554, hSM350gh, hSM352gh, hSM361gh, hSM362gh, hSM363gh, hSM364gh, hSM365gh, hSM366gh, hSM370gh. If the information in Wave 8 was available, then the derived variable is coded as their response in Wave 8. Otherwise, it is coded as their response in Wave 7.

hSM160gh, hSM162gh, hSM163gh, hSM164gh, hSM260gh, hSM262gh, hSM263gh, hSM259gh, hSM265gh, hSM266gh, hSM267gh, hSM268gh, hSM269gh, hSM270vgh, hSM310gh, hSM311gh, hSM314gh, hSM312gh, hSM313gh, hSM319gh, hSM315gh, hSM316gh, hSM317gh, hSM318gh, hSM329gh, hSM330vgh, hSM345gh, hSM458gh and hSM459gh, hSM380gh, SM462gh, hSM464gh, hSM538gh, hSM544gh, hSM373gh, hSM375gh. If the answers in previous variables meet the condition of wave 8, then *hSM*xxx*gh* is coded as their response in Wave8, if the answers in previous variables don't meet the condition of Wave 8, but meet the condition of Wave 7, then *hSM*xxx*gh* is coded as their response in Wave 7. Otherwise, *hSM*xxx*gh* is coded as "NA".

hSM401gh, hSM402gh, hSM403gh, hSM404gh, hSM410gh, hSM411gh, hSM412gh, hSM420gh, hSM421gh, hSM424gh, hSM422gh, hSM423gh, hSM429gh, hSM435gh, hSM436gh, hSM437gh, hSM438gh, hSM119gh, hSM121gh, hSM122gh,hSM123gh, hSM124gh, hSM130gh, hSM135gh, hSM140gh,hSM498gh, hSM499gh, hSM501gh, hSM502gh, hSM502gh, hSM504gh, hSM510gh, hSM511gh, hSM512gh, hSM520gh, hSM521gh, hSM524gh, hSM522gh, hSM523gh, hSM529gh, hSM535gh, hSM536gh, hSM537gh . If respondents said 'Yes' to either hSMxxx or *ggSM*xxx, then *hSM*xxx*gh* is coded as 'Yes' (1). If *hSM*xxx has invalid value (missing, "NA", "refused" or "don't know"), then *hSM*xxx*gh* is coded as their response in *ggSM*xxx. Otherwise, *hSM*xxx*gh* is coded as their response in *hSM*xxx.

Variables in CESSATION HELP

hCH801gh, hCH811gh, hCH811vgh, hCH812gh,hCH821gh, hCH821vgh, hCH813gh, hCH813vgh, hCH815gh, hCH815vgh , hCH817gh, hCH817vgh, hCH861gh, hCH865gh, and *hCH869gh.* These cessation help variables are followed the same algorithm to combine information from *hCH*xxx and *ggCH*xxx . If respondents said 'Yes' to either *hCH*xxx or *ggCH*xxx, then *hCH*xxxgh was coded as 'Yes' (1). If *hCH*xxx has invalid value (missing, "NA", "refused" or "don't know"), then *hCH*xxxgh is coded as their response in *ggCH*xxx. Otherwise, *hCH*xxxgh is coded as their response in h*CH*xxx.

Appendix 2: Analytic Dataset in ITC 4C Wave 1-8

Data Management Core - ITC Project
University of Waterloo
March, 2014

About ITC 4C Analytic Datasets

The ITC 4C analytic datasets include

- Itc4_core
- Itc4_formats
- Itc4_formats_a
- Itc4_spss_formats_a
- Itc4_wave1
- Itc4_wave2
- Itc4_wave3
- Itc4_wave4
- Itc4_wave5
- Itc4_wave6
- Itc4_wave7
- Itc4_wave7.5
- Itc4_wave8
- Itc4_wave8.5

Derived Variables

A few derived variables have been added to the analytic dataset.

- QA106v: how many times has respondent ever tried to quit. If the respondent has never tried to quit, QA106v is set to 0.

- QA112v: longest time off smoking in days (derived). It is calculated by combing QA111a, QA111b, QA111c, QA111d and QA111e. If the longest time off smoking is less than a day, QA112v is set to be 0. If it is more than 30 days, QA112v is rounded to be an integer.

- QA232v: number of days since the end of most recent quit attempt (derived). It is calculated by combing QA231a, QA231b and QA231c if they are valid and meant to be added.

- QA239v: whether respondent has any attempt to quit smoking in the last year (derived). If the respondent does not smoke (QA101=2), QA239v is set to be 0. If the respondent has quit smoking (eFR309v>3), QA239v is set to be 1. In other cases, if QA232v is less than or equal to 365 (days), QA239v is 1.

- QA242v: time of longest quit attempt ending in last 6 months in days (derived). It is calculated by combining QA241a, QA241b and QA241c if they are valid and meant to be added.

- QA442v: number of days since start of last quit attempt (derived). QA442v is set to be the time duration between survey date and the time of last quit (which is gotten by combining QA441e, QA441d and survey year) if the latter is non-missing. Otherwise, if at least one in QA441a, QA441b and QA441c is not missing, QA442v is equal to the sum of QA441a, QA441b and QA441c (with proper transformation). In other cases, QA442v is missing.

- QA553v: longest time smoke-free since LSD in days (derived). QA553v is derived from QA551a, QA551b, QA551c and QA551d by taking the sum (with proper transformation).

- QA662v: number of days since quit ended (since LSD) (derived). If both the survey date and the date when last quit ended (combine QA661d, QA661e and QA661f) are available, QA662v is the time duration between the survey date and the date of quit ended. Otherwise QA611a, QA611b and QA611c are combined to get QA662v.

- BQ146v: whether respondent have set a firm date to quit smoking (derived). BQ146v is derived from BQ146.

- WL221v: whether respondent read or looked closely at labels in LM (derived). If WL201=1 (did not notice labels in LM), WL221v=1 (did not read or looked closely at labels in LM). If WK201 is missing, WL221v is also set to be missing. In other cases, WL221 is equal to WL211 if it is valid.

- PU603o: other number of packs purchased. It derived from PU222o.

- PU605o: other number of cigarettes per pack. It derived from PU227o

- BR502v: how long respondent smokes the brand in months. BR502v is derived by taking the sum of BR501a, BR501b, BR501c and BR501d (with proper transformation).

- BR711v: last brand purchased (derived; CA). If BR701=1 (respondent purchased brand same as brand smoked most), BR711v is set to be the same as BR311. Otherwise, BR712v is the same as BR711.

- BR712v: Last brand purchased (derived; CA). If BR701=1 (respondent purchased brand same as brand smoked most), BR712v is set to be the same as BR312. Otherwise, BR712v is the same as BR712.

- BR322v: brand respondent smoke most (derived; US)

- BR322o: which other brand respondent smoke most (derived; US)

- BR722v: Last brand purchased (derived; US)

- BR722o: which other brand last purchased (derived; US)

- BR731v: last brand purchased (derived; UK). If BR701=1 (respondent purchased brand same as brand smoked most), BR731v is set to be the same as BR331. Otherwise, BR731v is the same as BR731.

- BR741v: last brand purchased (derived; AU). If BR701=1 (respondent purchased brand same as brand smoked most), BR741v is set to be the same as BR341. Otherwise, BR741v is the same as BR741.

- DE512v: ethnic background (derived; CA). DE512v combines DE511, DE513, DE515, DE517, DE521, DE525, DE533 and DE541 into one variable. The response options for DE512v are:
    - 1=White
    - 2=Chinese, Filipino, Southeast Asian, West Asian, Korean
    - 3=South Asian
    - 4=Black
    - 5=Latinamer
    - 6=Arab
    - 7=Aboriginal
    - 9=Other response

- DE572v: ethnic background (derived; US). DE572v combines DE551, DE553, DE555, DE557, DE559 and DE571 into one variable. The response options for DE572v are:
    - 1=White
    - 2=Black
    - 3=Hispanic
    - 4=Asian
    - 5=Native
    - 6=more than 1 ethnic group
    - 7=Other

- DE612v: ethnic background (derived; UK). DE612v combines DE611 and DE616 into one variable. The response options for DE612v are:
    - 1=White
    - 2=Chinese or other Asian
    - 3=Indian, Pakistani and Bangladeshi
    - 4=Black, Black British
    - 5=Mixed
    - 6=Other

- NR811v/CH811v: whether respondent received advice from doctor to quit, overall. It is derived from NR811/CH811.

- NR813v/CH813v: whether respondent received referral from doctor to quit, overall. It is derived from NR813/CH813.

- NR815v/CH815v: whether respondent received quitting Rx from doctor to quit, overall. It is derived from NR815/CH815.

- NR817v/CH817v: whether respondent received quitting pamphlet from doctor to quit, overall. It is derived from NR817/CH817.

- NR821v/CH821v: whether respondent received support from doctor for having quit, overall. It is derived from NR821/CH821.

- NR823v/CH823v: whether respondent received referral from doctor to help stay quit, overall. It is derived from NR823/CH823.

- NR825v/CH825v: whether respondent received quitting RX from doctor, overall. It is derived from NR825.

- NR827v/CH827v: whether respondent received pamphlet from doctor on how to stay quit, overall. It is derived from NR827/CH827.

- NR405v: how long used/using nicotine water in days. If NR405 is valid, NR405v is equal to NR405. Otherwise NR405v is set to be NR405a, NR405b or NR405c (transformed to days).

- NR415v: how long used/using gum in days. If NR415 is valid, NR415v is equal to NR415. Otherwise NR415v is set to be NR415a, NR415b or NR415c (transformed to days).

- NR425v: how long used/using patch in days. If NR425 is valid, NR425v is equal to NR425. Otherwise NR425v is set to be NR425a, NR425b or NR425c (transformed to days).

- NR435v: how long used/using lozenges in days. If NR435 is valid, NR435v is equal to NR435. Otherwise NR435v is set to be NR435a, NR435b or NR435c (transformed to days).

- NR445v: how long used/using tablets in days. If NR445 is valid, NR445v is equal to NR445. Otherwise NR445v is set to be NR445a, NR445b or NR445c (transformed to days).

- NR455v: how long used/using inhaler in days. If NR455 is valid, NR455v is equal to NR455. Otherwise NR455v is set to be NR455a, NR455b or NR455c (transformed to days).

- NR465v: how long used/using spray in days. If NR465 is valid, NR465v is equal to NR465. Otherwise NR465v is set to be NR465a, NR465b or NR465c (transformed to days).

- NR475v: how long used/using Zyban in days. If NR475 is valid, NR475v is equal to NR475. Otherwise NR475v is set to be NR475a, NR475b or NR475c (transformed to days).

- NR485v: how long used/using Wellbutrin in days. If NR485 is valid, NR485v is equal to NR485. Otherwise NR485v is set to be NR485a, NR485b or NR485c (transformed to days).

- NR495v: how long used/using other methods in days. If NR495 is valid, NR495v is equal to NR495. Otherwise NR495v is set to be NR495a, NR495b or NR495c (transformed to days).

- SB082v: how long between puffs in seconds. If SB082a is valid, SB082v is equal to SB082a. Otherwise, SB082v is equal to SB082b (transformed to seconds).

Non-response Codes

In 4C analytic datasets, instead of using 7, 8, 9 or 77, 88, 99… to indicate "not applicable", "refused to answer" and "do not know", ".n", ".r" and ".d" are used in SAS datasets.

Appendix 3: Available Analytic Datasets

Table A2. Available Analytic Datasets

Country	Wave
4C / ITC4	**Wave 1** **Wave 2** **Wave 3** **Wave 4** **Wave 5** **Wave 6** **Wave 7** **Wave 8**
Germany	**Wave 1** **Wave 2** **Wave 3**
Kenya	**Wave 1**
Mauritius	**Wave 1** **Wave 2** **Wave 3**
Mexico	**Wave 1**
Netherlands	**Wave 1** **Wave 2** **Wave 3** **Wave 4** **Wave 5** **Wave 6** **Wave 7** **Wave 8**
South Korea	**Wave 1**

	Wave 2 **Wave 3**
Uruguay	**Wave 1** **Wave 2** **Wave 3** **Wave 4**