

Missing Data and Other Measurement Quality Issues in Mobile Geolocation Sensor Data

Social Science Computer Review
1-24

© The Author(s) 2020



Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/0894439320944118
journals.sagepub.com/home/ssc



Sebastian Bähr¹, Georg-Christoph Haas^{1,2}, Florian Keusch²,
Frauke Kreuter^{1,2,3}, and Mark Trappmann^{1,4}

Abstract

As smartphones become increasingly prevalent, social scientists are recognizing the ubiquitous data generated by the sensors built into these devices as an innovative data source. Passively collected data from sensors that measure geolocation or movement provide an unobtrusive way to observe participants in everyday situations and are free from reactivity biases. Information on day-to-day geolocation could provide valuable insights into human behavior that cannot be collected via surveys. However, little is known about the quality of the resulting data. Using data from a 2018 German population-based probability app study, this article focuses on the measurement quality of geolocation sensor data, with a strong focus on missing measurements. Geolocation sensor data are an example of an available data type that is of interest to social science research. Our findings can be applied to the wider subject of sensor data. In our article, we demonstrate (1) that sensor data are far from error-free. Instead, device-related error sources, such as the manufacturer and operating system settings, design decisions of the research app, third-party apps, and the participant, can interfere with the measurement. To disentangle the different influences, we (2) apply a multistage error model to analyze and control the error sources in the specific missingness process of geolocation data. We (3) raise awareness of error sources in geolocation measurement, such as the use of GPS falsifier apps, or device sharing among participants. By identifying the different error sources and analyzing their determinants, we recommend (4) identification strategies for future research.

Keywords

app data collection, smartphones, measurement quality, geolocation, passive measurements

As smartphones become increasingly prevalent in Western societies (Pew Research Center, 2018), social scientists are recognizing the ubiquitous data generated by the sensors built into these devices as

¹ Institute for Employment Research, Nuremberg, Germany

² University of Mannheim, Germany

³ University of Maryland, College Park, MD, USA

⁴ University of Bamberg, Germany

Corresponding Author:

Sebastian Bähr, Institute for Employment Research, Nuremberg, Germany.

Email: sebastian.baehr@iab.de

an innovative source for behavioral and social research (Raento et al., 2009; Sugie, 2018). Passively collected data from sensors that measure geolocation or movement are hailed because they are allegedly generated “[...] without any participation or action from the subject” (Onnela & Rauch, 2016, p. 1692). The resulting data are seen as “[...] unobtrusive, naturalistic observational records that reduce the likelihood that participants will behave reactively” (Harari et al., 2017, p. 83). In particular, information on day-to-day geolocation could provide valuable insights into human behavior patterns (Kelly et al., 2013) that cannot be collected via surveys (Nitsche et al., 2012).

However, while the praise is intriguing, little is known about the overall quality of data resulting from smartphone sensors (Jäckle, Gaia, et al., 2019). We address this gap using unique data collected from German smartphone users through a research app (Kreuter et al., 2018). While some previous studies investigate potential coverage and nonresponse bias resulting from differential smartphone ownership and willingness to participate in such research studies (Antoun et al., 2018; Elevelt et al., 2019; Hargittai, 2018; Jäckle, Burton, et al., 2019; Keusch et al., 2020a, 2020b), our research is concerned with the measurement quality of smartphone sensor data. By measurement quality, we mean the absence of measurement error and a high degree of completeness of the data to be measured. Measurement error comprises any deviations of measurement from the true score, where systematic error (bias) is usually distinguished from random error (variance; Miller, 2008). While systematic error leads to biased estimates, random error leads to an increase in the variance of estimates. An assessment of this quality aspect is highly desirable. However, it requires the availability of validation data representing the true score and we can present only partial evidence on this aspect of measurement quality.

The focus of this article is on the second quality aspect, the missingness of data from units that generally participated in the study (called “item nonresponse” in the survey context). Like random measurement error, missing data lead to a loss of efficiency (i.e., an increase in the variance of estimates). If missingness does not occur completely at random, that is, missingness is uncorrelated to its unknown value and all other variables (Rubin, 1976), and the problem is not addressed by adequate methods (e.g., multiple imputations), this leads to additional bias in the estimate.

We focus on geolocation data as an example of a readily available data type that is of interest to a broad range of social science research. Geolocation sensors face technical challenges as environmental (e.g., indoor areas and densely populated urban areas), device, and user-related factors affect the measurement accuracy or the ability of a smartphone to establish geolocation (Blunck et al., 2011; Ranasinghe & Kray, 2018). However, most of these studies are technical tests based on convenience samples. By contrast, our study is based on a sample of the general German population and focuses on various processes resulting in missing data and user-introduced aspects of geolocation measurement quality in smartphone data collection for social science research. Our multistage approach provides insights into the quality of other passive measurements taken from smartphones.

At the first workshop on Mobile Apps and Sensors in Surveys, Couper (2019) called the notion that passively measured sensor data are flawless, “passive fantasy.” Aside from issues of selectivity, when inferring from a smartphone-owning participant sample, several factors can potentially influence whether measurements are recorded or valid. For example, built-in sensors from different manufacturers might differ in their measurement quality, different operating system (OS) settings affect geolocating accuracy, third-party apps stop data collection, and participants themselves affect measurements by their handling of the smartphone (e.g., handgrip and walking style; Blunck et al., 2011). However, even ignoring these issues, smartphone sensors might offer a limited perspective of human movement and physical activity. If we assume that some users do not always carry around their devices or turn them off from time to time, then immobility inferred from smartphone data might just represent a device forgotten at home or resting in a desk drawer at work, a purse, or a gym locker (see e.g., Casilari et al., 2016, p. 2).

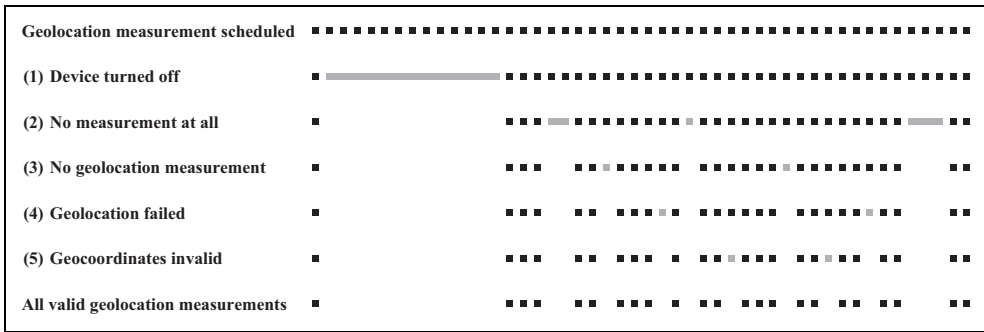


Figure 1. A five-stage model of smartphone geolocation data missingness. *Note.* This figure shows the missingness process. Episodes where the device was turned off (1) will generate missing geolocation data, as do episodes where the research app could not collect any measurements at all (2). For the remaining time, there are episodes where no geolocation measurements were collected (3) despite an otherwise active research app. Even for the times where a measurement was collected, there are several cases, where our attempt at geolocation was unsuccessful (4), leading to missing geocoordinates. Finally, some collected geolocations are invalid (5). In the context of our study, this includes cases of faked geolocations, shared devices, or the use of multiple concurrent devices.

In any given smartphone data collection of geolocation, a significant proportion of the measurements that were designed to be collected will not be successful. This missingness can be caused by intended or unintended user behavior, such as switching off the smartphone or running out of battery, by characteristics of the location (no signal), or by technical problems related to the hardware or OS of the smartphone (e.g., energy-saving modes).

To disentangle this plethora of factors interfering with data quality, we propose a multistage model of geolocation missingness. The proposed model includes general preconditions for data collection (i.e., device being turned on and allowing data collection) and processes specific to geolocation measurement that lead to valid geocoordinates in the context of a study. Each stage of the model is dependent on success in the previous stage, thus allowing us to rule out more general explanations for missingness and to focus on the occurrence of missing data and the duration of gaps in measurement at the current stage. We use data from the IAB-SMART study in Germany (Kreuter et al., 2018) and focus on four groups of explanatory variables: participant characteristics, device hardware, device state, and temporal and spatial context of the measurement.

A Stage Model of the Missingness of Mobile Geolocation Data

To disentangle the plethora of factors that can interfere with the successful measurement of geolocation on smartphones, we develop a multistage model that—on the first stage—analyzes general missingness processes unspecific to the type of data collected and works its way down to error sources specific to geolocation data.

By distinguishing the different processes, we can model the *occurrence*, defined as the identification of geolocation measurement gaps, and the *extent*, defined as the length of each identified gap in terms of the number of missing geolocation measurements, in each stage separately to avoid confounding effects of the previous stage. The separate modeling of gap occurrence and extent allows for factors to have differential effects on either. With this novel approach, we can systematically learn about the importance of different factors responsible for errors in each stage, which is a prerequisite for strategies to address these missing processes. Figure 1 illustrates the different stages of our multistage model and their effect on geolocation missingness.

The first four stages deal with the missingness of geolocation measurements we obtain, and the results are applicable to all research using geolocation smartphone data. In contrast, Stage 5 focuses on the measurement error of geolocation measurements once the data are collected. While addressing error at these stages is vital, it is to some extent dependent on the individual research aims.

In the first and most general stage, we analyze whether the *device is turned on* (1) at the scheduled time of a geolocation measurement. If a participant regularly turns off the device, this systematically affects our measurement and inference. Different reasons have been identified for turning off a smartphone: being in quiet locations such as at a university lecture (Ugur & Koc, 2015), at night (Chen et al., 2013), coping with overuse and its related problems (Lee et al., 2014), or consciously disengaging from work (MacCormick et al., 2012). Device downtime could also have technical reasons, most prominently when running out of battery. How fast smartphones run out of battery power and are subsequently turned off depends on factors such as the brand, model, and OS of the device (Kongsiriwattana & Gardner-Stephen, 2016; Zhang et al., 2013); the charging patterns of users (Ferreira et al., 2011), how intensively people use their smartphones (Oliver, 2010); what features and sensors of a smartphone are used (Hosio et al., 2016; Metri et al., 2012); and external factors such as outside temperature (Peltonen et al., 2016).

In the second stage, we check whether our app *could collect any measurements at all* (2) from the turned-on device. Energy-saving modes, restrictive permission management by the OS, or even third-party task killer apps could prevent an app from collecting data at the time of a scheduled geolocation measurement.

Missingness in the first two stages can also be due to participants temporarily uninstalling our research app. Approximately 12.5% of participants in our study installed the app more than once, so this is a nonnegligible factor driving data missingness in these stages. We control for this factor in our models. For participants who permanently uninstall our app, our analyses use their data only up to the point the app is uninstalled.

In the third stage, we focus on missing *geolocation measurement* (3) given that the device is turned on and the app is actively collecting data. Technical reasons for failed measurement include a low battery level and the availability and strength of GPS, Wi-Fi, and mobile network signals. An additional reason for measurement failure at this stage could be that users may manually disable location tracking in the general smartphone settings.

One could think that the existence of a geolocation measurement is a sufficient condition for substantive analyses. In fact, the existence of a measurement simply means that the app attempted to establish the geographic location of the device at the scheduled time; however, *geolocation can fail* (4), that is, although a measurement was initiated, a geolocation with no information is recorded. This differs from the third stage, where we miss a geolocation measurement, in that the measurement itself exists, but geographic coordinates are missing. A failure in recording the actual location means none of the built-in methods, GPS, mobile carrier network, Google's built-in Fused Location Provider application programming interface (API) that uses Wi-Fi information, or the cell tower location database could successfully establish geolocation. The main reason for failing to establish geolocation is indoor location(s), where reception depends on the number of walls, the wall material, and the closeness of surrounding buildings (Blunck et al., 2011; Kjærgaard et al., 2010). Also, the device type and OS version determine the quality of the built-in sensors, the computational resources, and the underlying algorithms (Ranasinghe & Kray, 2018, p. 10). User-related factors such as mobility or the placement of the smartphone on the body (Blunck et al., 2011) can also influence the quality of geolocation measurement.

A scheduled measurement that includes actual coordinates might still be *invalid* (5) for subsequent substantive analyses. For example, sharing devices with others or passing on study invitation codes to other household members makes the mapping of data to the corresponding targeted participant

impossible. Also, geolocation is a sensitive data type that could encourage participants to falsify their geolocation using third-party apps.

Data and Methods

We analyze the determinants of each of the five stages, using data from the IAB-SMART study (Kreuter et al., 2018). The following section provides background on the study, details about our variables of interest and the explanatory variables, and finally describes the analytical models.

The IAB-SMART Study

In January 2018, we invited 4,293 participants of the long-standing German panel study “Labour Market and Social Security” (PASS; Trappmann et al., 2019) to install the IAB-SMART app on their smartphones for 6 months. All invited participants had stated in PASS Wave 11 (2017) that they owned a smartphone with Android Version 4.0 or higher.

To ensure European Union General Data Protection Regulation-compliant sensor data collection, the IAB-SMART app requested consent to five different data collection functions: (1) *location and network connectivity data*, (2) *call and text message logs*, (3) *phonebook characteristics*, (4) *activity sensor data*, and (5) *app usage data*. By default, each function was deactivated. Participants needed to provide individual consent for each function and could withdraw consent and stop sensor data collection at any point in time using the app. In addition, we regularly administered in-app surveys during the study period from January to August 2018. Overall, 625 (14.6%) invited sample members downloaded the IAB-SMART app from the app store, went through the installation process, completed the short welcome survey in the app, and could be verified as eligible PASS Wave 11 participants (“verified installers”).

As this article is concerned with missing information in geolocation data collection, we focus on the first function of the IAB-SMART app, *location and network connectivity data*. If active, the function attempted to collect geolocation information using a two-step sampling process. First, to preserve the battery, our app checked whether geolocation already was available in the Android system. This was the case if the system or third-party apps recently had acquired geolocation. This geolocation could be established via GPS, the mobile carrier network, or Wi-Fi via Google’s built-in Fused Location Provider API. Only if no current location was available did our app try geolocation using GPS, the mobile carrier network, or a database of known cell tower locations proprietary to our app developer to establish at least a coarse location as a last resort. This second step did not make use of Wi-Fi information because the app developer custom-built this app feature according to one of their default templates, which—for compatibility reasons—only relies on native Android functions.

Each method achieves a different geolocation accuracy: GPS has the highest accuracy (median accuracy: 11 m), followed by the mobile carrier network (median accuracy: 21 m), Wi-Fi (median accuracy: 36 m), and the cell tower database (median accuracy: 967 m). These results are in line with those reported in earlier studies (e.g., Zandbergen, 2009). Except for GPS, the accuracy of geolocations from all methods varies considerably due to outliers with high uncertainty.

For each measurement, the app either recorded available locations or attempted to record the location with the highest accuracy of GPS, carrier network, or the cell tower database. From Table 1, we see that in more than half of all measurements, the carrier network was chosen because this technique offers reasonably precise geolocation without the device having to sync its location with multiple satellites. GPS, as the energetically most demanding option, was the third most frequent mode chosen, followed by Wi-Fi. Even though the cell tower database was used about 15% of the time as a last resort, no geolocation could be established in about 18% of all geolocation measurements. These shares of the different location providers are certainly the result of our two-step method. While

Table 1. Measurement Accuracy by Geolocation Source.

Geolocation Information Provider	% Share of Geolocation Measurements	Accuracy Vertical/Horizontal (in Meters)	
		Median	SD
Carrier network	55.22	20.54	34,569.30
GPS	9.54	11.00	27.16
Wi-Fi	2.77	36.36	519,809.84
Cell tower database	14.75	967.15	4,792.99
Not available	17.71	—	—
Total	100.00	19.56	90,737.05

locations based on GPS and the carrier network could be collected in both steps, locations based on Wi-Fi were only recorded if they were already available in the Android system. The developer's cell tower database is an additional source for geolocations in the second step and has certainly provided geolocation in cases where GPS or the carrier network would not have helped.

We implemented a 30-min sampling rate to address privacy issues and balance our data needs with battery consumption. Overall, 578 individuals were willing to share their geolocation ("network and location participants"). To identify gaps in the collection of data, we used every available data source collected during the IAB-SMART study (see below). For our analysis sample, we end up with 88,941,180 measurements from 499 participants, of which 1,031,165 are geolocation measurements. Table A1 in the Appendix displays the selection process of participants from participants of the PASS study to our analysis sample. Coverage and nonresponse analyses are beyond the scope of this article but are reported in the studies of Keusch et al. (2020a, 2020b).

Table 1 shows that even for the times when a measurement was collected, in several cases, *geolocation failed* (4) leading to missing geocoordinates. This could be a result of our two-step sampling process, where our app relied on GPS or carrier network but not Wi-Fi-information to establish geolocation. Finally, some collected *geocoordinates are invalid* (5). In the context of our study, this includes cases of faked geolocations, shared devices, and—depending on the research question and population—measurements outside of Germany. In our analyses, we separate each missingness process and analyze the occurrence and extent of each process individually, independent from the previous stages.

Variables

In the following, we list and briefly describe our variables of interest (the gaps) and the factors we hypothesized to be responsible for the occurrence and extent of the gaps (explanatory variables).

The gaps. Our data consist of measurements and survey responses with precise timestamps and paradata concerning, for example, information on the device, its operational state (i.e., battery level, display state), and the state of the mobile network connection. From these paradata, we create five variables of interest that align with the five stages of our model in Figure 1: (1) *turned off devices*, (2) *no measurements at all*, (3) *no geolocation measurements*, (4) *failed geolocation*, and (5) *invalid geocoordinates*. In the Appendix, we provide descriptive statistics for all dependent and independent variables at the participant level (Table A3) and measurement level (Table A4). Figure 2 illustrates the first three measurement gaps, which are related to the absence of scheduled geolocation measurements.

Some gaps in data collection occur naturally or by design. For the purpose of this study, we are interested in only those gaps that lead to a loss of scheduled geolocation measurements. Our planned

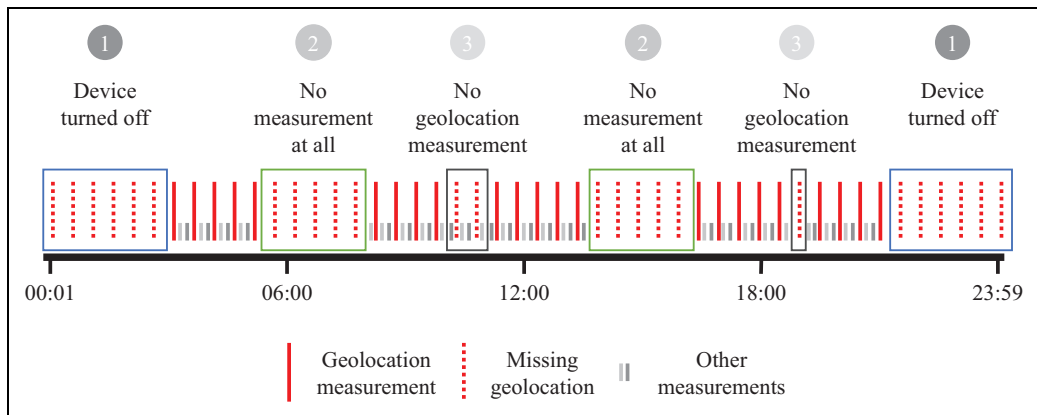


Figure 2. An example of missingness at Stages 1–3. Note. This figure shows the first three missingness processes. Episodes where the device was turned off (1) no data are collected, including no geolocation data. Similarly, there are episodes where the device is on, but the research app could not collect any measurements at all (2). For the remaining time, there are episodes where no geolocation measurement was collected (3) despite an otherwise actively collecting research app.

30-min geolocation sampling rate is satisfied, with a median actual sampling rate of approximately 30.7 min. However, the mean of 63.24 min and its standard deviation of 1,155.18 min point to considerable variation in the extent of gaps, which could be due to on-device processes that delay the geolocation measurement or device downtimes. Therefore, we choose a conservative measure to consider a geolocation measurement to be lost, namely, when the gap between two geolocation measurements is equal to or exceeds 60 min. In this way, we can be sure that at least one additional geolocation measurement was scheduled but is missing. Doing so, we analyze only a subgroup of all actual measurement gaps: Approximately 12% of all gaps between geolocation measurements are at least 60 min.

We identify (1) *turned off devices* from the information of the last device start. The span from the last time the IAB-SMART app could collect data until the recorded restart time constitutes the turned off device episode. We combine all data collected by the IAB-SMART app into a single data set, where each measurement episode and turned off device episode are represented as a data row with start and end time stamps.

This data structure provides an overview not only of the sequence of measurements but also the gaps in the collection of any data. From these gaps, we generate our second variable of interest, (2) *no measurements at all*. These gaps also constitute data rows, creating a gapless data set (at the 1-second level) from the first to the last observation of each participant.

We next identify (3) *no geolocation measurements* using the interval between location measurements. We define all intervals with a duration of at least 60 min as problematic and flag them as gaps. To rule out turned off devices and gaps in general data collection as explanations, we consider only the intervals between two location measurements where other measurements (e.g., app use data or activity sensor data) were collected.

Even if the app triggers a geolocation measurement as planned, this does not ensure that the geolocation, that is, the identification of the geographic location, was actually successful. In such a case, the information on longitude and latitude would be missing in our data set. According to Google (2019b), failed geolocations should rarely occur and is mainly due to disabled location in the device settings. However, in our study, 17.5% of all location measurements contain missing longitude and latitude coordinates. We identify such cases as (4) *geolocation failed*. Compared to (3), the difference

is that in (4), we did initiate a geolocation measurement as planned, nonetheless, the data are missing because of failed geolocating.

Finally, having collected functioning location data, it is possible that the (5) *coordinates are invalid* for use in substantive analyses, for example, when participants falsify their real location with third-party apps, pass on their invitation code to other people, share their device with others, or use more than one device with the research app installed. We obtained the information on device sharing from an in-app survey at the end of the study period. Participants who shared their device with others for more than 10 cumulative days over the study period, we define as problematic because the passively measured data cannot be attributed with certainty to the targeted participant (see Table A2 for details on the reasons for device sharing). From the paradata on device type, we could infer concurrent installations on multiple devices. We gained information on location faking from analyzing the app use data of participants. The use of apps that falsify the geolocation disqualifies the use of the participants' data. We identify approximately 2.6% of all available geolocations as potentially invalid for substantial analysis.

The explanatory variables. We group the explanatory variables into four major lines of explanation: *First*, we use a set of *sociodemographic variables* consisting of age, gender, household size, and employment status of the participant. These variables were surveyed during the IAB-SMART study period and supplemented with information from the PASS survey. For the models of Stages (1) through (4), we also control for participants with invalid measurements due to behavior such as device sharing, having the app installed on multiple devices, and location faking. Because the frequency of each of these behaviors is low, we group these participants together and call them "problematic participants." By including an indicator for these participants, we can parse out the effect of problematic behavior, freeing the other coefficients from its influence. At the same time, we gain at least summary insights into the direction of bias by this behavior from the coefficient of the indicator variable itself.

Second, we use a set of variables with information about the device hardware. The variables include the device manufacturer to account for device quality and user interface. We choose Samsung as the reference category because it is by far the largest manufacturer of smartphones in our data. We also control for the Android OS version, which provides an indication of the age of the device and the quality of algorithms for battery optimization and geolocating. In addition, newer Android versions contain more restrictive energy management systems that could prevent our app from collecting data in the background (e.g., Google, 2019a). The number of CPU cores and the size of the device's RAM control for the computational resources that influence system stability.

Third, we add a set of variables on the state of the device at the time of measurement. This includes battery status and level as important indicators for energy-saving modes, which restrict the data collection, to prevent the device from powering down. For the same reasons, we monitor whether task manager or task killer apps are being used before the gaps occur. The display state provides information about system-initiated doze modes and user-activated power saving modes. We also control for temporary uninstallation of our research app and temporary revoking of geolocation permissions because both factors provide unique explanations for geolocation missingness. All these factors can impede measurement and could be attributed to technical issues or user behavior.

Fourth, we consider the spatial and temporal context of the measurement. We control for activated roaming, mobile network location outside of Germany, and geolocation measurement outside of Germany to indicate stays in border regions or abroad that could interfere with the geolocation measurement. We use the last known location provider to establish the quality of measurement. Controlling for the time dimensions (weekday or weekend and time of day) enables analysis of behavioral patterns, such as switching off devices at night or during work hours.

Not every measurement source provides all paradata (e.g., on the device information or the device state), we, therefore, imputed missing values in two ways. We imputed time-constant information

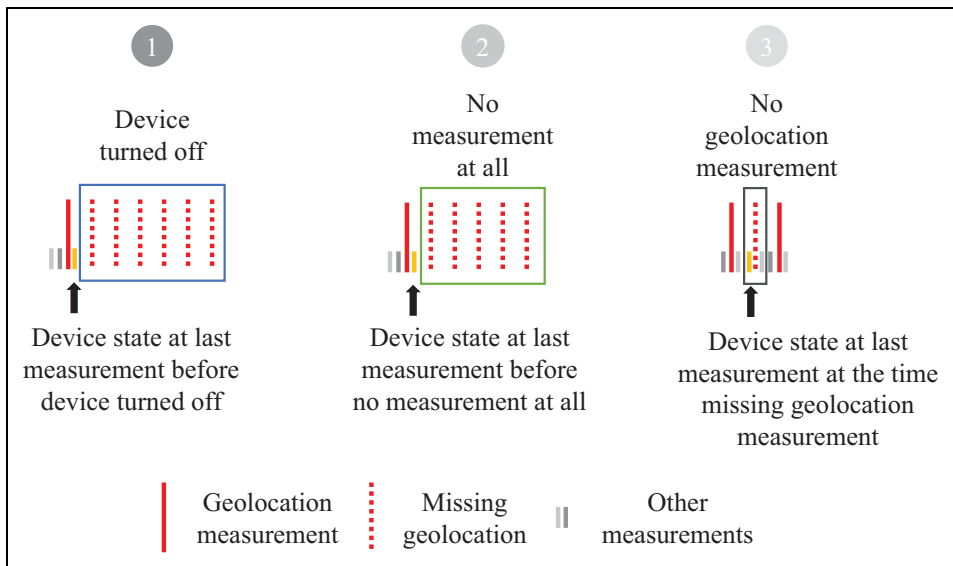


Figure 3. Analytical strategy for missingness at Stages 1–3.

(such as the device maker, RAM, etc.) with nonmissing preceding values if they were in accordance with the following non-missing information. For time-varying information (e.g., the battery level), we interpolated from values before and after the missing information. Including more elaborated imputation models in this complex set would exceed the scope of this article.

Models

For each of the Outcomes (1)–(3), we generate a variable identifying (a) the occurrence of such gaps and (b) their extent as a variable counting the number of geolocation measurements that are missing in each of these identified gaps. Figure 3 displays the analytical strategy for predicting the occurrence of the geolocation measurement due to one of the first three mechanisms. We use the measurement taken immediately before the gap occurred to inform us about the device state and the spatial and temporal context of the device when the gap occurred.

While the first three outcomes deal with the absence of geolocation measurements, that can span multiple planned measurements, *failed geolocation* (4) occurs only in existing measurements. Therefore, we analyze only the occurrence of this missingness. Measurement error due to *invalid coordinates* (5) is a collection of problematic behavior, which is heterogeneous, and the individual components rarely occur. We, therefore, limit ourselves to a descriptive presentation to inform future research.

Although our dependent variables deviate from the prerequisites of the linear model in that they are either binary outcomes or count variables, we follow the arguments of Angrist and Pischke (2009, p. 197) and Breen et al. (2018, p. 51) and provide all results as linear regression estimates since they are simple to compute, easily interpreted as marginal effects, and comparable across studies.¹

We include all explanatory variables that are not fixed over the study period as lags; thus, we consider the information from the measurement that immediately precedes the gap. In addition to the mentioned controls, we include indicator variables for the data permissions and data sources to capture the effects of varying smartphone use intensity and varying numbers or types of permissions on the frequency and extent of gaps. To provide a comparative overview, we use the same model specification on the occurrence (Models 1, 3, 5, 7 in Table 2) and extent (Models 2, 4, 6 in Table 2) of all gap phenomena.

Table 2. Multivariate Results.

Dependent Variables: Gap in Geolocation Measurement Due to . . .	(1) Device Turned Off		(2) No Measurement at All		(3) No Geolocation Measurement		(4) Failed Geolocation
	Occurrence Model 1 b/SE	Extent Model 2 b/SE	Occurrence Model 3 b/SE	Extent Model 4 b/SE	Occurrence Model 5 b/SE	Extent Model 6 b/SE	Occurrence Model 7 b/SE
Sociodemographics							
Age-group ref. 30–49	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
18–29	-0.001 (0.002)	8.524 (10.305)	-0.006 (0.007)	0.033 (0.433)	0.000 (0.001)	-0.580 (1.375)	0.018 (0.017)
50+	0.004 (0.003)	-4.385 (7.311)	0.015 (0.011)	0.068 (0.517)	0.001 (0.001)	-0.885 (1.004)	-0.027* (0.013)
Female	0.001 (0.002)	-3.459 (5.995)	0.001 (0.007)	-0.135 (0.383)	0.000 (0.000)	-1.207 (0.954)	-0.003 (0.011)
Size of household	0.001 (0.001)	-5.339* (2.421)	0.001 (0.003)	-0.308 (0.165)	0.000 (0.000)	0.227 (0.399)	0.000 (0.005)
Employment status ref. inactive	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Employed	0.000 (0.002)	1.902 (12.558)	0.005 (0.008)	0.875 (0.552)	0.000 (0.001)	1.202 (1.746)	-0.001 (0.017)
Unemployed	0.001 (0.003)	-1.057 (4.210)	-0.009 (0.011)	-0.291 (0.660)	0.001 (0.001)	-0.822 (1.298)	-0.006 (0.022)
Problematic participant behavior	-0.007** (0.002)	-34.875* (17.482)	-0.047 (0.025)	-1.651 (1.014)	0.000 (0.002)	-2.801 (1.452)	0.014 (0.031)
Device manufacturer: ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Device hardware							
Samsung							
HTC	0.000 (0.005)	93.945** (34.620)	0.029 (0.036)	2.813 (2.347)	0.002 (0.002)	3.718 (2.707)	-0.014 (0.059)
Huawei	-0.006* (0.002)	103.989** (16.052)	-0.025* (0.010)	1.930* (0.747)	0.002* (0.001)	4.948 (3.233)	-0.026 (0.018)
LG	-0.003 (0.005)	4.003 (19.666)	-0.035* (0.015)	-2.357** (0.771)	0.000 (0.001)	-0.217 (1.874)	-0.026 (0.041)
Motorola	-0.020** (0.003)	-7.973 (7.418)	-0.122** (0.022)	-2.420** (0.495)	0.000 (0.001)	-2.826* (1.189)	0.071 (0.054)
Other	-0.004 (0.004)	70.558* (31.964)	-0.028 (0.019)	2.224 (1.613)	-0.007* (0.001)	0.946 (2.125)	-0.007 (0.030)
Sony	-0.006** (0.002)	-10.378 (10.303)	-0.010 (0.010)	-1.819** (0.581)	-0.001 (0.001)	3.273 (1.712)	-0.055** (0.013)
Android OS version Ref. 7	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
4	-0.017* (0.008)	-27.299 (38.167)	-0.168** (0.040)	-5.787* (2.566)	-0.005* (0.002)	-1.645 (2.646)	0.044 (0.108)
5	-0.006 (0.004)	42.453 (22.430)	-0.123** (0.021)	5.944** (2.034)	0.001 (0.001)	2.864** (1.105)	0.031 (0.030)
6	-0.001 (0.003)	6.968 (9.077)	-0.056** (0.014)	0.879 (0.773)	0.000 (0.001)	3.944** (1.270)	0.000 (0.015)
8	-0.003* (0.001)	7.194 (5.793)	-0.020** (0.006)	0.034 (0.307)	0.001* (0.001)	-2.534 (1.355)	-0.009 (0.010)
Device no. CPU cores	0.001 (0.001)	-0.139 (1.395)	-0.001 (0.002)	-0.138 (0.108)	0.000 (0.000)	-0.003 (0.205)	0.003 (0.003)
Device RAM	-0.001 (0.001)	-5.664 (6.070)	-0.004 (0.006)	-0.591 (0.307)	0.000 (0.000)	-0.144 (0.606)	-0.008 (0.009)

(continued)

Table 2. (continued)

Device state	(1) Device Turned Off		(2) No Measurement at All		(3) No Geolocation Measurement		(4) Failed Geolocation	
	Occurrence Model 1 b/SE	Extent Model 2 b/SE	Occurrence Model 3 b/SE	Extent Model 4 b/SE	Occurrence Model 5 b/SE	Extent Model 6 b/SE	Occurrence Model 7 b/SE	
Lag geodata collection temporarily revoked	0.998*** (0.009)	44.525 (32.060)	0.464*** (0.086)	232.699 (223.611)			-0.076 (0.063)	
Lag app temporarily uninstalled	0.506*** (0.084)	692.253* (285.790)					-0.058*** (0.017)	
Lag task killer/manager app active	0.021 (0.056)	-33.096 (25.109)	-0.253*** (0.058)	0.955 (8.055)	-0.055*** (0.003)		0.602*** (0.077)	
Lag power save mode enabled	0.001 (0.004)	5.130 (7.644)	0.006 (0.014)	0.970 (0.707)	-0.005*** (0.001)	0.982 (1.085)	-0.004 (0.014)	
Lag device display state ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	
Display on								
Doze	-0.009*** (0.002)	-1.803 (7.691)	-0.020* (0.008)	3.139*** (0.595)	-0.003*** (0.001)	0.658 (1.962)	-0.028 (0.020)	
Doze suspend	-0.002 (0.002)	-10.802 (7.018)	0.086*** (0.009)	3.343*** (0.586)	-0.003*** (0.001)	-0.577 (1.091)	-0.015 (0.013)	
Display off	-0.006*** (0.001)	-16.192*** (5.631)	0.049*** (0.005)	-4.488*** (0.611)	-0.007*** (0.000)	-0.284 (0.584)	0.004 (0.005)	
Lag battery status ref. full	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	
Charging	0.004*** (0.001)	-1.190 (11.842)	0.059*** (0.004)	3.277*** (0.436)	-0.001*** (0.000)	6.748 (5.373)	-0.004 (0.006)	
Discharging	0.003*** (0.001)	4.405 (12.416)	0.105*** (0.006)	1.315*** (0.361)	-0.001* (0.000)	1.929 (1.318)	-0.003 (0.007)	
Not charging	0.003 (0.003)	-57.614* (27.394)	0.092*** (0.017)	0.287 (1.076)	-0.001 (0.001)	-0.693 (1.821)	0.012 (0.019)	
Lag battery level ref. 80%–100%	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	
0%–20%	0.029*** (0.004)	-21.559*** (5.400)	0.159*** (0.006)	-0.205 (0.547)	0.064*** (0.005)	1.489 (1.892)	-0.006 (0.006)	
21%–40%	0.003*** (0.001)	-7.781* (3.719)	0.011*** (0.003)	-0.646 (0.412)	0.001*** (0.000)	0.089 (0.796)	0.013*** (0.005)	
41%–60%	0.000 (0.001)	2.919 (5.955)	0.007* (0.003)	-1.057*** (0.204)	0.000 (0.000)	-0.571 (0.653)	0.003*** (0.003)	
61%–80%	-0.001* (0.000)	-3.626 (4.843)	-0.003 (0.002)	-0.905*** (0.181)	0.000 (0.000)	-0.189 (0.663)	0.003 (0.002)	

(continued)

Table 2. (continued)

Dependent Variables: Gap in Geolocation Measurement Due to . . .	(1) Device Turned Off		(2) No Measurement at All		(3) No Geolocation Measurement		(4) Failed Geolocation	
	Occurrence Model 1 b/SE	Extent Model 2 b/SE	Occurrence Model 3 b/SE	Extent Model 4 b/SE	Occurrence Model 5 b/SE	Extent Model 6 b/SE	Occurrence Model 7 b/SE	Geolocation Occurrence b/SE
Spatial and temporal context								
Lag mobile network roaming	-0.009 (0.008)	-1.800 (5.723)	-0.023 (0.033)	-1.072 (0.771)	0.004 (0.004)	-1.160 (1.028)	0.001 (0.021)	
Lag mobile network outside Germany	0.015 (0.009)	-0.794 (5.765)	0.071* (0.033)	2.446*** (0.713)	-0.003 (0.004)	1.440 (0.940)	-0.031 (0.028)	
Lag location outside Germany	-0.043 (0.039)	5.775 (16.192)	-0.045*** (0.010)	0.386 (0.700)	-0.236 (0.198)	10.239* (5.018)	0.027 (0.019)	
Lag location provider ref. network	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	
Fused	0.000 (0.017)	-33.071* (16.325)	-0.019*** (0.005)	-0.092 (0.337)	-0.507*** (0.144)		0.021** (0.008)	
GPS	0.018 (0.012)	9.574 (19.571)	-0.009* (0.003)	-0.387 (0.302)	0.362 (0.202)	-2.501 (1.877)	0.000 (0.003)	
Cell tower database	0.015 (0.010)	5.333 (18.505)	-0.002 (0.006)	0.502 (0.486)			-0.018* (0.007)	
Lag time of day Ref. Weekday (Mon-Fri)	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	
Weekday (Mon-Fri)	0.002** (0.001)	-1.4391 (8.734)	0.045*** (0.005)	0.181 (0.374)	0.001*** (0.000)	6.110 (3.668)	-0.002 (0.003)	
00:00-05:59	-0.003*** (0.001)	2.703 (6.004)	-0.015*** (0.002)	-0.496 (0.337)	0.000 (0.000)	0.195 (0.622)	-0.002 (0.002)	
Weekday (Mon-Fri)	0.002* (0.001)	-7.303 (5.204)	0.000 (0.003)	0.707 (0.361)	0.001* (0.000)	2.798** (0.922)	-0.001 (0.002)	
18:00-23:59	0.000 (0.001)	-6.501 (10.258)	0.001 (0.003)	-0.440 (0.425)	0.000 (0.000)	-3.654 (3.907)	0.000 (0.002)	
Weekend (Sa, Su)	0.000 (0.001)	5.004 (7.562)	0.002 (0.003)	0.296 (0.402)	0.000 (0.000)	1.769 (1.021)	-0.002 (0.002)	
Weekend (Sa, Su)	0.000 (0.001)	-8.762 (11.024)	0.001 (0.002)	-0.382 (0.453)	0.000 (0.000)	-1.735 (1.234)	0.001 (0.002)	
12:00-17:59	0.000 (0.001)	2.793 (9.442)	-0.005 (0.003)	-0.524 (0.428)	0.000 (0.000)	-2.289 (1.273)	0.002 (0.002)	
Weekend (Sa, Su)	0.001 (0.007)	45.486 (23.345)	-0.060* (0.030)	12.648*** (1.695)	-0.001 (0.002)	-0.430 (3.450)	0.186*** (0.038)	
Constant	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Lag data permissions	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Lag data source								
Measurements	1,041,598	10,444	1,273,172	242,018	1,037,827	6,673	2,779,003	
Participants	498	469	499	498	496	437	499	
R ²	.231	.101	.513	.011	.421	.007	.116	

Note. Average marginal effects (with 95% CI) based on linear regression with cluster robust standard errors.

* $p < .05$. ** $p < .01$. *** $p < .001$.

Table 3. Amount of Missingness of the Geolocation Measurements Per Participant.

	Mean	SD	Median	Share
Number of days with geolocation data	164	65	196	
Measurements of geolocation				
Target	9,359	2,790	10,219	100.00%
Actual	4,289	2,538	4,397	45.82%
Missing	5,070	2,450	5,017	54.18%
Missing (at least 60-min gap)				
All causes	4,006	2,541	3,827	100.00%
Device turned off	908	1,617	193	22.66%
No measurement at all	2,951	1,972	2,767	73.66%
No geolocation measurement	152	385	49	3.80%
Participants	578			

Results

Table 3 shows an overview of the amount of collected geoinformation over the data collection period. On average, the app collected at least one geolocation data point for 164 days for 578 participants who provided at least one geolocation measurement. As the median number of days is 196, which is greater than the mean, our data are left-skewed and have cases with only a few days of data collection. As our data collection interval for geolocation data was 30 min, the app should have collected 48 measures per day, leading to an average target number of 9,359 geolocations. The actual number of measured geolocations, however, is only 4,289 (45.8%), that is, 54.2% of the observations are missing.

Beside small variations in the actual sampling frequency, the following reasons should explain, why we have fewer observations than expected: turned off devices, gaps with no measurements at all, and gaps in geolocation measurement. When we concentrate on the geolocation gaps that last for at least 60 min, turned off devices are responsible for 22.7%, gaps in all data collection are responsible for 73.7%, and gaps in geolocation data are responsible for 3.8% of missing location measures. Table 3 also highlights our design decision to focus on gaps between geolocation measurements of at least 60 min. While these gaps constitute only 12% of all gaps in terms of time, they include 79% of all missing geolocation measurements.

Table 2 displays the results from our multivariate analyses reported as average marginal effects. For Outcomes (1)–(3), we display results for the *occurrence*, defined as the identification of geolocation measurement gaps, and the *extent*, defined as the length of each identified gap in terms of the number of missing geolocation measurements. For Outcome (4), we display the occurrence of failed geolocation.

Explaining Missingness Due to Turned Off Devices (1)

At the median, 193 missing geodata measurements per participant can be explained by turned off devices. Results for gap occurrence due to gaps in all measurements can be found in Table 2, Model 1. Model 2 explains the number of missing measurements due to turned off devices.

Gap occurrence. Model 1 shows no differences between sociodemographic groups in the likelihood of the device being turned off. Problematic participant behavior is negatively correlated ($-.007$) with the likelihood of gaps by turned off devices. We find some isolated effects of manufacturers. Compared to Samsung devices, smartphones from Huawei ($-.006$), Motorola ($-.020$), and Sony

(-.006) have a significantly lower likelihood of missing geolocation measurements due to device turnoffs. Compared to Android 7, the oldest version (Android 4) and the most frequent version (Android 8) in our sample are less prone to missing geolocations because of turned off devices. Other device indicators such as the number of CPUs or the device's RAM have no effect on gaps due to devices being turned off. Regarding the device's state, we find that temporarily revoking the geodata collection in the app settings menu (.998) and temporarily uninstalling the app (.506) are very likely to precede a turned off device period. The usage of the so-called task killer apps and enabling the power save mode do not influence the device being turned off. If the devices display state is not "on" but "doze" or "off," we see negative probabilities for turned off device episodes (-.009 and -.006). The battery states "charging" and "discharging" have a larger probability to precede a turned off device episode (.004 and .003) than the battery status "full." In addition, low battery levels, that is, 0%–20% (.029) and 20%–40% (.003) make immediate device turnoffs more likely while a battery state between 61% and 80% (-.001) makes it less likely for devices being turned off. Concerning spatial and temporal context factors, we find a higher probability that the device is turned off during the night hours on weekdays (.002) and lower probabilities for afternoons on weekdays (-.003). The location of the smartphone outside of Germany and the location provider have no effect on devices being turned off.

Gap Extent. For sociodemographic variables, we find no significant effects on the extent of gaps due to turned off devices. The number of missing geolocation measurements in a gap decreases with an increasing household size (-5.3) indicating the need to communicate with other household members. Here also problematic behavior by participants does not lead to more but to less missing geolocation data (-34.9). Compared to Samsung smartphones, we find three device manufacturers that have, on average, more missing geodata per gap: HTC (93.9), Huawei (104.0), and other manufacturers (70.6). We find no differences in gap extents by Android version, the number of CPU cores, or the device's RAM. Regarding the device state, we find longer gaps when the app was temporally uninstalled before the gap. We find no effects in the extent of gaps when considering that the consent to geodata collection was temporarily revoked and when the participant used a task killer app or enabled the power save mode. If the device's display was switched off before the device powered down, gaps were 16.2 geolocations shorter. Furthermore, if the device was actively charging, this decreased the extent of geodata measurement gaps by 57.6. Furthermore, we find that the extent of gaps decreases by 21.6 missing geolocation measurements when the device has a battery level of 0%–20% and by 7.7 when the battery level is between 21% and 40% compared to fully charged. Considering the spatial and temporal context, we see no effects on the extent of gaps due to turned off devices.

Explaining Missingness Due to No Measurement at All (2)

At the median, we find 2,767 missing geolocation measurements per participant due to episodes where the app was unable to collect any data at all. Results for gap occurrence due to gaps in all measurements can be found in Model 3. Results for the extent of gaps in all measurements (given a gap) can be found in Model 4.

Gap occurrence. We find that sociodemographic variables have no effect on the occurrence of gaps in all data collection (see Model 3). For device hardware factors, we find effects related to the device manufacturers and the OS: Huawei (-.025), LG (-.035), and Motorola (-.122) smartphones have fewer gaps than Samsung smartphones, and the number of gaps in all measurements decreases with older Android versions. Concerning the number of CPU cores and RAM, we find no effects. Shortly

after participants revoke their consent for geodata collection, we see an increased gap likelihood in missingness (.46) while we see a smaller likelihood for gaps if participants use a task killer app.

While one would expect that uninstalling the app would manifest itself in gaps in all measurements, in our data structure, this is covered by the gaps due to (1) *turned off devices*. We find gaps to be less likely if the device activated the “doze” mode (−.02) but more likely if the device activated the “doze suspend” mode (.086). This is in line with the finding that, the likelihood of gaps increased when the display was turned off (.049), indicating that most data collection happens when the device is somewhat inactive. Furthermore, if the battery has the status “full,” devices are less likely to have gaps in all measurements. A low battery level increases the occurrence of gaps compared to a fuller battery level which indicates that the Android OS may kill apps running in the background to save energy. For the spatial and temporal context factors, we are more likely to find gaps in all measurements (.015) if the device is logged into a non-German carrier network, while smartphones located outside of Germany are less likely to have gaps (−.045). This indicates first, connectivity issues in border regions, and second, more stable environments, once the device is abroad. We also find effects of the time of day with gaps being more likely during the night hours on weekdays (.045) and less likely on weekday afternoons (−.015).

Gap extent. For sociodemographic variables, we find no effects for the extent of gaps in all measurements. Compared to Samsung smartphones, we find three device manufacturers with, on average, two fewer missing geodata measurements per gap: LG (−2.4), Motorola (−2.4), and Sony (−1.8). Huawei smartphones produce on average 1.9 more missing geodata measurements per gap. Android 5 produces six additional missing measurements per gap, while Android 4 produces 5.8 fewer missing measurements per gap compared to Android 7. Regarding device states, we find no effects if the participants revoked the geodata collection, used a task killer app, or enabled the power saving mode. However, the extent of all measurement gaps is strongly related to the device states display status, battery status, and battery level. The device display state “doze” (−3.1), “doze suspend” (−3.3), and “off” (−4.5) as well as having a rather high battery level, that is, 41%–60% (−1.1) and 61%–80% (−0.9), decrease the extent of gaps. Charging (3.3) and discharging (1.3) the device increases the number of missing geolocations per gap compared to the status “full.” For our spatial and temporal factors, we find devices that use mobile networks outside of Germany have 2.4 more missing geolocation measurements per gap. For the time of day, we find no effect.

Explaining Missingness Due to No Geolocation Measurement (3)

At this stage, the app was able to collect other data but no geolocation data (see Figure 3). At the median, we find that 49 measurements per participant are lost due to this phenomenon. Results of the gap occurrence due to gaps in geolocation measurements can be found in Model 5. Results for the extent of gaps in geolocation measurements can be found in Model 6.

Gap occurrence. Model 5 shows that sociodemographic variables have no effect on the occurrence of geolocation measurement gaps. Users with a Huawei device (.002) are more likely to have geolocation measurement gaps than Samsung devices, while devices from other manufacturers (−.002) are less likely. Android 4 (−.005) devices have a smaller probability of geolocation data gaps, while smartphones with Android 8 (.001) have a higher probability for gaps compared to Android 7 devices. Gaps in geolocation data can be explained well by device state factors. We find gaps to be less likely when task killer apps were used (−.055) or the power saving mode was activated (−.005). Furthermore, we find that each display state decreases the probability for a gap compared to the display state “on.” Location data gaps are less likely to occur when the device is charging or discharging (−.001). Low battery levels of 21%–40% (.064) and 0%–20% (.001) increase the probability of a gap in geodata

measurement. For spatial and temporal context factors, we find that location data gaps are more likely during the night on weekdays.

Gap extent. Sociodemographic variables have no effect on the extent of geolocation measurement gaps (see Model 6). Motorola smartphones have 2.8 fewer missing geolocations per gap than Samsung smartphones. Android Versions 5 and 6 produce 2.8 and 3.9 more missing locations, respectively than Android 7. For device state factors, we find no effect on the extent of geolocation measurement gaps. Regarding spatial and temporal context variables, we find that smartphones located outside of Germany to have 10.2 more missing geolocations per gap than smartphones that are in Germany. Furthermore, we find 2.8 more missing geolocations per gap on weekday evenings.

Failed Geolocation (4)

In our study, the geolocating failed in about 17.5% of all nonmissing geolocation measurements. Results for the occurrence of failed geolocation can be found in Model 7. For sociodemographic variables, we find that participants 50+ years old are less likely to have failed geolocations ($-.027$) than the 30- to 49-year-olds. Regarding device hardware characteristics, we find that Sony devices are less likely to produce failed geolocation ($-.055$) than Samsung devices. Looking at the impact of device states, we find failed geolocations to be less likely if the participant temporally uninstalled the app ($-.058$). Task killer apps (.602) and battery levels of 21%–40% (.013) and 41%–60% (.008) seem to (strongly) increase the probability for failed geolocation. These last two results are in line with the technical nature of this phenomenon (Google, 2019b). If the last successful geolocation measurement before the gap was provided by Wi-Fi information, this increases the probability of failed geolocation (.021), which indicates that it is indoor situations, where Wi-Fi and mobile carrier network reception are weak, that produce these gaps (Kjærgaard et al., 2010). This also hints at the design decision not to implement Wi-Fi in the group of providers used for active geolocation, as a factor driving the extent of missingness at this stage. As indoor stays have a higher probability of missingness in our study, substantial analysis will have to deal with underrepresented indoor events. The occurrence of the gap is invariant over the time of day and week, so it affects indoor locations both at home and at work in the same way. In summary, Model 7 differs considerably from Model 5, underlining the distinctiveness of both missing phenomena.

Invalid Coordinates (5)

In the last stage, we subsume participant behavior that makes their geolocations in the context of our study invalid. At the end of the study period, we asked, how often participants shared their device with others (see Table 4). Three participants shared their devices for more than 10 cumulative days over the study period, which is the threshold we consider problematic, explicitly to enable others to be reachable on the go. Such sharing invalidates the geolocation data of the originally sampled study member; therefore, we need to exclude them from substantive analyses (see Table A2 for further details). All together 10 respondents provided measurements stemming from different devices at the same time; 4 had parallel measurements of only a few seconds, while the remaining 6 had on average 31 days of parallel measurement (the maximum was 118 days). We need to exclude these six from substantive analysis because it is unclear who was using the device and which locations belonged to the sampled participant. When further analyzing the app use data, we identified two participants, who used apps to falsify their geolocations. Close inspection, confirmed, for one of them, that geolocations were altered while the participant was using the app.² Thus, we need to exclude the affected observations for both individuals. All these problematic behaviors invalidate a total of 2.6% of all collected geolocation

Table 4. Failed Geolocation (Stage 4) and Invalid Coordinates (Stage 5).

	Geolocation Measurements	Share of Successful Geolocation Measurements (%)		Affected Participants
Location permission granted	2,478,181	100.00		578
4) Failed geolocation	433,872	17.51		434
Successful geolocation	2,044,309	82.49	100.00	576
5) Total invalid coordinates	53,483		2.62	11
Device sharing for location purposes (>10 cumulative days)	21,946		1.07	3
Multiple concurrent devices	37,405		1.83	6
Use of GPS fake apps	10,984		0.54	2
Remaining data for substantive analysis	1,990,826		97.38	565

measurements. Technically, we would have to count the 61 participants, whose gender or age provided in the initial in-app surveys differed from the sampled PASS respondents, here also. This is the case, when sampled respondents pass on their invitation to others, for example, household members. Because we ultimately cannot link their data to the PASS records, we excluded their data right from the start of our analysis.

Discussion and Conclusions

This article investigated the measurement quality of geolocations measured by smartphones. We showed that sensor data are far from objective and error-free. Instead, device-related error sources such as specific manufacturer and OS settings, design decisions of the research app, third-party apps, and the participants' behavior can greatly interfere with the measurement. To disentangle the different influences, we applied a multistage error model to first analyze and then control the general error source in the specific missingness processes of geolocation data.

The good news is that there are almost no effects of the (mostly sociodemographic) respondent attributes used in the models. The number of significant results does not exceed those we would expect by chance. However, user behavior like temporarily revoking the apps' access to geolocation data or temporarily uninstalling the app unsurprisingly has the largest effects on the occurrence and extent of gaps. Further research into this user behavior is needed to evaluate whether this is a threat to the unbiasedness of the resulting data.

Regarding device hardware, we find that Motorola and Sony smartphones produced fewer gaps and gaps of a smaller extent in our study than Samsung while Huawei smartphones produced fewer but larger gaps. Concerning the OS, with more recent Android versions, the number of instances where no measurement at all could be taken increases noticeably. This reflects the increasing restrictions imposed by Android on third-party apps with each new iteration of the OS. Extrapolating this trend, future research apps will have to manage an increasingly difficult environment for data collection.

The presence of task killer apps on the study smartphone has effects in different directions at different stages. While task killer apps even reduce the number of instances where no measurement at all or no geolocation measurement could be taken, they increase the likelihood of invalid geolocation measurements substantially.

Results for the device state are quite surprising. Doze modes and switched off displays before a scheduled measurement reduce the occurrence and extent of gaps. This challenges the intuition that these modes inhibit data collection.

Battery levels and charging status have large and consistent effects in the expected direction for gap occurrence. However, gaps due to turned off devices at low battery levels tend to be smaller than such gaps occurring at higher battery levels. This could be because faced with low battery status users turn off the smartphone to charge it (phones tend to charge faster when turned off), whereas if the smartphone is turned off at high battery status is might be for more substantial reasons, that is, for good night sleep. We have no data to examine this in the current study, but it could be an interesting next step for the exploration. Further analyses are needed to see what would predict low battery levels to see if there is an inherent threat to substantive conclusions derived from these data. To the extent that the charging behavior of respondents is unrelated to the variables of interest, this should not be a problem. We choose our geolocation sampling rate of 30 min with battery consumption in mind. Other studies might choose to measure geolocation at a higher frequency, in which case battery draining can become more of an issue. If the focus is on short periods (e.g., walking around the neighborhood), this can work fine. If like in our case, measurements are taken over 6 months, a different trade-off needs to be made to avoid participants turning off geolocation tracking, closing the app, or uninstalling the app altogether. By sampling geolocations actively only if they were not already available in the Android system, we emphasized battery efficiency over accuracy. Because GPS and mobile carrier networks were the only providers used for active geolocation, we missed locations that would have been available using Wi-Fi-information (Stage 4). Further research is needed to explore a suitable balance of data collection frequency and device performance.

Surprisingly, there are no consistent effects of locations and networks out of Germany and roaming. Using mobile networks outside of Germany increases the occurrence and extent gaps where no measurement was taken at all, while locations outside Germany reduce the occurrence of such gaps. This could point at two separate issues, living close to the border (or having a foreign mobile plan) and being temporarily abroad (e.g., on vacation). Another technical variable that can be shown to have a large effect is the location provider. If the previous geolocation measurement was fused, this leads to fewer and smaller gaps (although there is a slightly increased chance that geolocation failed).

Gaps in geolocation measurements are at all Stages (1–3) more likely to occur at night (0:00 to 5.59 a.m.) than at any other time of the day. It seems plausible that no activities are missed (when gaps coincide with sleep).

While very rare in our data, we nevertheless want to raise awareness of so far little known error sources in geolocation measurement, like the use of GPS falsifier apps or device sharing among participants. It seems advisable to collect this kind of information either through a survey among participants or where feasible directly by passive data collection from the smartphone (e.g., collect information on all apps installed). For shared devices, using machine learning algorithms might be another strategy to identify who used the device at which time.

Despite being a first comprehensive step to understand missingness in geolocation data collected by smartphones, our study left several questions unanswered and raised some new ones. A deeper understanding of the effects of doze modes related to smartphone use behavior will require more collaboration between computer scientists and social scientists. Such collaboration is generally encouraged as we move forward and technology changes (quickly). Here we already saw the effects of more recent OS on the occurrence of missing data. We highly recommend future studies to work with technology experts to ensure sparse battery consumption. Further research is needed to explore a suitable balance of data collection frequency and device performance.

Our study is longitudinal, and further analyses need to be done to understand the mechanisms leading to gaps in measurement through attrition. We cannot rule out that attrition is unrelated to our measurement of interest. It will be interesting to investigate how attrition relates to unit nonresponse, and if the potential biasing effect of unit nonresponse gets amplified or reduced.

GPS data will continue to be of importance for research in the social sciences. They can help us understand mobility patterns, commuting behavior, or the use of different means of transportation. By

combining geolocation data with open data about the locations visited, we can reconstruct day structures and infer about social participation. Connecting short surveys to geolocations provides opportunities for situation-specific measurements. Such short surveys do not need to be limited to customer satisfaction surveys (where we observe most applications until now) but could be used for research on interventions (like training programs or consulting meetings) where questions can be asked right after the location of the meeting is left (thus minimizing recall error) or for research on the effect of activities on happiness.

Appendix

Table A1. Sample Selection and Compositions.

	PASS Wave 11 (2017) Information												Geolocation Measurements	
	Persons	% Female		Age in Years		Years of Education		% Foreign		% Employed		% Unemployed		
		Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean		SD
(1) PASS participants in Wave 11 (2017)	13,703	.50	.50	45.84	16.80	12.22	2.93	.32	.47	.51	.50	.23	.42	
(2) Smartphone owners	10,677	.51	.50	41.77	15.39	12.39	2.99	.35	.48	.54	.50	.22	.41	
(3) Android smartphone owners	7,275	.49	.50	4.46	14.76	12.31	2.99	.35	.48	.53	.50	.23	.42	
(4) Invitees to the IAB-SMART study	4,293	.52	.50	4.16	12.75	12.35	2.83	.28	.45	.59	.49	.18	.39	
(5) IAB-SMART installer	686	.49	.50	38.45	12.16	12.83	2.84	.20	.40	.62	.48	.17	.37	
(6) Verified IAB-SMART installer	625	.49	.50	38.37	12.11	12.91	2.85	.19	.39	.63	.48	.16	.37	
(7) Location permission granted	578	.49	.50	38.56	12.11	12.92	2.82	.18	.39	.64	.48	.16	.36	2,478,180
(8) At least one successful geodata collected	576	.49	.50	38.53	12.12	12.92	2.81	.18	.39	.64	.48	.16	.36	2,477,440
(9) Analysis sample	496	.49	.50	38.37	12.09	12.89	2.77	.19	.39	.65	.48	.15	.35	1,518,260
t-test: p value (1) vs. (9)		.597		.000		.000		.000		.000		.000		

Table A2. End of Study Survey on Device Sharing.

Third-party smartphone use	Mean	SD
Third-party use % share yes	0.16	0.36
Cumulative duration of third-party use in days	1.68	11.27
Third-party use longer than 10 days % share yes	0.03	0.17

Reasons for third-party use	% Responses	% Participants
Use an app	21.24	82.09
Browse the internet	20.08	77.61
Listen to music or watch a video	16.22	62.69
Make a call	14.29	55.22
Use a messenger app (e.g., WhatsApp)	11.97	46.27
Be reachable on the go	6.56	25.37
Use text message (SMS)	6.18	23.88
Answer questionnaires in the IAB-SMART App	1.93	7.46
Get exercise	1.16	4.48
Something else	0.39	1.49

Table A3. Descriptive Statistics for Participant-Level Variables of the Multivariate Analysis Sample.

		Participants	Mean	SD	Median	Min	Max
Sociodemographics	Age-group						
	18–29	499	0.27	0.45	0	0	1
	30–49	499	0.51	0.50	1	0	1
	50 and older	499	0.22	0.42	0	0	1
	Gender–Female	499	0.48	0.50	0	0	1
	Size of household	499	2.60	1.27	2	1	7
	Employed in 2017	499	0.65	0.48	1	0	1
	Unemployed in 2017	499	0.15	0.35	0	0	1
Problematic participant behavior	499	0.02	0.15	0	0	1	

Table A4. Descriptive Statistics for Measurement-Level Variables of the Multivariate Analysis Sample.

		Measurements	Mean	SD	Median	Min	Max
Dependent variables	Dependent variables						
	(1) Device turned off–Occ.	1,041,598	0.01	0.10	0	0	1
	(1) Device turned off–Extent	10,444	39.49	174.10	13	1	6,660
	(2) No meas. at all–Occ.	1,273,172	0.19	0.39	0	0	1
	(2) No meas. at all–Extent	242,018	6.03	32.29	2	1	8,506
	(3) No geolocation meas.–Occ.	1,037,827	0.01	0.08	0	0	1
Data permissions	(3) No geolocation meas.–Extent	6,673	5.83	30.03	1	1	2,089
	(4) Geolocation failed–Occ.	1,290,289	0.14	0.35	0	0	1
	Data permissions						
	Location	1,290,289	0.94	0.24	1	0	1
	Activity tracking	1,290,289	0.92	0.27	1	0	1
	App usage	1,290,289	0.85	0.35	1	0	1
	Contacts	1,290,289	0.94	0.24	1	0	1
	Digital test	1,290,289	0.94	0.24	1	0	1
	Interactions	1,290,289	0.92	0.28	1	0	1
	Phone	1,290,289	0.94	0.24	1	0	1
	Smartphone usage	1,290,289	0.85	0.36	1	0	1
Device hardware	SMS	1,290,289	0.94	0.24	1	0	1
	Social contacts	1,290,289	0.91	0.29	1	0	1
	Device manufacturer						
	HTC	1,290,289	0.03	0.16	0	0	1
	Huawei	1,290,289	0.14	0.35	0	0	1
	LG	1,290,289	0.04	0.21	0	0	1
	Motorola	1,290,289	0.02	0.15	0	0	1
	Other	1,290,289	0.06	0.24	0	0	1
	Samsung	1,290,289	0.59	0.49	1	0	1
	Sony	1,290,289	0.12	0.32	0	0	1
	Android OS version						
	4	1,290,289	0.01	0.09	0	0	1
	5	1,290,289	0.06	0.24	0	0	1
	6	1,290,289	0.20	0.40	0	0	1
7	1,290,289	0.48	0.50	0	0	1	
8	1,290,289	0.25	0.43	0	0	1	
Device number of CPU cores	1,290,289	5.87	1.97	6	1	10	
Device RAM	1,290,289	3.07	1.03	3	1	8	

(continued)

Table A4. (continued)

		Measurements	Mean	SD	Median	Min	Max
Device state	Geodata collection temp. revoked ^a	1,290,289	0.00	0.00	0	0	1
	App temp. uninstalled ^a	1,290,289	0.00	0.00	0	0	1
	Task killer/manager app active ^a	1,290,289	0.00	0.00	0	0	1
	Power saving mode	1,279,514	0.06	0.24	0	0	1
	Display state						
	Doze	1,240,417	0.02	0.14	0	0	1
	Doze suspend	1,240,417	0.06	0.23	0	0	1
	Display off	1,240,417	0.69	0.46	1	0	1
	Display on	1,240,417	0.23	0.42	0	0	1
	Unknown	1,240,417	0.01	0.09	0	0	1
	Battery status						
	Charging	1,266,790	0.13	0.33	0	0	1
	Discharging	1,266,790	0.63	0.48	1	0	1
	Full	1,266,790	0.19	0.39	0	0	1
	Not charging	1,266,790	0.06	0.23	0	0	1
	Battery level, categories						
	0%–20%	1,246,090	0.03	0.17	0	0	1
	21%–40%	1,246,090	0.11	0.31	0	0	1
	41%–60%	1,246,090	0.18	0.38	0	0	1
	61%–80%	1,246,090	0.22	0.41	0	0	1
81%–100%	1,246,090	0.46	0.50	0	0	1	
Battery level	1,037,827	73.79	24.71	79	0	100	
Spatial and temporal context	Mobile network outside Germany	1,037,827	0.02	0.15	0	0	1
	Location outside Germany	1,112,005	2.11	3.80	0	0	9
	Location provider						
	Fused	1,290,289	0.02	0.16	0	0	1
	GPS	1,290,289	0.05	0.23	0	0	1
	Network	1,290,289	0.48	0.50	0	0	1
	Cell tower database	1,290,289	0.14	0.35	0	0	1
	Not available	1,290,289	0.11	0.31	0	0	1
	Time weekend (Sat., Sun.)	1,290,289	0.28	0.45	0	0	1
	Time part of the day						
	00:00–05:59	1,290,289	0.23	0.42	0	0	1
	06:00–11:59	1,290,289	0.25	0.43	0	0	1
12:00–17:59	1,290,289	0.25	0.44	0	0	1	
18:00–23:59	1,290,289	0.26	0.44	0	0	1	

^aThese behaviors each make up less than .002% of all geolocation measurements.

Data Availability

Data used in this article are available at the Institute for Employment Research (IAB). Up-to-date access information can be found here: <https://www.iab.de/en/daten.aspx>.

Software Information

For all analyses, author-originated code under Stata Version 15.1 was used. The code is available at the Institute for Employment Research (IAB) and will be delivered together with data access (see Section Data Availability).

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Georg-Christoph Haas' time was supported by the German Research Foundation (DFG) through the Collaborative Research Center SFB 884 "Political Economy of Reforms" (Project A9) [139943784 to Markus Frölich, Florian Keusch, and Frauke Kreuter].

Notes

1. This model could alternatively be stated as a zero-inflated count model. Given our numbers of observations and explanatory variables, such models are computationally very demanding, and maximum-likelihood estimation is often unstable. The chosen linear regression model produces less efficient but otherwise consistent estimates of our parameters without commitment to functional form and distributional assumptions (Angrist & Pischke, 2009, p. 197). We also estimated the occurrence variables using a logit specification and achieved comparable marginal effects to the coefficients from the linear regression model.
2. In the case of the app "Fake GPS with joystick," latitude and longitude values were simply switched.

References

- Angrist, J. D., & Pischke, J.-S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.
- Antoun, C., Conrad, F. G., Couper, M. P., & West, B. T. (2018). Simultaneous estimation of multiple sources of error in a smartphone-based survey. *Journal of Survey Statistics and Methodology*, 7(1), 93–117.
- Blunck, H., Kjærgaard, M. B., & Toftegaard, T. S. (2011). Sensing and classifying impairments of GPS reception on mobile devices. In K. Lyons, J. Hightower, & E. M. Huang (Eds.), *Proceedings of the 9th International Conference on pervasive computing* (pp. 350–368). Springer.
- Breen, R., Karlson, K. B., & Holm, A. (2018). Interpreting and understanding logits, probits, and other non-linear probability models. *Annual Review of Sociology*, 44(1), 39–54.
- Casilari, E., Santoyo-Ramón, J. A., & Cano-García, J. M. (2016). Analysis of a smartphone-based architecture with multiple mobility sensors for fall detection. *PLoS One*, 11(12), e0168069.
- Chen, Z., Lin, M., Chen, F., Lane, N. D., Cardone, G., Wang, R., Li, T., Chen, Y., Choudhury, T., & Campbell, A. T. (2013). Unobtrusive sleep monitoring using smartphones. *Proceedings of Seventh International Conference on Pervasive Computing Technologies for Healthcare and Workshops* (pp. 145–152). Institute for Computer Sciences, Social Informatics and Telecommunications Engineering.
- Couper, M. P. (2019). Mobile data collection: A survey researcher's perspective. *Mobile Apps and Sensors in Surveys (MASS) Workshop*. Mannheim University & Utrecht University, March 4.
- Elevelt, A., Lugtig, P., & Toepoel, V. (2019). Doing a time use survey on smartphones only: What factors predict nonresponse at different stages of the survey process? *Survey Research Methods*, 13(2), 195–213.
- Ferreira, D., Dey, A. K., & Kostakos, V. (2011). Understanding human-smartphone concerns—A study of battery life. In K. Lyons, J. Hightower, & E. M. Huang (Eds.), *Proceedings of the Ninth International Conference on Pervasive Computing* (pp. 19–33). Springer.
- Google. (2019a). *Android Oreo* (Version 8.0). <https://www.android.com/versions/oreo-8-0/>
- Google. (2019b). *Google's APIs for Android—Fused Location Provider Client: getLastLocation*. <https://developers.google.com/android/reference/com/google/android/gms/location/FusedLocationProviderClient.html#getLastLocation>
- Harari, G. M., Müller, S. R., Aung, M. S. H., & Rentfrow, P. J. (2017). Smartphone sensing methods for studying behavior in everyday life. *Current Opinion in Behavioral Sciences*, 18, 83–90.
- Hargittai, E. (2018). Potential biases in big data: Omitted voices on social media. *Social Science Computer Review*, 38(1), 10–24.
- Hosio, S., Ferreira, D., Goncalves, J., van Berkel, N., Luo, C., Ahmed, M., Flores, H., & Kostakos, V. (2016). Monetary assessment of battery life on smartphones. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 1869–1880). Association for Computing Machinery.

- Jäckle, A., Burton, J., Couper, M. P., & Lessof, C. (2019). Participation in a mobile app survey to collect expenditure data as part of a large-scale probability household panel: Coverage and participation rates and biases. *Survey Research Methods*, 13(1), 23–44.
- Jäckle, A., Gaia, A., Lessof, C., & Couper, M. P. (2019). *A review of new technologies and data sources for measuring household finances: Implications for total survey error* (Understanding Society Working Paper Series No. 2019-02). <https://www.understandingsociety.ac.uk/sites/default/files/downloads/working-papers/2019-02.pdf>
- Kelly, D., Smyth, B., & Caulfield, B. (2013). Uncovering measurements of social and demographic behavior from smartphone location data. *IEEE Transactions on Human-Machine Systems*, 43(2), 188–198.
- Keusch, F., Bähr, S., Haas, G.-C., Kreuter, F., & Trappmann, M. (2020a). Coverage error in data collection combining mobile surveys with passive measurement using apps: Data from a German national survey. *Sociological Methods & Research*. Advance Online Publication. <https://doi.org/10.1177/0049124120914924>
- Keusch, F., Bähr, S., Haas, G.-C., Kreuter, F., & Trappmann, M. (2020b). Participation rates and bias in a smartphone study collecting self-reports and passive mobile measurements using a research app. *AAPOR 75th Annual Conference, Virtual Conference*, June 11–12.
- Kjærgaard, M. B., Blunck, H., Godsk, T., Toftkjær, T., Christensen, D. L., & Grønbæk, K. (2010). Indoor positioning using GPS revisited. In P. Floréen, A. Krüger, & M. Spasojevic (Eds.), *In Pervasive computing* (pp. 38–56). Springer.
- Kongsiriwattana, W., & Gardner-Stephen, P. (2016). Smart-phone battery-life short-fall in disaster response: Quantifying the gap. *2016 IEEE Global Humanitarian Technology Conference* (pp. 220–225). Institute of Electrical and Electronics Engineers.
- Kreuter, F., Haas, G.-C., Keusch, F., Bähr, S., & Trappmann, M. (2018). Collecting survey and smartphone sensor data with an app: Opportunities and challenges around privacy and informed consent. *Social Science Computer Review*. Advance online publication.
- Lee, U., Yang, S., Ko, M., & Lee, J. (2014). Supporting temporary non-use of smartphones. *Proceedings of the 2014 CHI Workshop: Refusing, Limiting, Departing: Why We Should Study Technology Non-Use Workshop* (pp. 65–68). Association for Computing Machinery.
- MacCormick, J. S., Dery, K., & Kolb, D. G. (2012). Engaged or just connected? Smartphones and employee engagement. *Organizational Dynamics*, 41(3), 194–201.
- Metri, G., Agrawal, A., Peri, R., & Shi, W. (2012). What is eating up battery life on my SmartPhone: A case study. *Proceedings of the 2012 International Conference on Energy Aware Computing*. Institute of Electrical and Electronics Engineers.
- Miller, P. V. (2008). Measurement error. In P. J. Lavrakas (Ed.), *Encyclopedia of survey research methods* (pp. 457–461). Sage.
- Nitsche, P., Widhalm, P., Breuss, S., & Maurer, P. (2012). A strategy on how to utilize smartphones for automatically reconstructing trips in travel surveys. *Procedia—Social and Behavioral Sciences*, 48, 1033–1046.
- Oliver, E. (2010). The challenges in large-scale smartphone user studies. *Proceedings of the Second ACM International Workshop on Hot Topics in Planet-scale Measurement* (pp. 1–5). ACM.
- Onnela, J.-P., & Rauch, S. L. (2016). Harnessing smartphone-based digital phenotyping to enhance behavioral and mental health. *Neuropsychopharmacology*, 41, 1691–1696.
- Peltonen, E., Lagerspetz, E., Nurmi, P., & Tarkoma, S. (2016). Where has my battery gone? A novel crowd-sourced solution for characterizing energy consumption. *IEEE Pervasive Computing*, 15(1), 6–9.
- Pew Research Center. (2018). *Mobile fact sheet*. <https://www.pewinternet.org/fact-sheet/mobile/>
- Raento, M., Oulasvirta, A., & Eagle, N. (2009). Smartphones an emerging tool for social scientists. *Sociological Methods & Research*, 37(3), 426–454.
- Ranasinghe, C., & Kray, C. (2018). Location information quality: A review. *Sensors*, 18(11), E3999.
- Rubin, D. B. (1976). Inference and missing data. *Biometrika*, 63(3), 581–592.

- Sugie, N. F. (2018). Utilizing smartphones to study disadvantaged and hard-to-reach groups. *Sociological Methods & Research*, 47(3), 458–491.
- Trappmann, M., Bähr, S., Beste, J., Eberl, A., Frodermann, C., Gundert, S., Schwarz, S., Teichler, N., Unger, S., & Wenzig, C. (2019). Data resource profile: Panel study labour market and social security (PASS). *International Journal of Epidemiology*, 48(5), 1411–1411g.
- Ugur, N. G., & Koc, T. (2015). Time for digital detox: Misuse of mobile technology and phubbing. *Procedia—Social and Behavioral Sciences*, 195, 1022–1031.
- Zandbergen, P. A. (2009). Geocoding quality and implications for spatial analysis. *Geography Compass*, 3(2), 647–680.
- Zhang, J., Musa, A., & Le, W. (2013). A comparison of energy bugs for smartphone platforms. *First International Workshop on the Engineering of Mobile-Enabled Systems* (pp. 25–30). Institute of Electrical and Electronics Engineers.

Author Biographies

Sebastian Bähr is a senior researcher at the research department “Panel Study Labour Market and Social Security” (PASS) at the Institute for Employment Research (IAB) in Nuremberg, Germany. He holds a BA in international business from the University of Hull, United Kingdom, as well as a master’s degree in social economics and a PhD (Dr.rer.pol) from the School of Business and Economics of the Friedrich-Alexander University Erlangen-Nuremberg (FAU), Germany. His primary research interests are social inequality, labor markets, social networks, and unemployment. He studies social networks and their implications on labor market behavior (E-mail: sebastian.baehr@iab.de).

Georg-Christoph Haas is a research associate at the Statistical Methods Centre of the Institute for Employment Research (IAB) where he contributes to the project “Quality in Establishment Surveys” (QuEst) and “Mobile Device Measures” (MoDeM) and a researcher at the Collaborative Research Center 884 “Political Economy of Reforms” (SFB 884), University of Mannheim, Germany, where he contributes to the project A9 “Survey mode, survey technology and technology innovations in data collection.” He studied sociology (Diplom) at TU Dresden, Germany (E-mail: georg-christoph.haas@iab.de).

Florian Keusch is professor (interim) for statistics and methodology at the University of Mannheim, Germany and adjunct assistant professor at the joint program in survey methodology, University of Maryland, USA. He received a PhD in social and economic sciences (Dr.rer.soc.oec.) and an MSc in business (Mag.rer.soc.oec.) from WU, Vienna University of Economics and Business. His research focuses on nonresponse and measurement error in web and mobile web surveys, passive mobile data collection, and visual design effects in questionnaires (E-mail: f.keusch@uni-mannheim.de).

Frauke Kreuter is a professor in the joint program in survey methodology at the University of Maryland, USA, professor of methods and statistics at the University of Mannheim, Germany, and head of the statistical methods group at the German Institute for Employment Research (IAB), Nuremberg. She is the cofounder of the Coleridge Initiative and the International Program in Survey and Data Science (IPSDS). Her recent textbooks include *Big Data and Social Science: A Practical Guide to Methods and Tools*, and *Practical Tools for Designing and Weighting Survey Samples* (E-mail: fkreuter@umd.edu).

Mark Trappmann is in charge of the Panel Study “The Labour Market and Social Security” (PASS) at the Institute for Employment Research (IAB) at Nuremberg, Germany and a professor for sociology, especially survey methodology, at Otto-Friedrich-University Bamberg, Germany. He completed his first state teacher training examination (1. Staatsexamen für das Lehramt) in social sciences and mathematics after studying at the Gerhard-Mercator University in Duisburg, Germany, and the Rijks University in Groningen, The Netherlands, and he received his doctorate (Dr.Phil.) from the University of Essen, Germany (E-mail: mark.trappmann@iab.de).