Statistical Methods & Applications
https://doi.org/10.1007/s10260-024-00774-w

ORIGINAL PAPER

The use of multiple systems estimation to estimate the number of unattributed paintings by Modigliani

James Edward Jackson[1] · Brian Francis[2]

Accepted: 8 December 2024

Abstract
The number of unattributed paintings by Amedeo Modigliani is estimated, using the method of multiple systems estimation (MSE). Most major artists' works are listed in one catalogue raisonné, but there are five catalogues purporting to list Modigliani paintings. These can be treated as list sources from which MSE can be applied. We obtain estimates by following the classical MSE approach using log-linear models, and compare these with estimates obtained via a Bayesian non-parametric latent class approach. We also consider the impact of fake paintings through sensitivity analyses. Our estimates point to there being around 20–120 unattributed Modigliani paintings.

Keywords Multiple Systems Estimation · Capture-recapture · Amedeo Modigliani · Catalogue Raisonné · Art history

1 Introduction

Amedeo Clemente Modigliani (1884–1920) was an Italian painter and sculptor considered to be one of the pioneers of modern art. His works are greatly sought after, and have sold for over $100 million at auction. For world-renowned artists like Modigliani, there usually exists one *catalogue raisonné* (CR), i.e. an exhaustive list of all *bona fide* works, compiled by a leading authority. This is not the case for Modigliani, where there are at least five CRs purporting to cover his works. These are listed below.

James Jackson and Brian Francis have contributed equally to this work.

✉ James Edward Jackson
jjackson@turing.ac.uk; jejackson96@gmail.com

Brian Francis
b.francis@lancaster.ac.uk

1 The Alan Turing Institute, London NW1 2DB, UK

2 School of Mathematical Sciences, Lancaster University, Lancaster LA1 4YF, UK

Published online: 14 January 2025

- *Modigliani et son oeuvre: Étude critique et catalogue raisonné* by Arthur Pfannstiel (Pfannstiel 1956)
- *Modigliani, 1884–1920: catalogue raisonné: sa vie, son œuvre complet, son art* by Joseph Lanthemann (Lanthemann 1970)
- *I dipinti di Modigliani* by Ambrogio Ceroni (Piccioni and Ceroni 1970)
- *Amedeo Modigliani: catalogo generale* by Osvaldo Patani (Patani 1991)
- *Modigliani catalogue raisonné, Tome II* by Christian Parisot (Parisot 1991).

(Note: there are two further CRs currently being compiled by Marc Restellini and Kenneth Wayne, though these have not yet been published.)

Modigliani is considered to be a problematic painter in terms of attribution of his works. A *Vanity Fair* article (Esterow 2017) reported that Jean Cocteau stated that "[Modigliani] used to hand out his drawings like some gypsy fortuneteller, giving them away". Stewart (2005), writing in the Smithsonian Magazine, reports that Modigliani's landlord took paintings in lieu of rent, and used the paintings to patch old mattresses. Biographies of Modigliani (one by his daughter) (Modigliani 1958; Secrest 2011) state that he was impoverished all of his life, before dying at age 35, and sold his paintings for pittances. Moreover, scientific attribution of his work is in its infancy, with Genty-Vincent et al. (2021) stating that only two technical research projects have been carried out on his work, and his artistic process remains largely unknown. The potential for misattribution of his works is therefore greater than for many artists.

Coupled with this, there is a large number of fakes which exist of Modigliani's works. Marc Restellini, the author of one of the new CRs being compiled, is quoted by Esterow (2017) as saying there are at least 1000 Modigliani fakes in the world. Esterow goes on to credit De Hory as the most prolific Modigliani forger, with de Hory's assistant in possession of over 300 de Horys, many in the style of Modigliani. An art exhibition in Genoa's Ducal Palace exhibited 21 Modiglianis in 2017. The vast majority were declared to be fakes and the exhibition closed early (Serriano 2020). Similarly, three Modigliani heads found in a canal in Livorno, and exhibited, were found to be fakes (Bellandi 2024).

A consequence of the existence of multiple CRs is that the statistical technique of Multiple Systems Estimation (MSE) can be applied to estimate the number of unattributed Modigliani works—specifically, the number of paintings. Also known as capture-recapture, MSE is a method for estimating the size, N, of an unknown population from at least two incomplete lists of the population's subjects. The method has a remarkably rich history that dates back to at least the 17th century. John Graunt used an MSE-type approach to estimate the population of London in 1661 (see Hald 2005); likewise Pierre-Simon Laplace used this type of approach to estimate the population of France in 1802 (see Pollock 1991). In an ecological setting, MSE was used by the Danish biologist C.G. Johannes Petersen—after whom the two-list estimator is named (see Goudie and Goudie 2007)—to estimate fish populations (Petersen 1894).

In recent times, MSE methods have been increasingly used in epidemiological and demographical settings to estimate the size of hard-to-reach human populations (Bird and King 2018), such as war casualties (Manrique-Vallier et al. 2013), drug users (King et al. 2009, 2013a, b), victims of human trafficking and modern slavery (Cruyff et al. 2017; Silverman 2020), deaths in the European refugee and migrant crisis (Farcomeni 2022), prevalence of air- and water-borne diseases (Bhuyan and Chatterjee 2024), as well as census-like population estimates (Dunne and Zhang 2024). There are fewer examples of MSE being used in relation to non-living populations, but these include estimating the number of Italian businesses (Di Cecco et al. 2018) and the number of unused words in Shakespeare's vocabulary (Efron and Thisted 1976); note that this latter example uses a dataset based on frequency of captures (see McCrea and Morgan 2014).

However, there are limitations with MSE, which has resulted in a history of controversy. Fienberg (1972) noted that the standard approach, using log-linear models, "is analogous to, and has the same dangers as, fitting an arbitrary curve to a series of points (x, y), where $x > 0$, with the intention of estimating y at $x = 0$."; Cormack (1999) observed that many applications "give estimates which are not scientifically justified by the underlying data"; and Whitehead et al. (2019) and Binette and Steorts (2022) commented on its "unreliability". Yet the redeeming feature of MSE is that it can provide an insight into the size of a population when there are few alternative options available.

Given how widely forged Modigliani was, the use of MSE in this context brings with it the unique challenge of considering fake artworks. Fakes represent false captures (i.e. units that do not belong to the target population) and their inclusion is an example of misclassification, which ultimately leads to over-coverage. Note that Di Cecco et al. (2018) and Link et al. (2010) used latent class approaches when addressing the problem of over-coverage and misclassification.

With this in mind, our main two research questions are:

1. How reliably can we estimate the number of unattributed Modigliani paintings?
2. How does the presence of fakes affect these estimates?

The remainder of this paper is structured as follows. Section 2 gives an overview of MSE methods, with particular focus on the log-linear modelling approach. Section 3 describes the Modigliani data. Section 4 estimates the number of Modigliani works using log-linear models, and compares these estimates with ones obtained from a Bayesian latent class approach. In Sect. 5 we undertake sensitivity analyses to assess the impact of fake paintings. In Sect. 6 we give some concluding remarks.

2 The method of multiple systems estimation

2.1 The Petersen estimator for two lists

The Petersen estimator, also known as the Lincoln-Petersen estimator, is used to estimate the size of a closed population from two incomplete lists of the target population's subjects, say, lists 1 and 2. Every subject in the population observes one of four inclusion patterns: they are either included in lists 1 and 2, included in list 1 but not in 2, included in 2 but not in 1, or included in neither 1 nor 2. Let n_{11}, n_{10}, n_{01} and n_{00} denote the numbers belonging to these four respective categories, which can be a 2×2 contingency table (see Table 1). The quantity of interest is n_{00}, which is unobserved and often referred to as the "dark figure" (e.g. Silverman 2020).

The Petersen estimate for the size of the total population $\hat{N}$ is given as:

$$\hat{N} = \hat{n}_{00} + n_{01} + n_{10} + n_{11} = \frac{(n_{11} + n_{01})(n_{11} + n_{10})}{n_{11}}. \quad (1)$$

This two-list estimator makes three key assumptions (Manrique-Vallier et al. 2013):

1. Each list is targeting the same closed population.
2. Each subject's list inclusion probabilities are homogeneous (this probability can differ between lists).
3. Each list is independent.

2.2 The use of log-linear models

An equivalent way to obtain the estimate for n_{00}, and hence N, is via log-linear models (Fienberg 1972; Cormack 1989). The log-linear model, which can be represented as a Poisson generalised linear model (GLM), is fit to the three observed counts, n_{11}, n_{10} and n_{01}, i.e.

$$\begin{aligned} n_{ij} &\sim \text{Poisson}(\theta_{ij}), \quad \text{where} \\ \log(\theta_{ij}) &= \mu + \alpha_1 I(i = 1) + \alpha_2 I(j = 1). \end{aligned} \quad (2)$$

The parameter μ is the intercept and α_1 and α_2 are the main effects corresponding to lists 1 and 2, respectively. The dark figure is given as $\hat{n}_{00} = \exp(\hat{\mu})$, where $\hat{\mu}$ is the maximum likelihood estimate of μ.

The use of log-linear models becomes more efficient when there are more than two lists and closed form expressions such as that in (1) are less convenient. When there are

Table 1 The two-list case as a contingency table

List 1	List 2	
	Included	Not Included
Included	n_{11}	n_{10}
Not Included	n_{01}	n_{00}

$K \geq 3$ lists, assumptions (2) and (3) can, to a certain extent, be relaxed. The inclusion of interaction terms introduces some heterogeneity among inclusion probabilities, i.e. subjects' inclusion probabilities can depend on their previous inclusion history. Nevertheless, this approach still assumes that subjects have the same probability of observing a particular inclusion pattern.

When there are $K = 5$ lists, for example, the data can be expressed as a contingency table with $2^K = 32$ cells, where all but one of the cells are observed. Following Silverman (2020), suppose the lists are labelled 1, 2, 3, 4 and 5. We consider each subset A of $\{1, 2, 3, 4, 5\}$, assuming that the elements in A correspond to a particular capture pattern, e.g. $\{1, 5\}$ represents inclusion on lists 1 and 5 and exclusion from lists 2, 3 and 4. There are 32 subsets which correspond to the 32 possible capture patterns.

Unlike in the two-list case, there is now a sufficient number of degrees of freedom to include interactions in the model. The final model typically lies somewhere between the independence model, which includes all main effects but no interactions, and the maximal model, which includes all possible interactions excluding the highest order interaction. For the cell count corresponding to A, n_A, these two models are:

Independence model:

$$\begin{aligned} n_A &\sim \text{Poisson}(\theta_A), \quad \text{where} \\ \log(\theta_A) &= \mu + \sum_{r \in A} \alpha_r. \end{aligned} \tag{3}$$

Maximal model:

$$\begin{aligned} n_A &\sim \text{Poisson}(\theta_A), \quad \text{where} \\ \log(\theta_A) &= \mu + \sum_{r \in A} \alpha_r + \sum_{\substack{r,s \in A \\ r<s}} \beta_{rs} + \sum_{\substack{r,s,t \in A \\ r<s<t}} \gamma_{rst} + \sum_{\substack{r,s,t,u \in A \\ r<s<t<u}} \delta_{rstu}. \end{aligned} \tag{4}$$

The parameter μ is the intercept, $\{\alpha_r\}$ are the main effects, $\{\beta_{rs}\}$ are the two-way interactions, $\{\gamma_{rst}\}$ are the three-way interactions and $\{\delta_{rstu}\}$ are the four-way interactions. The dark figure is still given as $\exp(\mu)$.

Model selection is often based on some information criterion such as the Akaike Information Criterion (AIC) (Akaike 1974) or the Bayesian Information Criterion (BIC) (Schwarz 1978). Preferably, estimates are computed over a range—or even all—possible models. The maximal model typically results in estimates with a large amount of associated uncertainty, while the independence model is often too restrictive. Note, however, that all models, including the maximal model, make an identifying assumption of no-highest-order-interaction (NHOI) (Aleshin-Guendel et al. 2024).

2.3 Review of alternative MSE approaches

Over time, alternative MSE methods have developed. Bayesian methods (see Brooks et al. 2000; King et al. 2009) have been introduced, which provide scope for model averaging. For example, Madigan and York (1997) introduced a Bayesian approach which uses decomposable graphical models, and which can be implemented via the R package **dga** (Johndrow et al. 2021). Moreover, the Bayesian non-parametric latent class model (NPLCM) (Dunson and Xing 2009) can be used for MSE (Manrique-Vallier 2016). This model is a Dirichlet process mixture of product-Bernoulli distributions, described through a Bayesian hierarchical model, and which uses stick breaking priors. This approach also does not require the use of model selection; nor does the user have to make the non-trivial decision of finding an appropriate number latent classes. It can be implemented in the R package **LCMCR** (Manrique-Vallier 2021).

Models accounting for individual heterogeneity are a key branch of MSE methods, many of which evolved from ecological applications. Otis et al. (1978) introduced the model class M_h, which incorporates unobserved heterogeneity and covers an array of mixture models, including: finite binomial mixture models (Norris and Pollock 1996; Pledger 2000), infinite beta-binomial mixture models (Burnham and Rexstad 1993), normal-logistic-binomial models (Coull and Agresti 1999) and a mixture of binomial and beta-binomial models (Morgan and Ridout 2008) (see McCrea and Morgan (2014) for a summary of these approaches).

Recent areas of research also include dealing with: record linkage errors in MSE datasets (Liseo and Tancredi 2011; Tancredi and Liseo 2011; Tancredi et al. 2020), sparse datasets (Chan et al. 2021) and bootstrapping methods for computing MSE estimates' standard errors (Silverman et al. 2024).

3 The Modigliani data

We examine the five CRs listed in the introduction that purport to list Modigliani's works. We limit ourselves to estimating the number of Modigliani paintings as, unlike drawings and sculptures, these are covered by all five CRs. The data needed to be cross-classified, which involved establishing in which CRs each painting appears. Although incomplete, the Secret Modigliani (2023) website has cross-classified a sizeable number of Modigliani's paintings, while also helpfully providing catalogue reference numbers. This cross-classification was completed by physically going through the CRs and obtaining the required information.

There were several challenges when undertaking the cross-classification. Firstly, the CR by Pfannstiel (A) includes a description but not an image for every work, making it difficult to match paintings to descriptions. Secondly, it is not always straightforward to distinguish between a painting and a drawing, e.g. some works classified as a painting by Patani (D) are classified as a drawing by Parisot (E). The upshot is that there is the potential for human errors within the dataset.

We have some reservations about the quality of the Pfannstiel and Lanthemann lists. In addition to the lack of images in the Pfannstiel CR, the Lanthemann CR includes 64 paintings that are not included on any of the other CRs. These are also the oldest two CRs. We therefore consider versions of the data where these CRs are discounted in turn. That is, we consider (i) the full 5-list data, (ii) 4-list data with

Table 2 The full 5-list dataset

Pfannstiel	Lanthemann	Ceroni	Parisot	Patani	Count
*					23
	*				64
*	*				34
		*			1
*		*			2
	*	*			3
*	*	*			3
			*		5
*			*		0
	*		*		3
*	*		*		5
		*	*		14
*		*	*		5
	*	*	*		33
*	*	*	*		54
				*	10
*				*	0
	*			*	5
*	*			*	1
		*		*	0
*		*		*	0
	*	*		*	0
*	*	*		*	6
			*	*	2
*			*	*	0
	*		*	*	4
*	*		*	*	0
		*	*	*	2
*		*	*	*	0
	*	*	*	*	28
*	*	*	*	*	181
Total					488

The character * denotes inclusion, e.g. the top row gives the number of paintings (23) included in Pfannstiel only

Table 3 For the models with the lowest AIC, the estimates obtained when the standard log-linear model approach is used

CRs	Model with lowest AIC	Est. for dark figure	95% C.I.
5-list data	[*ABC*, *ABD*, *ABE*, *ACD*, *ADE*, *BCE*]	648	(82, 1702)
4-list data, exc. Pfann	[*BCD*, *CE*, *DE*]	56	(0, 119)
4-list data, exc. Lanth	[*ACD*, *CE*, *DE*]	82	(0, 168)

Table 4 For the models with the lowest BIC, the estimates obtained when the standard log-linear model approach is used

CRs	Model with lowest BIC	Est. for dark figure	95% C.I.
5-list data	[*ABD*, *ACD*, *BC*, *BE*, *DE*]	111	(28, 195)
4-list data, exc. Pfann	[*BCD*, *CE*, *DE*]	56	(0, 119)
4-list data, exc. Lanth	[*ACD*, *CE*, *DE*]	82	(0, 168)

Pfannstiel omitted and (iii) 4-list data with Lanthemann omitted. (The other two versions can be obtained by removing/combining relevant rows.)

The number of paintings included in each CR are: Pfannstiel, 314; Lanthemann, 424; Ceroni, 332; Parisot, 239; and Patani, 336. These totals differ slightly from the totals quoted in the CRs themselves, since we consider only paintings. The number of unique paintings is 488. The full 5-list dataset is given in Table 2.

4 Estimating the number of Modigliani paintings

4.1 Log-linear models

We first obtain estimates via the standard MSE approach of using log-linear models. Hierarchical log-linear models can be described using the [·] notation, e.g. [*ACD*, *CE*, *DE*] refers to the model with the three-way interaction A:C:D, the two-way interactions A:C, A:D, C:D, C:E, D:E and all main effects (whenever a three-way interaction is included, all two-way interactions relating to those variables are also included).

Rather than using some model selection routine (e.g. forward or backward selection), we compare estimates over all possible hierarchical log-linear models up to order 3 (i.e. we do not consider four-way interactions). This can be carried out via the `closedpMS` routine in the **Rcapture** package (Baillargeon et al. 2007; Rivest and Baillargeon 2022) in **R**. Log-linear models can suffer from issues of non-existence (Fienberg et al. 2012); while this specific application is not particularly hampered by such issues, some of the models that include higher-order interactions do trigger warning messages, hence the exclusion of four-way interactions.

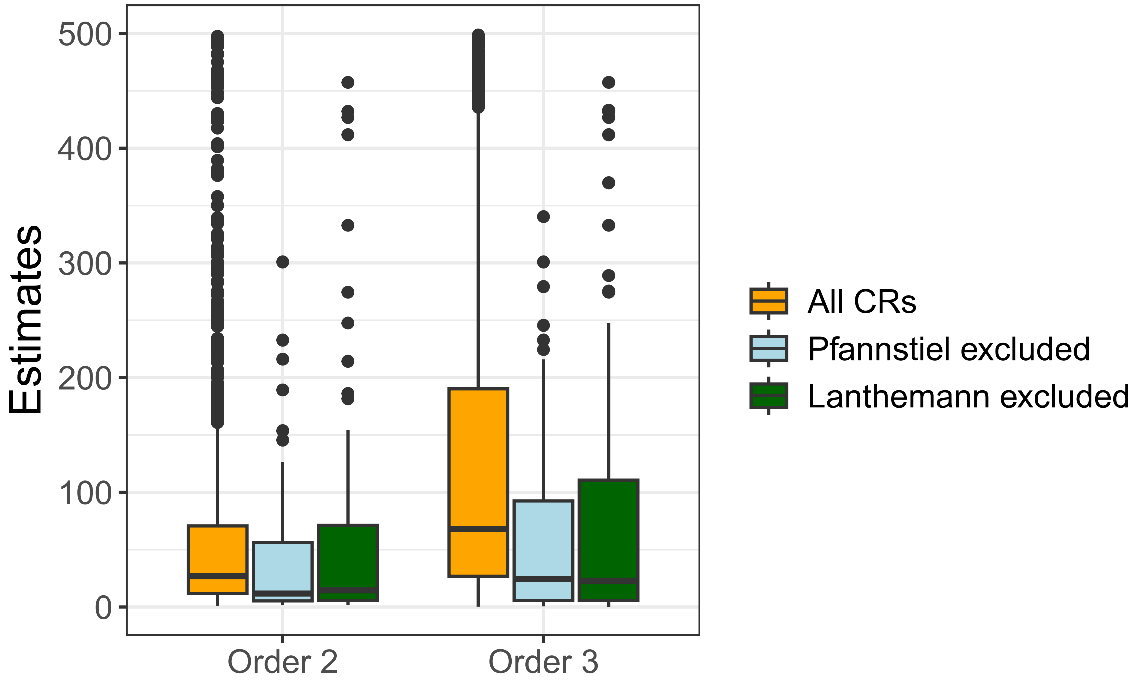

Fig. 1 Estimates for the number of unattributed Modigliani paintings estimated from all hierarchical log-linear models models of order 2 (left) and order 3 (right). We also look at three different versions of the Modigliani data

The estimates for the "dark figure", i.e. the number of unattributed Modigliani paintings, are given in the boxplots in Fig. 1 (note, these boxplots, as with the boxplots given later in the paper, have been truncated above to improve readability). We distinguish between estimates obtained from models where only two-way interactions are considered (order 2) and where two- and three-way interactions are considered (order 3). For the three versions of the data, all median estimates are below 80. There is little to separate the results using the versions of the data where the Pfannstiel and Lanthemann CRs omitted (blue and green boxplots, respectively), but the median and upper quartile are noticeably larger when considering the full 5-list version (orange boxplots), especially when considering three-way interactions, e.g. the upper quartile nearly reaches 200.

In Fig. 2 we plot estimates (for the number of unattributed paintings) against AIC (top plot) and BIC (bottom plot). Interestingly, in general it seems that the models with lower AIC and BIC values produce larger estimates. In Tables 3 and 4, we explicitly give the models (i.e. the interactions included) for the models with the lowest AIC and BIC values and their corresponding estimates. We also give the corresponding Wald confidence intervals for these estimates; MSE confidence intervals should be taken lightly, especially when model selection has taken place (Regal and Hook 1991; Whitehead et al. 2019). The estimates obtained from the models with the lowest AIC and BIC values (when using the full 5-list data)—which are 648 and 111, respectively—highlight one of the problems with MSE: estimates can differ wildly with hardly any difference in model-fit.

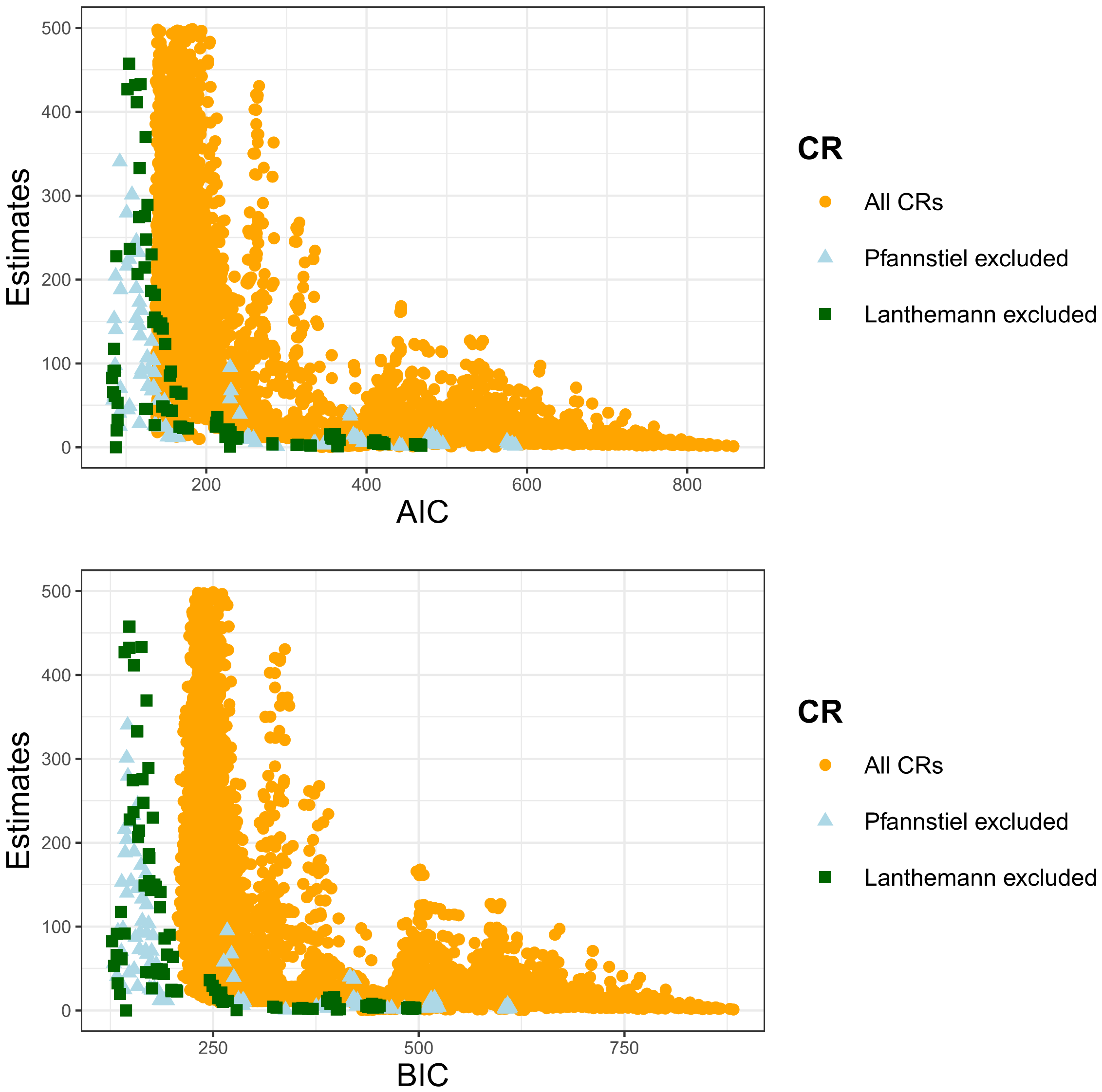

Fig. 2 The estimates for the number of unattributed Modigliani paintings plotted against AIC (top plot) and BIC (bottom plot)

Table 5 The median estimates for the number of unattributed Modigliani paintings (the dark figure) and their 95% credible intervals when the NPLCM approach is used

CRs	Post. Median Est. for dark figure	95% Cred. I.
5-list data	23	(47, 85)
4-list data, exc. Pfann	70	(46, 108)
4-list data, exc. Lanth	111	(87, 149)

4.2 Bayesian non-parametric latent class model

We also obtain estimates via the Bayesian NPLCM approach. We use the `lcmCR` function in the **LCMCR** R package (Manrique-Vallier 2021). We use the default settings, except for `thinning` increased to 100, e.g. we use the default settings for the hyperparameters of the stick-breaking process (`a_alpha=b_alpha`=0.25). The results are given in Table 5. These posterior median estimates range from 23 when considering the full 5-list version of the data, to 111 when considering the version of the data that excludes Lanthemann.

It is worth noting that all log-linear models, even when the maximal model is used, are making some kind of identifying assumption. For example, suppose that the all two-way interaction log-linear model is used. As the interaction terms in this model can be interpreted as conditional log-odds ratios, the model is assuming that these odds ratios are unaffected by a third variable. With this NPLCM approach, no such assumption is made, and hence the fact that the results for the 4-list versions of the data are broadly similar from the two approaches provides some reassurance.

5 Sensitivity analysis to assess the impact of fake paintings

Up to now we have implicitly made a key assumption: that all paintings included in each CR are genuine works by Modigliani, i.e. fakes have not been considered. We can assess the impact of fake paintings through sensitivity analyses. We construct various scenarios as to which paintings—and what proportion of paintings—could be fakes. In Scenario 1, we suppose that all paintings have the same probability of being a fake. In Scenario 2, we suppose that paintings appearing in only one CR could be a fake. In Scenario 3, we suppose that paintings appearing in the Lanthemann CR only could be fake.

5.1 Scenario 1: constant probabilities of being fake

In this first scenario, we suppose that any painting (i.e. a painting appearing in at least one of the CRs) could be a fake. From a practical perspective, a fake is a painting that should be removed from the dataset. Essentially, in this first scenario we randomly remove a proportion, τ_1, of paintings from the dataset. We consider five values of τ_1: 0, 0.05, 0.1, 0.25 and 0.5.

As previous, we then fit the log-linear MSE models to the datasets with these fakes removed and assess how this affects the resulting estimates for the number of unattributed Modigliani paintings. We consider all possible models of order 3.

The results are given in Fig. 3. Focusing first on the "All CRs" version of the data (orange boxplots), there is a downward trend in the estimates as τ_1 increases. For example, the median estimate goes down from roughly 70 when $\tau_1 = 0$ to 20 when $\tau_1 = 0.5$. The results from the versions of the data with Pfannstiel and Lanthemann excluded are somewhat surprising (light blue and green boxplots, respectively): the median, lower quartile and upper quartile estimates take almost a negative-parabola shape, i.e. they initially increase as τ_1 increases before beginning to fall again.

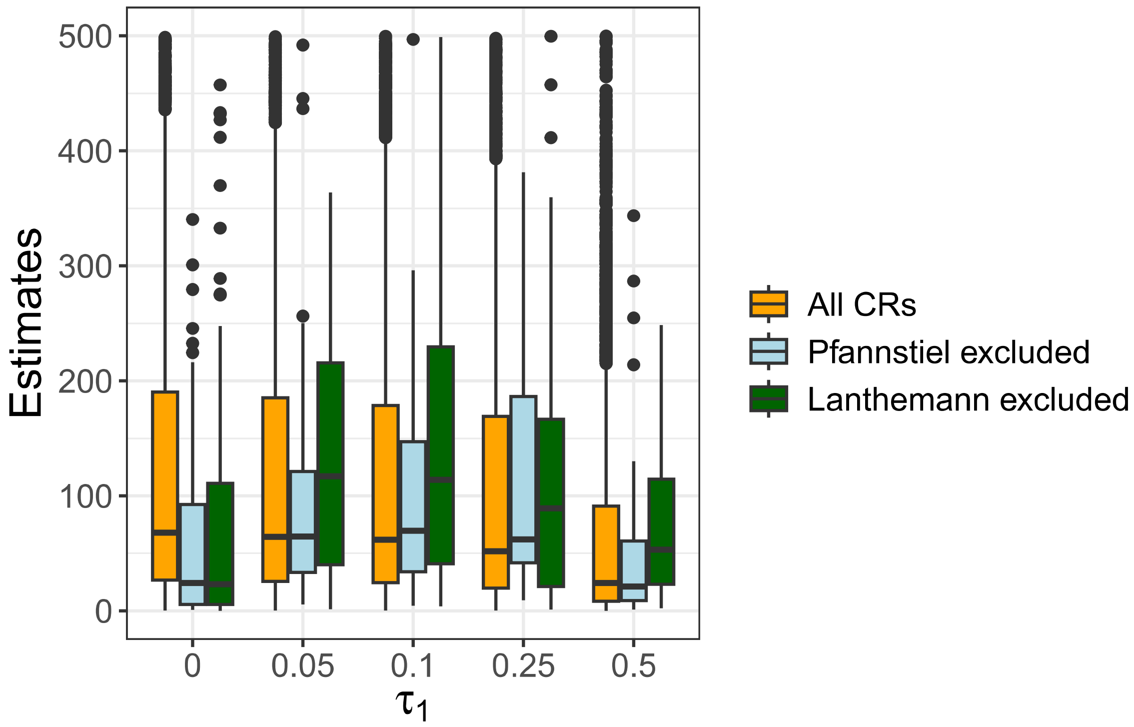

Fig. 3 Scenario 1: The effect of fakes on estimates when we suppose any painting could be a fake

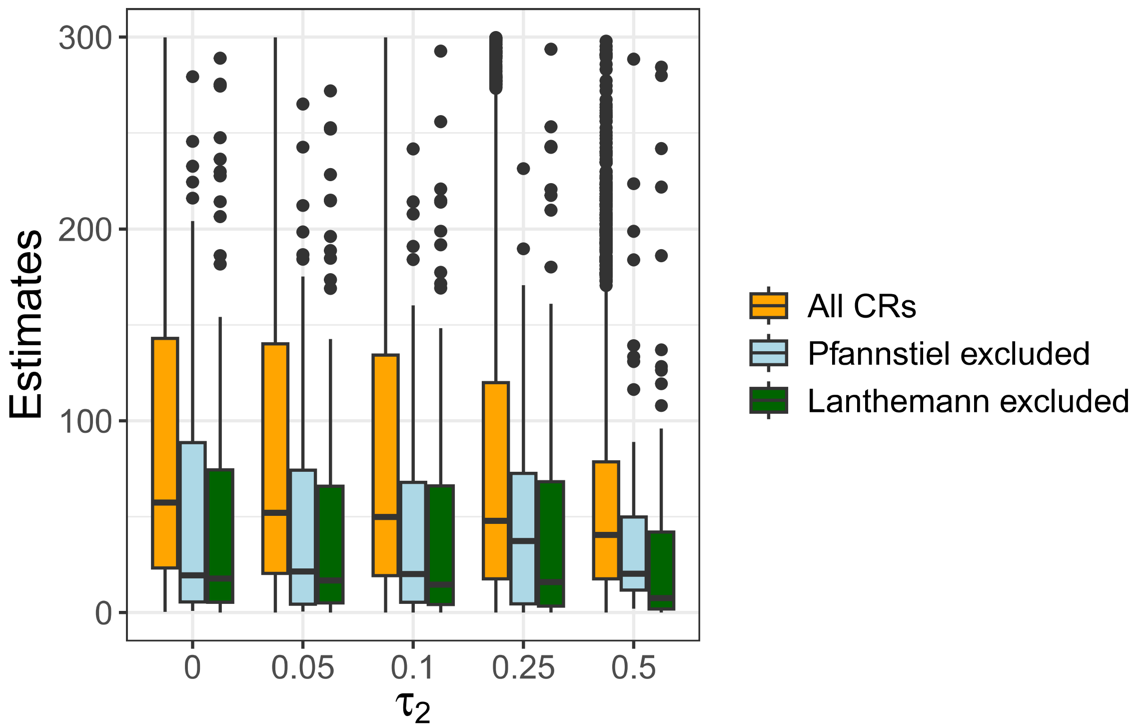

Fig. 4 Scenario 2: The effect of fakes on the estimates when only paintings appearing in one of the CRs have a non-zero probability of being fake

5.2 Scenario 2: only fakes when a painting appears in 1 CR

In this second scenario, we suppose that any painting that appears in two or more CRs is always genuine (i.e. zero probability of being a fake) and that only paintings appearing in one (and only one) CR have a non-zero probability of being a fake. In this instance, τ_2 gives the proportion of fakes among paintings appearing in just one CR.

The results are given in Fig. 4. In general, the estimates show a downward trend as τ_2 increases. For example, considering the "All CRs" case, there is a slight but steady fall in the median estimate as τ_2 increases and a more noticeable fall in the upper quartile estimate.

5.3 Scenario 3: the effect of fakes in Lanthemann

In this third scenario, we suppose that paintings appearing in Lanthemann only (of which there are 64) could be fake. Hence, in this instance, τ_3 gives the proportion of fakes among paintings appearing in Lanthemann only. The version of the CRs where Lanthemann is omitted is not relevant in this scenario. The results are given in Fig. 5. For both versions of the data, there is hardly any change in the estimates as τ_3 increases.

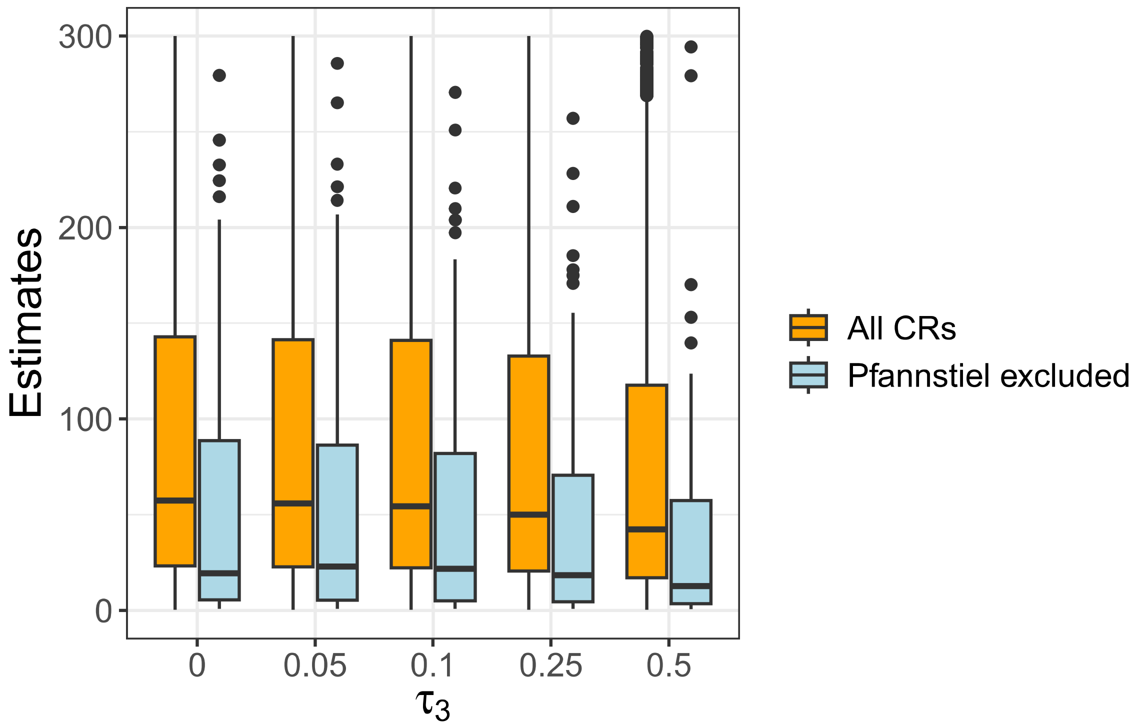

Fig. 5 Scenario 3: The effect of fakes on the estimate for the number of unattributed paintings when paintings appearing in just one of the CRs have a probability of τ of being fake

5.4 Summary of sensitivity analyses

The results from these sensitivity analyses, in a sense, provide reassurance to the possible effect of fakes and hence the validity of the estimates obtained in Sect. 4. Particularly in Scenarios 2 and 3, there is no nasty impact on estimates when fakes are introduced. In fact, in these scenarios, making an assumption of no fakes is essentially providing an upper bound on the number of unattributed paintings, i.e. estimates for the dark figure tend to reduce as the number of fakes increases. Scenario 1 is a little more complicated and is perhaps an area for further research. It clearly shows that estimates increase—at least, initially—as fakes are assumed. Yet, it is arguably also the least likeliest of the three scenarios considered, as fakes are arguably less likely among paintings that have been verified by multiple experts.

6 Discussion

In this paper we have focused on the use of MSE to obtain estimates for the number of unattributed Modigliani paintings, for which a figure of between 20 and 120 seems reasonable. This estimate could potentially be further improved through the inclusion of covariate information, e.g. the year of the painting, its provenance, etc, which would allow observed heterogeneity models to be fit (McCrea and Morgan 2014).

In specific relation to Modigliani, it is important to be aware of the existence of fakes. The sensitivity analysis results in Sect. 5 show that, in some scenarios, we can almost predict the effect that fake paintings will have on estimates, i.e. the presence of fakes tends to reduce the size of the estimates. The assumption of no fakes in these scenarios, therefore, is effectively leading to an upper bound for the number of paintings.

To conclude, not only is it rare that a problem in art history can be addressed by a statistical approach, it is rare that datasets can be obtained directly from literature. Although there are few instances of a major artist like Modigliani having as many as five CRs, there are instances of artists having more than one CR (e.g. Renoir). There is scope, therefore, to apply this same method to other artists' CRs and hence obtain similar estimates for the number of unattributed works.

Acknowledgements We would like to thank Kathryn and Francesca Jackson (the first author's mother and sister) for their help in building the Modigliani dataset. We would also like to thank the University of Leeds's Brotherton Library, who hold three of the five Modigliani catalogues raisonnés in their collection. Thank you to Niels Hagenbuch for his help with the proof reading. Finally, thank you to Francisco Garcia, the author of the Secret Modigliani website, whose listing and cross-referencing of the Modigliani catalogues raisonnés (and willingness to forego copyright) provided the data for this analysis.

Funding There is no funding to report.

Data availability The data is publicly available at https://doi.org/10.17635/lancaster/researchdata/696.

Declarations

Conflict of interest The authors declare no conflict of interest.

References

Akaike H (1974) A new look at the statistical model identification. IEEE Trans Autom Control 19(6):716–723. https://doi.org/10.1109/TAC.1974.1100705

Aleshin-Guendel S, Sadinle M, Wakefield J (2024) The central role of the identifying assumption in population size estimation. Biometrics. https://doi.org/10.1093/biomtc/ujad028

Baillargeon S, Rivest L-P et al (2007) Rcapture: loglinear models for capture-recapture in R. J Stat Softw 19(5):1–31. https://doi.org/10.18637/jss.v019.i05

Bellandi M (2024) Modigliani 1909: the 1984 Hoax. http://www.modigliani1909.com/la_beffa_del_1984.html

Bhuyan P, Chatterjee K (2024) Estimation of population size with heterogeneous catchability and behavioural dependence: applications to air- and water-borne disease surveillance. J R Stat Soc Ser A Stat Soc 187(1):110–131. https://doi.org/10.1093/jrsssa/qnad084

Binette O, Steorts RC (2022) On the reliability of multiple systems estimation for the quantification of modern slavery. J R Stat Soc Ser A Stat Soc 185(2):640–676. https://doi.org/10.1111/rssa.12803

Bird SM, King R (2018) Multiple systems estimation (or capture-recapture estimation) to inform public policy. Annu Rev Stat Appl 5:95–118. https://doi.org/10.1146/annurev-statistics-031017-100641

Brooks S, Catchpole E, Morgan B (2000) Bayesian animal survival estimation. Stat Sci 15(4):357–376. https://doi.org/10.1214/ss/1009213003

Burnham KP, Rexstad EA (1993) Modeling heterogeneity in survival rates of banded waterfowl. Biometrics 49(4):1194–1208. https://doi.org/10.2307/2532261

Chan L, Silverman BW, Vincent K (2021) Multiple systems estimation for sparse capture data: inferential challenges when there are nonoverlapping lists. J Am Stat Assoc 116(535):1297–1306. https://doi.org/10.1080/01621459.2019.1708748

Cormack RM (1989) Log-linear models for capture-recapture. Biometrics 45(2):395–413. https://doi.org/10.2307/2531485

Cormack RM (1999) Problems with using capture-recapture in epidemiology: an example of a measles epidemic. J Clin Epidemiol 52(10):909–914. https://doi.org/10.1016/s0895-4356(99)00058-x

Coull BA, Agresti A (1999) The use of mixed logit models to reflect heterogeneity in capture-recapture studies. Biometrics 55(1):294–301. https://doi.org/10.1111/j.0006-341X.1999.00294.x

Cruyff M, Van Dijk J, van der Heijden PG (2017) The challenge of counting victims of human trafficking: Not on the record: a multiple systems estimation of the numbers of human trafficking victims in the netherlands in 2010–2015 by year, age, gender, and type of exploitation. Chance 30(3):41–49. https://doi.org/10.1080/09332480.2017.1383113

Di Cecco D, Zio MD, Filipponi D, Rocchetti I (2018) Population size estimation using multiple incomplete lists with overcoverage. J Offl Stat 34(2):557–572. https://doi.org/10.2478/jos-2018-0026

Dunne J, Zhang L-C (2024) A system of population estimates compiled from administrative data only (with discussion). J R Stat Soc Ser A Stat Soc 187(1):3–21. https://doi.org/10.1093/jrsssa/qnad065

Dunson DB, Xing C (2009) Nonparametric Bayes modeling of multivariate categorical data. J Am Stat Assoc 104(487):1042–1051. https://doi.org/10.1198/jasa.2009.tm08439

Efron B, Thisted R (1976) Estimating the number of unseen species: How many words did Shakespeare know? Biometrika 63(3):435–447. https://doi.org/10.1093/biomet/63.3.435

Esterow M (2017) The art market's Modigliani forgery epidemic. https://www.vanityfair.com/style/2017/05/worlds-most-faked-artists-amedeo-modigliani-picasso

Farcomeni A (2022) How many refugees and migrants died trying to reach Europe? Joint population size and total estimation. The Ann Appl Stat 16(4):2339–2351. https://doi.org/10.1214/21-AOAS1593

Fienberg SE (1972) The multiple recapture census for closed populations and incomplete 2^k contingency tables. Biometrika 59(3):591–603. https://doi.org/10.1093/biomet/59.3.591

Fienberg SE, Rinaldo A et al (2012) Maximum likelihood estimation in log-linear models. Ann Stat 40(2):996–1023. https://doi.org/10.1214/12-AOS986

Genty-Vincent A, Laval E, Senot M-A, Menu M (2021) Modigliani's studio practice revealed by ma-xrf and non-invasive spectral imaging techniques. X-Ray Spectrom 50(4):375–383. https://doi.org/10.1002/xrs.3211

Goudie IB, Goudie M (2007) Who captures the marks for the Petersen estimator? J R Stat Soc A Stat Soc 170(3):825–839. https://doi.org/10.1111/j.1467-985X.2007.00479.x

Hald A (2005) A history of probability and statistics and their applications before 1750. John Wiley & Sons, New Jersey

Johndrow J, Lum K, Ball P, Binette O (2021) Package 'dga'. https://cran.r-project.org/web/packages/dga/index.html (R package version 2.0.1)

King R, Bird SM, Overstall AM, Hay G, Hutchinson SJ (2013a). Estimating prevalence of injecting drug users and associated heroin-related death rates in england by using regional data and incorporating prior information. J R Stat Soc Series A: Stat Soc 177(1):209-236, https://doi.org/10.1111/rssa.12011

King R, Bird SM, Overstall A, Hay G, Hutchinson SJ (2013b) Injecting drug users in Scotland, 2006: listing, number, demography, and opiate-related death-rates. Addict Res Theory 21(3):235–246. https://doi.org/10.3109/16066359.2012.706344

King R, Morgan B, Gimenez O, Brooks S (2009) Bayesian analysis for population ecology. CRC Press, Boca Raton, FL

Lanthemann J (1970) Modigliani, 1884–1920: catalogue raisonné: sa vie, son œuvre complet, son art. Gráficas Condal, Barcelona

Link WA, Yoshizaki J, Bailey LL, Pollock KH (2010) Uncovering a latent multinomial: analysis of mark-recapture data with misidentification. Biometrics 66(1):178–185. https://doi.org/10.1111/j.1541-0420.2009.01244.x

Liseo B, Tancredi A (2011) Bayesian estimation of population size via linkage of multivariate normal data sets. J Offl Stat 27(3):491–505

Madigan D, York JC (1997) Bayesian methods for estimation of the size of a closed population. Biometrika 84(1):19–31. https://doi.org/10.1093/biomet/84.1.19

Manrique-Vallier D (2016) Bayesian population size estimation using Dirichlet process mixtures. Biometrics 72(4):1246–1254. https://doi.org/10.1111/biom.12502

Manrique-Vallier D (2021) https://cran.r-project.org/web/packages/LCMCR/index.html (R package version 0.4.13)

Manrique-Vallier D, Price ME, Gohdes A (2013) Multiple systems estimation techniques for estimating casualties in armed conflicts. T Seybolt, B Fischhoff, and J Aronson (Eds.), Counting civilian casualties: an introduction to recording and estimating nonmilitary deaths in conflict (p 77-93). New York: Oxford University Press

McCrea RS, Morgan BJT (2014) Analysis of capture-recapture data. CRC Press, Boca Raton, FL

Modigliani J (1958) Modigliani: man and myth – biography and works of italian painter and sculptor Amedeo Modigliani. Pantianos Classics

Morgan BJ, Ridout M (2008) A new mixture model for capture heterogeneity. J R Stat Soc: Ser C: Appl Stat 57(4):433–446. https://doi.org/10.1111/j.1467-9876.2008.00620.x

Norris JL, Pollock KH (1996) Nonparametric MLE under two closed capture-recapture models with heterogeneity. Biometrics 52(2):639–649. https://doi.org/10.2307/2532902

Otis DL, Burnham KP, White GC, Anderson DR (1978) Statistical inference from capture data on closed animal populations. Wildlife Monogr 62:1–135

Parisot C (1991) Modigliani catalogue raisonné, Tome II. Graphis Arte, Livorno

Patani O (1991) Amedeo Modigliani: catalogo generale. Leonardo, Milano

Petersen CGJ (1894) On the biology of our flat-fishes and on the decrease of our flat-fish fisheries. Report of the Danish Biological Station 4:1893–1894

Pfannstiel A (1956) Modigliani et son œuvre: Étude critique et catalogue raisonné. Bibliotheque Des Arts, Paris

Piccioni L, Ceroni A (1970) I dipinti di modigliani. Rizzoli, Milano

Pledger S (2000) Unified maximum likelihood estimates for closed capture-recapture models using mixtures. Biometrics 56(2):434–442. https://doi.org/10.1111/j.0006-341X.2000.00434.x

Pollock KH (1991) Review papers: modeling capture, recapture, and removal statistics for estimation of demographic parameters for fish and wildlife populations: past, present, and future. J Am Stat Assoc 86(413):225–238. https://doi.org/10.1080/01621459.1991.10475022

Regal RR, Hook EB (1991) The effects of model selection on confidence intervals for the size of a closed population. Stat Med 10(5):717–721. https://doi.org/10.1002/sim.4780100506

Rivest L-P, Baillargeon S (2022) Package 'Rcapture'. https://cran.r-project.org/web/packages/Rcapture/index.html (R package version 1.4-4)

Schwarz G (1978) Estimating the dimension of a model. The Ann Stat 6(2):461–464

Secrest M (2011) Modigliani: A life. Alfred a Knopf Inc, New York

Secret Modigliani (2023) https://www.secretmodigliani.com

Serriano T (2020) https://itsartlaw.org/2020/07/17/remembering-modigliani-italys-ongoing-battle-against-forgery/

Silverman BW (2020) Multiple-systems analysis for the quantification of modern slavery: classical and Bayesian approaches (with discussion). J R Stat Soc Ser A Stat Soc 183(3):691–736. https://doi.org/10.1111/rssa.12505

Silverman BW, Vincent K, Chan L (2024) Bootstrapping multiple systems estimates to account for model selection. Stat Comput 34:44. https://doi.org/10.1007/s11222-023-10346-9

Stewart D (2005) https://www.smithsonianmag.com/arts-culture/modigliani-misunderstood-84411676/

Tancredi A, Liseo B (2011) A hierarchical Bayesian approach to record linkage and population size problems. The Ann Appl Stat 5(2B):1553–1585. https://doi.org/10.1214/10-AOAS447

Tancredi A, Steorts R, Liseo B (2020) A unified framework for de-duplication and population size estimation (with discussion). Bayesian Anal 15(2):633–682. https://doi.org/10.1214/19-BA1146

Whitehead J, Jackson J, Balch A, Francis B (2019) On the unreliability of multiple systems estimation for estimating the number of potential victims of modern slavery in the UK. J Human Traffick 7(1):1–13. https://doi.org/10.1080/23322705.2019.1660952