Title

A log-linear modelling approach to assessing the consistency of ego reports of dyadic outcomes with applications to fertility and sexual partnerships

Journal

Journal of the Royal Statistical Society. Series A: Statistics in Society, 178(2)

ISSN

0964-1998

Authors

Admiraal, R
Handcock, MS

Publication Date

2015-02-01

DOI

10.1111/rssa.12067

Peer reviewed

Journal of the Royal Statistical Society
Statistics in Society
Series A

J. R. Statist. Soc. A (2015)
178, *Part* 2, *pp.* 363–382

A log-linear modelling approach to assessing the consistency of ego reports of dyadic outcomes with applications to fertility and sexual partnerships

Ryan Admiraal

Murdoch University, Perth, Australia

and Mark S. Handcock

University of California at Los Angeles, USA

[Received September 2013. Revised February 2014]

Summary. We propose a log-linear model to assess the consistency of ego reports of dyadic outcomes. We do so specifically in the context where males and females report on shared events, and we demonstrate how inconsistencies can be assessed by using a log-linear model that estimates separate mixing totals for each set of reports. This modelling approach immediately allows us to determine where inconsistencies in reports occur. To demonstrate how our method can be easily implemented for survey data, we apply it to both the 1992 National Health and Social Life Survey and the 2002 National Survey of Family Growth. Our analysis identifies inconsistencies in male and female reports of concurrent partnerships and the number of biological children.

Keywords: Bipartite networks; Demography; Egocentric data; Social network models; Survey sampling

1. Introduction

Research centring on sex-specific fertility and sexual behaviour relies on the accuracy of male and female reports. Unfortunately, it is well documented that discrepancies between male and female reports are common in surveys that query males and females about fertility and sexual partnerships. In this paper, we propose the use of log-linear models to reconcile ego reports of dyadic outcomes.

In the case of fertility, male and female reports theoretically can be tested against birth registries, revealing whether one sex is overreporting or underreporting births or if the differences are due to sampling. Although possible, this approach is almost never employed, especially for large surveys where participants' anonymity is ensured, and nearly all examinations of differences between male and female reports assume that female reports are accurate. Under this condition, some studies have suggested that male reports are of poorer quality than female reports with 'multipartnered fertility' being most responsible for deficiencies in the number of births reported by males (Garfinkel *et al.*, 1998; Rendall *et al.*, 1999; Guzzo and Furstenberg, 2007). This phenomenon encompasses children produced through non-marital unions or previous marital relationships, which account for an increasingly large percentage of all births with

Address for correspondence: Ryan Admiraal, Department of Mathematics and Statistics, Murdoch University, 90 South Street, Murdoch, WA 6150, Australia.
E-mail: R.Admiraal@murdoch.edu.au

 0964–1998/15/178363

non-marital fertility estimated to be 36% at 2004 (Hamilton *et al.*, 2005) and an estimated 8.5% of unmarried American males having been divorced at 2010 (US Census Bureau, 2010).

Studies have shown deficiencies in the number of children from prior unions reported by males in the 1980 Current Population Survey (United States Bureau of the Census, 1984; Cherlin *et al.*, 1983) and deficiencies in the number of children reported by non-resident fathers in the 1987–1988 National Survey of Families and Households (Sweet and Bumpass, 1996) and 1990 Survey of Income and Program Participation (US Census, 2001; Seltzer and Brandreth, 1994; Garfinkel *et al.*, 1998; Sorensen, 1997). Rendall *et al.* (1999) demonstrated similar results for the Panel Study of Income Dynamics (Hill, 1992) and the British Household Panel Survey (Taylor *et al.*, 2010) with estimated total births by fathers for children from non-marital or previous marital relationships being approximately from a third to a half of that estimated from female reports. These deficiencies were believed to be due in part to non-reporting of children (possibly because some males were unaware that they had fathered children) as well as underrepresentation of males relative to females for these groups. Later, Rendall *et al.* (2006) examined age-specific reports by males in the 2002 National Survey of Family Growth (NSFG) (US National Center for Health Statistics, 2002) and found underreporting of births to be most significant for males of younger ages (18–21 years) when assuming that female reports accurately reflect population totals. They also showed greater underreporting of births by black or African-American males than males as a whole.

When considering sexual partnerships, generally it is not possible to determine whether one sex is underreporting or overreporting, and reports can only be compared for consistency. It is well documented that the number of lifetime female partners reported by males in surveys tends to exceed the number of lifetime male partners reported by females (Johnson *et al.*, 1992; Smith, 1992; Morris, 1993; Brown and Sinclair, 1999). Johnson *et al.* (1990), Wellings *et al.* (1990) and Morris (1993) suggested that this discrepancy in reports is primarily driven by those reporting large numbers of lifetime partners, with Morris recommending reducing the timeframe in which people are asked to report on sexual partnerships to reduce reporting bias for those with many partners. The argument is that people have a much more accurate memory of the number of sexual partnerships in which they are currently involved than the number of sexual partners that they have had in their lifetime. Laumann *et al.* (1994) and Lewontin (1995) provided another theory and suggested that discrepancies in male and female reports are the result of a larger social influence in which males feel compelled to inflate the number of sexual partners that they have had, whereas females tend to be more conservative in reporting their numbers of sexual partners. In contrast, Brown and Sinclair (1999) argued that such blatant misreporting is not the cause for such discrepancies, but instead a difference in how males and females produce their estimates is the culprit with females tending to take the approach of enumerating their partnerships and males tending to make rough estimates.

Before using survey data to address questions relating to sex-specific fertility or sexual behaviour, it is important to check such data for reliability by assessing the consistency of male and female reports. Although our focus is not on the causes of differences, the method that we propose not only enables us to assess the consistency of male and female reports but also to highlight where differences lie. Current methods are largely descriptive in nature with inferential approaches still being rather rudimentary.

In this paper, we propose the use of log-linear models to assess the consistency of male and female reports. The approach that we consider will be presented in the context of partnership totals, or ‘mixing’ totals, between males and females of different types, although we note that it can be applied to a variety of situations where multiple contingency tables are to be checked for consistency. This includes assessing the consistency of judges or raters, where cross-classified ratings for each judge or rater constitute a contingency table, or consistency of contingency tables over time. Our results are applied to log-linear models for mixing totals, where mixing totals are assumed to follow a Poisson distribution. We explain how to incorporate sample weights in standard error estimates for parameters and demonstrate our methodology on ongoing, or concurrent, partnership totals derived from the 1992 National Health and Social Life Survey (NHSLS) (Laumann *et al.*, 1992, 1994), showing inconsistencies between male and female reports in the reported number of partnerships between white males and white females. We also consider an application to fertility reports by using the 2002 NSFG and show widespread inconsistencies in male and female reports of children produced with partners born in specified ranges of years.

2. Sampling methods in surveys of dyadic outcomes

In this section, we review survey sampling approaches to collecting ego reports of dyadic outcomes, doing so in the context of heterosexual partnerships. We refer to the two modes of the nodes as $\mathcal{M}$ and $\mathcal{F}$, although the approach applies to general two-mode networks. Suppose that a population of size P consists of M males ($\mathcal{M}$) and F females ($\mathcal{F}$). The subpopulation of males can be partitioned into I different types of sizes $M_1, M_2, \ldots, M_I$, and the subpopulation of females can be partitioned into J different types of sizes $F_1, F_2, \ldots, F_J$.

2.1. *Dyad census or dyad samples*

If a census is carried out for all dyads consisting of a male and a female, partnership totals between males and females of given types can be represented in a 'mixing matrix' $N = \{N_{11}, N_{12}, \ldots, N_{IJ}\}$ given by the table on the left-hand side of Table 1, where N_{ij} denotes the observed number of partnerships between males of type i and females of type j. This mixing matrix is a two-way contingency table for which dyads consisting of a male and a female contribute to a cell only if a partnership exists between the two. Thus, it is a contingency table that conditions on the presence of a partnership and for which males and females can potentially contribute multiple times if they have more than one partner of the opposite sex. If the observed mixing totals are a realization from some underlying stochastic process, then corresponding to this mixing matrix of observed partnerships is a mixing matrix of expected partnership totals $\mu = \{\mu_{11}, \mu_{12}, \ldots, \mu_{IJ}\}$, which is represented by the table on the right-hand side of Table 1.

A dyad census typically is not feasible except for very small populations, although dyad

Table 1. Mixing matrix of observed partnership totals for a dyad census based on I levels of $\mathcal{M}$ and J levels of $\mathcal{F}$ (left), and corresponding mixing matrix of expected partnership totals (right)

$\mathcal{M}$	$\mathcal{F}$					$\mathcal{M}$	$\mathcal{F}$				
	1	2	$\cdots$	J			1	2	$\cdots$	J	
1	N_{11}	N_{12}	$\cdots$	N_{1J}	$N_{1\cdot}$	1	μ_{11}	μ_{12}	$\cdots$	μ_{1J}	$\mu_{1\cdot}$
2	N_{21}	N_{22}	$\cdots$	N_{2J}	$N_{2\cdot}$	2	μ_{21}	μ_{22}	$\cdots$	μ_{2J}	$\mu_{2\cdot}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
I	N_{I1}	N_{I2}	$\cdots$	N_{IJ}	$N_{I\cdot}$	I	μ_{I1}	μ_{I2}	$\cdots$	μ_{IJ}	$\mu_{I\cdot}$
	$N_{\cdot 1}$	$N_{\cdot 2}$	$\cdots$	$N_{\cdot J}$	$N_{\cdot\cdot}$		$\mu_{\cdot 1}$	$\mu_{\cdot 2}$	$\cdots$	$\mu_{\cdot J}$	$\mu_{\cdot\cdot}$

samples have been used to assess the consistency of male and female reports on sexual behaviour (Kinsey *et al.*, 1948; Julian *et al.*, 1992; Seal, 1997; Ochs and Binik, 1999). These consist of couple data, so analysis is restricted to samples where a partnership is known to exist, and, because the focus tends to be on sexual behaviour within the partnership, information about other sexual partners is rarely included. Consequently, such data tend not to be useful in estimating partnership totals for the population or subgroups in the population, so we turn our attention to more commonly employed sampling mechanisms for estimating partnership activity levels and fertility rates. In particular, we consider egocentric samples, where individuals are sampled and report information about their partners.

2.2. Egocentric census or egocentric samples

If an egocentric census is carried out, all males report their partnerships with females of the J different types, and females do likewise for males of the I different types. This produces two separate mixing matrices, $N^{\mathcal{M}} = \{N_{11}^{\mathcal{M}}, N_{12}^{\mathcal{M}}, \ldots, N_{IJ}^{\mathcal{M}}\}$ and $N^{\mathcal{F}} = \{N_{11}^{\mathcal{F}}, N_{12}^{\mathcal{F}}, \ldots, N_{JI}^{\mathcal{F}}\}$, corresponding to male reports and female reports respectively. These mixing matrices can be represented as off-diagonal regions of a larger mixing matrix, which is shown in Table 2. If there is no reporting error, these regions of the larger mixing matrix are symmetric ($N_{ij}^{\mathcal{M}} = N_{ji}^{\mathcal{F}}$), so inconsistencies in reporting can easily be assessed by examining the symmetry of this larger mixing matrix. Again assuming that the observed mixing totals were produced by some underlying stochastic process, let $\mu_{ij}^{\mathcal{M}}$ denote the expected number of partnerships between males of type i and females of type j, as derived from male reports, and let $\mu_{ji}^{\mathcal{F}}$ denote the expected number of partnerships between females of type j and males of type i, as derived from female reports. The consistency of male and female reports can then be assessed by comparing $\mu_{ij}^{\mathcal{M}}$ and $\mu_{ji}^{\mathcal{F}}$.

Rarely would we expect that such a census of all members of a population could be carried out, so we turn our attention to data obtained through egocentric samples. Although mixing totals reported by sampled males and females still take on the general form of Table 2, the population level reports $N^{\mathcal{M}}$ and $N^{\mathcal{F}}$ are replaced by their corresponding sample quantities $n^{\mathcal{M}} = \{n_{11}^{\mathcal{M}}, n_{12}^{\mathcal{M}}, \ldots, n_{IJ}^{\mathcal{M}}\}$ and $n^{\mathcal{F}} = \{n_{11}^{\mathcal{F}}, n_{12}^{\mathcal{F}}, \ldots, n_{JI}^{\mathcal{F}}\}$, based on sample totals $m = \{m_1, m_2, \ldots, m_I\}$

Table 2. Mixing matrix for $\mathcal{M}$ and $\mathcal{F}$ for an egocentric census, stratified on I types for $\mathcal{M}$ and J types for $\mathcal{F}$

	$\mathcal{M}$				$\mathcal{F}$			
	1	2	$\cdots$	I	1	2	$\cdots$	J
$\mathcal{M}$								
1					$N_{11}^{\mathcal{M}}$	$N_{12}^{\mathcal{M}}$	$\cdots$	$N_{1J}^{\mathcal{M}}$
2					$N_{21}^{\mathcal{M}}$	$N_{22}^{\mathcal{M}}$	$\cdots$	$N_{2J}^{\mathcal{M}}$
$\vdots$					$\vdots$	$\vdots$	$\ddots$	$\vdots$
I					$N_{I1}^{\mathcal{M}}$	$N_{I2}^{\mathcal{M}}$	$\cdots$	$N_{IJ}^{\mathcal{M}}$
$\mathcal{F}$								
1	$N_{11}^{\mathcal{F}}$	$N_{12}^{\mathcal{F}}$	$\cdots$	$N_{1I}^{\mathcal{F}}$				
2	$N_{21}^{\mathcal{F}}$	$N_{22}^{\mathcal{F}}$	$\cdots$	$N_{2I}^{\mathcal{F}}$				
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$				
J	$N_{J1}^{\mathcal{F}}$	$N_{J2}^{\mathcal{F}}$	$\cdots$	$N_{JI}^{\mathcal{F}}$				

Table 3. Example sociomatrix of partnerships for males $\mathcal{M}$ and females $\mathcal{F}$†

	$\mathcal{M}$						$\mathcal{F}$					
	1	*2*	*3*	$\cdots$	$M-1$	M	1	2	3	$\cdots$	$F-1$	F
$\mathcal{M}$												
1							1	1	0	$\cdots$	0	0
2							0	0	0		1	0
3							0	0	0		0	0
$\vdots$								$\vdots$		$\ddots$	$\vdots$	
$M-1$							0	0	0	$\cdots$	0	1
M							0	0	1		0	0
$\mathcal{F}$												
1	1	0	0	$\cdots$	0	0						
2	1	0	0		0	0						
3	0	0	0		0	1						
$\vdots$		$\vdots$		$\ddots$	$\vdots$							
$F-1$	0	1	0	$\cdots$	0	0						
F	0	0	0		1	0						

†Egocentric samples are highlighted in black with grey denoting the intersection of egocentric samples.

for males and $f=\{f_1, f_2, \ldots, f_J\}$ for females. Thus, models for expected mixing totals based on survey data must incorporate the sex-specific mixing matrices $n^{\mathcal{M}}$ and $n^{\mathcal{F}}$. With egocentric samples, partners nominated by respondents are likely to fall outside the sample, so the observed mixing totals that are reported by males need not match those given by females.

To make this clear, suppose that the network of partnerships in the population is given by the sociomatrix shown in Table 3. The egocentric sample is represented by the black rows and columns of the sociomatrix, and partnerships between respondents in the egocentric sample fall in the grey cells. Partnerships with people outside the sample fall in the black cells. For samples that are small relative to the population size, we would expect few (if any) partnerships to fall in the intersection of sampled males and females, and, consequently, we would not expect the mixing matrices corresponding to male and female reports to be replicates, even with no reporting error.

A sociomatrix of partnerships like that presented in Table 3 can be condensed into multiway contingency tables for males and females where each dimension of the table corresponds to each of the possible types for the opposite sex and each of these dimensions is stratified by the possible number of partnerships with this type. Such tables can account for the within-person dependence of partners, but, as discussed later, are also sparse when used in the context of fertility and sexual partnerships. Consequently, the standard approach has been to use mixing matrices of the form shown in Table 2. Such an approach makes the simplifying assumption that partnerships are independent and identically distributed for people of the same sex and type. This means that, for all males of type i (whether they have multiple partners or not), each partner has the same probability of being of type j. At the same time, for all males of type l (whether they have multiple partners or not), each partner has the same probability of being of type j, although this probability need not be the same as that for males of type i.

3. Log-linear models for expected mixing totals for egocentric samples

The distribution of $n^{\mathcal{M}}$ and $n^{\mathcal{F}}$ will in part determine appropriate models for expected mixing totals $\mu^{\mathcal{M}}$ and $\mu^{\mathcal{F}}$. Following the generative process that was developed in Morris (1991), we assume that the number of partnership opportunities for individuals is determined by a Poisson process, and the conditional distribution of the number of partnerships given a specific number of opportunities is binomial. Then the mixing totals can be shown to be distributed according to a Poisson distribution, and it is natural to model expected mixing totals by using a log-linear model.

We specifically consider models of the form

$$\begin{aligned} \log(\mu^{\mathcal{M}(S)}) &= X\lambda, \\ \log(\mu^{\mathcal{F}(S)}) &= X\lambda+\gamma \end{aligned} \tag{1}$$

where $\mu^{\mathcal{M}(S)}$ and $\mu^{\mathcal{F}(S)}$ are $IJ\times 1$ vectors representing expected mixing totals corresponding to sample mixing totals $n^{\mathcal{M}}$ and $n^{\mathcal{F}}$, X is an $IJ\times p$ design matrix, λ is a $p\times 1$ vector and γ is an $IJ\times 1$ vector. It should be clear that these separate models for the two mixing matrices can be modelled simultaneously through

$$\log(\mu^{(S)}) = X^*\lambda^* \tag{2}$$

where $\mu^{(S)} = (\mu^{\mathcal{M}(S)}, \mu^{\mathcal{F}(S)})$, $\lambda^* = (\lambda, \gamma)$, and X^* is a $2IJ\times(p+IJ)$ matrix given by

$$X^* = \left(\begin{array}{c|c} X & \mathbf{0} \\ \hline X & \mathbf{I}_{IJ} \end{array}\right) \tag{3}$$

where $\mathbf{I}_{IJ}$ is the identity matrix of size IJ and $\mathbf{0}$ is a zero matrix.

The design matrix X can take a variety of forms, but we shall assume a saturated model using dummy coding. Under such a specification and when considering reported partnerships between males of type i and females of type j, model (1) takes the form

$$\begin{aligned} \log(\mu_{ij}^{\mathcal{M}(S)}) &= \lambda+\lambda_i^{\mathcal{M}}+\lambda_j^{\mathcal{F}}+\lambda_{ij}^{\mathcal{M}F}, \\ \log(\mu_{ji}^{\mathcal{F}(S)}) &= \lambda+\lambda_i^{\mathcal{M}}+\lambda_j^{\mathcal{F}}+\lambda_{ij}^{\mathcal{M}F}+\gamma_{ij} \end{aligned} \tag{4}$$

subject to the constraints

$$\begin{aligned} \lambda_1^{\mathcal{M}} &= 0, \\ \lambda_1^{\mathcal{F}} &= 0, \\ \lambda_{i1}^{\mathcal{M}F} &= 0, \qquad i=1,\ldots,I, \\ \lambda_{1j}^{\mathcal{M}F} &= 0, \qquad j=1,\ldots,J. \end{aligned}$$

Under this parameterization, the λ-parameters are interpretable strictly in terms of male reports, whereas the γ-parameters are interpretable as comparisons of female reports with male reports. Specifically, λ denotes the expected sample partnership total between males and females of type 1 according to male reports, first-order effects ($\lambda^{\mathcal{M}} = \{\lambda_2^{\mathcal{M}}, \lambda_3^{\mathcal{M}}, \ldots, \lambda_I^{\mathcal{M}}\}$ and $\lambda^{\mathcal{F}} = \{\lambda_2^{\mathcal{F}}, \lambda_3^{\mathcal{F}}, \ldots, \lambda_J^{\mathcal{F}}\}$) are interpretable as conditional log-odds within type 1 for each sex for male reports only and second-order effects ($\lambda^{\mathcal{M}F} = \{\lambda_{22}^{\mathcal{M}F}, \lambda_{23}^{\mathcal{M}F}, \ldots, \lambda_{IJ}^{\mathcal{M}F}\}$) are interpretable as log-odds ratios and represent deviations from independence (again, specific to male reports). The parameters $\gamma = \{\gamma_1, \gamma_2, \ldots, \gamma_{IJ}\}$ provide cell-specific comparisons of expected sample mixing totals from female reports with the corresponding expected sample mixing totals from male reports through

$$\gamma_{ij} = \log(\mu_{ji}^{\mathcal{F}(S)} / \mu_{ij}^{\mathcal{M}(S)}).$$

Note that the saturated model that we have proposed produces cell estimates that are no different from those produced by a saturated model for a three-way table where the dimensions of the table correspond to the sex of the respondent, the type of the respondent and the type of the partner. In other words, model (4) produces expected mixing totals that are identical to those of the model given by

$$\log(\mu_{ijk}^{(S)}) = \lambda + \lambda_i^{\mathcal{X}} + \lambda_j^{\mathcal{Y}} + \lambda_k^{\mathcal{Z}} + \lambda_{ij}^{\mathcal{X}Y} + \lambda_{ik}^{\mathcal{X}Z} + \lambda_{jk}^{\mathcal{Y}Z} + \lambda_{ijk}^{\mathcal{X}YZ},$$

where $\mu_{ijk}^{(S)}$ denotes the expected sample total number of partnerships reported by people of sex i and type j with partners of type k, $\mathcal{X}$ denotes the sex of the respondent, $\mathcal{Y}$ denotes the type for the respondent and $\mathcal{Z}$ denotes the type for the partner. The benefit of the parameterization that we propose is that, when our models are represented in terms of expected (population) mixing totals (instead of sample mixing totals), γ provides direct comparisons of corresponding male and female reports for partnerships between males and females of specific types, so these parameters indicate where the inconsistencies in male and female reports occur and the magnitudes of those differences.

4. Incorporating sample weights

To show how this sample level modelling approach relates to a population level modelling approach and to generalize to sample designs other than simple random sampling, suppose that m_i of the M_i males of type i are sampled and f_j of the F_j females of type j are sampled, producing a total sample size of $p = \Sigma_{i=1}^{I} m_i + \Sigma_{j=1}^{J} f_j$. Then corresponding to males $1, 2, \ldots, m_i$ of type i are sample weights $w_{i1}^{\mathcal{M}}, w_{i2}^{\mathcal{M}}, \ldots, w_{im_i}^{\mathcal{M}}$, where the sample weights are the inverse of the inclusion probabilities $\pi_{i1}^{\mathcal{M}}, \pi_{i2}^{\mathcal{M}}, \ldots, \pi_{im_i}^{\mathcal{M}}$. Similarly, females $1, 2, \ldots, f_j$ have sample weights $w_{j1}^{\mathcal{F}}, w_{j2}^{\mathcal{F}}, \ldots, w_{jf_j}^{\mathcal{F}}$. These sample weights are typically normalized to sum to the overall sample size p, so

$$\sum_{i=1}^{I} \sum_{k=1}^{m_i} w_{ik}^{\mathcal{M}} + \sum_{j=1}^{J} \sum_{l=1}^{f_j} w_{jl}^{\mathcal{F}} = p.$$

How these sample weights are incorporated in analyses depends in part on the sample design.

4.1. Clogg and Eliason approach

Historically, the standard approach for incorporating sample weights in log-linear models for contingency tables has been that developed by Clogg and Eliason (1987), which takes the following approach. Suppose that each respondent reports one or fewer partnerships (i.e. contributes to no more than one cell of the mixing matrix). Corresponding to mixing totals $n_{ij}^{\mathcal{M}}$ and $n_{ji}^{\mathcal{F}}$ are mean sample weights $\overline{w}_{ij}^{\mathcal{M}}$ and $\overline{w}_{ji}^{\mathcal{F}}$ respectively, where $\overline{w}_{ij}^{\mathcal{M}}$ is the mean of the sample weights of all males of type i who report a female partner of type j, and $\overline{w}_{ji}^{\mathcal{F}}$ has a similar interpretation for females. Then $\overline{w}_{ij}^{\mathcal{M}} P/p$ gives the population-adjusted aggregate of the sample weights of all males of type i who report partnerships with females of type j, and $\overline{w}_{ji}^{\mathcal{F}} P/p$ gives the corresponding population-adjusted aggregate of the sample weights of all females of type j who report partnerships with males of type i. The expected (population) partnership totals $\mu_{ij}^{\mathcal{M}}$ and $\mu_{ji}^{\mathcal{F}}$ for male and female reports respectively are then given by

$$\mathbb{E}[(\overline{w}_{ij}^{\mathcal{M}} P/p) n_{ij}^{\mathcal{M}}] = \mu_{ij}^{\mathcal{M}},$$
$$\mathbb{E}[(\overline{w}_{ji}^{\mathcal{F}} P/p) n_{ji}^{\mathcal{F}}] = \mu_{ji}^{\mathcal{F}}, \tag{5}$$

so $\overline{w}_{ij}^{\mathcal{M}}$, $\overline{w}_{ji}^{\mathcal{F}}$ and P/p are offsets. This means that models (1), (2) and (4) can all be represented as population level models (i.e. in terms of $\mu^{\mathcal{M}}$, $\mu^{\mathcal{F}}$ and μ instead of $\mu^{\mathcal{M}(S)}$, $\mu^{\mathcal{F}(S)}$ and $\mu^{(S)}$) by incorporating these offsets, allowing for interpretation of parameters at the population level. Note that it can be easily shown that the scalar P/p only influences λ for model (4), so all first- and second-order effects remain unchanged, as does γ. This makes the inclusion of P/p as an offset optional if interest lies solely in comparisons between cells of the mixing matrices.

Extending this to the case where respondents can potentially report multiple partnerships, we recall the simplifying assumption that partnerships are independent and identically distributed for people of the same sex and type. Then sampled male k of type i can contribute to the mixing matrix multiple times, each time with a weight of $w_{ik}^{\mathcal{M}}$. This means that $\overline{w}_{ij}^{\mathcal{M}}$ and $\overline{w}_{ji}^{\mathcal{F}}$ are calculated as the mean of the sample weights corresponding to the partnerships, not people, contributing to a particular cell of the mixing matrix, and expression (5) still holds.

In general, the Clogg and Eliason approach produces unbiased estimates of expected mixing totals but incorrect standard errors, as it treats $\overline{w}_{ij}^{\mathcal{M}}$ and $\overline{w}_{ji}^{\mathcal{F}}$ as fixed (hence, making them offsets) when, in fact, they are random (Loughin and Bilder, 2011). Skinner and Vallet (2010) have shown that using mean cell weights as offsets produces both correct estimates and correct standard errors only when

(a) all sample weights are the same (as in simple random sampling) or
(b) the sample weights are the same for a given cell of the mixing matrix (i.e. there is no within-cell variability of sample weights) and this is true for each cell of the mixing matrix.

Under more complex sampling schemes, failure to account for the within-cell variability of sample weights leads to underestimation of standard errors.

4.2. Skinner and Vallet approach

Suppose that sampling is done with unequal probabilities (according to a Poisson sampling design), and sample weights corresponding to at least one cell of the mixing matrices are not all constant. Then the correct variance–covariance matrix can be obtained by using linearization. Consider model (2), which is simply a unified representation of model (1), and let $\mathbb{V}_{\mathrm{CE}}(\hat{\lambda}^*)$ denote the variance–covariance matrix of $\hat{\lambda}^*$ under the Clogg and Eliason approach. Skinner and Vallet (2010) derived the (large sample) variance–covariance matrix of $\hat{\lambda}^*$ as

$$\mathbb{V}_{\mathrm{SV}}(\hat{\lambda}^*) = \mathbb{V}_{\mathrm{CE}}(\hat{\lambda}^*) + \mathbb{V}_{\mathrm{CE}}(\hat{\lambda}^*) \left(\sum_{i=1}^{2IJ} \mu_i^{(S)} c_i^2 X_i^{*\prime} X_i^* \right) \mathbb{V}_{\mathrm{CE}}(\hat{\lambda}^*) \tag{6}$$

where $\mu_i^{(S)}$ is the ith expected sample mixing total (corresponding to $\mu^{(S)} = (\mu^{\mathcal{M}(S)}, \mu^{\mathcal{F}(S)})$), c_i^2 is the squared coefficient of variation for the sample weights corresponding to the cell of the mixing matrix for μ_i and X_i^* is the ith row of X^* as given in equation (3). Since $\mu^{(S)}$ is unknown, it is replaced with the observed mixing totals $n = (n^{\mathcal{M}}, n^{\mathcal{F}})$, and the square root of the diagonal of the resulting matrix produces the correct standard errors.

If the sample design is more complex, such as a stratified, cluster or multistage sampling, Skinner and Vallet advocated a pseudolikelihood approach based on the census likelihood. In this approach, mean cell weights are used to scale sample mixing totals to (or at least proportional to) population size, and the population level likelihood function (i.e. model (2) but now modelling μ, not $\mu^{(S)}$) is maximized for λ^* according to these weighted sample totals. Correct standard errors can then be obtained by using the jackknife or bootstrap, both of which are dependent on the particular sampling scheme.

5. Application: 1992 National Health and Social Life Survey

As an application of the methods that are developed in this paper, we assess the reports of concurrent sexual partnerships in the NHSLS, specifically examining the consistency of males and females in terms of their reports of partnerships with people of various ethnicities. The NHSLS is a cross-sectional survey of 3432 males and females in the USA aged 18–59 years, and we focus on the 1992 survey. Questions centred on sexual behaviours and attitudes, and detailed information was collected in regard to sexual partnerships, including demographic characteristics of partners (such as age and ethnicity) and when partnerships began and ended. Concurrent partnerships were determined by those partnerships that had not terminated at the date of the interview.

Sampling of individuals for this survey was done by using a stratified, cluster sample with blacks and African-Americans and Hispanics overrepresented relatively to the population as a whole (Laumann *et al.* (1994), appendix B). Variables that would enable us to replicate the design (such as geographic cluster) were not available to preserve respondents' anonymity, however, so we could not incorporate the exact sample design in our calculations of parameter standard errors. As the next best option, we instead treated partnership totals as being obtained through a Poisson sampling design. For such a design, only the mixing matrices, mean cell weights and corresponding within-cell variances are needed to obtain correct standard errors by using equation (6).

Of those who reported partners, one male failed to provide his ethnicity, one male and one female did not report their ages, one male and two females failed to report the sex of their partners, 54 males and 47 females did not report the age of partners and seven males and 12 females failed to report the ethnicity of their partners. In all cases, these observations were removed, and remaining observations were post stratified on the basis of similarity of respondent and nominated partner sex, respondent and nominated partner age (treated as a binary variable denoting whether or not the age was between 18 and 59 years), and respondent and nominated partner ethnicity to that of removed observations. For instance, the one female who failed to report her age was black or African-American and reported a partnership with a black or African-American male with unspecified age. Post-stratification in response to the removal of this observation was applied to all other black or African-American females reporting partnerships with black or African-American males. After post-stratification, one observation was removed because the respondent was outside the prescribed age range of the study, and an additional 164 observations were removed because nominated partners fell outside the prescribed age range. These exclusions were necessary to ensure a closed population better, and the resulting reported and used degree distributions for males and females are shown in Fig. 1.

Because of sparse mixing matrices when considering all ethnicities recorded in the NHSLS, we restricted our analysis to partnerships between ethnicities labelled as white and black or African-American. Male and female reports of concurrent partnerships, stratified by these two ethnicities, are presented in Table 4. For both male and female respondents, the overwhelming majority of reported partnerships are with partners of the same ethnicity as the respondent. Mean sample weights by cell corresponding to these partnership totals, along with corresponding within-cell variances of sample weights, are presented in Table 5 and show higher mean

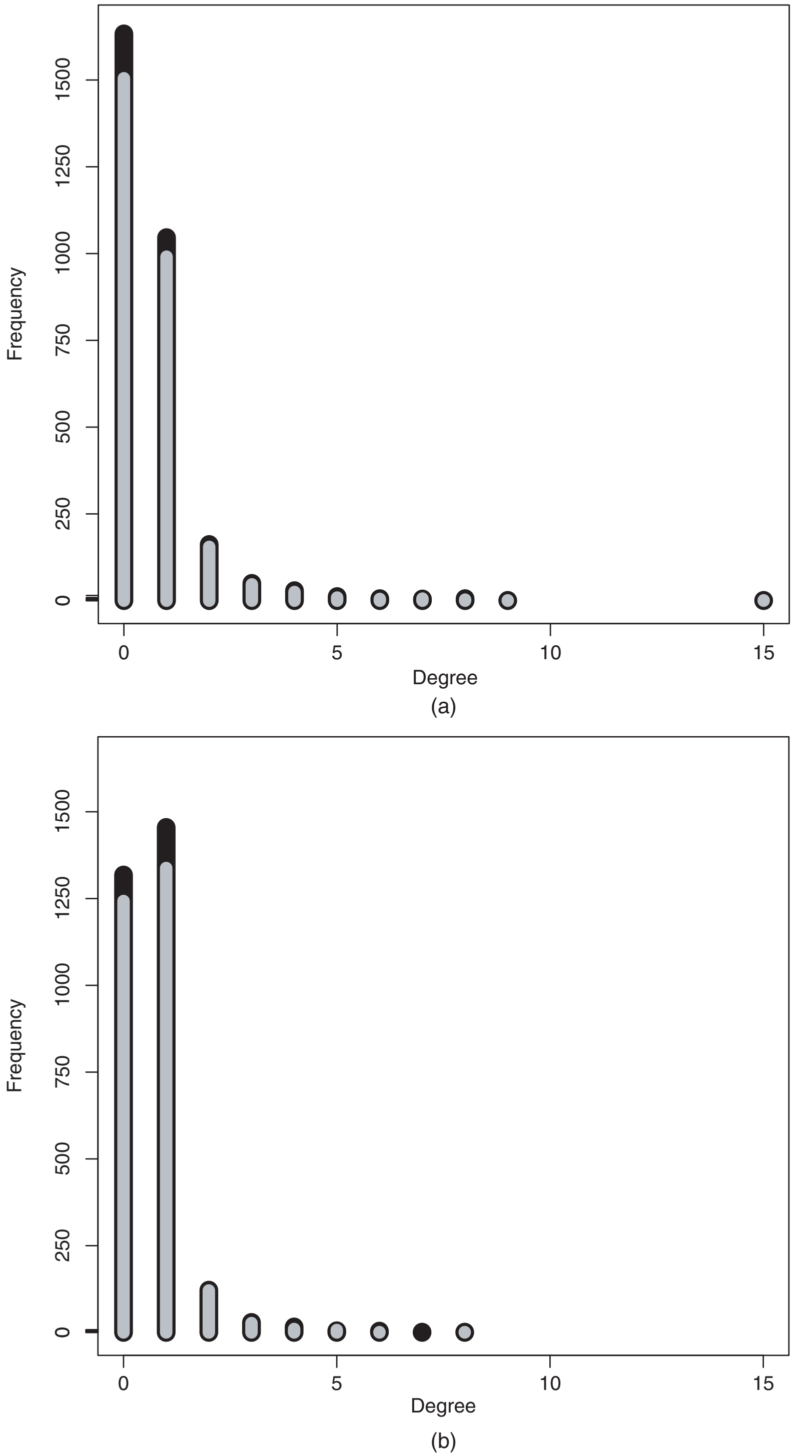


Fig. 1. Reported degree distributions (■) and resulting degree distributions when eliminating reports with incomplete respondent or partner information or invalid ages (▒) for (a) males and (b) females

Table 4. Mixing totals between white and black or African-American males and females, as reported by males (left) and females (right)

Males (respondent)	*Females*		*Females (respondent)*	*Males*	
	White	*Black*		*White*	*Black*
White	1110	6	White	1094	31
Black	26	246	Black	4	306

Table 5. Male reports (left) and female reports (right) of mean sample weights (and corresponding variances) corresponding to mixing totals between white and black or African-American males and females

Males (respondent)	*Females*		*Females (respondent)*	*Males*	
	White	*Black*		*White*	*Black*
White	1.235 (0.437)	1.293 (0.692)	White	1.022 (0.275)	0.935 (0.441)
Black	0.850 (0.241)	0.795 (0.247)	Black	0.716 (0.020)	0.655 (0.188)

sample weights (and corresponding variances) for partnerships reported by white males and females than for partnerships reported by their black or African-American counterparts. The higher mean sample weights are the direct result of the overrepresentation of blacks and African-Americans in the survey. Similarly, mean sample weights (and corresponding variances) for white and black or African-American males are higher than those of their female counterparts, which is indicative of oversampling of females. Recall that these mean sample weights and corresponding variances are based on assigning an individual's sample weight to each partnership reported by that person.

Fitting a log-linear model of the form (4) and adjusting standard errors according to equation (6) by using R (R Core Team, 2013), we obtain the parameter estimates, standard errors and p-values that are given in Table 6. The main effect $\lambda_{\mathrm{B}}^{\mathcal{M}}$ provides a comparison of the expected reported number of partnerships with white females by black or African-American males and the expected reported number of partnerships with white females by white males, whereas $\lambda_{\mathrm{B}}^{\mathcal{F}}$ provides a comparison of the expected reported number of partnerships with black or African-American females by white males and the expected reported number of partnerships with white females by white males. In both cases, we find strong evidence for higher numbers of partnerships between white males and white females on the basis of male reports. The parameter $\lambda^{\mathcal{M}F}$ provides a measure of the tendency for assortative mixing as opposed to disassortative mixing within male reports. The highly significant positive value of this interaction effect provides strong evidence of assortative mixing by ethnicity on the basis of male reports, which is consistent with what we see in the observed mixing matrix from male reports.

Primary interest lies in those parameters that provide a direct comparison of expected mixing totals corresponding to male and female reports, in the bottom part of Table 6. On the basis of the Wald tests for these parameters presented in Table 6, we find clear inconsistencies in white male and female reports of partnership totals with each other (p-value less than 0.001) with white males reporting an estimated 1.226 (1.116, 1.348) times as many partnerships with white females as what white females report with white males. In spite of these inconsistencies in white male and female reports of partnerships with each other, the results are largely in agreement with the

Table 6. Parameter estimates, standard errors and p-values for the model fitting separate mixing totals for male and female reports

Parameter	*Estimate*	*Standard error*	*z-value*	*p-value*
Reference category				
λ	7.223	0.034	212.179	<0.001
Main effects				
$\lambda_{\mathrm{B}}^{\mathcal{M}}$	−4.128	0.229	−18.028	<0.001
$\lambda_{\mathrm{B}}^{\mathcal{F}}$	−5.175	0.487	−10.634	<0.001
Interaction effect				
$\lambda^{\mathcal{MF}}$	7.355	0.542	13.572	<0.001
Cell mean comparisons				
γ_{WW} ($\mu_{\mathrm{WW}}^{\mathcal{F}}$ *versus* $\mu_{\mathrm{WW}}^{\mathcal{M}}$)	−0.204	0.048	−4.246	<0.001
γ_{WB} ($\mu_{\mathrm{WB}}^{\mathcal{F}}$ *versus* $\mu_{\mathrm{BW}}^{\mathcal{M}}$)	0.271	0.316	0.857	0.391
γ_{BW} ($\mu_{\mathrm{BW}}^{\mathcal{F}}$ *versus* $\mu_{\mathrm{WB}}^{\mathcal{M}}$)	−0.996	0.704	−1.415	0.157
γ_{BB} ($\mu_{\mathrm{BB}}^{\mathcal{M}}$ *versus* $\mu_{\mathrm{BB}}^{\mathcal{F}}$)	0.025	0.102	0.242	0.809

theory of Morris (1993) where we would expect male and female reports to be fairly consistent for concurrent partnerships, as all other Wald tests fail to suggest significant differences in male and female reports of partnership totals based on ethnicity.

6. Application: 2002 National Survey of Family Growth

As a second application, we consider fertility reports by males and females in the 2002 NSFG. The NSFG is a cross-sectional survey that has included a total of six cycles between 1973 and 2002 (before being conducted over 5-year spans since 2006). Cycle 6 in 2002 was the first to include both men and women and consisted of 4928 men and 7643 women 15–44 years of age. Participants were queried about biological children, behaviours that may impact on fertility, family structure and a variety of demographic characteristics. Participants also reported information on children and partners.

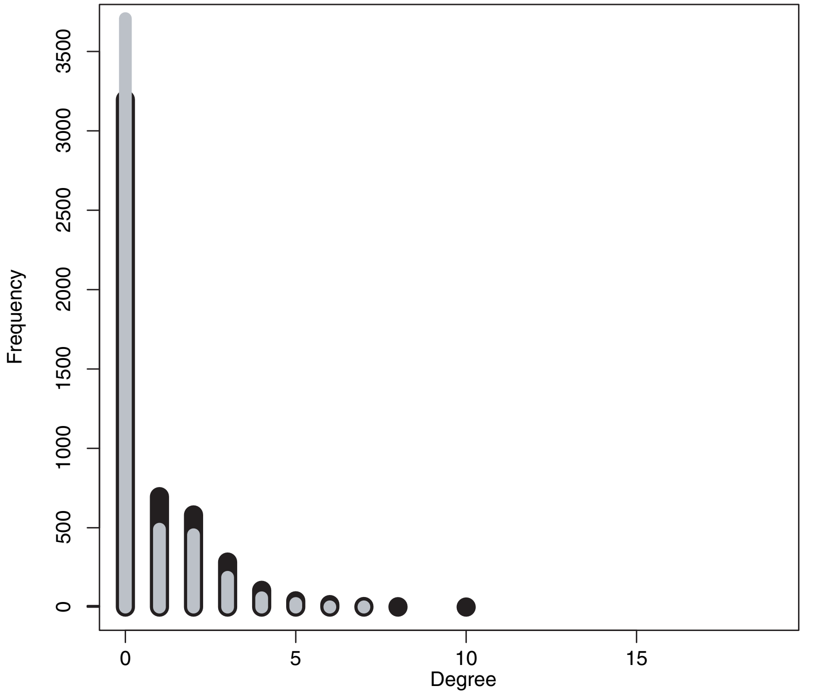
The actual design of the 2002 NSFG was quite involved. A national multistage cluster sample was used to select households from which an individual was randomly selected with probability related to sex, age, race and ethnicity, and household size. Complicating this was that the sample design was not effective in sampling Hispanic households (and, consequently, individuals), so a second multistage cluster sample with some overlap in sampled primary sampling units was obtained to produce greater Hispanic representation. (Full details of the sample design are provided in Lepkowski *et al.* (2006).) As with the NHSLS, not all variables that are needed to replicate the sample design were available, so we resorted to treating birth totals as being obtained through a Poisson sampling design.

Although both men and women were included in the survey, they were subjected to different questionnaires. Men were asked about children in the context of current and previous wives or partners, as this was believed to elicit the most accurate reporting of biological children (National Center for Health Statistics, 2004). However, information was collected for only the current wife or partner, three previous wives or partners, and first wife or partner for each man. Women, in contrast, were asked directly about their children, and information was collected about each of these children as well as the fathers corresponding to them. This difference in questionnaires led to complications with each sex. For men, the list of partners for whom they were queried did not necessarily include all women with whom they had produced a child, so information was missing for some children and their mothers. For women, in contrast, information was collected for each child, but the information that was collected for fathers lacked the level of detail that was achieved in males' reports about wives and partners.

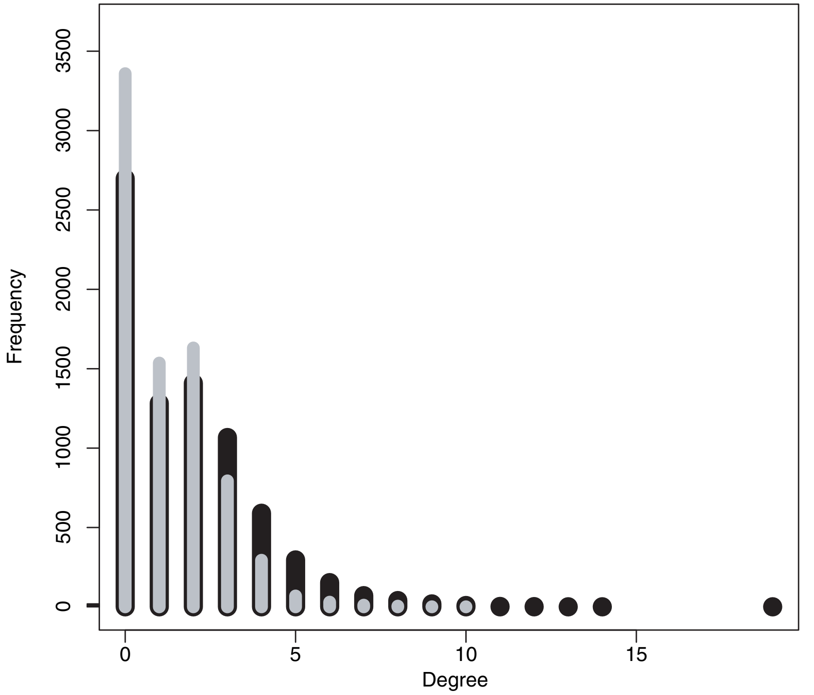
Given the limited information on female partners, we considered the consistency of males and females in terms of reported number of biological children produced with partners born between certain years. In particular, we considered males born between 1957 and 1985 and females born between 1957 and 1986 to ensure a closed population based on year of birth. Whereas birth years of respondents and birth years of mothers as reported by fathers were easily ascertained, the birth years of fathers as reported by mothers required certain assumptions, as females did not report birth years of the father but instead only the father's age at the birth of the child. For a given child, this meant that the birth year of the father could be narrowed down to only one of two years. In addition to this uncertainty, there was no unique indicator for the father. Consequently, it is possible that a woman reporting multiple children could have produced them with different fathers, even if the birth years for the fathers were all the same. Taking into consideration these issues, we first assumed that, for women reporting multiple children, if the possible birth years of the fathers for successive children were within 1 year of each other, then these fathers were the same person. Second, although multiple children reported for the same father could sometimes eliminate one of the two possible birth years, in some cases this simply made one of the two birth years more probable on the basis of the birth months of the children and the known age of the father when those children were born. In these cases (and in all cases where only one child was reported), the birth year of the father was imputed on the basis of a Bernoulli distribution with probabilities corresponding to the probability that the father was born in either of the two possible years.

Once birth years for the father had been determined or imputed, mixing matrices could be constructed for both male and female reports of fertility. These matrices consisted of total numbers of children produced by males and females born within the years under consideration for each sex and for which birth year information could be ascertained for both the mother and the father. For males, only 64.7% of the 1731 males reporting having had any children provided complete birth year information for the mothers of all their children, and 71.6% of the 3492 reported childbirths contained complete information and could be used in our analysis. For females, 53.5% of the 5033 females reporting having had any children provided complete information for the fathers of all of their children, and 66.4% of the 13593 reported childbirths contained complete information and could be used. The actual number of cases used was less than this, however, because we restricted our analyses to males born between 1957 and 1985 and females born between 1957 and 1986. For both male and female reports, weights were post stratified by respondent's birth year to adjust for the missingness of the partner's birth year.

Although we could have restricted our analyses to individuals for whom complete information was available for all children, we opted to use all reported births where complete information for the father and mother was available. The resulting degree distributions (where we now use 'degree' to refer to the number of reported biological children) for males and females based on children for whom complete parent birth year information is available shows a noticeable shift from reported degree distributions in terms of the number of biological children that respondents claimed. These comparisons of the reported degree distributions and degree distributions from births used in computing mixing totals from both male and female reports are presented in Fig. 2. Since only respondents who report children contribute to the mixing totals, only degrees of 1 or higher are of interest, and, not surprisingly, we note that there is substantial missingness in partner information for both males and females reporting large numbers of children.



(a)



(b)

Fig. 2. Reported degree distributions (■) and resulting degree distributions when eliminating reports with incomplete child or partner information (▒) for (a) males and (b) females

Table 7. Male reports (top) and female reports (bottom) of the number of children produced by males and females born between specified years

	Females		
	1957–1966	*1967–1976*	*1977–1986*
Males (respondent)			
1957–1966	904	339	12
1967–1976	89	747	106
1977–1985	2	25	120
	Males		
Females (respondent)			
1957–1966	2571	238	3†
1967–1976	1179	2473	102†
1977–1986	59	523	510†

†1977–1985.

Table 8. Mean sample weights and corresponding variances for male reports (top) and female reports (bottom) of the number of children produced by males and females born between specified years

	Females		
	1957–1966	*1967–1976*	*1977–1986*
Males (respondent)			
1957–1966	2.937 (9.793)	2.043 (2.864)	3.629 (17.888)
1967–1976	2.377 (2.741)	2.351 (3.552)	2.590 (8.031)
1977–1985	2.139 (0)	1.951 (1.205)	2.326 (3.143)
	Males		
Females (respondent)			
1957–1966	1.629 (4.151)	1.187 (0.685)	0.817 (0.049)†
1967–1976	1.147 (0.679)	1.039 (0.619)	1.213 (1.555)†
1977–1986	1.007 (0.764)	1.123 (0.872)	1.208 (0.680)†

†1977–1985.

We wanted to examine mixing based on birth year, but, owing to sparse mixing matrices when considering mixing based on individual birth years, we chose to group years into 10-year blocks for females and similar blocks for males (with the exception of the last block, which consisted of 9 years). This produced the mixing totals and corresponding mean sample weights and variances that are given in Table 7 and Table 8 respectively. As we might expect, most births were reported with partners within the same age category as the respondent, and partnerships outside the respondent's age category typically involved a male who was older than the female. Also, the much higher mean cell weights corresponding to the mixing matrix that was produced by male respondents is as expected, given the oversampling of females relative to males.

Table 9. Parameter estimates, standard errors and p-values for the model fitting separate mixing totals for male and female reports†

Parameter	*Estimate*	*Standard error*	*z-value*	*p-value*
Reference category				
λ	7.884	0.033	236.905	< 0.001
Main effects				
$\lambda^{\mathcal{M}}_{67-76}$	−2.530	0.111	−22.715	< 0.001
$\lambda^{\mathcal{M}}_{77-85}$	−6.431	0.708	−9.085	< 0.001
$\lambda^{\mathcal{F}}_{67-76}$	−1.344	0.064	−21.082	< 0.001
$\lambda^{\mathcal{F}}_{77-86}$	−4.110	0.308	−13.355	< 0.001
Interaction effects				
$\lambda^{\mathcal{M}F}_{67-76,67-76}$	3.461	0.129	26.778	< 0.001
$\lambda^{\mathcal{M}F}_{67-76,77-86}$	4.371	0.340	12.858	< 0.001
$\lambda^{\mathcal{M}F}_{77-85,67-76}$	3.778	0.738	5.119	< 0.001
$\lambda^{\mathcal{M}F}_{77-85,77-86}$	8.288	0.777	10.673	< 0.001
Cell mean comparisons				
$\gamma_{57-66,57-66}$	0.456	0.046	9.935	< 0.001
$\gamma_{57-66,67-76}$	0.669	0.065	10.278	< 0.001
$\gamma_{57-66,77-86}$	0.311	0.351	0.885	0.376
$\gamma_{67-76,57-66}$	0.289	0.132	2.183	0.029
$\gamma_{67-76,67-76}$	0.381	0.044	8.563	< 0.001
$\gamma_{67-76,77-86}$	0.761	0.113	6.731	< 0.001
$\gamma_{77-85,57-66}$	−0.557	0.926	−0.602	0.547
$\gamma_{77-85,67-76}$	0.930	0.246	3.776	< 0.001
$\gamma_{77-85,77-86}$	0.792	0.106	7.467	< 0.001

†The γ-parameters represent comparisons of the corresponding means. For example, $\gamma_{57-66,57-66}$ compares $\mu^{\mathcal{F}}_{57-66,57-66}$ with $\mu^{\mathcal{M}}_{57-66,57-66}$.

Again fitting a model of the form (4) and using the Skinner and Vallet adjustment for standard errors, we obtain the parameter estimates, standard errors and p-values that are shown in Table 9. The parameters of interest, again at the bottom of Table 9, represent direct comparisons of corresponding cells of the mixing matrix for male reports and the mixing matrix for female reports, and Wald tests on these parameters provide strong evidence for incongruities in these reports, except for two cases (reports of children produced by males born 1957–1966 and females born 1977–1986, and males born 1977–1985 and females born 1957–1966, both of which correspond to cells with low counts in both mixing matrices). In all cases where Wald tests are significant, parameter estimates are positive, corresponding to females reporting significantly higher numbers of children. This appears to be in line with what we might expect, given the observed differences between males and females in both reported and used distributions shown in Fig. 2.

7. Discussion

The modelling approach that we have described can be easily implemented for any situation where two disjoint sets separately report shared events, and it provides a mechanism to highlight quickly whether the reports are inconsistent and, if so, where these inconsistencies occur. We demonstrated the use of our approach for male and female reports of sexual partnerships in the NHSLS and fertility in the NSFG and, in both instances, showed inconsistencies in male and female reports. In the case of fertility, inconsistencies were more widespread. Although our method highlights inconsistencies, it does not explain the reasons for them. Nevertheless, it provides a mechanism by which to determine where the researcher should further investigate to determine the causes for inconsistent reporting. For instance, although the focus may be on reporting error, differences could also appear due to sampling bias. In the case of our analysis of fertility reports in the NSFG, imputation of fathers' birth years also comes into play. Additionally, for a fair comparison of the two sets of reports, this should be done on a closed population. Although we have restricted analyses to males and females of certain ages (in the case of the NHSLS) and males and females who were born in certain years (in the case of the NSFG) in an attempt to achieve this, there are segments of the population, such as those who are incarcerated, who may be nominated by respondents as partners or fathers but themselves cannot be sampled. Should they behave significantly differently from the population as a whole and represent a significant portion of a population or subpopulation (such as a particular sex or race), they could substantially impact on these comparisons.

An admitted weakness of our modelling approach is its treatment of multiple reports of partnerships or children by a respondent as being independent and identically distributed, thereby enabling us to represent mixing totals through a two-way contingency table. Although, undoubtedly, modelling these multiple reports by a respondent through multiway contingency tables provides an improvement in terms of capturing the dependence structure for reports by an individual, it introduces its own problems with significantly more involved modelling and sparse contingency tables.

To illustrate these issues, we return to our application of assessing the consistency of male and female reports of concurrent partnerships in the NHSLS. There, we considered partnerships between white and black or African-American males. Rather than allow respondents to contribute multiple times to the mixing matrix, we could instead consider separate three-way tables to represent reported partnership totals from male reports and female reports. An example of such a table for male reports is provided in Fig. 3. The first dimension represents the ethnicity of the respondent, and each successive dimension represents a different possible partner ethnicity. Interpreting one of the expected cell counts, $\mu^{\mathcal{M}}_{\mathrm{W}k1}$ denotes the expected number of males who are white and have k white partners and one black or African-American partner. It should be clear that there is no expected cell count for female reports that can be compared with $\mu^{\mathcal{M}}_{\mathrm{W}k1}$ to check for consistency. To assess consistency, it is necessary to examine the equivalence of linear combinations of expected cell counts. For instance, to obtain the expected number of partnerships between white males and white females, as estimated from male reports, one must calculate

$$\sum_{i=1}^{k} \sum_{j=1}^{l} i\mu^{\mathcal{M}}_{\mathrm{W}ij}, \tag{7}$$

and a similar calculation must be carried out to produce the corresponding expected partnership total from female reports. To obtain direct comparisons of male and female reports, then, a model for these multiway contingency tables must be able simultaneously to model marginal distributions to allow us to assess consistency of reports, and log-linear marginal models provide a means to do this. (Aitchison and Silvey (1958), Haber (1985), Haber and Brown (1986) and Lang and Agresti (1994) developed much of the theory and algorithms for fitting such models, and Bergsma *et al.* (2009) have provided a recent comprehensive examination of the theory, applications and implementation of these models.)

In spite of the existence of these models, properties of estimators and standard errors for these models under complex sample designs are not well understood, so extensions to even a Poisson sampling design are not straightforward. Even if they were, this does not address the more

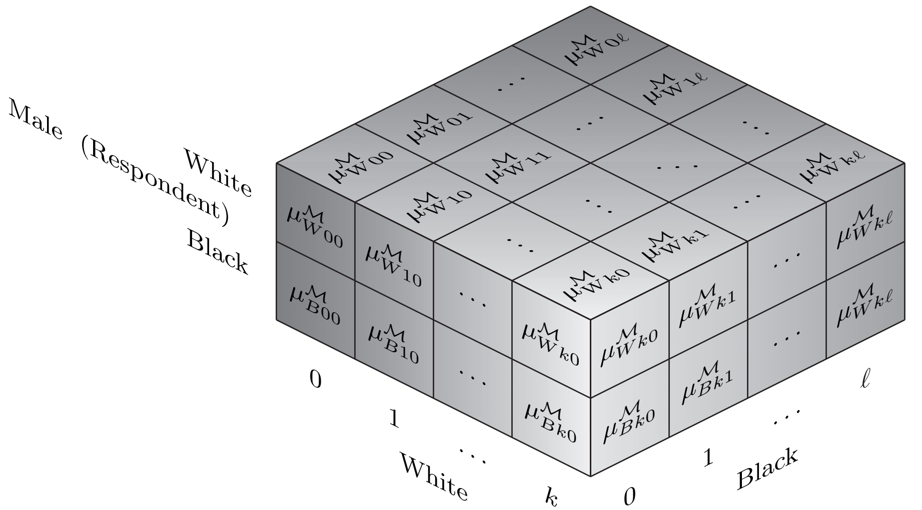


Fig. 3. Three-way table for reported mixing totals from male reports

pressing problem that tables like Fig. 3 tend to be sparse, especially when considering fertility and sexual partnerships. In fact, we would expect the majority of cells to be 0s, and approaches such as artificially inflating zero count cells create issues in terms of assessing the consistency of reports, especially for sparse tables. In the types of application that we consider, it seems most appropriate to treat 0s as random or sampling 0s, and parameters corresponding to such 0s are not estimable in saturated models like those considered, compromising the validity of Wald tests. They may be estimable for other models, although this depends both on the sparsity of the table and on the specific models being considered, and, in general, we recommend caution in interpreting Wald tests for sparse tables.

To address sparsity in a table like Fig. 3, we might consider collapsing cells. This introduces a new problem, however, as linear combinations of the form (7) can no longer be computed. In other words, it is no longer possible to assess the consistency of reports. Artificially inflating empty cells also creates issues by introducing error in these linear combinations. Consequently, although not ideal, the assumption that partnerships are independent and identically distributed at least produces a tractable solution.

Acknowledgements

We thank the referees and Joint Editor for their thoughtful comments and suggestions which led to substantial improvements in the paper.

This work was supported by National Institutes of Health awards R21 HD063000-01 and R21 HD075714, National Science Foundation award MMS-0851555 and Office of Naval Research award N00014-08-1-1015. M. S. Handcock benefited from facilities and resources provided by the California Center for Population Research at the University of California at Los Angeles, which receives core support (R24-HD041022) from the Eunice Kennedy Shriver National Institute of Child Health and Human Development.

References

Aitchison, J. and Silvey, S. D. (1958) Maximum likelihood estimation of parameters subject to restraints. *Ann. Math. Statist.*, **29**, 813–828.

Bergsma, W., Croon, M. and Hagenaars, J. A. (2009) *Marginal Models for Dependent, Clustered, and Longitudinal Categorical Data*. Berlin: Springer.
Brown, N. R. and Sinclair, R. C. (1999) Estimating number of lifetime sexual partners: men and women do it differently. *J. Sex Res.*, **36**, 292–297.
Cherlin, A., Griffth, J. and McCarthy, J. (1983) A note on maritally-disrupted men's reports of child support in the June 1980 Current Population Survey. *Demography*, **20**, 385–389.
Clogg, C. C. and Eliason, S. R. (1987) Some common problems in log-linear analysis. *Sociol. Meth. Res.*, **16**, 8–44.
Garfinkel, I., McLanahan, S. S. and Hanson, T. L. (1998) A patchwork portrait of nonresident fathers. In *Fathers under Fire: the Revolution in Child Support Enforcement* (eds I. Garfinkel, S. S. McLanahan, D. Meyer and J. Seltzer), pp. 31–60. New York: Sage.
Guzzo, K. B. and Furstenberg, F. F. (2007) Multipartnered fertility among American men. *Demography*, **44**, 583–601.
Haber, M. (1985) Maximum likelihood methods for linear and log-linear models in categorical data. *Computnl Statist. Data Anal.*, **81**, 1–10.
Haber, M. and Brown, M. B. (1986) Maximum likelihood methods for log-linear models when expected frequencies are subjected to linear constraints. *J. Am. Statist. Ass.*, **81**, 477–482.
Hamilton, B. E., Ventura, S. J., Martin, J. A. and Sutton, P. D. (2005) Preliminary births for 2004: health E-stats.
Hill, M. S. (1992) *The Panel Study of Income Dynamics: a User Guide*, vol. 2. Newbury Park: Sage.
Johnson, A. M., Wadsworth, J., Field, J., Wellings, K. and Anderson, R. M. (1990) Surveying sexual attitudes. *Nature*, **343**, 109.
Johnson, A. M., Wadsworth, J., Wellings, K., Bradshaw, S. and Field, J. (1992) Sexual lifestyles and HIV risk. *Nature*, **360**, 410–412.
Julian, D., Bouchard, C., Gagnon, M. and Pomerleau, A. (1992) Insider's views of marital sex: a dyadic analysis. *J. Sex Res.*, **29**, 343–360.
Kinsey, A. C., Pomeroy, W. B. and Martin, C. E. (1948) *Sexual Behavior in the Human Male*. Philadelphia: Saunders.
Lang, J. B. and Agresti, A. (1994) Simultaneously modeling joint and marginal distributions of multivariate categorical responses. *J. Am. Statist. Ass.*, **89**, 625–632.
Laumann, E. O., Gagnon, J. H., Michael, R. T. and Michaels, S. (1992) National Health and Social Life Survey. Inter-University Consortium for Political and Social Research, Ann Arbor.
Laumann, E. O., Gagnon, J. H., Michael, R. T. and Michaels, S. (1994) *The Social Organization of Sexuality: Sexual Practices in the United States*. Chicago: University of Chicago Press.
Lepkowski, J. M., Mosher, W. D., Davis, K. E., Groves, R., Van Hoewyk, J. and Willem, J. (2006) National Survey of Family Growth, Cycle VI: sample design, weighting, imputation, and variance estimation. *Vitl Hlth Statist.*, **2**, no. 142.
Lewontin, R. C. (1995) Sex, lies, and social science. *New Yrk Rev. Books*, 24–29.
Loughin, T. M. and Bilder, C. R. (2011) On the use of a log-rate model for survey-weighted categorical data. *Communs Statist. Theor. Meth.*, **40**, 2661–2669.
Morris, M. (1991) A log-linear modeling framework for selective mixing. *Math. Biosci.*, **2**, 349–377.
Morris, M. (1993) Telling tails explain the discrepancy in sexual partner reports. *Nature*, **365**, 437–440.
National Center for Health Statistics (2004) *Public Use Data File Documentation, National Survey of Family Growth Cycle 6: 2002, User's Guide*. Hyattsville: National Center for Health Statistics.
Ochs, E. P. and Binik, Y. M. (1999) The use of couple data to determine the reliability of self-reported sexual behavior. *J. Sex Res.*, **36**, 374–384.
R Core Team (2013) *R: a Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.
Rendall, M. S., Clarke, L., Peters, H. E., Ranjit, N. and Verropoulou, G. (1999) Incomplete reporting of men's fertility in the United States and Britain: a research note. *Demography*, **36**, 135–144.
Rendall, M. S., Joyner, K., Peters, H. E., Yang, F., Handcock, M. S. and Admiraal, R. (2006) Combining population and survey data for male fertility estimation. *Technical Report 2006-2*. Transition to Fatherhood Project, Cornell University, Ithaca.
Seal, D. W. (1997) Interpartner concordance of self-reported sexual behavior among college dating couples. *J. Sex Res.*, **34**, 39–55.
Seltzer, J. A. and Brandreth, Y. (1994) What fathers say about involvement with children after separation. *J. Famly Iss.*, **15**, 49–77.
Skinner, C. and Vallet, L.-A. (2010) Fitting log-linear models to contingency tables from surveys with complex sampling designs: an investigation of the Clogg-Eliason approach. *Sociol. Meth. Res.*, **39**, 83–108.
Smith, T. W. (1992) Discrepancies between men and women in reporting number of sexual partners: a summary of four countries. *Socl Biol.*, **39**, 203–211.
Sorensen, E. (1997) A national profile of nonresidential fathers and their ability to pay child support. *J. Marr. Famly*, **59**, 785–797.
Sweet, J. A. and Bumpass, L. L. (1996) The National Survey of Families and Households—Waves 1 and 2: data description and documentation. Center for Demography and Ecology, University of Wisconsin—Madison, Madison. (Available from `http://www.ssc.wisc.edu/nsfh/home.html`.).

Taylor, M. F., Brice, J., Buck, N. and Prentice-Lane, E. (2010) *British Household Panel Survey User Manual*, vol. A, *Introduction, Technical Report, and Appendices*. Colchester: University of Essex.
United States Bureau of the Census (1984) *Current Population Survey, June 1980*. Washington DC: United States Bureau of the Census.
US Census (2001) Survey of income and program participation (SIPP) 1990 panel: waves 1-5 longitudinal micro-data file. United States Department of Commerce, Bureau of the Census, Washington, DC.
US Census Bureau (2010) Table MS-1: Marital status of the population 15 years old and over, by sex and race: 1950 to present. US Census Bureau, Washington DC.
US National Center for Health Statistics (2002) National Survey of Family Growth, Cycle VI, 2002. United States Department of Health and Human Services, National Center for Health Statistics, Hyattsville.
Wellings, K., Wadsworth, J., Field, J., Johnson, A. M., Bradshaw, S. A. and Anderson, R. M. (1990) Sexual lifestyles under scrutiny. *Nature*, **348**, 276–278.