ITC Netherlands Survey Team
Dr. Marc Willemsen (Principal Investigator, University of Maastricht)
Dr. Gera Nagelhout (University of Maastricht)
Dr. Hein de Vries (University of Maastricht)
Karin Hummel (University of Maastricht)
Dr. Ciska Hoving (University of Maastricht)
Dr. Bas van den Putte (University of Amsterdam)
Matty Crone (Leiden University Medical Centre)

International Planning Group
Dr. Geoffrey T. Fong (Principal Investigator, University of Waterloo)
Dr. Mary Thompson (Dept. of Statistics and Actuarial Science, University of Waterloo)
Dr. Christian Boudreau (Dept. of Statistics and Actuarial Science, University of Waterloo)
Grace Li (Data Analyst, Dept. of Psychology, University of Waterloo)
Sophie He (Data Analyst, Dept. of Psychology, University of Waterloo)
Thomas Agar (Project Manager Wave 8, Dept. of Psychology, University of Waterloo)
Lorraine Craig (Project Manager Waves 1-7, Dept. of Psychology, University of Waterloo)
Dr. Anne C. K. Quah (Research Scientist, Dept. of Psychology, University of Waterloo)

Funding
The ITC Netherlands Project has received funding support from:
- ZonMw (The Netherlands Organisation for Health Research and Development)
- Dutch Cancer Foundation (KWF)

Suggested Citation
## Contents

Preface to ITC Netherlands Technical Report Waves 1-8 ...................................................... iv

I. Introduction .......................................................................................................................... 1
   1.1 Background .................................................................................................................. 1
   1.2 Main Objectives ......................................................................................................... 1
   1.2 Survey Design ............................................................................................................ 2
   1.3 The Research Team .................................................................................................. 3

2. Survey Development and Content .................................................................................. 6
   2.1 Survey Development and Translation ....................................................................... 6
   2.2 Types of Survey Instruments .................................................................................... 6
   2.3 Survey Content ......................................................................................................... 6
This report documents the methodology used in the eight waves of the International Tobacco Control Policy Evaluation Survey carried out between 2008 and 2014 in The Netherlands. Survey waves were conducted approximately one year apart starting with Wave I in 2008.
I. Introduction

I.1 Background

The International Tobacco Control (ITC) Project is a multi-country prospective cohort study designed to measure the psychosocial and behavioural impact of key policies of the World Health Organization (WHO) Framework Convention on Tobacco Control (FCTC).

To evaluate the effect of the FCTC, the ITC Project is conducting parallel prospective cohort surveys with adult smokers in 22 countries—Canada, United States, Australia, United Kingdom, Ireland, Thailand, Malaysia, South Korea, China, New Zealand, Mexico, Uruguay, Germany, France, the Netherlands, Bangladesh, Brazil, Mauritius, Bhutan, India, Kenya, and Zambia.

The ITC Netherlands Survey has been conducted over nine waves. The information contained in this report, however, relates to the first eight waves of the ITC Netherlands Project.

- Wave 1: March - April of 2008
- Wave 2: November - December 2008
- Wave 3: March – May 2009
- Wave 4: May - June 2010
- Wave 5: May – June 2011
- Wave 6: May – June 2012
- Wave 7: May – June 2013
- Wave 8: May – July 2014
- Wave 9: November – December 2015

I.2 Main Objectives

The objectives of the ITC Netherlands Survey are:

- To examine patterns of behaviour and opinions associated with the use of tobacco by adults in The Netherlands.

  The study will provide very detailed information about the behaviour and the opinions of smokers, as well as their consumption patterns and other important aspects of tobacco use.

- To examine the impact of specific tobacco control policies that have been, and will be, implemented in The Netherlands from 2008 onward.
The ITC survey has several sections that are intended to evaluate the impact of certain tobacco control policies, such as the health warnings on cigarette packages, advertising campaigns that promote quitting, and cigarette tax increases. As a result, the survey will be able to examine the extent to which policies can change smoking behaviour and attitudes toward smoking.

- **To compare smoking behaviour and the impact of policies between The Netherlands and other ITC countries.**

The ITC Survey asks similar questions in more than 20 countries, thus allowing a comparison of the policies and the pattern of tobacco use between Uruguay and the other participating countries.

### I.2 Survey Design

The ITC Netherlands Survey is a national survey conducted by the Maastricht University in Maastricht, The Netherlands, in collaboration with the ITC Netherlands Project team, based at the University of Waterloo, Canada.

The ITC Survey is a longitudinal cohort study. Thus, the respondents who participated in previous waves were re-contacted for a follow-up survey in subsequent waves. Respondents who had been recruited as smokers, but who had quit smoking by the time of the next wave, were still surveyed using a ‘Quitter’ version of the survey instrument. Any respondents who were lost to follow-up or unable to complete the subsequent survey wave were replaced with new, randomly selected respondents from within the sample areas. Eligible new respondents in The Netherlands were cigarette smokers (smoked at least one cigarette in the last week) aged 15 years and older who have smoked more than 100 cigarettes in their lifetimes.

The ITC Netherlands Survey contains questions that examine the impact of key policies of the FCTC. These questions consist of common measures of tobacco use found in across ITC country surveys; including historical and current behaviour, addiction/dependence, beliefs and attitudes about smoking, knowledge about smoking related diseases, and perceived risk. Additional information about the survey content is provided in Chapter 2.

A detailed description of the sampling weights and weights construction the latter part of this report.

Figure 1 shows the timeline of the ITC Netherlands Project. It shows the dates of selected tobacco control policies implemented in The Netherlands since 2005, as well as the ITC Survey fieldwork dates.
1.3 The Research Team

The ITC Netherlands Survey has been conducted in Netherlands by researchers from the Maastricht University, the University of Amsterdam (ASCoR), Leiden University Medical Centre, and the former STIVORO (the Dutch Expert Centre on Tobacco Control). The research team in The Netherlands collaborates with the team in the University of Waterloo in Canada.
Figure 1. ITC Netherlands Project Timeline

**NETHERLANDS**
Timeline of Tobacco Control Policies and ITC Surveys

**2005**
- Jan 2005: FCTC ratification.
- July 2008: Ban suspended in small owner run cafes.

**2006**
- April 2008 - Jan 2009: Mass media campaign to encourage quitting smoking.

**2007**
- July 2008: Smoking ban in hospitality industry, public transportation, workplaces, and sport venues.
- Smoking rooms allowed.
- €0.25 tax increase on cigarettes and €0.31 increase on rolling tobacco.

**2008**
- Jan 2011: Reimbursement of smoking cessation treatments accompanied by a mass media campaign to encourage quitting smoking from Dec 2010 to Mar 2011.

**2009**
- Feb 2010: Ban reinstated in small owner run cafes.
- Nov 2010: Dutch gov't announced intention to reverse smoking ban in small bars and cafes.
- Nov 2011: €0.22 tax increase on cigarettes.

**2010**
- July 2011: Ban suspended in small owner run cafes and bars (< 70 sq. m or 750 sq. feet).
- Merul: 2011: €0.35 tax increase on cigarettes, €0.50 tax increase on shag.
- Reimbursement of smoking cessation treatments reinstated.

**2011**
- Feb 2013: Ban to be reinstated in hospitality sector in 2015.

**Survey Mode:** Telephone (CATII), Web (CAWI)
**Respondent Types:** Smoker

*Updated April 2016*
NETHERLANDS
Timeline of Tobacco Control Policies and ITC Surveys

Jan 2014
Legal age to buy tobacco increased from 16 to 18, accompanied by a mass media campaign to de-normalize smoking among youth.

July 2014
Bill to ban smoking in small bars and cafés still waiting Parliamentary approval, potential enforcement by January 2015.

May 2016 - TPD 2014
- Content yields removed, instead to carry 50% general HW’s on sides.
- Combined HW’s increased from 35% to 65%.
- Combined HW’s to display: text warning, pictorial warning and cessation info.
- HW’s to be rotated per annum.
- HW’s min. dimensions increased to 44x52mm, effectively restricting lipstick-style packs.
- Herbal tob. packages to carry one-sided 30% HW’s.
- Min. combined HW’s increased from 15% to 35% for all tob. products.
- Characterizing flavours banned.
- Extensive list of additives banned.
- Flavoured components banned.

Nov. 20, 2016 - TPD 2014
- Max. nicotine concentration for e-liquids set at 20mg/ml.
- HW’s required on e-cigs & e-liquid packs, 30% of front and back. HW’s are specific to e-cigs.
- Comprehensive advertising ban for e-cigs & e-liquids.

May 2019 - TPD 2014
- Packaging banned from using U/M descriptors and vitality claims.
- Unit packs of cigarettes to be 20 ± units.
- Unit packs of RYO tobacco to be 30 ± grams.
- Factory-made cig. and RYO tobacco packs to include unique identifiers to mitigate illicit trade and employ track & trace measures.

May 2020 - TPD 2014
- Products with a market share ≥ 3% are required to comply with characterizing flavour, additive and flavoured component ban.

May 2024 - TPD 2014
- All tobacco products to include unique identifiers to counter illicit trade and employ track & trace measures.

Survey Mode: Telephone (CATI), Web (CAWI)
Respondent Types: Smoker

NL Gold Magic (GM):
Survey Mode: Web (CAWI)
Respondent Types: Smoker & Non-Smoker

Updated April 2016
2. Survey Development and Content

2.1 Survey Development and Translation

The Wave 1 ITC Netherlands Survey questionnaire was based on that of the ITC 4-Country and ITC Germany Surveys. The surveys were revised in the original English and were translated into Dutch. The translated surveys were then reviewed by team members who were bilingual in English and Dutch. Some revisions were made at each following wave to best capture the information that was important at the respective time frames.

2.2 Types of Survey Instruments

Following the first Wave, two versions of the questionnaire were used for fieldwork from Waves 2-8. One was designated for those who had already participated in the previous wave: The Recontact Survey, including questions for both smokers and those who quit after the previous Wave. The second was for those who were being recruited to replace the respondents from the previous wave who were lost to attrition: the Replenishment Questionnaire.

Copies of all ITC Netherlands Survey questionnaires are available at www.itcproject.org. The two survey instrument types, the participant type to whom the survey would be administered, and the average length of each type of survey is provided in Table 5.

Table 5: Wave 8 Survey Characteristics

<table>
<thead>
<tr>
<th>Types of Survey at Waves 2 through 8</th>
<th>Participant Characteristics</th>
<th>Average Time (Mins)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Recontact Survey</td>
<td>Smokers who participated in previous waves and were still smoking at the time of subsequent waves, as well as ex-smokers who participated in previous waves as smokers.</td>
<td>30-55</td>
</tr>
<tr>
<td>2. Replenishment Smoker Survey</td>
<td>Smokers who were newly recruited into the cohort at later waves to replace a participant from a previous wave who had dropped out or become ineligible.</td>
<td>55</td>
</tr>
</tbody>
</table>

2.3 Survey Content

Respondents who were smokers were asked the following types of survey questions:
1. Smoking- and cessation-relevant questions: smoking history and frequency, as well as current smoking behaviour and dependence, and quitting behaviours.

2. Knowledge and basic beliefs about smoking: knowledge of the health effects of smoking and important beliefs relevant to smoking and quitting, perceived risk, and perceived severity of tobacco-related diseases.

3. Policy-relevant questions: awareness of, impact of, and beliefs relevant for each of the FCTC demand reduction policy domains (warning labels, taxation/price, advertising/ promotion, smoke-free policies, light/mild descriptors).

4. At Wave 7, additional questions on the use of electronic cigarette devices (e-cigarettes) were added.

5. Other important psychosocial predictors of smoking behaviour and potential moderator variables (e.g., normative beliefs, self-efficacy, intentions to quit).

6. Individual difference variables relevant to smoking (e.g., depression, stress, time perspective).

6. Demographics (e.g., age, gender, marital status, income, education).

Respondents who were ex-smokers were asked parallel survey questions from the categories listed above. Question phrasing was revised where necessary for the ex-smoker context.

Replenishment smokers were questions that were very similar to the Recontact smokers.

Between Waves 1 to 8, each questionnaire type was updated to ensure that it was relevant for the target respondent within the context of the tobacco control landscape in the Netherlands.
Sampling Design and Weight Construction for the International Tobacco Control (ITC) Netherlands Survey

C. Boudreau¹,², M. Thompson¹,² and Y. Li²,³
Waves 1–8 (Sep. 2015)

This technical report details the sampling design and weight construction for waves 1–8 of the International Tobacco Control (ITC) Netherlands Survey. The ITC Netherlands Survey is a prospective longitudinal survey of a nationally representative random sample of youth (ages 15–17) and adult (18 years and older) smokers. More than 2200 smokers were first interviewed in Mar.–Apr. 2008, and the most recent wave took place in May-Jun. 2014. Respondents lost to follow-up between waves 2 and 8 were replenished by new randomly selected respondents.

This technical report is organized as follows: Section 1 describes the sampling design of the ITC Netherlands Survey, and section 2 details the construction of the sampling weights from wave 1 (section 2.2), to wave 8 (section 2.9).

1 Sampling design

The ITC Netherlands Survey is a prospective longitudinal study, and its sampling design was chosen to yield a representative random sample of smokers residing in that country. Fieldwork has been conducted by the Dutch survey firm, TNS NIPO, and used a dual sampling frame. The first frame consists of the traditional random digit dialling (RDD) telephone survey, with computer assisted telephone interviews (CATI). The second frame is the web portion of the TNS NIPObase, which consists of over 140000 respondents who have agreed to participate in TNS NIPO research on a regular basis. All respondents sampled from the second frame were invited to complete the Survey using computer assisted web interviews (CAWI). It should be mentioned that members of the TNS NIPObase were randomly selected (mostly by mail and RDD), and are thus not a panel of self-selected volunteers.

¹Dept. of Statistics & Actuarial Science, University of Waterloo, Waterloo, Ontario, Canada.
²Data Management Core (DMC) – ITC Project, University of Waterloo.
³Dept. of Psychology, University of Waterloo.
1 SAMPLING DESIGN

1.1 Wave 1

At wave 1, a sample of 404 respondents was selected from the RDD frame, whereas 1820 respondents were sampled from the web frame (i.e., the TNS NIPObase); for a total of 2224 respondents.

1.1.1 RDD frame

The sampling design called for the population to be stratified into 12 geographic strata corresponding to the 12 Dutch provinces; see figure 1. The numbers of respondents or quotas to be sampled in each of the strata were assigned using proportional allocation to the estimated size of the adult population in each of the strata. Within each stratum, households were randomly called using RDD and adult smokers interviewed until the corresponding quota was met. This process was repeated independently for each of the 12 strata. To qualify for the RDD sample, respondents must be 18 years or older, have smoked more than 100 cigarettes in their life and smoked at least once in the 30 days prior to being recruited. In households with multiple eligible respondents, the Next-Birthday method (Binson et al. (2000)) was used to select a single one. No substitution within household was allowed, except when it was known that the selected respondent would be absent for the entire fieldwork period.

1.1.2 Web frame

The TNS NIPObase was first screened to identify individuals that met the criteria for the web sample; i.e., are 15 years or older, have smoked more than 100 cigarettes in their life and smoked at least once in the 30 days prior to being recruited. This yielded a sampling frame of about 40000 potential respondents. This corresponds to a smoking prevalence of about 28%, which is in agreement with current smoking prevalence estimates for the Netherlands.

The sampling design then called for the population to be stratified into 2 strata: 15 ≤ age ≤ 30 and age > 30. Quotas were then assigned to each stratum so that young smokers were over-sampled and formed about 44% of the web sample; whereas, they only represented about 25% of Dutch smokers. Within each stratum, respondents were randomly selected from the screened TNS NIPObase until the corresponding quota was met. As mentioned before, they were then invited to complete a computer assisted web interview (CAWI), which was referred to as computer assisted self interview (CASI) at wave 1.
1 SAMPLING DESIGN

Figure 1: Strata of the ITC Netherlands Survey RDD frame.

1.2 Wave 2

Of the 1820 respondents who completed the wave 1 CAWI survey, a subset of 744 (337 aged \( \leq 30 \) and 407 aged \( > 30 \)) were randomly selected to complete the wave 2 survey. Out of these, 643 (310 aged \( \leq 30 \) and 333 aged \( > 30 \)) were successfully recontacted; yielding a retention rate of 86.4\% (92.0\% for those aged \( \leq 30 \), and 81.2\% for those aged \( > 30 \)). None of the wave 1 respondents who completed the CATI survey (i.e., RDD sample) were interviewed at wave 2. Figure 2 shows the attrition and replenishment of the various ITC Netherlands samples over the first 6 waves of the Survey.
1 SAMPLING DESIGN

1.3 Wave 3

Of the 404 respondents who completed the wave 1 CATI survey (i.e., RDD sample), 296 were successfully recontacted at wave 3, yielding a retention rate of 73.3%. In addition, 586 of the respondents who completed the wave 2 CAWI survey (i.e., Web sample) were successfully recontacted at wave 3; for a retention rate of 91.1%. Of the 1177 respondents who completed the wave 1 CAWI survey but not the wave 2 survey, 861 were successfully recontacted at wave 3; for a retention rate of 73.2%. These numbers are summarized in figure 2.

To compensate for the attrition of the Web sample, 270 additional respondents were randomly sampled from the screened TNS NIPObase; for a total of 2013 respondents interviewed at wave 3. This sample of 270 respondents is generally referred to as the wave 3 replenishment sample or cohort #3. As with other ITC surveys, replenishment for wave 3 of the ITC Netherlands Survey was carried out using the same sampling design and interview protocol as in wave 1 (see section 1.1.2). Hence, the replenishment sample is representative of the population at the time of data collection (i.e., March – May 2009), rather than those lost to follow-up/attrition. The RDD sample was not replenished at wave 3.

1.4 Wave 4

Respondents from the RDD frame were not contacted after wave 3. Of the 1717 respondents who completed the wave 3 CAWI survey, 1399 were successfully recontacted at wave 4; yielding a retention rate of 81.5%. In addition, 77 respondents who completed the wave 1 CAWI survey, but not the wave 2 survey, were successfully recontacted at wave 4. To compensate for this attrition and the end of follow-up for all respondents from the RDD frame, 579 new respondents were randomly sampled from the screened TNS NIPObase; for a total of 2060 respondents interviewed at wave 4. This sample of 579 respondents is generally referred to as the wave 4 replenishment sample or cohort #4. As with other ITC surveys, replenishment for wave 4 of the ITC Netherlands Survey was carried out using the same sampling design and interview protocol as in wave 1 (see section 1.1.2). Hence, the replenishment sample is representative of the population at the time of data collection (i.e., May – June 2010), rather than those lost to follow-up/attrition.

1.5 Wave 5

Of the 2060 respondents who completed the wave 4 survey, 1567 were successfully recontacted at wave 5. In addition, 52 respondents (43 recruited at wave 1 and 9 recruited at wave 3) who
completed the wave 3 survey, but skipped the wave 4 survey, were successfully recontacted at wave 5, totally yielding a retention rate of (1619/2060) or 78.6%. To compensate for attrition, 482 new respondents aged between 18 and 30 years old were randomly sampled from a newly screened pool of respondents in the TNS NIPObase, for a total of 2101 respondents interviewed at wave 5. This sample of 482 respondents is generally referred to as the wave 5 replenishment sample or cohort #5. Again, the replenishment sample is representative of the population at the time of data collection (i.e., May – June 2011), rather than those lost to follow-up/attrition.

1.6 Wave 6

Of the 2101 respondents who completed the wave 5 survey, 1736 were successfully recontacted at wave 6; yielding a retention rate of 82.6%. To compensate for attrition, 286 new respondents were randomly sampled from the screened TNS NIPObase; for a total of 2022 respondents interviewed at wave 6. This sample of 286 respondents is generally referred to as the wave 6 replenishment sample or cohort #6. Again, the replenishment sample is representative of the population at the time of data collection (i.e., May – June 2012), rather than those lost to follow-up/attrition.

1.7 Wave 7

Of the 2022 respondents who completed the wave 6 survey, 1677 were successfully recontacted at wave 7; yielding a retention rate of 82.9%. To compensate for attrition, 293 new respondents were randomly sampled from the screened TNS NIPObase; for a total of 1970 respondents interviewed at wave 7. This sample of 293 respondents is generally referred to as the wave 7 replenishment sample or cohort #7. Again, the replenishment sample is representative of the population at the time of data collection (i.e., May – June 2013), rather than those lost to follow-up/attrition.

1.8 Wave 8

Of the 1970 respondents who completed the wave 7 survey, 1607 were successfully recontacted at wave 8; yielding a retention rate of 81.6%. To compensate for attrition, 404 new respondents were randomly sampled from the screened TNS NIPObase; for a total of 2011 respondents interviewed at wave 8. This sample of 404 respondents is generally referred to as the wave 7 replenishment sample or cohort #8. Again, the replenishment sample is representative of the population at the time of data collection (i.e., May – June 2014), rather than those lost to follow-up/attrition.
Figure 2: Attrition and replenishment in the ITC Netherlands Survey.
2 CONSTRUCTION OF THE SURVEY WEIGHTS

Note: In Figure 2, the number in a box is the number of members of the specified cohort present in the data set in the specified wave. Curved arrows indicate situations where at least one respondent rejoined the sample after dropping out for one or more waves.

2 Construction of the survey weights

2.1 General comments about weight construction

As with most survey weights, the ITC Netherlands weights are constructed to correct and adjust for sample mis-representation caused by unequal sampling probabilities, frame error (i.e., under-coverage and multiplicity) and non-response, as well as improving precision of estimates through the use of auxiliary information (e.g., smoking prevalences). In addition, conservative weight trimming has been performed to prevent extreme weight variation arising from a few respondents having very large sampling weights. We briefly describe these key concepts of weight construction in this section, but refer the reader to Levy & Lemeshow (2008), chapter 16, for more detailed information.

At their base, sampling weights are defined as the inverse of selection probabilities, and thus adjust for sample mis-representation caused by unequal sampling probabilities. For example, a smoker residing alone has a probability of selection twice that of a smoker residing with another smoker.

Great efforts are made to create a complete/perfect sampling frame (i.e., a frame that includes all members of the target population, without duplication and without any erroneous inclusions\(^1\)). However, this is seldomly achieved and, consequently, some members of the target population are not part of the sampling frame (i.e., have a 0 probability of being selected). This is referred to as frame under-coverage, and can result in non-coverage bias. To reduce such bias in the ITC Netherlands Survey, post-stratification adjustments were performed on the sampling weights to ensure that, for each province and age/sex/district, totals of the sampling weights equal known benchmarks; see steps 4 & 5 in section 2.2.1. Note that these benchmark figures are also referred to as calibration or target figures, and thus such post-stratification adjustment is also referred to as weight calibration.

If non-respondents behave differently than respondents, then inference based solely on the sample of respondents will be biased unless adjustments are made. The greater the expected proportion of non-response, the greater this bias can be. In the ITC Netherlands Survey, the

---

\(^1\)Erroneous inclusions refers to units that are not part of the target population, but included in the sampling frame.
post-stratification adjustments described in the above paragraph also adjust for non-coverage bias. It should be noted that if data are missing completely at random (MCAR, see Little & Rubin (2002)) within each age/sex/district group, then non-response bias will be completely eliminated. Realistically though, non-response bias is greatly reduced, but not eliminated in the ITC Netherlands Survey.

The distribution of sampling weights is often skewed to the right, echoing the fact that most populations are composed of many average/typical members and of few atypical ones. Average members have a fairly high probability of selection, and thus most sampling weights are fairly small. There are however few members of the population that have a much smaller probability of selection, and consequently have sampling weights that are quite large. These few large weights can be the source of high weight variation, which increases the variability of estimators and thus decreases precision. To correct for this, large weights are often trimmed in the weight construction process. This must be done with care and conservatively, as trimming can increase bias. There are various ways of trimming sampling weights. In the ITC Netherlands Survey, trimming was done by capping the number of adult smokers in each household at 2 (see step 2 in section 2.2.1), as well as the number of phone lines at 2 (see step 1). Capping is a fairly conservative weight trimming technique and, since it is done at the beginning of weight construction, helps minimize potentially biasing estimates.

It is well known, from survey sampling theory that, in the vast majority of cases, the ratio estimator has much greater precision than the commonly used Horvitz-Thompson estimator. Heuristically, this is due to the fact that the ratio estimator utilizes auxiliary (i.e., additional) information in addition to the sampling weights, whereas the Horvitz-Thompson estimator does not. As mentioned above however, smoking prevalence figures were used to calibrate/benchmark the ITC Netherlands sampling weights in order to reduce biases from frame errors and non-response. Our calibrating procedure yields the (so-called) ratio weights, which enable all estimators to inherit the increased precision of the ratio estimator.

All weights for the ITC Netherlands Survey were computed using the statistical software R (http://www.r-project.org).

### 2.2 Wave 1 weights

Two sets of weights were computed at wave 1:

i) Section 2.2.1 describes the computation of the **cross-sectional wave 1 CATI weights** for the 404 respondents who completed the wave 1 survey via CATI.

ii) Section 2.2.2 describes the computation of the **cross-sectional wave 1 CAWI weights** for the 1820 respondents who completed the wave 1 survey via CAWI.


2 CONSTRUCTION OF THE SURVEY WEIGHTS

The RDD and Web samples were obtained through different sampling frames and designs, and thus require different weight calculation procedures. However, most analyses will utilize all respondents, instead of respondents from a single frame. Hence, the sampling weights of the 404 RDD respondents and those of the 1820 Web respondents were combined into a single variable, labelled aDE42919v. Since both the CATI and CAWI weights were rescaled to sum to sample size (see step 6 in section 2.2.1 and step 4 in section 2.2.2), variable aDE42919v is simply the set of all those rescaled weights.

It should be noted that all ITC Netherlands survey weights were calibrated to smoking prevalences (see step 5 of section 2.2.1) and rescaled to have a mean equal to 1 (see step 6 of section 2.2.1). Consequently, these weights should not be used to estimate population totals (e.g., the total number of daily smokers). However, all weights can obviously be used to estimate population means and proportions/percentages, as well as in various statistical models (e.g., logistic and linear regressions). Rescaling the weights also has the advantage of greatly simplifying comparisons between the various countries part of the ITC Project (http://www.itcproject.org).

2.2.1 CATI weights

The RDD sample was interviewed by CATI, and their sampling weights are thus referred to as the CATI weights or the RDD weights. Computation of these weights proceeded as follows.

Step 1: Each of the 404 respondents was first assigned an initial weight defined as

\[ w_i^{(1)} = \begin{cases} 
1 & \text{if his/her household has a single residential phone line} \\
1/2 & \text{if his/her household has multiple residential phone lines} 
\end{cases} \]

where \( i \) stands for the \( i^{th} \) respondent.

Step 2: These \( w_i^{(1)} \) weights were then adjusted for the number of adults smokers in the household. If the \( i^{th} \) respondent is the sole smoker in the household, then his/her \( w_i^{(1)} \) weight is multiplied by 1, otherwise it is multiplied by 2. The resulting weights are labelled \( w_i^{(2)} \), and can be summarized by the following formula:

\[ w_i^{(2)} = w_i^{(1)} \times \left(1 + I(\# \text{ smokers}_i \geq 2)\right) , \]

where \( \# \text{ smokers}_i \) is the number of adult smokers in the household of the \( i^{th} \) respondent and \( I(\cdot) \) is the indicator function.
Step 3: A post-stratification adjustment was then performed to calibrate the \( w_i^{(2)} \) weights to known family size benchmarked proportions. This consisted in multiplying the \( w_i^{(2)} \) weights by a factor \( f_k \) to yield the \( w_i^{(3)} \) weights, such that the sum of these \( w_i^{(3)} \) weights for a given family size divided by the sum of all \( w_i^{(3)} \) weights equals the corresponding proportion in table 1. In other words,

\[
w_i^{(3)} = w_i^{(2)} \times f_k,
\]

where

\[
f_k = \frac{p_k}{\sum_{i \in F_k} w_i^{(2)} / \sum_{i=1}^n w_i^{(2)}},
\]

with \( p_1 = 0.19, \ldots , p_4 = 0.28 \) and where \( F_k \) is the set of all respondents living in a family of size \( k \) (\( k = 1, \ldots , 4 \)).

<table>
<thead>
<tr>
<th>Family size</th>
<th>Proportion</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.19</td>
</tr>
<tr>
<td>2</td>
<td>0.34</td>
</tr>
<tr>
<td>3</td>
<td>0.19</td>
</tr>
<tr>
<td>4+</td>
<td>0.28</td>
</tr>
</tbody>
</table>

Table 1: Proportion of families of size 1 to 4+ in the Dutch population.

Step 4: A second post-stratification adjustment was performed, and the \( w_i^{(3)} \) weights were calibrated to the proportions of adults residing in each of the 12 Dutch provinces. Hence,

\[
w_i^{(4)} = w_i^{(3)} \times \frac{\text{prov}_k}{\sum_{i \in P_k} w_i^{(3)} / \sum_{i=1}^n w_i^{(3)}},
\]

where \( \text{prov}_1, \ldots , \text{prov}_{12} \) are given in table A.2 and \( P_k \) is the set of all respondents residing in province \( k \).

Step 5: The \( w_i^{(4)} \) weights were then calibrated to smoking prevalences by age/sex/district groups using the same post-stratification technique as in steps 4 & 5. To this end, age was divided into 5 groups: 15-17, 18-24, 25-39, 40-54 and 55+; whereas, provinces were grouped into 4 districts: East, West, North and South (see table A.1). Some of the resulting 40 age/sex/district cells contained too few respondents, and were collapsed with cells of similar characteristics. The resulting \( w_i^{(5)} \) weights thus sum
to estimated number of adult smokers in each of the 22 age/sex/district cells of table A.3. In other words,

\[ w_i^{(5)} = w_i^{(4)} \times \sum_{i \in C_k} w_i^{(4)} , \]

where cell1, . . . , cell22 are given in table A.3 and \( C_k \) is the set of all respondents in cell \( k \).

The prevalence figures of table A.3, were obtained by combining population estimates as of Jan. 1, 2008 from the Centraal Bureau voor de Statistiek (i.e., Statistics Netherlands) to prevalence estimates from the 2007 Continuous Survey of Smoking Habits.

Step 6: To facilitate comparisons across multiple ITC countries, the \( w_i^{(5)} \) weights were rescaled to have a mean equal to 1 or, equivalently, to sum to \( n = 404 \) (the number of respondents who completed the wave 1 CATI survey); that is,

\[ \sum_{i=1}^{n} w_i^{(6)} = \sum_{i=1}^{n} w_i^{(5)} \times RF = n = 404 , \]

where RF is the rescaling factor.

2.2.2 CAWI weights

Respondents from the Web sample were interviewed by CASI (later referred to as CAWI), and their sampling weights are thus referred to as the CAWI weights or Web weights.

TNS NIPO provided two sets of weights for the Web sample (which they labelled \textit{GEW1} and \textit{GEW2}). Both sets adjusted for unequal sampling probabilities, but were not calibrated to smoking prevalences. More importantly, both sets of weights had some adjustments to account for the fact that the TNS NIPObase is not an unbiased frame of the Dutch adult population; e.g., Web users tend to be younger males with a higher level of education and income. TNS NIPO partially corrected for this bias by post-stratifying their weights on age, sex, province of residence, family size and key variables related to web usage. Unfortunately, because of proprietary issues, it was impossible for the DMC to find the exact details of the latter adjustment on variables related to web usage. Nonetheless, these weight adjustments are valuable, and it was decided to build on NIPO’s weights instead of starting from scratch. We opted to build from the \textit{GEW2} weights, as that they presented slightly better properties than the \textit{GEW1} weights.
Starting with \( w^{(0)}_i \), the GEW2 weight for the \( i \)th respondent, computation of the CAWI weights proceeded as follows.

Step 1: The \( w^{(0)}_i \) were first adjusted to account for the oversampling of smokers aged \( \leq 30 \) (see section 1.1.2). This was done using a post-stratification method as the one described in step 3 of section 2.2.1. The resulting \( w^{(1)}_i \) weights are thus given by

\[
 w^{(1)}_i = w^{(0)}_i \times \frac{\text{age}_k}{\sum_{i \in A_k} w^{(0)}_i / \sum_{i=1}^n w^{(0)}_i},
\]

where \( \text{age}_k = 0.765 \) if the \( i \)th respondent is > 30 years old and \( \text{age}_k = 0.235 \) if his/her age \( \in [15, 30] \); and where \( A_k \) is the set of all respondents aged \( [15, 30] \) for \( k = 1 \) and \( (30, 100) \) for \( k = 2 \). The proportions of Dutch aged \( [15, 30] \) and > 30 was obtained from the population estimates, as of Jan. 1 2008, published by the Centraal Bureau voor de Statistiek.

Step 2: A post-stratification adjustment for province was carried out, same as step 4 of section 2.2.1.

Step 3: The weights were then calibrated to smoking prevalences by age/sex/district cells, same as step 5 of section 2.2.1.

Step 4: The weights were rescaled to sum to sample size \( n = 1820 \), same as step 6 of section 2.2.1.

### 2.3 Wave 2 weights

As mentioned in section 1.2, a sub-sample of 744 Web respondents were randomly selected to complete the wave 2 CAWI survey and 643 of them actually completed the survey. Because waves 1 and 2 were only 6 months apart, all the calibration figures used in steps 1–4 of section 2.2.2 were also used to compute the wave 2 weights. For this reason, the waves 1–2 longitudinal weights and the wave 2 cross-sectional weights are equal, and simply referred to as the wave 2 CAWI weights (variable \texttt{bDE42919v} - formerly \texttt{bDE42915v}). As mentioned at the beginning of section 2.2, all ITC Netherlands survey weights were calibrated to smoking prevalences and rescaled to have a mean equal to 1. Consequently, these weights should not be used to estimate population totals (e.g., the total number of daily smokers).

Starting with \( w^{(0)}_i \), the wave 1 CAWI weight for the \( i \)th respondent (computed in section 2.2.2), calculation of the wave 2 CAWI weights proceeded as follows.
2 CONSTRUCTION OF THE SURVEY WEIGHTS

Step 1: The weights were first adjusted for oversampling of smokers aged $\leq 30$ using post-stratification, same as step 1 of section 2.2.2.

Step 2: A post-stratification adjustment for province was carried out, same as step 4 of section 2.2.1.

Step 3: The weights were then re-calibrated to smoking prevalences by age/sex/district cells, same as step 5 of section 2.2.1.

Step 4: The weights were rescaled to sum to sample size $n = 643$, same as step 6 of section 2.2.1.

2.4 Wave 3 weights

Six sets of weights were computed for wave 3 data. Waves 1–3 longitudinal CATI weights were computed for the 296 RDD respondents recruited at wave 1 that were retained and interviewed at wave 3. Wave 3 cross-sectional CATI weights were also computed for those respondents. Waves 1–3 longitudinal CAWI weights were computed for the 1447 Web respondents recruited at wave 1 that were retained and interviewed at wave 3 (regardless whether or not they were interviewed at wave 2). Wave 3 main survey cross-sectional CAWI weights were also computed for those respondents. Waves 2–3 longitudinal CAWI weights (variable cDE42923v) were computed for the 586 Web respondents recruited at wave 1 that were also interviewed at both waves 2 & 3. Wave 3 replenishment cross-sectional CAWI/CATI weights were computed for the 270 respondents recruited at wave 3 from the Web frame.

As was done at wave 1, waves 1–3 longitudinal CATI and CAWI were combined into a single variable, labelled cDE42921v. Similarly, all wave 3 cross-sectional weights (i.e., CATI, CAWI for main survey and CAWI for new recruits) were combined into a single variable, labelled cDE42919v (formerly labelled cDE42915v). As mentioned at the beginning of section 2.2, all ITC Netherlands survey weights were calibrated to smoking prevalences and rescaled to have a mean equal to 1. Consequently, these weights should not be used to estimate population totals (e.g., the total number of daily smokers).

2.4.1 CATI weights

Starting with $w_i^{(0)}$, the wave 1 RDD/CATI weight for the $i^{th}$ respondent (computed in section 2.2.1), calculation of the waves 1–3 longitudinal CATI weights proceeded as follows.
2 CONSTRUCTION OF THE SURVEY WEIGHTS

Step 1: The $w_i^{(0)}$ weights were first re-calibrated to the proportions of adults residing in each of the Dutch provinces. This was done the same way as step 4 of section 2.2.1, with the exception that some provinces were collapsed because too many of their wave 1 respondents were lost to attrition; see footnote of table A.2.

Step 2: The weights were then re-calibrated to smoking prevalences by age/sex/district cells. This was done the same way as step 5 of section 2.2.1, with the exception that the 22 cells of table A.3 were collapsed into the 8 cells of table A.4 because of attrition.

Step 3: The weights were rescaled to sum to sample size $n = 296$, same as step 6 of section 2.2.1.

The wave 3 cross-sectional CATI weights were constructed the same way as the waves 1–3 longitudinal CATI weights, with the following 2 exceptions.

1. Two of the 296 respondents moved to another province between waves 1 & 3. For these respondents, $w_i^{(0)}$ is no longer the weight from the previous wave, but the average wave 1 weight of respondents in the same age-sex group living in their new province. For example, one of the two movers was a 42 year old female who moved from Friesland to Limburg. Her $w_i^{(0)}$ weight is thus the average wave 1 weight of all 40–55 year old female living in Limburg at wave 1.

   Note that for longitudinal weights, a respondent is always associated with the province in which he/she resided at the time of recruitment. Hence, no special adjustment is required. In addition, respondents who moved out of the Netherlands between waves are dropped from the study.

2. Updated smoking prevalences (by age/sex/district) figures were used to calibrate the weights. This was done so respondents are weighted to represent the Dutch adult smoking population at wave 3 and not at wave 1. This is essential when computing cross-sectional weights.

   The prevalence figures are given in table ??, and were obtained by combining population estimates as of Jan. 1, 2008 (same as in waves 1–2) to prevalence estimates from the 2008 Continuous Survey of Smoking Habits.

2.4.2 CAWI weights

Starting with $w_i^{(0)}$, the wave 1 CAWI weight for the $i^{th}$ respondent (computed in section 2.2.2), calculation of the waves 1–3 longitudinal CAWI weights proceeded as follows.
2 CONSTRUCTION OF THE SURVEY WEIGHTS

Step 1: The $w_i^{(0)}$ weights were first re-adjusted for oversampling of smokers aged $\leq 30$ using post-stratification, same as step 1 of section 2.2.2.

Step 2: A post-stratification adjustment for province was carried out, same as step 4 of section 2.2.1.

Step 3: The weights were then re-calibrated to smoking prevalences by age/sex/district cells, same as step 5 of section 2.2.1.

Step 4: The weights were rescaled to sum to sample size $n = 1447$, same as step 6 of section 2.2.1.

Computation of the waves 2–3 longitudinal CAWI weights proceeded as for the waves 1–3 longitudinal CAWI weights, with the exception that $w_i^{(0)}$ is the wave 2 CAWI weight for the $i^{th}$ respondent (computed in section 2.3).

Computation of the wave 3 main survey cross-sectional CAWI weights also proceeded as for the waves 1–3 longitudinal CAWI weights, with the exception that updated smoking prevalences (by age/sex/district) figures were used to calibrate the weights. These are the same figures as those used in section 2.4.1, though fewer cells needed to be collapsed; see table A.7.

Lastly, computation of the wave 3 replenishment cross-sectional CAWI weights proceeded along the same lines as the wave 1 CAWI weights (see section 2.2.2), with the following 2 exceptions.

1. As with the wave 3 CATI weights of section 2.4.1 some provinces were collapsed because they contained too few respondents.

2. Updated smoking prevalences (by age/sex/district) figures were used to calibrate the weights, as these respondents are weighted to represent the Dutch adult smoking population at wave 3.; see table A.7.

2.5 Wave 4 weights

Five sets of weights were computed for wave 4 data. Waves 1–4 longitudinal weights were computed for the 1275 Web respondents recruited at wave 1 that were retained and interviewed at wave 4 (regardless whether or not they were interviewed at wave 2). Waves 2–4 longitudinal CAWI weights were computed for the 506 Web respondents recruited at
wave 1 that were retained and also interviewed at both waves 2 & 4. Waves 3–4 longitudinal CAWI weights (variable dDE42925v) were computed for the 1382 Web respondents that were interviewed at both wave 3 & 4. Wave 4 main survey cross-sectional CAWI weights were computed for all continuing wave 4 respondents. Wave 4 replenishment cross-sectional CAWI weights were computed for the 579 respondents recruited at wave 4 from the Web frame.

As was done at wave 1, waves 1–4 longitudinal CAWI were labelled as dDE42921v. Similarly, all wave 4 cross-sectional weights (CAWI for main survey and new recruits) were combined into a single variable, labelled dDE42919v. As mentioned at the beginning of section 2.2, all ITC Netherlands survey weights were calibrated to smoking prevalences and rescaled to have a mean equal to 1. Consequently, these weights should not be used to estimate population totals (e.g., the total number of daily smokers).

2.5.1 CAWI weights

Starting with \( w_i^{(0)} \), the waves 1–3 rescaled longitudinal CAWI weight for the \( i \)th respondent (computed in section 2.4.2), calculation of the waves 1–4 longitudinal CAWI weights proceeded as follows.

Step 1: The \( w_i^{(0)} \) weights were first re-adjusted for oversampling of smokers aged \( \leq 30 \) using post-stratification, same as step 1 of section 2.2.2.

Step 2: A post-stratification adjustment for province was carried out, same as step 4 of section 2.2.1.

Step 3: The weights were then re-calibrated to smoking prevalences by age/sex/district cells, same as step 5 of section 2.2.1.

Step 4: The weights were rescaled to sum to sample size \( n = 1275 \), same as step 6 of section 2.2.1.

Computation of the waves 2–4 longitudinal CAWI weights proceeded as for the waves 1–4 longitudinal CAWI weights, with the exception that \( w_i^{(0)} \) is the waves 2–3 rescaled longitudinal CAWI weight for the \( i \)th respondent (computed in section 2.4.2).

Computation of the waves 3–4 longitudinal CAWI weights proceeded as for the waves 1–4 longitudinal CAWI weights, with the exception that \( w_i^{(0)} \) is the wave 3 CAWI rescaled cross-sectional weight for the \( i \)th respondent (computed in section 2.4.2).
Computation of the wave 4 main survey cross-sectional CAWI weights also proceeded as for the waves 1–4 longitudinal CAWI weights, with the exception that updated smoking prevalences (by age/sex/district) figures were used to calibrate the weights. The prevalence figures of table A.7, were obtained by combining population estimates as of Jan. 1, 2010 from the Centraal Bureau voor de Statistiek (i.e., Statistics Netherlands) to prevalence estimates from the 2009 Continuous Survey of Smoking Habits, though fewer cells needed to be collapsed; see table ??.

Lastly, computation of the wave 4 replenishment cross-sectional CAWI weights proceeded along the same lines as the wave 1 CAWI weights (see section 2.2.2).

### 2.6 Wave 5 weights

Six sets of weights were computed for wave 5 data. Waves 1–5 longitudinal CAWI weights were computed for the 1012 Web respondents recruited at wave 1 that were retained and interviewed at wave 5 (regardless whether or not they were interviewed at wave 2). Waves 2–5 longitudinal CAWI weights were computed for the 409 Web respondents recruited at wave 1 that were retained and also interviewed at both wave 2 and wave 5. Waves 3–5 longitudinal CAWI weights were computed for the 1104 Web respondents that were interviewed at both wave 3 and wave 5. Waves 4–5 longitudinal CAWI weights (variable eDE42927v) were computed for the 1567 Web respondents that were interviewed at both wave 4 and wave 5. Wave 5 main survey cross-sectional CAWI weights were also computed for all 1619 wave 5 continuing respondents. Wave 5 replenishment cross-sectional CAWI weights were computed for the 482 respondents recruited at wave 5 from the Web frame.

As was done at wave 1, waves 1–5 longitudinal were labelled as eDE42921v. Similarly, all wave 5 cross-sectional weights (CAWI for main survey and new recruits) were combined into a single variable, labelled eDE42919v. As mentioned at the beginning of section 2.2, all ITC Netherlands survey weights were calibrated to smoking prevalences and rescaled to have a mean equal to 1. Consequently, these weights should not be used to estimate population totals (e.g., the total number of daily smokers).

#### 2.6.1 CAWI weights

Starting with \( w_i^{(0)} \), the waves 1–4 rescaled longitudinal weight for the \( i^{th} \) respondent (computed in section 2.5.1), calculation of the waves 1–5 longitudinal weights proceeded as follows.
2 CONSTRUCTION OF THE SURVEY WEIGHTS

Step 1: The $w_i^{(0)}$ weights were first re-adjusted for oversampling of smokers aged $\leq 30$ using post-stratification, same as step 1 of section 2.2.2.

Step 2: A post-stratification adjustment for province was carried out, same as step 4 of section 2.2.1.

Step 3: The weights were then re-calibrated to smoking prevalences by age/sex/district cells, same as step 5 of section 2.2.1.

Step 4: The weights were rescaled to sum to sample size, same as step 6 of section 2.2.1.

For those 43 cohort #1 respondents missing at wave 4 but presenting at wave 3 and wave 5, $w_i^{(0)}$ is replaced by the “attrition adjusted” wave 1-3 rescaled longitudinal weights cDE42921v.

Computation of the waves 2–5 longitudinal CAWI weights proceeded as for the waves 1–5 longitudinal CAWI weights, with the exception that $w_i^{(0)}$ is the waves 2–4 CAWI rescaled longitudinal CAWI weight for the $i^{th}$ respondent (computed in section 2.5.1). For those cohort #1 respondents missing at wave 4 but presenting at wave 3 and wave 5, $w_i^{(0)}$ is replaced by the “attrition adjusted” wave 2-3 rescaled longitudinal weights cDE42923v. Please note this only applied to the subset of 12 respondents who completed the wave 2 survey.

Computation of the waves 3–5 longitudinal CAWI weights proceeded as for the waves 1–5 longitudinal CAWI weights, with the exception that $w_i^{(0)}$ is the waves 3–4 CAWI rescaled longitudinal CAWI weight for the $i^{th}$ respondent (computed in section 2.5.1). For those 43 cohort #1 respondents missing at wave 4 but presenting at wave 3 and wave 5, $w_i^{(0)}$ is replaced by the “attrition adjusted” wave 3 rescaled cross-sectional weights cDE42919v. For those 9 cohort #3 respondents who were missing at wave 4 but presenting at wave 5, $w_i^{(0)}$ is also replaced by the “attrition adjusted” wave 3 rescaled cross-sectional weights cDE42919v.

Computation of the waves 4–5 longitudinal CAWI weights proceeded as for the waves 1–5 longitudinal weights, with the exception that $w_i^{(0)}$ is the wave 4 CAWI rescaled cross-sectional weight for the $i^{th}$ respondent (computed in section 2.5.1).

Computation of the wave 5 main survey cross-sectional CAWI weights also proceeded as for the waves 1–5 longitudinal weights, with the exception that updated smoking prevalences (by age/sex/district) figures were used to calibrate the weights. The prevalence figures of table A.7 were obtained by combining population estimates as of January 1, 2011 from the Centraal Bureau voor de Statistiek (i.e., Statistics Netherlands) with prevalence estimates from the 2010 Continuous Survey of Smoking Habits, though fewer cells needed to be collapsed; see table ??.
Lastly, computation of the wave 5 replenishment cross-sectional CAWI weights proceeded along the same lines as the step 2 to step 4 of wave 1 CAWI weights (see section 2.2.2), because the new replenishment cohort are all under 30 years old. The smoking prevalence (by age/sex/district) figures of table A.8 used in calibration, were obtained by combining population estimates as of January 1, 2011 from the Centraal Bureau voor de Statistiek (i.e., Statistics Netherlands) with prevalence estimates from the 2010 Continuous Survey of Smoking Habits.

2.7 Wave 6 weights

Seven sets of weights were computed for wave 6 data. Waves 1–6 longitudinal CAWI weights were computed for the 842 Web respondents recruited at wave 1 that were retained and interviewed at wave 6 (regardless whether or not they were interviewed at wave 2). Waves 2–6 longitudinal CAWI weights were computed for the 334 Web respondents recruited at wave 1 that were retained and also interviewed at both wave 2 and wave 6. Waves 3–6 longitudinal CAWI weights were computed for the 915 Web respondents that were interviewed at both wave 3 and wave 6. Waves 4–6 longitudinal CAWI weights were computed for the 1320 Web respondents that were interviewed at both wave 4 and wave 6. Waves 5–6 longitudinal CAWI weights (variable fDE42929v) were computed for the 1736 Web respondents that were interviewed at both wave 5 and wave 6. Wave 6 main survey cross-sectional CAWI weights were also computed for those respondents. Wave 6 replenishment cross-sectional CAWI weights were computed for the 286 respondents recruited at wave 6 from the Web frame.

As was done at wave 1, waves 1–5 longitudinal CAWI were labelled as eDE42921v. Similarly, all wave 5 cross-sectional weights (CAWI for main survey and new recruits) were combined into a single variable, labelled eDE42919v. As mentioned at the beginning of section 2.2, all ITC Netherlands survey weights were calibrated to smoking prevalences and rescaled to have a mean equal to 1. Consequently, these weights should not be used to estimate population totals (e.g., the total number of daily smokers).

2.7.1 CAWI weights

Starting with $w_i^{(0)}$, the waves 1-5 CAWI rescaled longitudinal weight for the $i^{th}$ respondent (computed in section 2.6.1), calculation of the waves 1–6 longitudinal CAWI weights proceeded as follows.

Step 1: The $w_i^{(0)}$ weights were first re-adjusted for oversampling of smokers aged $\leq 30$ using post-stratification, same as step 1 of section 2.2.2.
2 CONSTRUCTION OF THE SURVEY WEIGHTS

Step 2: A post-stratification adjustment for province was carried out, same as step 4 of section 2.2.1.

Step 3: The weights were then re-calibrated to smoking prevalences by age/sex/district cells, same as step 5 of section 2.2.1.

Step 4: The weights were rescaled to sum to sample size, same as step 6 of section 2.2.1.

Computation of the waves 2–6 longitudinal CAWI weights proceeded as for the waves 1–6 longitudinal weights, with the exception that \( w_i^{(0)} \) is the waves 2–5 CAWI rescaled longitudinal weight for the \( i \)th respondent (computed in section 2.6.1)

Computation of the waves 3–6 longitudinal CAWI weights proceeded as for the waves 1–6 longitudinal weights, with the exception that \( w_i^{(0)} \) is the waves 3–5 CAWI rescaled longitudinal weight for the \( i \)th respondent (computed in section 2.6.1)

Computation of the waves 4–6 longitudinal CAWI weights proceeded as for the waves 1–6 longitudinal weights, with the exception that \( w_i^{(0)} \) is the waves 4–5 CAWI rescaled longitudinal weight for the \( i \)th respondent (computed in section 2.6.1)

Computation of the waves 5–6 longitudinal CAWI weights proceeded as for the waves 1–6 longitudinal weights, with the exception that \( w_i^{(0)} \) is the wave 5 CAWI rescaled cross-sectional weight for the \( i \)th respondent (computed in section 2.6.1)

Computation of the wave 6 main survey cross-sectional CAWI weights also proceeded as for the waves 1–6 longitudinal weights, with the exception that updated smoking prevalences (by age/sex/district) figures were used to calibrate the weights. The prevalence figures of table A.7, were obtained by combining population estimates as of January 1, 2012 from the Centraal Bureau voor de Statistiek (i.e., Statistics Netherlands) to prevalence estimates from the 2011 Continuous Survey of Smoking Habits, though fewer cells needed to be collapsed; see table ??.

Lastly, computation of the wave 6 replenishment cross-sectional CAWI weights proceeded along the same lines as the wave 1 CAWI weights (see section 2.2.2), except that following the post-stratification adjustment for province, the weights were then re-calibrated to smoking prevalences by age/sex cells, collapsing over all districts.

2.8 Wave 7 weights

Eight sets of weights were computed for wave 7 data:
2 CONSTRUCTION OF THE SURVEY WEIGHTS

i) Waves 1–7 longitudinal weights were computed for the 712 Web respondents recruited at wave 1 that were retained and interviewed at wave 7 (regardless whether or not they were interviewed at wave 2).

ii) Waves 2–7 longitudinal weights were computed for the 284 Web respondents recruited at wave 1 that were retained and also interviewed at both wave 2 and wave 7.

iii) Waves 3–7 longitudinal weights were computed for the 769 Web respondents that were interviewed at both wave 3 and wave 7.

iv) Waves 4–7 longitudinal weights were computed for the 1116 Web respondents that were interviewed at both wave 4 and wave 7.

v) Waves 5–7 longitudinal weights were computed for the 1453 Web respondents that were interviewed at both wave 5 and wave 7.

vi) Waves 6–7 longitudinal weights (variable gDE42931v) were computed for the 1677 Web respondents that were interviewed at both wave 6 and wave 7.

vii) Wave 7 main survey cross-sectional weights were also computed for those respondents.

viii) Wave 7 replenishment cross-sectional weights were computed for the 293 respondents recruited at wave 7.

As was done at wave 1, waves 1–7 longitudinal were labelled as gDE42921v. Similarly, all wave 7 cross-sectional weights (for main survey and new recruits) were combined into a single variable, labelled eDE42919v. As mentioned at the beginning of section 2.2, all ITC Netherlands survey weights were calibrated to smoking prevalences and rescaled to have a mean equal to 1. Consequently, these weights should not be used to estimate population totals (e.g., the total number of daily smokers).

2.8.1 CAWI weights

Starting with \( w_i^{(0)} \), the waves 1-6 rescaled longitudinal weight for the \( i \)th respondent (computed in section 2.7.1), calculation of the waves 1–7 longitudinal weights proceeded as follows.

Step 1: The \( w_i^{(0)} \) weights were first re-adjusted for oversampling of smokers aged \( \leq 30 \) using post-stratification, same as step 1 of section 2.2.2.

Step 2: A post-stratification adjustment for province was carried out, same as step 4 of section 2.2.1.
2 CONSTRUCTION OF THE SURVEY WEIGHTS

Step 3: The weights were then re-calibrated to smoking prevalences by age/sex/district cells, same as step 5 of section 2.2.1.

Step 4: The weights were rescaled to sum to sample size, same as step 6 of section 2.2.1.

Computation of the waves 2–7 longitudinal weights proceeded as for the waves 1–7 longitudinal weights, with the exception that $w_i^{(0)}$ is the waves 2–6 rescaled longitudinal weight for the $i^{th}$ respondent (computed in section 2.7.1)

Computation of the waves 3–7 longitudinal weights proceeded as for the waves 1–7 longitudinal weights, with the exception that $w_i^{(0)}$ is the waves 3–6 rescaled longitudinal weight for the $i^{th}$ respondent (computed in section 2.7.1)

Computation of the waves 4–7 longitudinal weights proceeded as for the waves 1–7 longitudinal weights, with the exception that $w_i^{(0)}$ is the waves 4–6 rescaled longitudinal weight for the $i^{th}$ respondent (computed in section 2.7.1)

Computation of the waves 5–7 longitudinal weights proceeded as for the waves 1–7 longitudinal weights, with the exception that $w_i^{(0)}$ is the waves 5–6 rescaled longitudinal weight for the $i^{th}$ respondent (computed in section 2.7.1)

Computation of the waves 6–7 longitudinal weights proceeded as for the waves 1–7 longitudinal weights, with the exception that $w_i^{(0)}$ is the wave 6 rescaled cross-sectional weight for the $i^{th}$ respondent (computed in section 2.7.1)

Computation of the wave 7 main survey cross-sectional weights also proceeded as for the waves 1–7 longitudinal weights, with the exception that updated smoking prevalences (by age/sex/district) figures were used to calibrate the weights. The prevalence figures of table A.7, were obtained by combining population estimates as of January 1, 2013 from the Centraal Bureau voor de Statistiek (i.e., Statistics Netherlands) to prevalence estimates from the 2012 Continuous Survey of Smoking Habits, though fewer cells needed to be collapsed; see table ??.

Lastly, computation of the wave 7 replenishment cross-sectional weights proceeded along the same lines as the wave 1 weights (see section 2.2.2), except that following the post-stratification adjustment for province, the weights were then re-calibrated to smoking prevalences by age/sex cells, collapsing over all districts.

2.9 Wave 8 weights

Nine sets of weights were computed for wave 8 data:
2 CONSTRUCTION OF THE SURVEY WEIGHTS

i) Waves 1–8 longitudinal weights were computed for the 606 Web respondents recruited at wave 1 that were retained and interviewed at wave 8 (regardless whether or not they were interviewed at wave 2).

ii) Waves 2–8 longitudinal weights were computed for the 245 Web respondents recruited at wave 1 that were retained and also interviewed at both wave 2 and wave 8.

iii) Waves 3–8 longitudinal weights were computed for the 642 Web respondents that were interviewed at both wave 3 and wave 8.

iv) Waves 4–8 longitudinal weights were computed for the 940 Web respondents that were interviewed at both wave 4 and wave 8.

v) Waves 5–8 longitudinal weights were computed for the 1210 Web respondents that were interviewed at both wave 5 and wave 8.

vi) Waves 6–8 longitudinal weights were computed for the 1390 Web respondents that were interviewed at both wave 6 and wave 8.

vii) Waves 7–8 longitudinal weights (variable gDE42931v) were computed for the 1607 Web respondents that were interviewed at both wave 7 and wave 8.

viii) Wave 8 main survey cross-sectional weights were also computed for those respondents.

ix) Wave 8 replenishment cross-sectional weights were computed for the 404 respondents recruited at wave 8 from the Web frame.

As was done at wave 1, waves 1–8 longitudinal were labelled as hDE42921v. Similarly, all wave 8 cross-sectional weights (for main survey and new recruits) were combined into a single variable, labelled eDE42919v. As mentioned at the beginning of section 2.2, all ITC Netherlands survey weights were calibrated to smoking prevalences and rescaled to have a mean equal to 1. Consequently, these weights should not be used to estimate population totals (e.g., the total number of daily smokers).

2.9.1 CAWI weights

Starting with \( w^{(0)}_i \), the waves 1-7 rescaled longitudinal weight for the \( i \)th respondent (computed in section 2.2.2), calculation of the waves 1–8 longitudinal weights proceeded as follows.

Step 1: The \( w^{(0)}_i \) weights were first re-adjusted for oversampling of smokers aged \( \leq 30 \) using post-stratification, same as step 1 of section 2.2.2.
2 CONSTRUCTION OF THE SURVEY WEIGHTS

Step 2: A post-stratification adjustment for province was carried out, same as step 4 of section 2.2.1.

Step 3: The weights were then re-calibrated to smoking prevalences by age/sex/district cells, same as step 5 of section 2.2.1.

Step 4: The weights were rescaled to sum to sample size, same as step 6 of section 2.2.1.

Computation of the waves 2–8 longitudinal weights proceeded as for the waves 1–8 longitudinal weights, with the exception that $w^{(0)}_i$ is the waves 2–7 rescaled longitudinal weight for the $i^{th}$ respondent (computed in section 2.8.1)

Computation of the waves 3–8 longitudinal weights proceeded as for the waves 1–8 longitudinal weights, with the exception that $w^{(0)}_i$ is the waves 3–7 rescaled longitudinal weight for the $i^{th}$ respondent (computed in section 2.8.1)

Computation of the waves 4–8 longitudinal weights proceeded as for the waves 1–8 longitudinal weights, with the exception that $w^{(0)}_i$ is the waves 4–7 rescaled longitudinal weight for the $i^{th}$ respondent (computed in section 2.8.1)

Computation of the waves 5–8 longitudinal weights proceeded as for the waves 1–8 longitudinal weights, with the exception that $w^{(0)}_i$ is the waves 5–7 rescaled longitudinal weight for the $i^{th}$ respondent (computed in section 2.8.1)

Computation of the waves 6–8 longitudinal weights proceeded as for the waves 1–8 longitudinal weights, with the exception that $w^{(0)}_i$ is the waves 6–7 rescaled longitudinal weight for the $i^{th}$ respondent (computed in section 2.8.1)

Computation of the waves 7–8 longitudinal weights proceeded as for the waves 1–8 longitudinal weights, with the exception that $w^{(0)}_i$ is the wave 7 rescaled cross-sectional weight for the $i^{th}$ respondent (computed in section 2.8.1)

Computation of the wave 8 main survey cross-sectional weights also proceeded as for the waves 1–8 longitudinal weights, with the exception that updated smoking prevalences (by age/sex/district) figures were used to calibrate the weights. The prevalence figures of table A.7, were obtained by combining population estimates as of January 1, 2014 from the Centraal Bureau voor de Statistiek (i.e., Statistics Netherlands) to prevalence estimates from the 2013 Continuous Survey of Smoking Habits, though fewer cells needed to be collapsed; see table ??.

Lastly, computation of the wave 8 replenishment cross-sectional weights proceeded along the same lines as the wave 1 weights (see section 2.2.2), except that following the
post-stratification adjustment for province, the weights were then re-calibrated to smoking prevalences by age/sex cells, collapsing over all districts.

Acknowledgements

Core funding for the ITC Project is provided by the U.S. National Cancer Institute (P50 CA111236 & P01 CA138389), and by the Canadian Institutes of Health Research (grant #79551 & MOP-115016). Major funding for the ITC Netherlands Survey provided by the Nederlandse organisatie voor gezondheidszorg en zorginnovatie (ZonMw; The Netherlands Organisation for Health Research & Development).

References


REFERENCES


Appendix: Benchmark/calibration figures

The estimated number of smokers given in tables A.3 and A.5 were obtained by combining population estimates (as of January 1, 2008) from the Centraal Bureau voor de Statistiek (CBS) to prevalence estimates from the Continuous Survey of Smoking Habits (CSSH). To this end, the population estimate for a given age/sex/district combination were simply multiplied by the smoking prevalence for the same age/sex/district combination. Some age/sex/district combinations were collapsed because they contained too few respondents. The 2007 CSSH figures were used for table A.3, whereas the 2008 CSSH figures were used for table A.5.

<table>
<thead>
<tr>
<th>District</th>
<th>Provinces</th>
</tr>
</thead>
<tbody>
<tr>
<td>North</td>
<td>Groningen, Friesland, Drenthe</td>
</tr>
<tr>
<td>East</td>
<td>Overijssel, Gelderland, Flevoland</td>
</tr>
<tr>
<td>West</td>
<td>Utrecht, Noord-Holland, Zuid-Holland</td>
</tr>
<tr>
<td>South</td>
<td>Zeeland, Noord-Brabant, Limburg</td>
</tr>
</tbody>
</table>

Table A.1: Grouping of 12 provinces into 4 districts.

<table>
<thead>
<tr>
<th>Province</th>
<th>Proportion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Friesland</td>
<td>3.93%</td>
</tr>
<tr>
<td>Groningen</td>
<td>3.51%(^a)</td>
</tr>
<tr>
<td>Drenthe</td>
<td>2.97%(^a)</td>
</tr>
<tr>
<td>Overijssel</td>
<td>6.82%(^b)</td>
</tr>
<tr>
<td>Flevoland</td>
<td>2.29%(^b)</td>
</tr>
<tr>
<td>Gelderland</td>
<td>12.10%</td>
</tr>
<tr>
<td>Utrecht</td>
<td>7.28%</td>
</tr>
<tr>
<td>Noord-Holland</td>
<td>15.97%</td>
</tr>
<tr>
<td>Zuid-Holland</td>
<td>21.12%(^c)</td>
</tr>
<tr>
<td>Zeeland</td>
<td>2.33%(^c)</td>
</tr>
<tr>
<td>Noord-Brabant</td>
<td>14.79%</td>
</tr>
<tr>
<td>Limburg</td>
<td>6.89%</td>
</tr>
</tbody>
</table>

Provinces sharing the same letter (a, b, etc.) were collapsed, when computing some of the wave 3 weights.

Table A.2: Dutch population by province as of January 1, 2008.
## APPENDIX

<table>
<thead>
<tr>
<th>District</th>
<th>Sex</th>
<th>Age</th>
<th># smokers</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>male</td>
<td>[15,18)</td>
<td>46780.2</td>
</tr>
<tr>
<td>All</td>
<td>female</td>
<td>[15,18)</td>
<td>47033.4</td>
</tr>
<tr>
<td>All</td>
<td>male</td>
<td>[18,25)</td>
<td>229105.7</td>
</tr>
<tr>
<td>All</td>
<td>female</td>
<td>[18,25)</td>
<td>175068.9</td>
</tr>
<tr>
<td>West</td>
<td>male</td>
<td>[25,40)</td>
<td>253629.8</td>
</tr>
<tr>
<td>North</td>
<td>male</td>
<td>[25,40)</td>
<td>86335.5</td>
</tr>
<tr>
<td>East</td>
<td>male</td>
<td>[25,40)</td>
<td>117340.2</td>
</tr>
<tr>
<td>South</td>
<td>male</td>
<td>[25,40)</td>
<td>139460.8</td>
</tr>
<tr>
<td>West</td>
<td>female</td>
<td>[25,40)</td>
<td>229105.7</td>
</tr>
<tr>
<td>West</td>
<td>male</td>
<td>[40,55)</td>
<td>287517.0</td>
</tr>
<tr>
<td>North</td>
<td>male</td>
<td>[40,55)</td>
<td>65344.7</td>
</tr>
<tr>
<td>East</td>
<td>male</td>
<td>[40,55)</td>
<td>94251.9</td>
</tr>
<tr>
<td>South</td>
<td>female</td>
<td>[40,55)</td>
<td>104402.1</td>
</tr>
<tr>
<td>West</td>
<td>female</td>
<td>[40,55)</td>
<td>239515.7</td>
</tr>
<tr>
<td>East</td>
<td>female</td>
<td>[40,55)</td>
<td>76525.7</td>
</tr>
<tr>
<td>South</td>
<td>female</td>
<td>[40,55)</td>
<td>119260.9</td>
</tr>
<tr>
<td>All</td>
<td>male</td>
<td>[55,100)</td>
<td>457924.0</td>
</tr>
<tr>
<td>All</td>
<td>female</td>
<td>[55,100)</td>
<td>404283.2</td>
</tr>
</tbody>
</table>

Table A.3: Estimated # of smokers, per age/sex/district, used for calibration at waves 1–2.

<table>
<thead>
<tr>
<th>Sex</th>
<th>Age</th>
<th># smokers</th>
</tr>
</thead>
<tbody>
<tr>
<td>female</td>
<td>[18,25)</td>
<td>175069.0</td>
</tr>
<tr>
<td>male</td>
<td>[18,25)</td>
<td>229105.7</td>
</tr>
<tr>
<td>female</td>
<td>[25,40)</td>
<td>436458.8</td>
</tr>
<tr>
<td>male</td>
<td>[25,40)</td>
<td>596766.5</td>
</tr>
<tr>
<td>female</td>
<td>[40,55)</td>
<td>594317.0</td>
</tr>
<tr>
<td>male</td>
<td>[40,55)</td>
<td>631670.4</td>
</tr>
<tr>
<td>female</td>
<td>[55,100)</td>
<td>404283.2</td>
</tr>
<tr>
<td>male</td>
<td>[55,100)</td>
<td>457924.0</td>
</tr>
</tbody>
</table>

Table A.4: Collapsed version of table A.3, used in computing the wave 3 CATI weights.
## Table A.5: Estimated # of smokers, per age/sex/district, used for calibration at wave 3.

<table>
<thead>
<tr>
<th>District</th>
<th>Sex</th>
<th>Age</th>
<th># smokers</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>male</td>
<td>[15,18)</td>
<td>48089.9</td>
</tr>
<tr>
<td>All</td>
<td>female</td>
<td>[15,18)</td>
<td>38475.7</td>
</tr>
<tr>
<td>All</td>
<td>male</td>
<td>[18,25)</td>
<td>227530.5</td>
</tr>
<tr>
<td>All</td>
<td>female</td>
<td>[18,25)</td>
<td>164565.4</td>
</tr>
<tr>
<td>West</td>
<td>male</td>
<td>[25,40)</td>
<td>247162.2</td>
</tr>
<tr>
<td>North</td>
<td>male</td>
<td>[25,40)</td>
<td>78345.0</td>
</tr>
<tr>
<td>East</td>
<td>male</td>
<td>[25,40)</td>
<td>113118.6</td>
</tr>
<tr>
<td>South</td>
<td>male</td>
<td>[25,40)</td>
<td>133094.1</td>
</tr>
<tr>
<td>West</td>
<td>female</td>
<td>[25,40)</td>
<td>194911.5</td>
</tr>
<tr>
<td>North</td>
<td>female</td>
<td>[25,40)</td>
<td>61381.4</td>
</tr>
<tr>
<td>East</td>
<td>female</td>
<td>[25,40)</td>
<td>70736.1</td>
</tr>
<tr>
<td>South</td>
<td>female</td>
<td>[25,40)</td>
<td>101503.8</td>
</tr>
<tr>
<td>West</td>
<td>male</td>
<td>[40,55)</td>
<td>269398.1</td>
</tr>
<tr>
<td>North</td>
<td>male</td>
<td>[40,55)</td>
<td>73031.3</td>
</tr>
<tr>
<td>East</td>
<td>male</td>
<td>[40,55)</td>
<td>119474.9</td>
</tr>
<tr>
<td>South</td>
<td>male</td>
<td>[40,55)</td>
<td>166066.3</td>
</tr>
<tr>
<td>West</td>
<td>female</td>
<td>[40,55)</td>
<td>244826.3</td>
</tr>
<tr>
<td>North</td>
<td>female</td>
<td>[40,55)</td>
<td>65150.2</td>
</tr>
<tr>
<td>East</td>
<td>female</td>
<td>[40,55)</td>
<td>118664.9</td>
</tr>
<tr>
<td>South</td>
<td>female</td>
<td>[40,55)</td>
<td>161835.9</td>
</tr>
<tr>
<td>All</td>
<td>male</td>
<td>[55,100)</td>
<td>439003.6</td>
</tr>
<tr>
<td>All</td>
<td>female</td>
<td>[55,100)</td>
<td>383253.3</td>
</tr>
</tbody>
</table>
### Table A.6: Proportions of Dutch population by province.

<table>
<thead>
<tr>
<th>Province</th>
<th>as of 01/01/2008</th>
<th>as of 01/01/2009</th>
<th>as of 01/01/2011</th>
<th>as of 01/01/2013</th>
<th>as of 01/01/2014</th>
</tr>
</thead>
<tbody>
<tr>
<td>Groningen</td>
<td>3.51</td>
<td>3.48</td>
<td>3.48</td>
<td>3.47</td>
<td>3.46</td>
</tr>
<tr>
<td>Friesland</td>
<td>3.93</td>
<td>3.91</td>
<td>3.90</td>
<td>3.87</td>
<td>3.84</td>
</tr>
<tr>
<td>Drenthe</td>
<td>2.97</td>
<td>2.97</td>
<td>2.96</td>
<td>2.93</td>
<td>2.91</td>
</tr>
<tr>
<td>Overijssel</td>
<td>6.82</td>
<td>6.83</td>
<td>6.82</td>
<td>6.80</td>
<td>6.77</td>
</tr>
<tr>
<td>Flevoland</td>
<td>2.29</td>
<td>2.33</td>
<td>2.34</td>
<td>2.36</td>
<td>2.38</td>
</tr>
<tr>
<td>Gelderland</td>
<td>12.10</td>
<td>12.08</td>
<td>12.06</td>
<td>12.02</td>
<td>12.00</td>
</tr>
<tr>
<td>Utrecht</td>
<td>7.28</td>
<td>7.34</td>
<td>7.37</td>
<td>7.40</td>
<td>7.45</td>
</tr>
<tr>
<td>Noord-Holland</td>
<td>15.97</td>
<td>16.05</td>
<td>16.10</td>
<td>16.20</td>
<td>16.29</td>
</tr>
<tr>
<td>Zeeland</td>
<td>2.33</td>
<td>2.31</td>
<td>2.30</td>
<td>2.28</td>
<td>2.26</td>
</tr>
<tr>
<td>Noord-Brabant</td>
<td>14.79</td>
<td>14.77</td>
<td>14.75</td>
<td>14.73</td>
<td>14.73</td>
</tr>
<tr>
<td>Limburg</td>
<td>6.89</td>
<td>6.81</td>
<td>6.77</td>
<td>6.71</td>
<td>6.66</td>
</tr>
</tbody>
</table>
### # smokers

<table>
<thead>
<tr>
<th>District</th>
<th>Sex</th>
<th>Age</th>
<th>Waves 1–2</th>
<th>Wave 3</th>
<th>Wave 4</th>
<th>Wave 5</th>
<th>Wave 6</th>
<th>Wave 7</th>
<th>Wave 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>female</td>
<td>[15, 18)</td>
<td>47,033</td>
<td>38,475</td>
<td>72,156</td>
<td>62,720</td>
<td>56,356</td>
<td>47,720</td>
<td>45,290</td>
</tr>
<tr>
<td>All</td>
<td>male</td>
<td>[15, 18)</td>
<td>46,780</td>
<td>48,089</td>
<td>93,472</td>
<td>67,829</td>
<td>57,669</td>
<td>57,860</td>
<td>47,513</td>
</tr>
<tr>
<td>All</td>
<td>female</td>
<td>[18, 25)</td>
<td>175,068</td>
<td>164,565</td>
<td>192,614</td>
<td>211,515</td>
<td>195,258</td>
<td>213,379</td>
<td>226,078</td>
</tr>
<tr>
<td>All</td>
<td>male</td>
<td>[18, 25)</td>
<td>229,105</td>
<td>227,530</td>
<td>230,217</td>
<td>271,853</td>
<td>257,426</td>
<td>262,848</td>
<td>224,916</td>
</tr>
<tr>
<td>East</td>
<td>female</td>
<td>[25, 40)</td>
<td>94,251</td>
<td>70,736</td>
<td>87,957</td>
<td>84,060</td>
<td>70,958</td>
<td>93,070</td>
<td>92,951</td>
</tr>
<tr>
<td>North</td>
<td>female</td>
<td>[25, 40)</td>
<td>51,476</td>
<td>61,381</td>
<td>58,316</td>
<td>51,190</td>
<td>36,934</td>
<td>50,584</td>
<td>40,487</td>
</tr>
<tr>
<td>South</td>
<td>female</td>
<td>[25, 40)</td>
<td>104,402</td>
<td>101,503</td>
<td>100,575</td>
<td>112,273</td>
<td>90,481</td>
<td>89,739</td>
<td>78,251</td>
</tr>
<tr>
<td>West</td>
<td>female</td>
<td>[25, 40)</td>
<td>186,328</td>
<td>194,911</td>
<td>193,601</td>
<td>193,593</td>
<td>175,277</td>
<td>174,968</td>
<td>186,320</td>
</tr>
<tr>
<td>East</td>
<td>male</td>
<td>[25, 40)</td>
<td>117,340</td>
<td>113,118</td>
<td>132,506</td>
<td>125,526</td>
<td>112,428</td>
<td>118,730</td>
<td>106,999</td>
</tr>
<tr>
<td>North</td>
<td>male</td>
<td>[25, 40)</td>
<td>86,335</td>
<td>78,345</td>
<td>68,298</td>
<td>63,487</td>
<td>59,508</td>
<td>54,523</td>
<td>63,639</td>
</tr>
<tr>
<td>South</td>
<td>male</td>
<td>[25, 40)</td>
<td>139,460</td>
<td>133,094</td>
<td>145,329</td>
<td>127,500</td>
<td>127,265</td>
<td>143,098</td>
<td>106,990</td>
</tr>
<tr>
<td>West</td>
<td>male</td>
<td>[25, 40)</td>
<td>253,629</td>
<td>247,162</td>
<td>242,553</td>
<td>244,243</td>
<td>223,339</td>
<td>215,599</td>
<td>239,498</td>
</tr>
<tr>
<td>East</td>
<td>female</td>
<td>[40, 55)</td>
<td>119,260</td>
<td>118,664</td>
<td>111,126</td>
<td>102,231</td>
<td>101,162</td>
<td>104,361</td>
<td>102,264</td>
</tr>
<tr>
<td>North</td>
<td>female</td>
<td>[40, 55)</td>
<td>765,525</td>
<td>65,150</td>
<td>59,303</td>
<td>60,782</td>
<td>45,692</td>
<td>59,508</td>
<td>61,002</td>
</tr>
<tr>
<td>South</td>
<td>female</td>
<td>[40, 55)</td>
<td>159,014</td>
<td>161,835</td>
<td>163,449</td>
<td>166,957</td>
<td>135,722</td>
<td>157,799</td>
<td>120,684</td>
</tr>
<tr>
<td>West</td>
<td>female</td>
<td>[40, 55)</td>
<td>239,515</td>
<td>244,826</td>
<td>229,218</td>
<td>239,089</td>
<td>189,477</td>
<td>207,709</td>
<td>215,852</td>
</tr>
<tr>
<td>East</td>
<td>male</td>
<td>[40, 55)</td>
<td>117,620</td>
<td>119,474</td>
<td>114,451</td>
<td>119,103</td>
<td>95,792</td>
<td>104,918</td>
<td>93,402</td>
</tr>
<tr>
<td>North</td>
<td>male</td>
<td>[40, 55)</td>
<td>65,344</td>
<td>73,031</td>
<td>73,889</td>
<td>70,199</td>
<td>80,586</td>
<td>73,976</td>
<td>64,040</td>
</tr>
<tr>
<td>South</td>
<td>male</td>
<td>[40, 55)</td>
<td>161,187</td>
<td>166,066</td>
<td>154,683</td>
<td>171,282</td>
<td>156,741</td>
<td>130,882</td>
<td>138,524</td>
</tr>
<tr>
<td>West</td>
<td>male</td>
<td>[40, 55)</td>
<td>287,517</td>
<td>269,398</td>
<td>273,903</td>
<td>255,665</td>
<td>236,037</td>
<td>266,007</td>
<td>250,883</td>
</tr>
<tr>
<td>All</td>
<td>female</td>
<td>[55, 100)</td>
<td>404,283</td>
<td>383,253</td>
<td>494,762</td>
<td>492,662</td>
<td>455,847</td>
<td>541,199</td>
<td>527,158</td>
</tr>
<tr>
<td>All</td>
<td>male</td>
<td>[55, 100)</td>
<td>457,924</td>
<td>439,003</td>
<td>459,507</td>
<td>453,033</td>
<td>440,529</td>
<td>438,711</td>
<td>469,998</td>
</tr>
</tbody>
</table>

Table A.7: Estimated # of smokers, per age/sex/district, used for calibration.
<table>
<thead>
<tr>
<th>District</th>
<th>Sex</th>
<th>Age</th>
<th># smokers</th>
</tr>
</thead>
<tbody>
<tr>
<td>West</td>
<td>male</td>
<td>[15,18)</td>
<td>29,927</td>
</tr>
<tr>
<td>North</td>
<td>male</td>
<td>[15,18)</td>
<td>6,936</td>
</tr>
<tr>
<td>East</td>
<td>male</td>
<td>[15,18)</td>
<td>16,615</td>
</tr>
<tr>
<td>South</td>
<td>male</td>
<td>[15,18)</td>
<td>14,352</td>
</tr>
<tr>
<td>West</td>
<td>female</td>
<td>[15,18)</td>
<td>26,756</td>
</tr>
<tr>
<td>North</td>
<td>female</td>
<td>[15,18)</td>
<td>7,032</td>
</tr>
<tr>
<td>East</td>
<td>female</td>
<td>[15,18)</td>
<td>18,513</td>
</tr>
<tr>
<td>South</td>
<td>female</td>
<td>[15,18)</td>
<td>10,419</td>
</tr>
<tr>
<td>West</td>
<td>male</td>
<td>[18,25)</td>
<td>129,962</td>
</tr>
<tr>
<td>North</td>
<td>male</td>
<td>[18,25)</td>
<td>35,848</td>
</tr>
<tr>
<td>East</td>
<td>male</td>
<td>[18,25)</td>
<td>51,500</td>
</tr>
<tr>
<td>South</td>
<td>male</td>
<td>[18,25)</td>
<td>54,543</td>
</tr>
<tr>
<td>West</td>
<td>female</td>
<td>[18,25)</td>
<td>88,963</td>
</tr>
<tr>
<td>North</td>
<td>female</td>
<td>[18,25)</td>
<td>36,349</td>
</tr>
<tr>
<td>East</td>
<td>female</td>
<td>[18,25)</td>
<td>39,758</td>
</tr>
<tr>
<td>South</td>
<td>female</td>
<td>[18,25)</td>
<td>46,446</td>
</tr>
<tr>
<td>West</td>
<td>male</td>
<td>[25,40)</td>
<td>244,243</td>
</tr>
<tr>
<td>North</td>
<td>male</td>
<td>[25,40)</td>
<td>63,487</td>
</tr>
<tr>
<td>East</td>
<td>male</td>
<td>[25,40)</td>
<td>125,526</td>
</tr>
<tr>
<td>South</td>
<td>male</td>
<td>[25,40)</td>
<td>127,500</td>
</tr>
<tr>
<td>West</td>
<td>female</td>
<td>[25,40)</td>
<td>193,593</td>
</tr>
<tr>
<td>North</td>
<td>female</td>
<td>[25,40)</td>
<td>51,190</td>
</tr>
<tr>
<td>East</td>
<td>female</td>
<td>[25,40)</td>
<td>84,060</td>
</tr>
<tr>
<td>South</td>
<td>female</td>
<td>[25,40)</td>
<td>112,273</td>
</tr>
<tr>
<td>West</td>
<td>male</td>
<td>[40,55)</td>
<td>255,665</td>
</tr>
<tr>
<td>North</td>
<td>male</td>
<td>[40,55)</td>
<td>70,199</td>
</tr>
<tr>
<td>East</td>
<td>male</td>
<td>[40,55)</td>
<td>119,103</td>
</tr>
<tr>
<td>South</td>
<td>male</td>
<td>[40,55)</td>
<td>171,282</td>
</tr>
<tr>
<td>West</td>
<td>female</td>
<td>[40,55)</td>
<td>239,089</td>
</tr>
<tr>
<td>North</td>
<td>female</td>
<td>[40,55)</td>
<td>60,782</td>
</tr>
<tr>
<td>East</td>
<td>female</td>
<td>[40,55)</td>
<td>102,231</td>
</tr>
<tr>
<td>South</td>
<td>female</td>
<td>[40,55)</td>
<td>166,957</td>
</tr>
<tr>
<td>West</td>
<td>male</td>
<td>[55,100)</td>
<td>183,999</td>
</tr>
<tr>
<td>North</td>
<td>male</td>
<td>[55,100)</td>
<td>47,089</td>
</tr>
<tr>
<td>East</td>
<td>male</td>
<td>[55,100)</td>
<td>84,463</td>
</tr>
<tr>
<td>South</td>
<td>male</td>
<td>[55,100)</td>
<td>137,482</td>
</tr>
<tr>
<td>West</td>
<td>female</td>
<td>[55,100)</td>
<td>230,242</td>
</tr>
<tr>
<td>North</td>
<td>female</td>
<td>[55,100)</td>
<td>43,406</td>
</tr>
<tr>
<td>East</td>
<td>female</td>
<td>[55,100)</td>
<td>88,979</td>
</tr>
<tr>
<td>South</td>
<td>female</td>
<td>[55,100)</td>
<td>130,034</td>
</tr>
</tbody>
</table>

Table A.8: Estimated # of smokers, per age/sex/district, used in computing the wave 5 replenishment cross-sectional weights.