

Calibrating Items With Time Use Diaries: A Refined Method

Ettore Scappini¹ 

[1] *Department of Education Studies, University of Bologna, Bologna, Italy.*

Methodology, 2025, Vol. 21(3), 197–219, <https://doi.org/10.5964/meth.13215>

Received: 2023-11-08 • **Accepted:** 2025-07-02 • **Published (VoR):** 2025-09-30

Handling Editor: Jochen Mayerl, Chemnitz University of Technology, Chemnitz, Germany

Corresponding Author: Ettore Scappini, via Filippo Re, 6, 40129 Bologna, Italy. E-mail: ettore.scappini@unibo.it

Supplementary Materials: Materials [see [Index of Supplementary Materials](#)]



Abstract

The aim of the article is to refine a calibration method already presented and used to improve the information provided by the scales of frequency in questionnaires by combining it with data from time use diaries. In other words, this study proposes improvements to an existing calibration method aiming at “adjusting” the data gathered through items – which is useful for the analysis of phenomena with relatively long time cycles, but also notoriously subject to bias – with the data gathered through daily diaries – which are less subject to distortion, but generally only suitable for analysing phenomena with short or very short time cycles. In some cases, in fact, the calibration model already proposed may be problematic since, as we shall see, it could introduce another possible cause of bias. Such distortion could become relevant in certain situations and can be remedied by the proposed refinement with the new calibration model under consideration here. Finally, to highlight the advantages of the proposed method, we will develop with practical applications, comparisons by applying the presented models on data on religious practice collected in a large survey conducted in Italy in 2008. It should be specified, however, that the applicability of the proposed model is not limited to this example and can be extended to other contexts and types of data.

Keywords

time-use research, daily diaries, measured presence, stylized presence, calibrating items, stylized items, ogive

As is well known, current research identifies two main data collection instruments to obtain information on the intensity with which an activity is carried out: the scale of



frequency or *stylized item* (hereinafter also simply referred to as *item*) and the *time use diaries* (hereafter simply referred to as *diaries*). In the former case, respondents are generally asked to indicate how frequently they perform a given activity within an established time, which could be a week, a month, or more commonly a year. In the latter case, typically the respondents compile a daily diary in which they note down which activity, or activities, they carry out at established intervals, typically 10 to 15 minutes, and in which place. Both data collection methods have their advantages and disadvantages.

Using items, it is possible to carry out surveys on the distribution of the intensity of a given activity over relatively long time cycles. However, such information can be subject to bias, and, in fact, for many surveys is one of the most important sources of error (Biemer, 2010, p. 823).¹ Often the use of items doesn't allow to collect data in an accurate manner (Kan & Pudney, 2008) because recalling what activity took place, how often it occurred, and how long it lasted after some time has elapsed is rather difficult. Moreover, they are disproportionately prone to social desirability or demonstration effects because the indication or omission of having carried out a given activity involves, at most, passive action (Gershuny, 2003, 2012). Finally, the problem of the low precision level of the stylized item is exacerbated when the study addresses specific issues, such as those related to some form of obligation or expectation – also in moral terms – regarding the anticipated behaviour, such as in the case of religious behaviour (Presser & Stinson, 1998).

Instead, diaries can be used to collect more reliable information. Their chronological structure makes it easier to record the timing and recollection of events (Belli, 1998). Moreover, any bias due to memory gaps during the compilation phase is generally limited (Al Baghal et al., 2014), a distortion which can be further reduced through containment strategies such as, for example, providing options to log activities as soon as possible, using smartphones and tablets (te Braak et al., 2023), and properly training interviewers on techniques for “retrieving” memory lapses (Kirchner et al., 2018). Notably, unlike with items, falsification requires episodes to be actively invented. Consequently, as honesty is the easiest behaviour for respondents, there are fewer desirability effects (Gershuny, 2012). However, the data covers a limited reference time, as almost nothing is known about the distribution of the intensity with which the given activity is carried out in the period when the diary is not updated (Scappini, 2010).

To overcome the difficulties arising from these two data collection methods, researchers have suggested finding a model that can calibrate the values obtained from items with those collected from diaries in such a way as to obtain data as complete and reliable

1) This type of error is referred to in various ways, such as validity or measurement error (Andrews, 1984). Furthermore, the total survey error can be due to multiple factors, here we will deal only with this specific aspect (Groves & Lyberg, 2010).

as possible. This requirement has led to many attempts by several scholars — above all by Gershuny — to make combined use of the qualities of the two data collection tools. Kan and Gershuny (2009) showed that it is possible to calibrate items by combining two datasets: One derived from a survey that collected questionnaire and diary data from the same respondents, and the other from a questionnaire-based survey. This method was then developed using the latest matching techniques (Borra et al., 2013; Walthery & Gershuny, 2019).

The disadvantage of these approaches is the use of unusual theoretical concepts and relatively complex regression techniques. In contrast, a simpler solution is the one presented in Scappini (2021), which is also, however, not without its problems. While the *Linear*, and conclusive, model presented may be attractive in that it allows, with the information derived from the diaries, for the easy calibration of items by means of *a function that has a gradual and not discontinuous development*, as will be seen, it also introduces a possible bias factor. Such distortion could become relevant in certain situations, a problem for which a solution is proposed with the refinement of the model under consideration here.

In order to present some practical applications of the method, data from the Time Use Surveys, conducted in Italy in 2008, in which there is a questionnaire that includes an item on frequency at mass and a daily diary, will be used.² Indeed, it is well known that surveys that include items that detect religious behaviour are particularly affected by distortion that can be both very high and very inconstant (Hadaway et al., 1993; Presser & Chaves, 2007; Presser & Stinson, 1998; Rossi & Scappini, 2012; Scappini, 2018). In contrast, diaries in the face of substantially correct information do not allow us to identify, or delimit, within the community surveyed, the subgroup of regular practitioners, as variously defined.³ The application of this method, to this kind of data, has the peculiarity of being able to show the important level of bias that can be generated by the items and allows to overcome the diary's limit.

The paper is organized as follows. The following section presents the data and discusses the different characteristics of the indicators. The next one reviews the existing models and the reason why their application may produce calibrations subject to bias, then goes on to present two new models, which are the focus of the present paper. Finally, in the last section, some comparisons will be made using real data.⁴

2) It should be noted that calibration is a similar procedure that leads to the construct of an *ogive* (Klugman et al., 2019). The goal of this study, however, is not only to formalize a specific model for making a discrete function continuous, but also to better investigate the validity of its assumptions and to calibrate, with the use of diaries, the stylized items.

3) Typically, the rule is to divide between regular churchgoers (those who go to church at least once a month or more) and irregular or non-churchgoers (all the others). This is not the only solution and there are variations in which regular churchgoers are established as those who attend a service at least every two weeks (Lechner, 1996) or also every week (Knippenberg, 2015).

Diaries and Stylized Items: Indicators With Different Characteristics

The Indicator Provided Via the Diary

The following is a description of the indicator that is derived from the use of a daily diary. After determining the total number of subjects who will have to keep a diary (N), we will build sub-samples, each composed of N/D individuals. The number of sub-samples is equal to D , which is the period (defined in days) during which the diaries are kept. Typically, the period D is equal to 364 days.⁵

We can now organize the data in the form of a matrix composed of N/D rows and D columns and calculate the ratio $P = \frac{1}{N} \cdot \sum_{j=1}^D \sum_{i=1}^{N/D} x_{i,j}$ with $x = \{0,1\}$ – i.e. the proportion between *positive events* ($\sum_{j=1}^D \sum_{i=1}^{N/D} x_{i,j}$) and *possible events* (N) – (Rossi & Scappini, 2014). As will be seen, there will be a need to decompose the overall value P into I subgroups. In this case we will identify different values of p_i with $p_i = \{p_1, \dots, p_I\}$. We will call the statistics P and p_i with the term *measured presence* (hereinafter also just *presence*).

However, P is a “poor” indicator of information because, as we have seen, the subjects surveyed on the various days belong to different sub-samples. It is therefore not possible to select the part of the population that carries out a given activity within a specific range of intensity for periods longer than the extension of the diary.

The Indicator Provided Via the Stylized Item

Daily diaries are not always sufficient to provide an adequate answer, as there are activities with relatively long time cycles. In such cases, it would be necessary to select the part of the population that carries out a given activity over a longer period of time: for example, one week, one month, or longer intervals.

The typical solution to this problem is to employ a questionnaire with a suitable item that can be used to determine how frequently each subject performs a specific activity over a given period, which is usually one year. Ideally, if n is the number of days in the given period (generally $n = 364$), the $n + 1$ values f_t can be calculated, each of which provides the number of people attending mass t times, with t ranging from 0 to 364. Each ratio f_t/N , where N is the size of the sample, provides the daily attendance rate for each single value of t . This can also take the form of a cumulative rate to

4) Associated with this paper is a program which enables the production of the Table and Figures to be presented. Further information can be found in [Appendix A](#) of the [Appendices](#).

5) It should be noted that in [Scappini \(2021\)](#), weekly diaries were used, whereas the present study employs daily diaries. From a strictly formal point of view, nothing has changed, as the calibration models are identical.

indicate the proportion of people who perform an activity at least t times per year – $CF(t) = \sum_{t=0}^{364} f_t / N, \forall t \geq t$.

Having defined the indicators derived from the use of the two survey instruments, we now need to make these measures comparable.

The Conversion From Frequency to Presence

As is known, although the presence values provided by diaries cannot be converted into frequencies provided by items, the reverse process is possible. Using a similar approach to other authors (Presser & Chaves, 2007), to perform this conversion we have to add together the number of people and the relative typical frequency t – thereby identifying the positive events $\sum_{t=0}^{364} (f_t \cdot t)$ with $\sum_{t=0}^{364} f_t = N$ – and dividing the result by the number of possible events – $(N \cdot 364)$. In formal terms, $P(t) = (\sum_{t=0}^{364} (f_t \cdot t) / (N \cdot 364)) \cdot 100$.

However, it is unrealistic to ask respondents for such precise occurrence about their attendance at mass over the course of a year. In general, as in this case, it is preferable to offer a limited number of answer options I , that correspond to the frequency t_i for each option i of the item. If now we set $s_i = (t_i / 364) \cdot 100$ with $s_i = \{s_1, \dots, s_I\}$, then $S = \sum_{i=1}^I f_i \cdot s_i / N$. We will call the statistics S and s_i with the term *stylized presence*. To make this conversion, we must tackle an additional problem: identifying the values of frequency t_i .

Before carrying out this task, we need to present the data.

The Data

The rationale that follows in the next two paragraphs will be developed with the use of two datasets. The first dataset consists of simulated data which will be used to present the different calibration models. The second dataset is the *Time Use Survey* conducted in Italy in 2008 (henceforth TUS, 2008), which will be used to present an empirical example of the benefits that can be achieved with calibration.

The Simulated Data

The first dataset consists of targeted constructed data in order to clearly highlight the differences in the application of the four models and will not be used to present a real application.⁶ The criteria that guided the construction of this data are twofold. The first was to visualize graphically in a more distinguishable manner the outcome of the application of the calibration models, a result not achievable with the real data. The

6) For data and figures see Supplementary Materials A (Scappini, 2025a).

second was to highlight the situations in which bias may emerge due to the application of the models to be discussed below.

The TUS 2008 Data

The second dataset is made up of TUS 2008 data belonging to the more general ISTAT Multipurpose Survey System which was generally conducted every five years. The survey used here was carried out in the timeframe February 2008–January 2009 (TUS, 2008). The respondents kept a diary over a 1-day period and recorded what they were doing (every 10 minutes) and where they were. In addition, they answered a detailed stylized questionnaire.

The sample consisted of 18,240 families with response rate equal to 73.96% (American Association for Public Opinion Research response RR1). A further selection due to the non-response diaries must be added to this sample dropout. In this case, of the 43,460 eligible diaries, including only subjects with 3 years or more, 40,944 were collected, broken down as follows: 14,787 relate to a weekday (Monday–Friday), 13,286 relate to a Saturday, and 12,871 relate to a Sunday. For the purposes of this study, we will single out respondents aged between 18 and 74, with a final sample of 30,673 people. To minimize the potential bias, the analysis was weighed by day of the week, gender, age, level of urbanization and multi-regional area.

As highlighted earlier, to make a comparison the same activity (here, religious practice) needs to be surveyed using both a diary and a suitable item.

For the diary, we used the codes regarding religious practice in places of religious worship.⁷ The minimum period of time considered in diaries is 10 minutes and is associated with the main activity carried out in that timeframe. A subject was counted as “present at mass” if there were at least *two* minimum-length episodes in their compiled daily diary, thus corresponding to attendance for a time equal or greater than 15 minutes.

Regarding the stylized item, the question used is: “How often do you usually go to church or another place of worship?”. For the available options the respective values of s_i are identified with the following values applied t_i where: 0 times for the option “never” with $s_1 = 0.00$; 6 times for “a few times a year” with $s_2 = 6/364 \cdot 100 = 1.65$; 24 times a year for “a few times a month (but less than four times)” with $s_3 = 6.59$; 52 times a year for “once a week” with $s_4 = 14.29$; 182 times a year for “a few times a week” with $s_5 = 50.00$; and, finally, 364 times a year for “every day” with $s_6 = 100.00$.⁸

7) For additional information see [Appendix B](#) of the [Appendices](#).

8) In keeping with the strict monotonic ascending order necessary for the program to function, we have reversed the original order of the options.

The Calibration Models

Let us now introduce the four models of calibration pertinent to our discussion and the problems associated with their application. As previously mentioned, in this section we will only use simulated data.

The Uniform Model

We will now consider an item administered to a sample of N individuals, where the possible response options I correspond to data values or frequency ranges. Let us now identify the sub-sample n_i that selected the response option i , from which we will derive probability that will be equal to $d_i = n_i/N$, with $i = \{1, \dots, I\}$. If we now set $p_i = \{p_1, \dots, p_I\}$ with $p_{i-1} < p_i$, and $d_i = \{d_1, \dots, d_I\}$, then $P(X = p_i) \implies D = d_i$, where d_i is the fraction of the population that carries out a given activity with a measured presence in the sub-sample i equal to p_i . We note that the assumption $p_{i-1} < p_i$ is important because it guarantees that there is a reasonable link between declared behaviour – as expressed in the item – and measured behaviour – as noted in the diary (Scappini, 2021).

If, instead of a categorical variable, we assume that X is a continuous variable, with area $D = d_i$ and $X \sim U\left[\frac{p_{i-1} + p_i}{2}, \frac{p_i + p_{i+1}}{2}\right]$, then $P(X = x_i) \implies D = y_i$, where y_i is now the population density with measured presence equal to x_i with $x_i \in \left[\frac{p_{i-1} + p_i}{2}, \frac{p_i + p_{i+1}}{2}\right]$. We will now calculate the coordinates of $(x_i, y_i) \forall i = \{2, \dots, I-1\}$.

Starting from the *abscissa* values (x_i), the corresponding *ordinate* values will be $y_i = \frac{d_i}{(p_{i+1} - p_{i-1})/2}$. For the two tails, if $i = 1$, we have $y_1 = \frac{d_1}{(p_1 + p_2)/2 - p_0}$ with $p_0 \in [0, p_1)$, while if $i = I$, we have $y_I = \frac{d_I}{p_{I+1} - (p_{I-1} + p_I)/2}$ with $p_{I+1} \in (p_I, 100]$. Given these assumptions, we can now build the calibration function. The *uniform PDF*, hereinafter called $u(x)$, will be defined as follows:

$$u(x) = \begin{cases} \frac{d_1}{(p_1 + p_2)/2 - p_0}, & p_0 \leq x_1 < \frac{p_1 + p_2}{2} \\ \frac{d_i}{(p_{i+1} - p_{i-1})/2}, & \frac{p_{i-1} + p_i}{2} \leq x_i \leq \frac{p_i + p_{i+1}}{2} \\ \frac{d_I}{p_{I+1} - (p_{I-1} + p_I)/2}, & \frac{p_{I-1} + p_I}{2} < x_I \leq p_{I+1} \end{cases}$$

Then the relative uniform *CDF*, hereinafter $U(x)$, is equal to:

$$U(x) = \begin{cases} d_1 \cdot \frac{x_1 - p_0}{(p_1 + p_2)/2 - p_0}, & p_0 \leq x_1 < \frac{p_1 + p_2}{2} \\ \sum_{j=1}^{i-1} d_j + d_i \cdot \frac{x_i - (p_{i-1} + p_i)/2}{(p_{i+1} - p_{i-1})/2}, & \frac{p_{i-1} + p_i}{2} \leq x_i \leq \frac{p_i + p_{i+1}}{2} \\ \sum_{j=1}^{I-1} d_j + d_I \cdot \frac{x_I - (p_{I-1} + p_I)/2}{p_{I+1} - (p_{I-1} + p_I)/2}, & \frac{p_{I-1} + p_I}{2} < x_I \leq p_{I+1} \end{cases}$$

defined as $\forall i = \{1, 2, \dots, I\}$.

In this way, we obtained an initial result that is much less subject to bias than the one derived solely from the reliance on an item. However, the assumption of uniform distribution is extremely improbable in practice. It is unlikely the PDF pattern would feature break at the transition between the different values of $u(x)$. A more reasonable assumption is that the development is more progressive. In the next section, we will describe a solution to this problem.

The Linear Model

If we use the values y_i defined with the uniform distribution and set $m_i = \frac{y_i - y_{i-1}}{p_i - p_{i-1}}$ and $q_i = \frac{p_i \cdot y_{i-1} - p_{i-1} \cdot y_i}{p_i - p_{i-1}} \forall i = \{2, 3, \dots, I\}$, we can now develop a model that better responds to the above mentioned criteria of progression: the *linear PDF*, hereinafter called $l(x)$, will be defined as follows:

$$l(x) = \begin{cases} \frac{d_1}{(p_1 + p_2)/2 - p_0}, & p_0 \leq x_1 < p_1 \\ m_i \cdot x_i + q_i, & p_{i-1} \leq x_i \leq p_i \\ \frac{d_I}{p_{I+1} - (p_{I-1} + p_I)/2}, & p_I < x_{I+1} \leq p_{I+1} \end{cases}$$

The relative linear CDF, hereinafter $L(x)$, will be equal to:

$$L(x) = \begin{cases} d_1 \cdot \frac{x_1 - p_0}{(p_1 + p_2)/2 - p_0}, & p_0 \leq x_1 < p_1 \\ d_1 \cdot \frac{p_1 - p_0}{(p_1 + p_2)/2 - p_0} + m_2 \cdot \frac{x_2^2 - p_1^2}{2} + q_2 \cdot (x_2 - p_1), & p_1 \leq x_2 < p_2 \\ \sum_{j=1}^{i-2} d_j + d_{i-1} \cdot \frac{p_{i-1} - p_{i-2}}{p_i - p_{i-2}} + m_i \cdot \frac{x_i^2 - p_{i-1}^2}{2} + q_i \cdot (x_i - p_{i-1}), & p_{i-1} \leq x_i \leq p_i \\ \sum_{j=1}^{I-1} d_j + d_I \cdot \frac{x_{I+1} - (p_{I-1} + p_I)/2}{p_{I+1} - (p_{I-1} + p_I)/2}, & p_I < x_{I+1} \leq p_{I+1} \end{cases}$$

defined as $\forall x_i$ with $i = \{1, 2, \dots, I + 1\}$.⁹

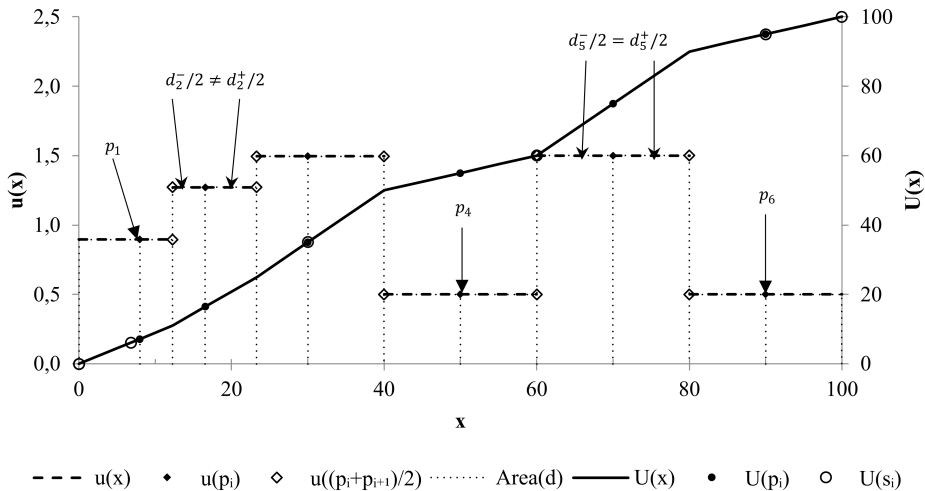
This formulation has the advantage of a better graduality in the development of the values of $l(x)$ and, therefore, in the development of the values of $L(x)$, as well as producing a non-discontinuous function. Let us now analyse why the application of the two models presented may be subject to bias.

The Bias in the Models

We are now going to introduce the information entered in Figure 1. Let's start from $u(x)$, which is a probability density function, whose trend is determined by the values of p_i and d_i and is marked as $u(p_i)$ on the graph. The values indicated with $u((p_i + p_{i+1})/2)$ together with the line marked as Area (d) are useful to delimit the relevant reference areas. The continuous line shows the trend of the calibrated values $U(x)$, while the associated symbols on the same line, $U(p_i)$ and $U(s_i)$, correspond to the specific calibrated values. In the first case, the calibration will be calculated for the values of p_i , and therefore with reference to the measured presence, values which, we note, are usually helpful only to aid the reading of the graph, while in the second case the calibration will be calculated for the values of s_i , and therefore with reference to the stylized presence.

Figure 1

Calibrated Uniform Model, $u(x)$ and $U(x)$



9) It is easy to demonstrate that $L(p_i) = U(p_i)$.

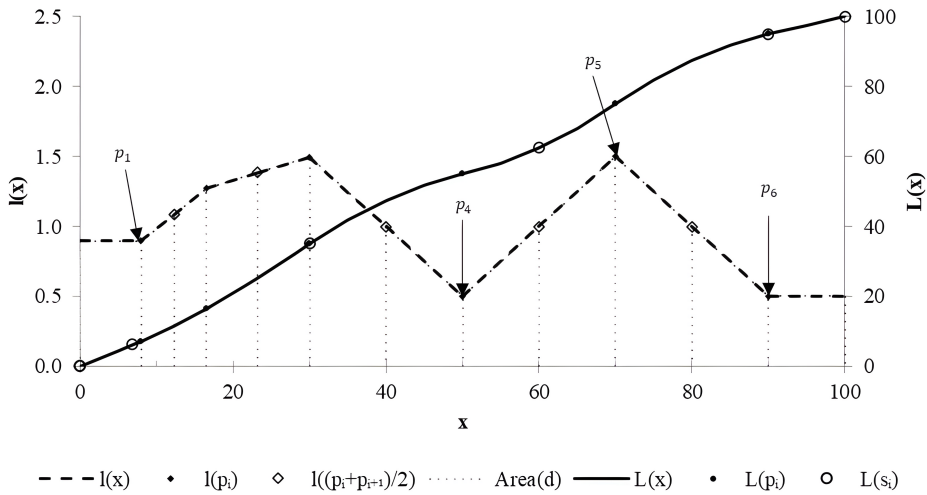
Before continuing, we would like to point out that it is possible to calibrate $F(x)$, $\forall x \in [p_0, p_{I+1}]$, while the comparisons between $CF(x)$ and $F(x)$ are feasible only for $x = s_i$.

From the comparison between Figure 1 and Figure 2, the improvement in terms of smoothness of linear vs uniform calibration is evident. The problem now arising is that neither of these formulations – $u(x)$ and $l(x)$ – guarantees that:

$$E[f(X)|X = x_i] = p_i \text{ with } x_i \in \left[\frac{p_{i-1} + p_i}{2}, \frac{p_i + p_{i+1}}{2} \right], \forall i = \{2, \dots, I - 1\}$$

Figure 2

Calibrated Linear Model, $l(x)$ and $L(x)$



In general, with the uniform calibration this doesn't happen, as normally $E[u(X)|x_i] \neq p_i$. In addition, the assumption of a gradual development of the function may generate a further asymmetry in the distribution of the probabilities $f(x)$. Therefore, even though $E[u(X)|x_i] = p_i$, in general it will still be the case that $E[l(X)|x_i] \neq p_i$.

Since it is not possible to generalize the attractive assumption discussed above, we propose as an alternative to take only the part that contributes in terms of CDF for a value related to the one given by $E[f(X)|x_i] = p_i$, basically ignoring what happens after the value of p_i . Therefore, as the average value divides the area of the part in two so that:

$$F(p_i) - F\left(\frac{p_{i-1} + p_i}{2}\right) = \frac{d_i}{2}$$

it is possible to consider equally attractive the following occurrence of equality:

$$F(p_i) = \sum_{j=1}^{i-1} d_j + \frac{d_i}{2} \tag{1}$$

Figure 1 shows an example of an ideally non-problematic situation, in which Equation (1) is verified, for $u(x_5)$ with $d_5^- = d_5^+$.¹⁰ However, if Equation (1) is not verified, then it is possible to regard this breach as a bias factor due to the calibration model. The Figure also shows the example of this situation for $u(x_2)$ with $d_2^- \neq d_2^+$.

To solve all the above problems, we need to look at an alternative model which we are going to illustrate.

The New Uniform Model

If we assume that X is a continuous variable, with $X \sim U\left[\frac{p_{i-1} + p_i}{2}, p_i\right)$ and $X \sim U\left[p_i, \frac{p_i + p_{i+1}}{2}\right)$ with an area $D = d_i/2$, we can now calculate the respective coordinates of (x_i^-, y_i^-) defined $\forall i = \{2, 3, \dots, I\}$ and (x_i, y_i) defined $\forall i = \{1, 2, \dots, I - 1\}$. Starting from the *abscissa* values x_i , with $x_i \in [p_0, p_{I+1}]$, the corresponding *ordinate* will be equal to:

$$y_i^- = \frac{d_i/2}{(p_i - p_{i-1})/2} \text{ if } (p_{i-1} + p_i)/2 \leq x_i^- < p_i,$$

and

$$y_i = \frac{d_i/2}{(p_{i+1} - p_i)/2} \text{ if } p_i \leq x_i < (p_i + p_{i+1})/2.$$

Regarding the two tails, we have if $i = 1$, $y_1^- = \frac{d_1/2}{p_1 - p_0}$ with $p_0 \in [0, p_1)$, and if $i = I + 1$, $y_{I+1} = \frac{d_I/2}{p_{I+1} - p_I}$ with $p_{I+1} \in [p_I, 100]$. We can now develop the modified calibration.

The *new uniform PDF*, hereinafter $nu(x)$, is then equal to:

$$nu(x) = \begin{cases} \frac{d_1/2}{p_1 - p_0}, & p_0 \leq x_1^- < p_1 \\ \frac{d_i/2}{(p_{i+1} - p_i)/2}, & p_i \leq x_i < (p_i + p_{i+1})/2 \\ \frac{d_i/2}{(p_i - p_{i-1})/2}, & (p_{i-1} + p_i)/2 \leq x_i^- < p_i \\ \frac{d_I/2}{p_{I+1} - p_I}, & p_I \leq x_{I+1} \leq p_{I+1} \end{cases}$$

10) The one presented here is a particular case in which also $E[I(X)|x_i] = p_i$, see Figure 2 for $i = 5$, which happens only if $p_i - p_{i-1} = p_{i+1} - p_i$ and if $m_i = -m_{i+1}$.

If we now set $d_0 = 0$, then the relative CDF, hereinafter $nU(x)$, is equal to:

$$nU(x) = \begin{cases} \frac{d_1}{2} \cdot \frac{x_1 - p_0}{p_1 - p_0}, & p_0 \leq x_1^- < p_1 \\ \sum_{j=0}^{i-1} d_j + \frac{d_i}{2} \cdot \left(1 + \frac{x_i - p_i}{(p_{i+1} - p_i)/2} \right), & p_i \leq x_i < (p_i + p_{i+1})/2 \\ \sum_{j=1}^{i-1} d_j + \frac{d_i}{2} \cdot \frac{x_i - (p_{i-1} + p_i)/2}{(p_i - p_{i-1})/2}, & (p_{i-1} + p_i)/2 \leq x_i^- < p_i \\ \sum_{j=1}^{I-1} d_j + \frac{d_I}{2} \cdot \left(1 + \frac{x_{I+1} - p_I}{p_{I+1} - p_I} \right), & p_I \leq x_{I+1} \leq p_{I+1} \end{cases}$$

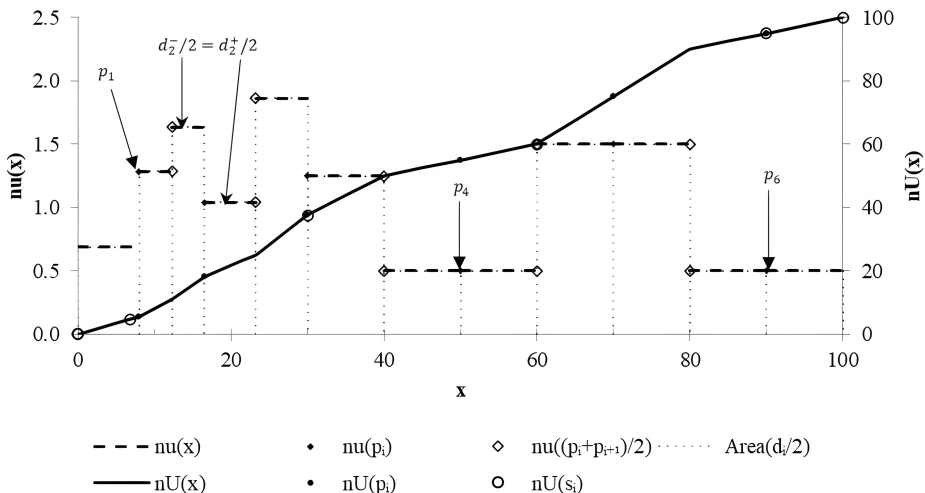
defined as $\forall i = \{1, 2, \dots, I + 1\}$.

We note that there are two special cases, where $p_0 = p_1$, and $p_{I+1} = p_I$. In both these situations calibration for values of x_1 and x_I is not possible. Then, the values of $nU(x)$ will be calculated assuming that x_i is discrete and we will put respectively $nU(X = p_1) = d_1/2$ in the first case, and $nU(X = p_I) = d_I/2$ in the second.

While it is true that this model is not subject to bias since by definition $nU(p_i) = \sum_{j=0}^{i-1} d_j + \frac{d_i}{2}$, the function $nU(x)$ is discontinuous (see Figure 3), and the resulting values are less smoothed out compared to those shown in Figure 2.

Figure 3

Calibrated New Uniform Model, $nu(x)$ and $nU(x)$



To sum up, we have now achieved a first result: a calibration model that is not subject to bias. However, it is also true that the assumption of uniform distribution is, in practice, very unlikely. Similarly to what we have already pointed out to justify the transition from $n(x_i)$ to $l(x_i)$, it can be considered unrealistic to have “breaks” between adjacent values of $nu(x_i)$ in the trend of the PDF, while it would seem more reasonable to assume that the trend from $nu(x_i)$ to $nu(x_{i+1})$ is more progressive. We will address the issue in the next section.

The New Linear Model

As in the case of the Linear model, if we use the values y_i defined with the new uniform distribution and we place $m_i^* = \frac{y_i - y_{i-1}}{p_i - p_{i-1}}$ and $q_i^* = \frac{p_i \cdot y_{i-1} - p_{i-1} \cdot y_i}{p_i - p_{i-1}} \forall i = \{2, 3, \dots, I\}$, while leaving the two tails unchanged, a model can be developed that better meets the progressivity criteria now mentioned.

The *new linear PDF*, hereinafter $nl(x)$, is then equal to:

$$nl(x) = \begin{cases} \frac{d_1/2}{p_1 - p_0}, & p_0 \leq x_1 < p_1 \\ m_i^* \cdot x_i + q_i^*, & p_{i-1} \leq x_i < p_i \\ \frac{d_I/2}{p_{I+1} - p_I}, & p_I \leq x_{I+1} \leq p_{I+1} \end{cases}$$

If we now set $d_0 = 0$, then the relative CDF, henceforth $nL(x)$, will be equal to:

$$nL(x) = \begin{cases} \frac{d_1}{2} \cdot \frac{x_1 - p_0}{p_1 - p_0}, & p_0 \leq x_1 < p_1 \\ \sum_{j=0}^{i-2} d_j + \frac{d_{i-1}}{2} + m_i^* \cdot \frac{x_i^2 - p_{i-1}^2}{2} + q_i^* \cdot (x_i - p_{i-1}), & p_{i-1} \leq x_i < p_i \\ \sum_{j=1}^{I-1} d_j + \frac{d_I}{2} \cdot \left(1 + \frac{x_{I+1} - p_I}{p_{I+1} - p_I}\right), & p_I \leq x_{I+1} \leq p_{I+1} \end{cases}$$

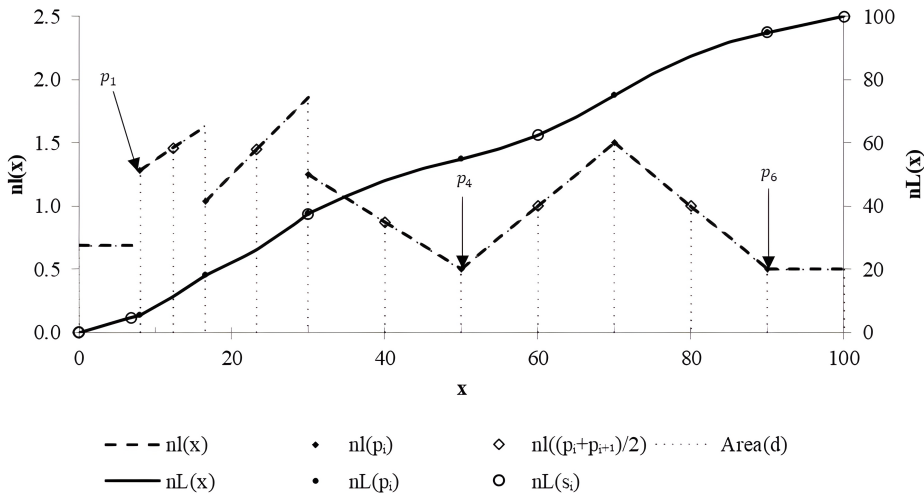
defined as $\forall i = \{1, 2, \dots, I + 1\}$.

Similar to the previous method, we observe that in the two special cases, those in which $p_0 = p_1$, and $p_I = p_{I+1}$, the values of $nL(x)$ will be calculated without calibration: in the first case we will assume that $nL(X = p_1) = d_1/2$, while in the second that $nL(X = p_I) = d_I/2$.

If we take a look at [Figure 4](#), we find we have a more attractive calibration model than the previous ones. While this is not, in general, a continuous model, like the Linear – $L(x)$ – it is nevertheless a correct model and more smoothed out than the new Uniform – $nU(x)$.¹¹

Figure 4

Calibrated new Linear Model, $nl(x)$ and $nL(x)$



We will now present the results of applying the calibration models to the TUS 2008 data.

Empirical Study

It has been shown that models named $U(x)$ and $L(x)$ may be subject to bias because they do not guarantee that $E[f(X)|X = x_i] = p_i$. Next, it was shown that the $U(x)$ and $nU(x)$ models have probably unreliable assumptions since it can be considered unrealistic to have “breaks” between adjacent values of $f(x_i)$ in the trend of the PDF. It follows from this reasoning that the most interesting models are those that assume a more progressive trend and thus those denoted by $L(x)$ and $nL(x)$. However, the former, as has been shown, can be affected by bias, while the latter does not exhibit this problem. Consequently, in the comparisons we will carry out we will use only the most advanced calibration models, $L(x)$ and $nL(x)$, assuming the latter as the correct one.¹²

11) The maximum hypothetical value of the bias is for $\lim_{p_{i-1} \rightarrow p_i} (nL(X \geq p_{i-1}) - L(X \geq p_{i-1}))$, equal to $\frac{d_i}{2}$ if $p_0 = p_1, \frac{d_i}{2}$ if $p_i = p_{i+1}$ and $\frac{d_{i-1} + d_i}{2}$ otherwise.

Moreover, in Supplementary Material A (Scappini, 2025a), there is the file “Transparencies SimpleExample four models.pdf”, which contains the four graphs useful for comparing the aforementioned models.

12) It is specified that the $nL(x)$ model is preferred for the application that has been investigated here. Calibration, however, can be extended to other situations. For example, it is possible to calibrate a stylized item even without the

Let us now go on to apply the calibration to a real survey. The data and item related to the example we are going to propose, namely religious practice in Italy in 2008, lend themselves well to highlighting the important aspects we have drawn attention to.

We will carry out the discussion in two parts. First, we will reconfirm what is already known about the important overestimation of the retro-cumulated values calculated using the stylized items alone compared to the calibrated values. Second, we will compare a series of calibrated CDF values from the two models. As will be seen, beyond the formal aspects discussed, in practical use, or at least in the exemplification presented here, the values obtained are not very different from each other. Only in one situation among those elaborated, which however, is potentially re-presentable, did we detect a level of bias that can be considered relevant.

Regarding overestimation, an aspect that typically characterizes surveys on religious practice, we point out that the bias “produced” by stylized items takes on considerable values. To give an example (see [Table 1](#) and [Figure 5](#)) if we consider those who say they go to Mass once a week (Option 4, $s_4 = 14.29$), compared with a value of $CF(X \geq s_4) = 30.2\%$, we have that $L(X \geq s_4) = 8.9\%$ and $nL(X \geq s_4) = 8.4\%$. Very large differences in both absolute (> 20 percentage points) and relative EI $> 250\%$ terms.¹³ The situation is not much better if we consider the values of $L(X \geq s)$ and $nL(X \geq s)$ in the other options, with EI varying, respectively, from a minimum of 33/44%, in Option 2, to a maximum of 680/995%, in option 5.¹⁴

use of diaries. It is beyond the scope of this study to delve into this aspect, let us just say that in such cases the most correct model to use is the second of those presented – $L(x)$.

13) With $EI = ((CF(s) - F(s))/F(s)) \cdot 100$ and $EC = ((L(x) - nL(x))/nL(x)) \cdot 100$.

14) See also Supplementary Materials B ([Scappini, 2025b](#)).

Table 1

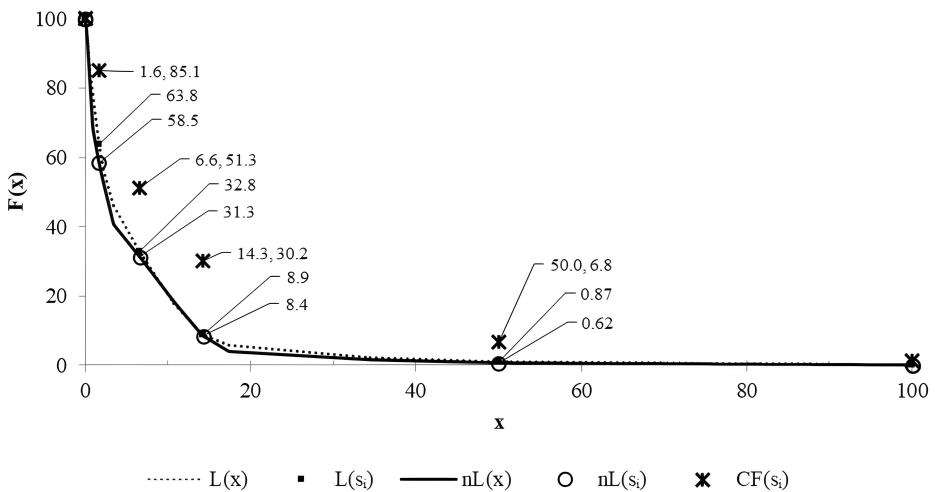
Calibrated TUS 2008 Data, Stylized CF(s), Linear Model L(x) and New Linear Model nL(x)

Options (i)	1	2	3	4	5	6	Total
Measured Presence (p_i) %	0.24	0.88	3.39	10.74	17.32	51.49	5.124
Stylized Presence (s_i) %	0.00	1.65	6.59	14.29	50.00	100.00	9.360
Sample %	14.9	33.8	21.2	23.4	5.6	1.1	100.0
N	4,565	10,372	6,488	7,172	1,728	348	30,673
<i>Retro-cumulative population %</i>							
Stylized CF($X \geq s_i$)	100.0	85.1	51.3	30.2	6.8	1.1	
Calibrate $L(X \geq s_i)$	100.0	63.8	32.8	8.9	0.87	0.0	
Calibrate $nL(X \geq s_i)$	100.0	58.5	31.3	8.4	0.62	0.0	

Note. Stylized question: “How often do you usually go to church or another place of worship?”; frequency options: 1. Never, 2. A few times a year, 3. A few times a month (but fewer than four times), 4. Once a week, 5. A few times a week, 6. Every day.

Figure 5

Calibrated TUS 2008 Data: Stylized CF(s), Linear Model L(x) and New Linear Model nL(x), Retro-Cumulative Function



To better highlight the total size of the errors, a measure of fit between the Stylized $CF(X \geq s_i)$ and the $L(X \geq s_i)$ and $nL(X \geq s_i)$ distributions can be used. This measure,

defined as the weighted Adjustment Indicator (*wAI*), is derived from the weighted Mean Absolute Error (*wMAE*).¹⁵ Intuitively, the *wAI* indicates the degree of similarity between the calibrated distributions and the stylized one: higher values suggest closer alignment and a reduced effect – or usefulness – of applying the model. Comparing the obtained *wAI* values, we observe percentages of 84% and 81% for the linear and new linear models and suggest a relatively large distance between the distributions.

It should be noted that the comparison between uncalibrated and calibrated values is also relevant for theoretical discussion. While using the items it can be assumed that religious practice constitutes a relevant phenomenon in Italy in 2008 as regular practitioners are an important fraction of the population (i.e., 30.2%), differently with the use of calibrated values it can be inferred that religion is a relatively minor phenomenon (i.e., 8.9/8.4%).¹⁶

Let us now turn to the comparison between the two calibration models studied. In this case the differences detected in their application are relatively small.¹⁷ Only in Option 2 ($s_2 = 1.65$) do we have a discrepancy, which can be relevant, with an overestimation of $L(X \geq s_2) = 63.8$ with respect to $nL(X \geq s_2) = 58.5$ equal to 5.3 percentage points and with EC = 9%. In the other options, the deviations are not as important, with errors of less than two percentage points and with EC < 6%. Only in relative terms does Option 5 ($s_5 = 50.00$) show considerable overestimation (EC = 40%) but we are dealing with very small values so that the absolute differences are quite negligible – in this case equal to 0.25 percentage points – i.e., $L(X \geq s_5) = 0.87\% - nL(X \geq s_5) = 0.62\%$.

The comparisons now presented show that the differences *are generally* not relevant and thus almost negligible in the theoretical discussion. The $nL(X)$ model, however, remains preferable not only because of the attractive fact that it is non-biased, but also because, in given situations, it allows us to better delimit the size of particular or specific subgroups, such as those who practice relatively intensively – i.e., $nL(X \geq s_5)$ or those who participate very rarely or never¹⁸ – i.e., $L(X < s_2) = 36.2\%$ while $nL(X < s_2) = 41.5\%$.

15) Where $wAE = 1 - wMAE$, with $wMAE = \frac{\sum_{i=1}^I w_i |CF(s_i) - F(s_i)|}{\sum_{i=1}^I w_i}$ and $\sum_{i=1}^I w_i = 1$, see Cleger-Tamayo et al. (2012).

16) It should be noted that in diachronic or demographic sub-samples comparisons, the stylized items could have different levels of error (Scappini, 2021). It is beyond the scope of this study to delve into this aspect as well, let's just say that through calibration even in these cases it is possible to make assessments and comparisons while avoiding the formulation of spurious relationships.

17) This consideration is derived from the high value of the *wAI* ($\approx 98\%$) and suggests a limited distance between the two distributions.

18) Called by Bruce "The penumbra of occasional attenders" (Bruce, 2016, p. 614). The Author points out that this is something that has not yet been explored enough, probably partly due to the fact that it is difficult to accurately estimate the size of the population that rarely or never attends religious services.

Conclusion

We now summarize the results. I think there are two points that are relevant and need to be focused on. The first concerns the choice of the most appropriate model to calibrate the data; the second pertains to the prerogatives of calibration.

Of the four calibration methods, we can summarize that the linear method, while attractive because of the contiguity of the functions $l(x)$ and $L(x)$ describing the trend, may have limitations related to the bias discussed above. This method is superseded by the new Uniform model, which is less attractive than the Linear model because it can introduce major discontinuities in the transition from one option to another. We believe, therefore, that the last model presented — $nL(x)$ — is undoubtedly preferable because, while generally it still shows discontinuities, it does not have the disadvantages of the Linear model in terms of bias or even those of the new Uniform model in terms of smoothness of the results.

Subsequently, I applied the models using data on religious practice. It should be noted, however, that calibration has the distinctive advantage of being applicable in many other areas. We will now examine some — though by no means all — of the possible fields of application.

First, time use surveys often include a questionnaire with stylized items, alongside diaries: this is seen in studies of Mass attendance in Canada (Brenner, 2011) and work hours in Germany (Otterbach & Sousa-Poza, 2010). The model could also be used in surveys measuring transport usage. In this case, the need for diaries covering many weeks could be simplified by joint use of diaries and questionnaires (Axhausen et al., 2002).

Furthermore, this method could be extended to the psychological/medical sphere, such as studies on the consumption of alcohol (Townshend & Duka, 2002) or food (Vereecken & Maes, 2003). In this case, the two tools are often used interchangeably. Using them together could increase precision and simplify data collection in cases where the analysis needs to be extended over the long term.

In short, regarding the prerogatives of calibration, we have already discussed enough about the “advantages” of being able to perform unbiased analysis on phenomena that have long time cycles. Here we just want to point out that the application presented was used for demonstration purposes only. In other words, the models are independent of the specific field of substantive research and in fact, with the appropriate data, can be applied to a wide variety of social phenomena.

Future Work

However, this article does not fully address several important topics that require further investigation. While the implemented applications effectively demonstrate the model by meeting its minimum criteria, further research would be useful. For example, a study

assessing the adequacy of the overlap between the survey items and diary-recorded activities, as some activities may not fully satisfy these assumptions. Additionally, refining model fit measures and calculating confidence intervals for parameters are necessary steps. The current approach to model fit is not completely satisfactory, but no better alternative has yet been identified.

Future research will prioritize resolving these issues to significantly improve the model's robustness and broaden its applicability.

Funding: The author has no funding to report.

Acknowledgments: I sincerely thank the reviewer for the thoughtful and constructive comments, which greatly contributed to improving the clarity and overall quality of this article. I am also grateful for the time and consideration dedicated to the review process.

Competing Interests: The author has declared that no competing interests exist.

Supplementary Materials

Type of supplementary material	Availability/Access
Data	
Data for this study are not publicly available.	—
Preregistration	
Study was not preregistered.	—
Code	
No code was provided for the study.	—
Material	
a) Input file for CaSty.2.0.exe program, corresponding output with figures from simulated data, additional files with figures to support discussion of models.	Scappini (2025a)
b) Input file for CaSty.2.0.exe program, corresponding output with the figures obtained from TUS 2008 data.	Scappini (2025b)
Software	
CaSty - Calibrating Stylized Items, Version 2.0.	Scappini (2025c)

References

- Al Baghal, T., Belli, R. F., Phillips, A. L., & Ruther, N. (2014). What are you doing now? Activity level responses and recall failures in the American Time Use Survey. *Journal of Survey Statistics and Methodology*, 2(4), 519–537. <https://doi.org/10.1093/jssam/smu020>

- Andrews, F. M. (1984). Construct validity and error components of survey measures: A structural modeling approach. *Public Opinion Quarterly*, 48(2), 409–442. <https://doi.org/10.1086/268840>
- Axhausen, K. W., Zimmermann, A., Schönfelder, S., Rindsfuser, G., & Haupt, T. (2002). Observing the rhythms of daily life: A six-week travel diary. *Transportation*, 29(2), 95–124. <https://doi.org/10.1023/A:1014247822322>
- Belli, R. F. (1998). The structure of autobiographical memory and the event history calendar: Potential improvements in the quality of retrospective reports in surveys. *Memory*, 6(4), 383–406. <https://doi.org/10.1080/741942610>
- Biemer, P. P. (2010). Total survey error: Design, implementation, and evaluation. *Public Opinion Quarterly*, 74(5), 817–848. <https://doi.org/10.1093/poq/nfq058>
- Borra, C., Sevilla, A., & Gershuny, J. (2013). Calibrating time-use estimates for the British Household Panel Survey. *Social Indicators Research*, 114(3), 1211–1224. <https://doi.org/10.1007/s11205-012-0198-2>
- Brenner, P. S. (2011). Exceptional behavior or exceptional identity? Overreporting of church attendance in the U.S. *Public Opinion Quarterly*, 75(1), 19–41. <https://doi.org/10.1093/poq/nfq068>
- Bruce, S. (2016). The sociology of late secularization: Social divisions and religiosity. *British Journal of Sociology*, 67(4), 613–631. <https://doi.org/10.1111/1468-4446.12219>
- Cleger-Tamayo, S., Fernández-Luna, J. M., & Huete, J. F. (2012). *On the use of Weighted Mean Absolute Error in Recommender Systems* (pp. 24–26). Workshop on Recommendation Utility Evaluation: Beyond RMSE (RUE 2011). <https://ceur-ws.org/Vol-910/paper5.pdf>
- Gershuny, J. (2003). *Changing times: Work and leisure in postindustrial society*. Oxford University Press. <https://doi.org/10.1093/oso/9780198287872.001.0001>
- Gershuny, J. (2012). Too many zeros: A method for estimating long-term time-use from short diaries. *Annals of Economics and Statistics*, 105/106, 247–270. <https://doi.org/10.2307/23646464>
- Groves, R. M., & Lyberg, L. (2010). Total survey error: Past, present, and future. *Public Opinion Quarterly*, 74(5), 849–879. <https://doi.org/10.1093/poq/nfq065>
- Hadaway, C. K., Marler, P. L., & Chaves, M. (1993). What the polls don't show: A closer look at U.S. church attendance. *American Sociological Review*, 58(6), 741–752. <https://doi.org/10.2307/2095948>
- Kan, M. Y., & Gershuny, J. (2009). Calibrating stylised time estimates using UK diary data. *Social Indicators Research*, 93(1), 239–243. <https://doi.org/10.1007/s11205-008-9365-x>
- Kan, M. Y., & Pudney, S. (2008). Measurement error in stylized and diary data on time use. *Sociological Methodology*, 38(1), 101–132. <https://doi.org/10.1111/j.1467-9531.2008.00197.x>
- Kirchner, A., Belli, R. F., Cordova-Cazar, A. L., & Deal, C. E. (2018). Memory gaps in the American Time Use Survey: Are respondents forgetful or is there more to it? *Survey Research Methods*, 12(3), 231–245. <https://doi.org/10.18148/srm/2018.v12i3.7257>
- Klugman, S. A., Panjer, H. H., & Willmot, G. E. (2019). *Loss models: From data to decisions* (5th ed.). Wiley.

- Knippenberg, H. (2015). Secularization and transformation of religion in post-war Europe. In S. D. Brunn (Ed.), *The changing world religion map: Sacred places, identities, practices and politics: Vol. IV* (pp. 2101–2127). Springer. https://doi.org/10.1007/978-94-017-9376-6_111
- Lechner, F. J. (1996). Secularization in the Netherlands? *Journal for the Scientific Study of Religion*, 35(3), 252–264. <https://doi.org/10.2307/1386556>
- Otterbach, S., & Sousa-Poza, A. (2010). How accurate are German work-time data? A comparison of time-diary reports and stylized estimates. *Social Indicators Research*, 97(3), 325–339. <https://doi.org/10.1007/s11205-009-9504-z>
- Presser, S., & Chaves, M. (2007). Is religious service attendance declining? *Journal for the Scientific Study of Religion*, 46(3), 417–423. <https://doi.org/10.1111/j.1468-5906.2007.00367.x>
- Presser, S., & Stinson, L. (1998). Data collection mode and social desirability bias in self-reported religious attendance. *American Sociological Review*, 63(1), 137–145. <https://doi.org/10.2307/2657486>
- Rossi, M., & Scappini, E. (2012). How should Mass attendance be measured? An Italian case study. *Quality & Quantity*, 46(6), 1897–1916. <https://doi.org/10.1007/s11135-011-9655-2>
- Rossi, M., & Scappini, E. (2014). Church attendance, problems of measurement, and interpreting indicators: A study of religious practice in the United States, 1975–2010. *Journal for the Scientific Study of Religion*, 53(2), 249–267. <https://doi.org/10.1111/jssr.12115>
- Scappini, E. (2010). Daily diaries in time use surveys. A solution to overcome measurement problems in single-activity events with long characteristic rhythms. *Quality & Quantity*, 44(5), 915–939. <https://doi.org/10.1007/s11135-009-9246-7>
- Scappini, E. (2018). Problems in measuring diachronic religious behavior, or using indicators to ‘make a virtue of necessity’: The case of the Netherlands (1975–2005). *Review of Religious Research*, 60(1), 133–151. <https://doi.org/10.1007/s13644-017-0314-5>
- Scappini, E. (2021). Calibrating questionnaires with weekly diaries: An application in religious behavior, Netherlands 1975 to 2005. *Sociological Methodology*, 51(1), 166–187. <https://doi.org/10.1177/0081175020927438>
- Scappini, E. (2025a). *Supplement for: Calibrating items with time use diaries: A refined method* [Contains: Input file to submit to CaSty.2.0.exe program, corresponding output with figures obtained from simulated data, and additional files with figures to support discussion of models]. PsychOpen GOLD. <https://doi.org/10.23668/psycharchives.21260>
- Scappini, E. (2025b). *Supplement for: Calibrating items with time use diaries: A refined method* [Contains: Input file to submit to CaSty.2.0.exe program, corresponding output with the figures obtained from TUS 2008 data]. PsychOpen GOLD. <https://doi.org/10.23668/psycharchives.21261>
- Scappini, E. (2025c). *CaSty - Calibrating Stylized Items, Version 2.0*. [Software]. AMSActa. <https://doi.org/10.6092/unibo/amsacta/8365>
- te Braak, P., van Tienoven, T. P., Minnen, J., & Glorieux, I. (2023). Data quality and recall bias in time-diary research: The effects of prolonged recall periods in self-administered online time-use surveys. *Sociological Methodology*, 53(1), 115–138. <https://doi.org/10.1177/00811750221126499>

- Townshend, J. M., & Duka, T. (2002). Patterns of alcohol drinking in a population of young social drinkers: A comparison of questionnaire and diary measures. *Alcohol and Alcoholism*, 37(2), 187–192. <https://doi.org/10.1093/alcalc/37.2.187>
- TUS. (2008). *Time Use Survey 2008-09*. Italian National Institute for Statistics (ISTAT) [Microdata available on request]. <https://www.istat.it/>
- Vereecken, C. A., & Maes, L. (2003). A Belgian study on the reliability and relative validity of the Health Behaviour in School-Aged Children Food-Frequency Questionnaire. *Public Health Nutrition*, 6(6), 581–588. <https://doi.org/10.1079/PHN2003466>
- Walthery, P., & Gershuny, J. (2019). Improving stylised working time estimates with time diary data: A multi study assessment for the UK. *Social Indicators Research*, 144(3), 1303–1321. <https://doi.org/10.1007/s11205-019-02074-3>

Appendices

Appendix A

Program: CaSty.2.0.exe

Manual: CaSty.2.0.pdf

Web: <https://amsacta.unibo.it/id/eprint/8365>

doi: <https://doi.org/10.6092/unibo/amsacta/8365>

See Scappini (2025c)

Preliminary Descriptions of the Program

The program CaSty.2.0.exe aims to calibrate data from questionnaires with those gathered from diaries. The data needed to obtain useful statistics are relatively few in number and the essential commands very simple. The processing output provides the tables and figures typically needed to present a research report or paper.

Appendix B

Activity codes for religious attendance and related descriptions:

(4321) *Religious practice, services, and prayer in a place of worship*

Definition: attend services, pray in a place of worship, catechism, etc.

Examples: I attend Mass; I attend catechism lessons in preparation to my Confirmation; I pray in a mosque

Notes:

1. Choir singing in a church is coded as 7120.
2. Visiting a church or another place of worship as a tourist is included in 5290.
3. Attending religious ceremonies such as weddings, christenings, etc. is coded as 4323.

Place codes for religious practice and related descriptions:

(38) *Places of religious worship and connected areas* (church, mosque, synagogue, parish recreation center).

Table B1*Descriptive Statistics Diary Data, TUS 2008*

Attendance at Mass	Working day	Saturday	Sunday	Total
Unweighted sample – All				
Yes	97.57	94.29	75.64	89.61
No	2.43	5.71	24.36	10.39
Total	100.00	100.00	100.00	100.00
(N)	(14,787)	(13,286)	(12,871)	(40,944)
Unweighted sample – Age 18–74				
Yes	98.12	95.20	77.41	90.69
No	1.88	4.80	22.59	9.31
(N)	100.00	100.00	100.00	100.00
Total	(11,132)	(9,934)	(9,607)	(30,673)
Weighted sample – Age 18–74				
Yes	98.21	95.31	77.72	94.87
No	1.79	4.69	22.28	5.13
(N)	100.00	100.00	100.00	100.00
Total	(21,930)	(4,371)	(4,372)	(30,673)



Methodology (METH) is the official journal of the European Association of Methodology (EAM).



Leibniz-Institut für
Psychologie

PsychOpen GOLD is a publishing service provided by the Leibniz Institute for Psychology (ZPID), Germany.