WP5 – Deliverable 3

Final methodological report discussing the use of mobile device sensors in ESS surveys

Sensor data for ESS surveys: a first inventory

Date: February 4, 2019

Ole Mussmann (CBS)
Barry Schouten (CBS)

WP5: Challenges for phone and tablet respondents within CAWI

SUMMARY: WP5 of the MIMOD project investigates the employment of mobile devices in ESS surveys. In particular, it explores fitness of ESS surveys for smartphones and it explores the utility of mobile device sensors to replace and/or supplement survey data. This third WP5 deliverable considers the second objective, the use of mobile device sensors. We run by all current ESS surveys and provide an inventory of potential sensor measurements. The inventory is the starting point for further exploration.

1. Introduction

Mobile devices carry a range of sensors that support apps and other functionality of the devices in every day operation and use. This paper explores the utility of such sensors to enrich or replace survey data.

With mobile devices, we refer to smartphones and tablets. We see a mobile device as a computing device small enough to hold, operate in the hand with a flatscreen interface, that has many of the following functions:

- It has a mobile operating system – which is not a full-fledged PC operating system – that allows third party apps to be installed and run;
- It can place and receive phone calls;
- It can connect to Internet;
- It can interconnect with other devices through one or more wireless channels (Wi-Fi, Bluetooth, GSM or NFC);
- It has an integrated camera;
- It has a Global Positioning System (GPS).

Mobile devices have become standard tools for communication within the time span of just a decade and have a high population coverage worldwide. Because of their high population coverage and daily life use and because of the availability of sensors, mobile devices have become tools that may supplement surveys with automated data from sensors. Some of these sensor data may also replace survey data.

Naturally, the potential to record sensor data is not sufficient reason to also do so. Criteria are needed that identify combinations of survey topics and sensor data that are promising. This paper is about identification of such combinations in the ESS context.

We focus on self-initiated sensor measurements. However, mobile device sensor data may already exist on a subset of the population and a survey may seek access to such data through respondent consent. Existing sensor data are often big data and often are privacy sensitive. Nonetheless, hybrid forms of survey data and secondary data are promising.

Within the ESS context, the exploration of sensors and sensor data is not new. Eurostat initiated taskforces for the Household Budget Survey (HBS) and Harmonized European Time Use Survey (HETUS). The two taskforces have identified relevant secondary (big) data that may potentially be linked. Furthermore, mobile devices apps have been advocated as a possible instrument to record diary data for the two surveys. The step from app-based data collection to mobile device sensors is small, and, under a recent Eurostat call, a number of projects has started to investigate the utility of sensors. The suggested options for the HBS and HETUS in this deliverable are in line with the task forces and projects.

In the subsequent sections, we present a number of criteria, describe the mobile device sensors, discuss the utility of sensors in ESS surveys and end with a discussion on future research and investigation.

2. Criteria for sensor data

Obviously, the mere possibility to collect sensor data is not enough reason to also do so, as it requires a separate and new architecture and infrastructure and respondents may not be willing to share the data. On the data collection side, sensor data, especially when they are collected on one's own initiative (i.e. primary data collection), demand for new data collection channels. These channels demand for new and/or additional processing tools and skills, for expansion of existing monitoring and analysis tools and for a redesign of survey estimation methodology. Such changes are costly and time consuming. On the respondent side, sensor data may still be burdensome and/or may be privacy intrusive. Hence, a strong business case for mobile device sensor data is needed and respondents need to benefit as well.

It is important to first note that sensor data do not necessarily have to come from a single sensor but may arise from a fusion of multiple sensors. An example is the echo of an ultrasonic sound produced by a smartphone that is captured by the microphone of the same device.

From the surveyor point of view, survey topics may be candidates for enrichment or replacement with sensor data when they satisfy at least one of the following criteria:

- Burden: The survey topic(s) are burdensome for a respondent, either in terms of time or in terms of cognitive effort;
- Centrality: The survey topic(s) are non-central to respondents, i.e. the average respondent does not understand the question or does not know the answer;
- Non-survey type: The survey topic(s) do not lend themselves to a survey question-answer approach to begin with;

To provide examples: Examples that satisfy the first criterion are topics that require keeping a diary for a specified time period, say a week or a month, and provide details about all time periods. Other examples are surveys that require consultation of personal information and archives, for example about assets and finances. The second criterion is satisfied, for instance, for travel surveys where respondents need to provide exact coordinates of locations they have visited and health surveys where respondents need to describe sleeping patterns. The third criterion applies to complex socio-economic or psychological topics such as happiness, health or wealth, where many questions are needed to derive latent constructs.

From the sensor point of view, the main criteria are:

- Omnipresence: The sensor(s) are available in most, if not all, contemporary devices;
- Data access: Data generated by the sensor(s), as well as metadata about the properties and accuracy of the sensor data, can be accessed and processed;
- Quality: The sensor data is comparable, reproducible and accurate;
- Costs: Any costs associated with the sensor(s) are affordable in most surveys;

The four criteria all link to the utility of the resulting sensor data. The omnipresence criterion refers to the coverage, and, hence, also price, of the sensors. In theory, tailored instruments can be developed that record complex phenomena and behaviours, but these are until now used only in lab settings. Smartphone sensors are examples of omnipresent sensors, whereas sensors in wearables have a much lower population coverage and pose challenges with regard to data access. The data access criterion means that sensor data can be stored, manipulated and interpreted. For instance, location data can be stored and processed, but it is not always clear what sensor, GSM, Wi-Fi or GPS, produced the data and how accurate the data are. The quality criterion originates from the statistical objective to derive accuracy of statistics and to be able to compare statistics between persons and in time. For instance, location data can be used to estimate travel distances but are subject to missing data, measurement errors and potentially also device effects. As such, sensors are just like other data collection instruments. The final criterion applies when sensors need to be provided to respondents and refer to costs associated with their use.

From the respondent point of view, sensors may vary in their intrusiveness. Four criteria follow:

- Respondent willingness: Respondents are willing to consent to provide the sensor data;
- Data handling: Respondents can retrieve, revise and delete sensor data on demand;
- Burden: Respondents are willing to devote the effort needed to collect and handle the sensor data;
- Feedback: Respondents may retrieve useful knowledge about themselves;

In order to employ sensors, respondents need to be asked for consent to activate sensors and to store and send data. Most mobile device sensors require consent by default. Exceptions are the various motion sensors that can be activated in Android without consent. However, even without the technical necessity to ask for consent, there are legal and ethical reasons why consent is imperative. Willingness to consent varies per type of sensor and depends on the context and purpose of the measurements. Recent literature, see Struminskaya et al 2018 for an overview of studies, has investigated willingness and confirmed differences between sensors and settings. Obviously, the more intrusive a sensor measurement is, the more respondents will refuse and the larger the potential damage of missing sensor data. Recent European legislation require that respondents can get copies of their data and can request deletion of their data at any time[1]. This requirement puts constraints on the storage of and access to sensor data. Next, sensor measurement themselves, such as photos or sound recordings, require some respondent effort. This effort may be too great for respondents so that missing data and/or lower data quality result. Finally, the sensor data may be fed back to respondents in an aggregated form and may provide valuable information to them.

Sensor data can be collected passively or actively. Passive sensor data is collected without respondent intervention or feedback, apart from consent. In active sensor data, respondents are asked to check, revise, accept and/or supplement sensor data, i.e. the respondent is involved in data collection. Motives for active sensor data collection are increased response rates, increased data quality and hybrid forms of survey and sensor data. An example is a travel survey in which respondents' locations are stored whenever a device is in motion. These data can be collected passively. The data may also be shown to respondents for quality checks and for enrichment of sensor data with stop motives and other context information. Active data collection is much more demanding as it requires real-time data handling and a careful design of a user interface.

Summarizing, there are various criteria from the respondent, sensor and surveyor points of view that need to be confronted with costs and logistics of sensor data collection and processing.

3. Mobile device sensors

Smartphones and their sensors have become a commodity[2]. It is, therefore, the right timing to consider the potential of sensors for surveys within MIMOD project.

There are two main mobile device operating platforms, Android and iOS[3], that basically support the same range of functions. However, the two platforms show differences in what sensors are supported and how the sensors can be employed. An example of the latter is location: Both platforms support positioning through a mix of GPS, Wi-Fi and GSM[4]. In this section, we focus on sensors that are shared by most devices.

One function of mobile devices is very relevant for surveys; they can communicate with various other devices, in particular with wearables. We will also consider them as far as deemed relevant and affordable in a survey context.

---

[1] See https://gdpr-info.eu/issues/right-of-access/ and https://gdpr-info.eu/issues/right-to-be-forgotten/

[2] See http://techsuplex.com/2014/06/09/smartphones-gotten-boring-part-one/

[3] See http://gs.statcounter.com/os-market-share/mobile/worldwide

[4] See https://www.lifewire.com/assisted-gps-1683306

In the following subsections, we discuss the three forms of sensor data: self-initiated sensor data, wearable sensor data and secondary sensor data.

3.1 Mobile device sensors

The following elements and sensors are supported by many contemporary smartphones and tablets:

- 3D touch: This sensor measures the pressure exerted on the screen. Small objects up to 385 grams can be weighted. Used from iPhone 6S (plus) until iPhone X(s). Could be phased out in the future[5].
- Accelerometer, gyroscope (motion sensors): A set of sensors measuring motion, acceleration and position of the device, as well as gravity. Used for position tracking or step counters.
- Ambient Light: Measures the intensity of the ambient light. More advanced versions also determine the light colour or –temperature. Commonly used to adapt the screen brightness and colour to the ambient conditions.
- Bluetooth: Wireless communication protocol. Can connect to small low energy devices, such as wireless headphones, key fobs, smartwatches or smart scales. Also detects the presence of Bluetooth beacons.
- Camera: Takes pictures or videos, or measures light intensity. Usable for image- or pattern recognition, scanning of QR- or barcodes and colour analysis. Can also coarsely measure gamma rays, a form of radioactivity, and heart rate in combination with the camera flash.
- Camera Flash: Usually used as a flashlight or as illumination for pictures or videos. Can create stimuli that can be picked up by other sensors, for example heart rate measurements in combination with the camera.
- Cellular: The core of all cell phones. Used to make and receive calls and text messages. Strength and ID of cell tower broadcasts can be measured. With the knowledge of tower positions the user location can be determined with a precision of ~500 meters. More advanced cell phones – almost all cell phones today – can also connect to the internet. The presence of the internet connection as well as its speed (upload, download, responsiveness/ping) can be determined. Not often found in wearables yet.
- Fingerprint: Some devices are capable of detecting fingerprints. The raw data is not accessible, but it can be used as identification or simply as a button or – in some devices – as a small touchpad.
- GPS: The Global Positioning System. Dozens of GPS satellites circle the earth and broadcast beacon signals. By measuring the time-of-flight of the satellite signals, the distance to that satellite can be calculated. With four or more satellites visible, the position on earth can be triangulated[6]. The precision is ~5 meters outdoors under clear sky. Tall houses or a forest will reduce the accuracy. Indoor performance is poor. The satellite signals also includes time information.
- Heart Rate: Measures the heart rate, usually optically on the finger (cell phone), wrist (smartwatch) or in-ear (headphones).
- Humidity: Measures ambient humidity. Not very widely used yet.
- Magnetic Field: Usually used as a compass, but can also measure the strength of magnetic fields or can be used, within limits, as a metal detector.
- Microphone: Detects speech and sounds that can be saved, streamed or analysed. Can also determine loudness and detect ambient noise. Multiple microphones in one device allow for determining the directionality or distance of sound sources. This is used to filter out ambient noise in phone calls. Microphones can also be used to record the heartbeat or estimate lung function/spirometry.
- NFC (near field communication): The same technology as contactless payments. Can be used to pay with the cell phone or smart watch, as "contactless QR code" / "NFC tags" to change phone settings (muting)

---

[5] See https://www.theverge.com/circuitbreaker/2018/9/13/17854864/iphone-xs-max-3d-touch-waste-features-apple

[6] See https://www.forbes.com/sites/quora/2017/01/18/why-are-four-gps-satellites-required-to-locate-someones-position/#5d872bd14fc3

or trigger the start of certain apps. A phone can be a tag as well, so two devices can identify each other and initiate a data channel for communication.

- Pressure: This sensor measures the ambient air pressure and functions as a barometer. The precision is so high, that height differences of a few meters (ground floor vs first floor) can be detected (Li, Harvey & Gallagher 2013). Is also used in combination with GPS for more precise height determination.
- Proximity: Measures the presence of objects close to the screen, usually binary (object present or not). Switches off screen and touchscreen during phone calls.
- Screen: Display of static or dynamic images. Can also illuminate the surroundings. Commonly used for user interaction.
- Speaker: Plays sounds or speech. Can be used for feedback to the user or in combinations with other sensors such as microphones.
- Thermometer: Usually placed in or near the battery to prevent overheating. Measures the battery/cell phone temperature which might be higher than the ambient temperature.
- Vibration: Feedback mechanism, induces vibration in the device. Can be used as feedback or combined with other sensors such as the accelerometer.
- Wi-Fi: Usually used for internet access. Can detect presence and strength of different wireless networks in different frequency bands. Measuring of connection speed (upload, download, responsiveness/ping) is possible. Also used for location measurements with a precision of ~50 meters in the proximity of WiFi routers/access points.
- Wireless Charging: Neither sensor nor feedback, strictly speaking. Some devices can be charged wirelessly over very short distances (1-2 millimeters). There are competing non-compatible standards, such as Qi and PMA for cell phones. The device can detect the presence of a certain charging station.

3.2 Mobile devices and wearables

Wearable devices are mobile devices that one can wear or attach to clothes. Their functionality is more specific and targeted than smartphones and tablets in order to reduce size and weight. The motivation to wear a device is often health-related, but wearables do support other functions. Most common wearables are activity trackers, fitness bands and smart watches, which are intended for continuous use. Less common wearables are shoe clips, smart glasses, clothes with sensors and jewellery with sensors, which are intended for temporary use.

Due to their more specific functionality, wearables often require other mobile devices to communicate with users. A dedicated app needs to be installed on a smartphone or tablet that provides a user interface to set or alter wearable settings and to read summary data and statistics.

Wearable sensor data can be sent to the user directly or indirectly through the producer of the wearable. In the latter case, often only edited and aggregated data are available and the raw sensor data remain at the side of the producer. Consequently, wearable sensor data lie somewhere in between primary and secondary data collection. They are mostly self-initiated but often are maintained and owned by another party.

Due to the closer proximity to a respondent's body and the fact that they can be worn 24 hours per day and 7 days per week, wearables can measure data that smartphones and tablets cannot. Examples include detailed sleeping patterns, calorie usage, more detailed activity tracking or electrocardiography.

Typical sensors that are included in wearables are the following:
- Accelerometer, gravity, gyroscope
- Bluetooth
- GPS
- Heart rate
- Screen

- Vibration

Less common sensors are:
- Blood oxygen levels
- Camera
- Light sensor
- Magnetic Field
- NFC (near field communication)
- Speaker
- Thermometer
- Wi-Fi

3.3 Secondary sensor data

Sensor data may already be collected and owned by various parties and not necessarily for statistical purposes. We term these data secondary sensor data. We distinguish data based on user-owned sensors and data based on the public Internet-of-Things (IoT) sensors (i.e. not owned by private users). Although a clear distinction is hard, we view privately owned IoT type sensors, such as weather stations or burglar protection systems, as user-owned sensors. So the main distinction is between individual and public use. In this project, we consider linkage of secondary sensor data to samples of a target population after respondent consent.

Within user-owned sensors, a distinction can be made between commercial and non-commercial data collectors. A range of companies produces apps for mobile devices, in particular wearable devices, to provide paid services to individual customers. Other companies, such as Google, also provide unpaid services in exchange for the right to use the resulting data for commercial purposes. These services are usually continuous and, consequently, create a dynamic, vast stream of sensor data. The resulting data are mostly based on self-selection, i.e. the initiative is with the user, and not based on invitation. Non-commercial parties may collect mobile device sensor data for research or policy making motives. Such data collection often has a finite time horizon and is invitation-only. In all these cases, however, the sensor data are self-initiated by users, but the data is stored, handled and owned by others.

Apart from privacy and legal constraints, such user-owned sensor data can potentially be linked to individual respondents in surveys. After linkage, a hybrid data collection follows where part of the data may be asked through questions, part of the data may be measured on respondent mobile devices and part of the data may be linked. An example is activity tracker data owned by a third party linked to persons in the sample supplemented by survey data on health perceptions and health determinants (e.g. Bankova et al 2018). Another example is budget expenditure diary data linked to supermarket scanner data and supplemented by scanned receipts for other types of stores. The latter is currently explores in Eurostat project @HBS.

IoT sensors are usually owned by government authorities and implemented for very specific purposes, such as monitoring of weather, pollution oror traffic. The sensors are attached to objects, often have a fixed location, operate continuously and measure activities of multiple persons. Data is stored, handled and owned by the same authorities independently of persons' consent, but subject to privacy constraints and legislation.

IoT sensor data do not necessarily include identifying information about the persons to which data correspond. As a consequence, linkage of IoT sensor data to individual respondents may be hard or even infeasible without additional information such as time stamps and/or location coordinates. Nonetheless, hybrid forms of data collection may arise where respondents are asked for additional information that enables linkage.

4. ESS surveys

In analogy to the MIMOD survey conducted on EU NSIs, we consider the following ESS surveys: Labour Force Survey (LFS), European Statistics on Income and Living Conditions (EU-SILC), European Health Interview Survey (EHIS), Adult Education Survey (AES), Information and Communication Technology survey (ICT), Household Budget Survey (HBS) and Harmonized European Time Use Survey (HETUS). For each survey, we evaluate the sensor data criteria and we consider potential sensor measurements. We restrict ourselves to the Eurostat model questionnaires/guidelines; ESS NSI's have often extended these surveys with country-specific modules.

4.1 Survey topic criteria

First, we consider the survey-specific criteria: burden, centrality and non-survey type. Table 4.1 contains marks for the three criteria for each survey, where a mark means that the criterion applies. The flags are subjective; other researchers may come to different conclusions.

Table 4.1: Assessment of survey topic criteria.

	Burden	Centrality	Non-survey type
LFS	×		
EU-SILC	×	×	×
EHIS	×	×	×
AES		(×)	
ICT		×	
HBS	×	×	
HETUS	×	×	

The Labour Force Survey collects survey data on (self-)employment, unemployment and educational attainment for all household members of 16 years and older (the labour force population). It is implemented in most ESS countries as a rotating panel design with five to ten waves within a time span of one to two years. The LFS survey topics are central to respondents, lend themselves for a survey-type data collection and a single wave is not burdensome. However, given the household and panel nature, the LFS is considered burdensome as a whole.

The European Statistics on Income and Living Conditions has survey data on income, poverty, social exclusion and living conditions. The survey has a cross-sectional component and a four year longitudinal component. The living conditions concern health and labour conditions, housing conditions and childcare arrangements. Derived statistics concern material deprivation, risk of poverty and social exclusion, monetary poverty and in work poverty. EU-SILC is considered burdensome due to the range of topics and the longitudinal nature. The detailed income and living conditions are non-central to part of the population. The main statistics about living conditions, poverty and deprivation concern topics that are hard to measure through surveys and require a wide set of questions.

The European Health Interview Survey has three substantive modules on health status, health care use and health care determinants and a fourth background module on demographics and socio-economics. Health status involves topics such as self-perceived health, chronic conditions, (functional) limitations and disease specific morbidity. Health care use deals with hospitalisation, consultations, unmet needs, medicine use and other forms of health care use. Health care determinants involve for example height and weight, consumption of fruits and vegetables, smoking and alcohol consumption, and daily fitness activities. EHIS is burdensome due to its lengthy questionnaire and wide range of topics. The survey concerns topics that are non-central

such as consultations, medicine use, chronic conditions, limitations, consumption and daily activities. Part of the EHIS topics are ill-suited for a survey as they tend to measure complex latent variables such as health, limitations in daily activities, fitness, and life style.

The Adult Education Survey measures participation of adults of 25 to 65 years in education and training, characteristics of learning activities, number of instruction hours, reasons for and obstacles to participating, access to information on learning possibilities, employer financing and costs, and self-reported language skills. The AES is not burdensome or non-central, except for language skills, and lends itself to a survey type instrument, despite the difficulties in harmonizing educational attainment across the EU and across time.

The Information and Communication Technology Survey collects survey data on the frequencies, durations and purposes of use of all forms of contemporary ICT (TV, internet, social media, mobile devices). The survey is not burdensome and a survey type instrument is well-suited due to the lack of complex latent constructs. However, ICT use, especially frequency and duration, and details about the respondent ICT facilities are non-central.

The Household Budget Survey records all household expenditures and purchases during a week and large household expenditures and purchases during a longer period up to a month. The HBS is very burdensome due to the duration of the diary keeping. Detailed expenditures and purchases are non-central to respondents. The HBS does not involve complex latent concepts that relate to many survey questions.

The Harmonised European Time Use Survey asks respondents to keep a diary for a week in which activities at a resolution of ten minutes need to be provided from a specified list of classifications. The diary is preceded by an individual and a household questionnaire. The HETUS is very burdensome due to the duration of the diary keeping. The detailed time keeping and detail of activities are non-central to respondents. Like the HBS, the HETUS does not involve complex latent concepts.

Summarizing, when treating the criteria as equally important or severe, the EU-SILC, EHIS, HBS and HETUS would profit the most from automated sensor measurements; they score on at least two criteria. ICT may profit from sensor data due to its non-central topics. LFS may profit due to its household and panel setting. AES benefits the least from sensor data. We have to remark that the criteria may be viewed as having different weights. Such weights would reflect the magnitude and variation of measurement errors, which are inherently hard to measure or quantify. An approach to introduce weights is to consult a small group of questionnaire experts involved in cognitive testing and derive a consensus on the relative importance of the criteria.

4.2 Mobile device sensors

Next, we move to potential sensor measurements for each of the ESS surveys. We omit the AES as it did not score on any of the survey topic criteria. We must remark beforehand that the suggestions are by no means exhaustive. The search for potential sensor measurements depends on the creativity of the designers, in particular, because multiple sensors may be combined. Table 4.2 presents sensors that may be useful for each survey.

LFS: Researchers have suggested to employ passive sensor data to derive whether persons are employed or unemployed (Haas et al 2018). One option to do so, is to consider their daily life style patterns such as the locations that they visit, and the times and types of social media that they use. Another option is to consider the content of their social media messages and internet searches. While these sensor data may be related to employment and pursuits to get employed, they are oriented at a potential centrality problem and not the burden problem of the LFS. The LFS as a household survey would still require passive sensor data for all of its members. Furthermore, per LFS wave, the sensor measurements would need to be collected for a longer time period that corresponds to the LFS reference periods. Essentially, sensor measurements would need to

run for the full duration of the LFS waves. Hence, the sensor measurements do not necessarily reduce burden.

Table 4.2: Potential sensor measurements.

	Sensors
LFS	Time-location, mobile device use
EU-SILC	Camera, microphone, time-location, mobile device use
EHIS	Time-location, motion, heart rate, wearables, camera
ICT	Mobile device use, mobile device properties
HBS	Time-location, camera, mobile device use
HETUS	Time-location, motion, mobile device use, NFC, Bluetooth, wearables

EU-SILC: The EU-SILC topics have some similarity to the LFS and may be viewed as providing more context and depth. As for the LFS, EU-SILC is a panel, but the duration is much longer. However, although EU-SILC requires household information, it is essentially a person survey. Given the similarity, the sensor measurements that are proposed in the LFS could also be proposed for the EU-SILC. One may derive life style patterns by looking at visited locations, social media use, internet use, content of social media and content of internet searches. Again this may not solve the burden problem of the EU-SILC, especially because the duration of the survey is multiple years. It is even very likely that the respondent will switch devices in the course of the EU-SILC. However, other than the LFS, EU-SILC aims at broader concepts as poverty, social isolation and deprivation. These concepts are hard to measure through questions and the selected sensor measurements may shed an insightful light on the respondents' life styles. Nonetheless, the respondent would need to allow sensor measurements repeatedly for certain time periods over a few years, An additional source of sensor data may come from camera measurements. Photos may be taken of the house and neighbourhood, and of specific parts of the house. Living conditions and wealth may be indirectly derived from such images. Furthermore, the camera can be used to measure the size and volume of the house and possible garden/outside areas. Dedicated apps exist that can measure such properties using the camera. Finally, the camera or the ambient light sensor can be employed to measure light intensity in the house. As a supplementary measure, the microphone may be used to measure noise/sound levels in the house.

EHIS: The EHIS topics are physical and mental, and, thus, lend themselves for wearables. In fact, the most frequently used wearables are aiming at health, fitness, life style indicators. Activity trackers, smart scales and smart watches may be employed for at least a specified duration in order to derive physical conditions, and, possibly, also mental conditions. Wearable sensor data may replace parts of the EHIS questionnaire but may also provide important insight into the underlying concepts that the EHIS attempts to measure. Instead of wearables, also mobile device sensors may be activated that measure similar data such as activity, heart rate and locations. Obviously, mobile devices are usually not attached to one's body, so that, in practice, the physical activity of the device is measured. As a supplemental type of sensor data, camera photos may be included that give an impression of a person's physical condition.

ICT: As mobile devices are ICT by themselves and connect to other ICT, they can provide sensor data about the use of the mobile devices and connected devices. Obviously, the type of device and operating system (OS) may be derived. Furthermore, the measurements may consist of the frequency and type of use of the mobile device, such as social media, online browsing, SMS texting, gaming, video streaming, taking pictures and phone calls. Apart from the presence of apps on the device that are used for these purposes, one may also consider the frequency and amount of use. However, access to app (meta)data is, of course, limited for privacy reasons and the potential amount and detail varies between apps, even when respondents would consent. The mobile device may also provide insight into other ICT through its Wi-Fi connection(s) and

through Bluetooth connections. The mobile device can, for example, measure the type, speed and strength of the Wi-Fi at home or provide a list of Bluetooth connected devices. Although it is not the purpose of the ICT, the mobile use data also present a view on the general ICT profile of a person.

HBS: The HBS deals with all kinds of purchases and expenditures, both small and large, and both frequent and infrequent. Some of these purchases are done on site, such as shops, restaurants and cinemas. Linked to the Eurostat HBS taskforce, some sensor measurements have been identified. Time-location sensor data may be employed to assist respondents in memorizing or recalling locations where products or services have been purchased. Some purchases are done online and part of those may be done through a mobile device. The use of certain online shopping apps may be tracked to again assist the respondent. In all of these cases, direct access to the type, amount and cost of products and services will, generally, not be possible to privacy restrictions on the apps. Another option is to use the camera to scan shopping receipts. This is mostly useful for purchases that involve many products/services simultaneously and that are burdensome to insert into a diary/questionnaire.

Table 4.3: Assessment of the sensor criteria for the suggested survey-sensor pairs. A flag means that the pair scores positively on the criterion.

Survey	Sensor	Omnipresence	Data access	Quality	Costs
LFS	Location	ν	ν	ν	ν
	Device use	ν		ν	ν
EU-SILC	Camera	ν	ν		ν
	Microphone	ν	ν		ν
	Location	ν	ν	ν	ν
	Device use	ν		ν	ν
EHIS	Location	ν	ν	ν	ν
	Motion	ν	ν	ν	ν
	Heart rate	ν	ν	ν	ν
	Wearables			ν	(ν)
	Camera	ν	ν		ν
ICT	Device use	ν		ν	ν
	Device properties	ν		ν	ν
HBS	Location	ν	ν	ν	ν
	Camera	ν	ν		ν
	Device use	ν		ν	ν
HETUS	Location	ν	ν	ν	ν
	Motion	ν	ν	ν	ν
	Device use	ν		ν	ν
	Wearables			ν	
	NFC		ν	ν	ν
	Beacons		ν		

HETUS: The HETUS has some similarity to ICT but is much broader and covers all possible daily activities, not just ICT related. Within the existing Eurostat HETUS task force, various forms of sensor measurements have been suggested. Since part of the activities are location related, such as work, shopping or dining, an obvious set of sensor data consist of time-location measurements. Such data can be enriched using geographical context information obtained from Google street maps data and other sources. The time-location data may also just function as a time roster for respondents. Part of the activities are physical, such as walking, jogging or cycling, and may be derived from activity trackers and other wearables, or from mobile device motion sensors. Parallel to the ICT survey, similar mobile device use data may be extracted directly to get information on ICT related activities. All these sensor data may cover many outdoor activities

and some indoor activities when they are linked to the mobile device, but most indoor activities remain out of reach. A possible way to assist respondents in indoor activities is to use Bluetooth beacons for indoor localization. This would, however, require the installation of one or more beacons. Another way to assist respondents is through a set of NFC tags. NFC tags initiate actions on a mobile device or wearable once they are held against the tags. The tags may be stuck to objects in various rooms, so that respondents can quickly indicate what they are doing.

We have suggested a range of sensors for the different ESS surveys. The important question is how these sensor measurement score on the sensor criteria and the respondent criteria. The sensor criteria are omnipresence, data access, quality and costs. The respondent criteria are respondent willingness, data handling and burden. We present simple flags in tables 4.3 and 4.4, respectively, when the sensors score positively on the criteria. We discuss only those pairs and criteria that score negatively on both tables.

In table 4.3, we indicated all wearables as not omnipresent. In addition, NFC tags and Bluetooth beacons need to be provided to respondents. Data access is scored negatively for device use and device properties, as part of the apps is not accessible to other apps. Furthermore, wearables' sensor data are scored negatively, as their data are often owned by other parties. Quality is the hardest criterion to assess. For now, we scored camera and microphone recordings as negative, since quality depends strongly on the respondent. We also scored beacons as negatively as their positioning depends on the respondent. When it comes to costs, wearables may not be affordable for the HETUS as they cover only part of the activities. They may be affordable for EHIS as they are closely related to fitness. Bluetooth beacons may be too expensive in the HETUS as well.

Table 4.4: Assessment of the respondent criteria for the suggested survey-sensor pairs. A flag means that the pair scores positively on the criterion.

Survey	Sensor	Willingness	Data handling	Burden	Feedback
LFS	Location	۷		۷	
	Device use	?		۷	۷
EU-SILC	Camera	?	۷		
	Microphone	?	۷		
	Location	۷		۷	
	Device use	?		۷	۷
EHIS	Location	۷		۷	۷
	Motion	۷		۷	۷
	Heart rate	?		۷	۷
	Wearables	?			۷
	Camera	?	۷		
ICT	Device use	?		۷	۷
	Device properties	?		۷	
HBS	Location	۷		۷	
	Camera	?	۷		۷
	Device use			۷	۷
HETUS	Location	۷		۷	۷
	Motion	۷		۷	۷
	Device use	?		۷	۷
	Wearables	?			۷
	NFC	?		۷	۷
	Beacons	?			

The respondent criteria in table 4.4 often score negatively or unknown. For the respondent willingness criterion, it is very hard to predict whether sensors measurements will be approved. From experience we

know that location and motion measurements are often accepted, but for all other sensor data it yet has to be tested empirically. The data handling criterion is scored negatively for most of the sensor data because the devices themselves do not generally provide access to them. As a consequence, any data handling has to go through the survey institute which imposes a logistical and procedural burden. The exception are camera and microphone recordings that can usually be viewed or listened to before they are stored and submitted. All sensor data are scored negatively on burden when they involve effort from the respondent other than accepting the measurements for a certain time period. The exception to this is the set of NFC tags in the HETUS. The tags do require effort but only to avoid effort in completing the diary. Although, wearables collect data passively, they are still scored as negatively, because they need to be worn for a longer time period. Finally, for the feedback criterion, all location measurements, including indoor localization using beacons, are viewed as weakly informative to respondents in the LFS, EU-SILC and HBS. In these surveys, we anticipate that respondents learn little from an overview of locations they have visited. The measurements related to living conditions in the EU-SILC are also scored negatively, as we expect that respondents know this information themselves. For the same reason, the photos in the EHIS are not deemed valuable as respondents know themselves best. This holds true also for device properties.

Summarizing, there are no survey-sensor pairs that score positively on all sensor and respondent criteria, but there are a few that score on many. The sensor data that score best are the time-location data.

4.3 Secondary sensor data

Finally, we discuss potential secondary sensor data for each of the ESS surveys, except, again, for the AES. Within the Eurostat taskforces for HBS and HETUS, similar explorations have been conducted. We include and extend these explorations.

Table 4.5: Potential secondary sensor data.

	Sensor data
LFS	Social media data, Mobile phone provider data, Internet provider data
EU-SILC	Social media data, Mobile phone provider data, Internet provider data, Smart energy use meters data (electricity, water)
EHIS	Wearable sensor data
ICT	Social media data, Mobile phone provider data, Internet provider data
HBS	Scanner data from shops, Bank transaction data, Loyalty card data
HETUS	Wearable sensor data

Secondary sensor data pose the challenge of collaborating with a third party company, as well as legal and ethical challenges. It does, however, offer access to a trove of data without – or very little – respondent burden and is, thus, worth investigating. The potential impact on the privacy of the respondents suggests that secondary sensor data should be offered as a choice to the respondents with a clear explanation of the benefits and risks involved.

The LFS, the EU-SILC, and the ICT surveys could benefit from social media data or mobile phone provider data or internet provider data. Social media data would be provided by global companies like Facebook, Twitter or Instagram. Data access could be requested from the respondent for public data by providing user names or handles, or – with permission from the respondents – from the companies for private data. Mobile phone provider data would be contributed by a mobile phone provider local to the country of the respondents. Possible data include phone call, SMS and internet usage as well as time-location data with a coarse resolution in space and time. The respondent's internet provider could supply data covering the internet use at home.

EHIS and HETUS could make use of wearable sensor data. Companies like Fitbit collect data via their own wearables, Strava collects data via apps, and companies like Apple or Garmin have both data channels at their disposal. Respondents could allow access to their already existing (historical) sensor data, or start collecting at the start of the survey.

The HBS might make use of scanner data from shops, as well as bank transaction data and loyalty card data. Scanner data would be limited to large chains and can be matched to a respondent via time-stamp and purchase amount. Bank transaction data can cover purchases independent of shops but have varying coverage depending on the penetration of electronic payments in the target country. Requesting bank transaction data is of course more intrusive to the respondent's privacy. In project @HBS, financed by Eurostat, the utility of these data sources is explored.

We evaluate the options in tables 4.6 and 4.7 for the sensor and respondent criteria, respectively.

Table 4.6: Assessment of the sensor criteria for the suggested survey-sensor data pairs. A flag means that the pair scores positively on the criterion.

Survey	Sensor data	Omnipresence	Data access	Quality	Costs
LFS	Social media	✓		?	✓
	Mobile provider	✓		?	✓
	Internet provider	✓		?	✓
EU-SILC	Social media	✓		?	✓
	Mobile provider	✓		?	✓
	Internet provider	✓		?	✓
	Energy meters			?	✓
EHIS	Wearables			?	✓
ICT	Social media	✓		?	✓
	Mobile provider	✓		?	✓
	Internet provider	✓		?	✓
HBS	Scanner data	✓		✓	✓
	Bank transactions	✓		✓	✓
	Loyalty card			✓	✓
HETUS	Wearables			?	✓

Table 4.7: Assessment of the respondent criteria for the suggested survey-sensor data pairs. A flag means that the pair scores positively on the criterion.

Survey	Sensor data	Willingness	Data handling	Burden	Feedback
LFS	Social media	?		✓	
	Mobile provider	?		✓	✓
	Internet provider	?		✓	✓
EU-SILC	Social media	?		✓	
	Mobile provider	?		✓	✓
	Internet provider	?		✓	✓
	Energy meters	?	✓	✓	✓
EHIS	Wearables	?		✓	✓
ICT	Social media	?		✓	
	Mobile provider	?		✓	✓
	Internet provider	?		✓	✓
HBS	Scanner data	?	✓	✓	✓
	Bank transactions	?	✓	✓	✓
	Loyalty card	?	✓	✓	✓
HETUS	Wearables	?		✓	✓

The sensor criteria in table 4.6 vary greatly in scores. The costs are scored positively for all options, since data have been collected by other parties. That does not mean that there are no costs; data holders my want to be reimbursed for the processing and preparation of the data either by payments or by receiving data in return. The data access criterion is scored negatively for all options as data is not owned by the NSI. The lack of data ownership implies that linkage and storage of data have to adhere to strict legislation procedures. The other two criteria, omnipresence and quality, show a diverse picture. Omnipresence is positive for all options except smart energy meter data and wearable sensor data. Quality is scored positively for scanner data, bank transactions data and loyalty card data. These data are crucial to the data holders. For the other options quality is set at unknown, and most likely depends on the type of information that is extracted.

The respondent criteria in table 4.7 show great variation as well. The data handling criterion is mostly set at a negative score. This is done because it will be logistically complex to show to respondents what data are used at any given time and to offer them the option to alter or retract data. On the positive side, the burden criterion is always scored as positive; the respondent will, in general, only have to consent. Furthermore, the feedback criterion is also scored positively for most options. The willingness criterion is set to unknown for all options, as there is (to our current knowledge) little empirical evidence about consent to linkage to the secondary data in a survey context.

Summarizing, the secondary data options have strong features from the viewpoints of omnipresence, costs, burden and feedback, but weaker features when it comes to data access and data handling. Quality and respondent willingness to link data are mostly unknown and need to be assessed in real studies.

5. Discussion

We contribute to existing literature in two ways: We propose a set of criteria to support cost-benefit assessments of sensor measurements and sensor data and we make a first inventory of sensor options for all ESS surveys. The criteria are constructed from three viewpoints. The first viewpoint is from the perspective of the survey itself; does the survey contain topics or questions that may benefit from automated measurements. The second viewpoint is that of the sensor; What are accuracy and costs of the sensor options. The final viewpoint is the respondent; How does the respondent react to a request for sensor data. For each ESS survey, pairs of topics and sensor measurements are identified. All pairs are evaluated against the three types of criteria.

Two side remarks are in place before we conclude our evaluation. First, the pairs of survey topics and sensor measurements that we identified in this deliverable, are by no means exhaustive. Others may come to new or alternative options. Also in time, new options may emerge. There is some literature on promising pairs of survey topics and sensors, but this is scattered and not complete. One cause is that developments go faster than academic journals can follow; it is mostly at conferences where references to new applications can be found. Furthermore, commercial parties are involved in part of the projects, which are less inclined to publish details of their studies. For these reasons, we decided to restrict ourselves to an enumeration of what we conjecture to be potential applications. Nevertheless, a follow-up to this paper should give a more in-depth account of existing literature and case studies. Second, our assessments of the survey, sensor and respondent criteria are subjective and, in some cases, even speculations. This is partly for the very reason that they are new and have not been tried in practice. It is very hard, for example, to predict respondent willingness to provide sensor data (or to consent to sensor data linkage) independently of the context. Another reason is that sensors and wearables by themselves show variety in accuracy and costs, even within mobile devices, so that it is hard to judge about quality. It is imperative that the assessments are made more rigorously by consulting multiple experts. Also this exercise may be part of an in-depth follow-up paper.

In our inventory, we distinguish sensor measurements initiated by the survey institute and existing secondary sensor data. Although the data may originate from the same type of sensors, the context of the two is very different. Self-initiated sensor data require a data collection infrastructure and lead to direct data collection costs. Obviously, other aspects such as data processing, data storage, privacy and legislation are very different as well. In this paper, we consider secondary data as complementary to survey data, i.e. we ask respondents to consent to linkage. Consequently, a hybrid form of data collection arises. However, such a combination of data sources is no goal by itself; secondary data may provide the sole source of data for specific topics. The on-going ESSnet Big Data attempts to identify a range of areas where this is the case for official statistics. It is likely that from the big data perspective, survey data may be used as complementary data in some settings. In the near future, the two starting points may, in fact, lead to similar hybrid forms of data collection.

The most promising combinations of sensor data and survey data are those that score well on all three sets of criteria. For all ESS surveys, except AES, we found pairs of survey topics and sensor measurements that potentially have a positive business case. These examples deserve further exploration and elaboration. That being said, we also concluded that there may often be one or more obstacles in exploiting the sensor measurements or sensor data. These obstacles are related to respondent willingness, data access, data handling or unknown quality.

We make two recommendations: First, we advise to replicate our assessments of the various criteria with experts in mobile device and wearable sensors, especially for quality. Second, we propose to empirically test respondent willingness to provide sensor data and to consent to linkage to existing secondary data. Such experiments have emerged, but are yet at early stages.

References

Bankova, G., Kloft, M., Maneshi, P., Tran, T. (2018), Data integration from fitness apps and wearable devices, Technical report, Statistics Netherlands, The Hague.

Haas, G.C., Kreuter, F., Keusch, F., Trappmann, M., Bähr, S. (2018), What do researchers have to invest for collecting smartphone data?, Paper presented at International Workshop on Household Survey Nonresponse, August 22-24, Budapest, Hungary.

Li, B.; Harvey, B.; Gallagher, T. (2013), Using barometers to determine the height for indoor positioning. In: Indoor Positioning and Indoor Navigation (IPIN), 2013 International Conference on. IEEE, 1 – 7.

Struminskaya, B., Lugtig, P., Schouten, B., Toepoel, V., Haan, M., Dolmans, R., Giesen, D., Luiten, A., Meertens, V. (2018), Collecting Smartphone Sensor Measurements in the General Population: Willingness and Nonparticipation, Paper presented at BigSurv 2018, Oct 25 – 27, Barcelona, Spain.