The Stata Journal (2019)
19, Number 2, pp. 390–434 DOI: 10.1177/1536867X19854018

Statistical analysis of the item-count technique using Stata

Chi-lin Tsai
Department of Political Science
National Chengchi University
Taipei, Taiwan
chi.lin.tsai.clt@gmail.com

Abstract. In this article, I review recent developments of the item-count technique (also known as the unmatched-count or list-experiment technique) and introduce a new package, `kict`, for statistical analysis of the item-count data. This package contains four commands: `kict deff` performs a diagnostic test to detect the violation of an assumption underlying the item-count technique. `kict ls` and `kict ml` perform least-squares estimation and maximum likelihood estimation, respectively. Each encompasses a number of estimators, offering great flexibility for data analysis. `kict pfci` is a postestimation command for producing confidence intervals with better coverage based on profile likelihood. The development of the item-count technique is still ongoing. I will continue to update the `kict` package accordingly.

Keywords: st0559, kict, kict deff, kict ls, kict ml, kict pfci, item-count technique, unmatched-count technique, list experiment, sensitive question

1 Introduction

The item-count technique (also known as the unmatched-count or list-experiment technique) is a questioning technique for eliciting truthful responses to sensitive survey questions. The standard design for the item-count technique randomizes a survey sample into two groups. One group receives a list of items (that is, statements), while the other receives the same list plus an item that addresses a sensitive issue of interest to researchers. Instead of answering each item separately and directly, respondents report the number of items that fit certain criteria (for example, counting the items with which a respondent agrees). The item-count technique ensures the privacy and confidentiality of responses to the sensitive issue, thereby reducing respondents' motives for deliberate misreporting. Furthermore, it still provides researchers with sufficient information for statistical inferences about the sensitive issue.

The item-count technique is becoming increasingly popular in various disciplines as a promising means of studying sensitive issues (Von Hermanni [2016]; Wolter and Laier [2014]; compare Gelman [2014]). However, data analysis for this technique is not straightforward and requires special methods that have not been built in most statistical software suites. I aim to fill this gap by introducing a new package—`kict`. This article proceeds as follows. In section 2, I elaborate on the item-count technique and review commonly used methods for data analysis. Many of those methods were developed for the standard item-count technique. I modify them for nonstandard item-count techniques. In section 3, I introduce the package `kict`. In section 4, I demonstrate its usage by replicating several studies of the item-count technique. I also highlight some practical considerations that are crucial to the applications of the item-count technique but have not received adequate attention in the literature. Finally, I conclude with remarks on the future development of the `kict` package.

2 The item-count technique

2.1 Basic idea

The item-count technique is essentially an encryption scheme that allows survey respondents to encode their answers to a sensitive question. Once encrypted, their answers can be deciphered only at a certain aggregate level for analysis (for example, the mean of a group of respondents' answers). Therefore, it is safe for respondents to answer the sensitive question truthfully.[1]

The 1991 National Race and Politics Survey (Sniderman, Tetlock, and Piazza 1991) is a classic example for illustrating the item-count technique. To measure the prevalence of racial prejudice among white Americans, the survey randomly divides a sample into two. Half the respondents (hereafter the "short-list group") are prompted with a question as follows:

> "Now I'm going to read you [several] things that sometimes make people angry or upset. After I read [them] all, just tell me HOW MANY of them upset you. I don't want to know which ones, just how many.
>
> - the federal government increasing the tax on gasoline;
> - professional athletes getting million-dollar-plus salaries;
> - large corporations polluting the environment.
>
> How many, if any, of these things upset you?"

The other half (hereafter the "long-list group") are prompted with the same list plus an item of interest (hereafter a "sensitive item" or "key item" as opposed to the three nonkey items above):

> - "a black family moving in next door."

1. Another well-known questioning technique for encrypting respondents' answers to a sensitive question is the randomized response technique. A community-contributed command, `rrlogit` (Jann 2005), is available online for Stata users to fit the maximum-likelihood logistic regression for the randomized response data.

Racism is a sensitive issue in most modern societies. People who are upset about having new black neighbors may be reluctant to admit it publicly. Therefore, the item-count question above asks respondents not to answer each item separately but to report the number of items that upset them, thereby allowing respondents to encrypt their answers to the sensitive item. Suppose that a respondent in the long-list group gives an answer of "two"; no one—not even the interviewer—can possibly know whether the answer counted in the key item (because it could be either two nonkey items or one nonkey item plus the key item).[2]

The simplest tool to decrypt the item-count data is a difference-in-means estimator. For example, if the long-list group reports that, on average, 2.20 items upset them, and the short-list group's average count is 2.13, then the estimated prevalence of racial prejudice is 2.20 minus 2.13—that is, 7% of white Americans would be upset about a black family moving in next door.

2.2 Methods for data analysis

Consider a simple random sample of n respondents. Let T_i be the group indicator for respondent i, where $T_i = 1$ if the respondent is assigned to the long-list group and $T_i = 0$ otherwise. Let S_i and $R_{i,j}$ be respondent i's potential answers to the key item and to the jth nonkey item, respectively (where there are J nonkey items). Take the study of racial prejudice as an example: $S_i = 1$ if a black family moving in next door angers respondent i and $S_i = 0$ otherwise. Similarly, $R_{i,2} = 1$, if professional athletes getting million-dollar-plus salaries angers respondent i, and $R_{i,2} = 0$ otherwise. By design, S_i and $R_{i,j}$ are unobserved. The observed variable is the number of affirmative answers: $Y_i = T_i S_i + R_i$, where $R_i = \sum_{j=1}^{J} R_{i,j}$.

Under three assumptions—treatment randomization, no design effect, and no liar—any difference between the two groups' average counts is attributed to the key item.[3] This justifies the use of the difference-in-means estimator, (1), to estimate the prevalence of the key item in the population. (Note that all estimators reviewed in this article rest on at least these three assumptions.)

$$E(S_i) = P(S_i = 1) = \frac{\sum_{i=1}^{n} Y_i T_i}{\sum_{i=1}^{n} T_i} - \frac{\sum_{i=1}^{n} Y_i (1 - T_i)}{\sum_{i=1}^{n} (1 - T_i)} \tag{1}$$

The difference-in-means estimate is identical to the slope coefficient of a simple linear regression of Y_i on T_i. Holbrook and Krosnick (2010, 53–54) generalize this connection from univariate analysis to multivariate modeling. To model S_i, they regress Y_i on T_i, a set of covariates $\mathbf{X}_i$, and interaction terms between T_i and each of $\mathbf{X}_i$. Equation (2)

2. A notable exception is that when a respondent in the long-list group reports that no item or all items make him or her upset, his or her answer to the key item is no longer encrypted. A well-designed item-count question should minimize this problem.
3. "Treatment randomization" means that the sample is split at random. "No liar" means that respondents answer the key item truthfully. "No design effect" means that respondents do not give different answers to nonkey items depending on whether they are in the long-list group. (See Imai [2011, 408–409] for further discussion.)

illustrates this method by a single covariate case. The coefficient of the interaction term δ_β estimates the relative prevalence of the key item among the subpopulation of $X = x$.

$$Y_i = \gamma_\alpha + \gamma_\beta X_i + T_i(\delta_\alpha + \delta_\beta X_i) + \epsilon_i \tag{2}$$

NOTES:

a) $\gamma_\alpha + \gamma_\beta X_i = E(Y_i|X_i, T_i = 0)$.
$\delta_\alpha + \delta_\beta X_i = E(Y_i|X_i, T_i = 1) - E(Y_i|X_i, T_i = 0) = E(S_i|X_i) = P(S_i = 1|X_i)$.

b) Additional assumption: $E(\epsilon_i|X_i, T_i) = 0$.

Holbrook and Krosnick's (2010) method is essentially a linear probability model and so may produce nonsensical predicted values—that is, the predicted probability of answering the key item affirmatively, $\widehat{P}(S_i = 1|\mathbf{X}_i = \mathbf{x})$, may be outside the interval between 0 and 1. Imai (2011, 409) overcomes this drawback using the nonlinear least-squares regression. Equation (3), for example, is a nonlinear version of (2), in which the logistic function restricts predicted probabilities within the unit interval.

$$Y_i = J\left(1 + e^{-\gamma_\alpha - \gamma_\beta X_i}\right)^{-1} + T_i\left(1 + e^{-\delta_\alpha - \delta_\beta X_i}\right)^{-1} + \epsilon_i \tag{3}$$

NOTES:

a) $(1 + e^{-\gamma_\alpha - \gamma_\beta X_i})^{-1} = E(Y_i|X_i, T_i = 0)/J$.
$(1 + e^{-\delta_\alpha - \delta_\beta X_i})^{-1} = E(S_i|X_i) = P(S_i = 1|X_i)$.

b) Additional assumption: $E(\epsilon_i|X_i, T_i) = 0$.

Imai (2011, 409–411) also derives a maximum likelihood estimator as a more statistically efficient alternative to the nonlinear least-squares estimator. Note that S_i and R_i, though unobserved, are mutually identifiable (Glynn 2013, 163). For instance, given $Y_i = 2$ and $T_i = 1$, (S_i, R_i) must be either $(1, 1)$ or $(0, 1)$, and the probabilities of these combinations are also identifiable. Therefore, Imai proposes modeling the joint probability of S_i and R_i. Because S_i is the primary focus, Imai factorizes the joint probability as $P(R_i|S_i, \mathbf{X}_i)P(S_i|\mathbf{X}_i)$—the estimates of the second term represent the association between respondents' characteristics and their answers to the key item. Equation (4) shows the likelihood function of this estimator, where $\boldsymbol{\psi}_s$ and $\boldsymbol{\delta}$ are coefficients to be estimated:

$$\begin{aligned}\prod_{i=1}^{n}\sum_{s=0}^{1}\{&P(R_i = Y_i - sT_i|S_i = s, \mathbf{X}_i = \mathbf{x}, \boldsymbol{\psi}_s)P(S_i = s|\mathbf{X}_i = \mathbf{x}, \boldsymbol{\delta})\\ &\times (\mathbf{1}_{Y_i \neq J+1})^{1-s}(\mathbf{1}_{Y_i \neq 0 \cup T_i = 1})^{s}\}\end{aligned} \tag{4}$$

NOTES:

a) $\mathbf{1}_{\text{Condition}} = 1$ if Condition holds; $\mathbf{1}_{\text{Condition}} = 0$ otherwise.

b) Additional assumption: the distribution of R_i.
For example, specifying a binomial distribution for R_i is equivalent to assuming $(R_{i,j} \perp R_{i,k}|S_i, \mathbf{X}_i)$ and $P(R_{i,j} = 1|S_i, \mathbf{X}_i) = P(R_{i,k} = 1|S_i, \mathbf{X}_i)$, where $j \neq k$.

2.3 Nonstandard designs

The dual-list item-count technique

The standard item-count technique does not require the short-list group to provide information about the key item, so the item-count estimates tend to be less precise (that is, have higher standard errors) than the estimates based on direct questioning. Droitcour et al. (1991, 189) partly overcome this limitation with a dual-list design. Consider the study of racial prejudice again. Suppose that, in addition to the original item-count question (labeled as Q^A), we use another two nonkey items, together with the same key item, to form a second item-count question (Q^B). As illustrated in table 1, there are still two random subgroups: group 0 answers to the short list of Q^A and then the long list of Q^B; in contrast, group 1 answers to the long list of Q^A and then the short list of Q^B. This design prompts respondents with the key item in either Q^A or Q^B. All respondents, regardless of group, have to provide information about the key item.

Table 1. An example of the dual-list design

<table>
<tr><th></th><th>Group 0 ($T_i = 0$)</th><th>Group 1 ($T_i = 1$)</th></tr>
<tr><td>Q^A</td><td>I'm going to read you three things that sometimes make people angry or upset. After I read all three, just tell me how many of them upset you. I don't want to know which ones, just how many.
• the federal government increasing the tax on gasoline;
• professional athletes getting million-dollar-plus salaries;
• large corporations polluting the environment.
How many, if any, of these things upset you?</td><td>I'm going to read you four things that sometimes make people angry or upset. After I read all four, just tell me how many of them upset you. I don't want to know which ones, just how many.
• the federal government increasing the tax on gasoline;
• professional athletes getting million-dollar-plus salaries;
• large corporations polluting the environment;
• a black family moving in next door.
How many, if any, of these things upset you?</td></tr>
<tr><td>Q^B</td><td>I'm going to read you another three things that sometimes make people angry or upset. After I read them, tell me how many of them upset you.
• the state government installing more speed cameras;
• fast food restaurant chains using horse meat in products;
• a black family moving in next door.
How many, if any, of these things upset you?</td><td>I'm going to read you another two things that sometimes make people angry or upset. After I read them, tell me how many of them upset you.
• the state government installing more speed cameras;
• fast food restaurant chains using horse meat in products;
How many, if any, of these things upset you?</td></tr>
</table>

For data analysis, Droitcour et al. (1991) propose applying the difference-in-means estimator to Q^A and Q^B separately to produce two estimates for the key item and then taking their arithmetic mean to obtain a more statistical efficient estimate. Formally, let Y_i^A and Y_i^B be respondent i's answers to Q^A and Q^B, respectively. T_i is still the group indicator, but now $T_i = 0$ if respondent i is assigned to group 0 and $T_i = 1$ if he or she is in group 1. Equation (5) shows the difference-in-means estimator for the dual-list item-count technique.

$$P(S_i = 1) = \left[\left\{\frac{\sum_{i=1}^{n} Y_i^A T_i}{\sum_{i=1}^{n} T_i} - \frac{\sum_{i=1}^{n} Y_i^A (1-T_i)}{\sum_{i=1}^{n}(1-T_i)}\right\} + \left\{\frac{\sum_{i=1}^{n} Y_i^B (1-T_i)}{\sum_{i=1}^{n}(1-T_i)} - \frac{\sum_{i=1}^{n} Y_i^B T_i}{\sum_{i=1}^{n} T_i}\right\}\right] /2 \tag{5}$$

There is a lack of methods for multivariate analysis. To meet this need, I extend the methods reviewed in section 2.2 to the dual-list item-count technique (see the appendix).

The partial item-count technique

Corstange (2009) was among the first to develop a regression method for the item-count technique. However, unlike those reviewed previously, Corstange's (2009) method requires a nonstandard item-count design. His design (hereafter "partial item count") requires respondents in the short-list group to answer each nonkey item separately and directly (while those in the long-list group still answer to the question in the item-count format). Information on each nonkey item is then used to resolve the issue of model identification.

Unlike Imai's (2011) maximum likelihood estimator that models the joint probabilities of S_i and R_i, the partial item-count technique allows Corstange (2009) to model the marginal probabilities of S_i and each $R_{i,j}$ simultaneously (the primary focus is still S_i). Corstange (2009, 51) originally devised his estimator based on an approximate likelihood function. Blair and Imai (2012, 57) improve it by deriving an exact likelihood function as shown in (6) ($\boldsymbol{\theta}_j$ and $\boldsymbol{\delta}$ are coefficients to be estimated):

$$\prod_{i\in(T_i=0)} \prod_{j=1}^{J} P(R_{i,j} = r_j | \mathbf{X}_i = \mathbf{x}, \boldsymbol{\theta}_j) \times \prod_{i\in(T_i=1)} \sum_{u\in U^{Y_i}} \left\{ \prod_{j=1}^{J} P(R_{i,j} = r_{u,j} | \mathbf{X}_i = \mathbf{x}, \boldsymbol{\theta}_j) \right\} P(S_i = s_u | \mathbf{X}_i = \mathbf{x}, \boldsymbol{\delta}) \tag{6}$$

NOTES:

1. U^{Y_i} is a set of combinations of S_i and $R_{i,j}$ that satisfy $S_i + R_i = Y_i$.
2. u is one of the combinations in U^{Y_i}.
   $s_u = 1$ if $S_i = 1$ in u, and $s_u = 0$ otherwise.
   $r_{u,j} = 1$ if $R_{i,j} = 1$ in u, and $r_{u,j} = 0$ otherwise.
3. Additional assumption:
   S_i, $R_{i,j}$, and $R_{i,k}$ (where $j \neq k$) are independent after controlling for $\mathbf{X}_i$.

The item-sum technique

All the item-count techniques reviewed thus far require researchers to phrase the sensitive question as a dichotomous item (yes or no, agree or disagree, etc.) Trappmann et al. (2014) overcome this limitation using the item-sum technique. The item-sum technique differs from the item-count technique in two aspects: first, the key and nonkey items can be continuous variables; second, each respondent reports the sum of his or her answers to the items on the list. For example, Trappmann et al. (2014) in their study, prompted the short-list group with two items as follows:

> "Please answer each of the following questions truthfully. However, please keep the pencil and the piece of article ready. Please note the answer to each question on your piece of article. Afterward, please add the numbers from both answers together and tell me the total result. However, please do not tell me how you answered the individual questions so that I, as interviewer, do not know how you came to your result.
>
> - How many hours did you watch TV last week?
> - How high are your monthly costs for your apartment or your house? Monthly costs can include rent, utilities, coop or condo fees, and mortgage.
>
> Thank you very much. What is your result?"

The long-list group was prompted with the same questions plus a key item:

- "On average, how much do you earn per month from undeclared work?"

Suppose that a respondent in the long-list group watched television 10.5 hours last week, spent €430 for his or her house per month, and made €100 from undeclared work per week; he or she is expected to give an answer of 540.5. The difference-in-means estimator (1) and the linear least-squares estimator (2) are directly applicable to the item-sum technique. For instance, if the short-list group's average response is 516.4, and the long-list group's is 664.3, the estimated earnings per month from undeclared work is €147.9.

2.4 Auxiliary information

Aronow et al. (2015) consider a scenario where a survey measures a sensitive issue not only by the item-count technique but also by direct questioning. For example, besides the item-count question mentioned at the beginning of section 2.1, the 1991 National Race and Politics Survey might have also asked every respondent a direct question on the key item, such as "Does a black family moving in next door make you upset? Yes/No." Thus, the long-list group would have answered the key item twice. With this design, Aronow et al. (2015) propose using the answers to the direct question as auxiliary information to increase the statistical efficiency of the item-count estimation.

Formally, let A_i be respondent i's answer to the direct question, where $A_i = 1$ for the affirmative answer and $A_i = 0$ otherwise. Aronow et al. (2015) make two assumptions about this variable: 1) A_i is independent of T_i, and 2) A_i is monotonically related to S_i. As illustrated in table 2, positive monotonicity means that respondents do not answer the key item affirmatively to the item-count question while answering the direct question negatively; that is, $P(S_i = 1, A_i = 0) = 0$; by contrast, negative monotonicity means $P(S_i = 0, A_i = 1) = 0$.

Table 2. Assumptions about the relationship between S and A

	Positive monotonicity		Negative monotonicity	
	$A = 0$	$A = 1$	$A = 0$	$A = 1$
$S = 0$	$P(S = 0, A = 0)$	$P(S = 0, A = 1)$	$P(S = 0, A = 0)$	0
$S = 1$	0	$P(S = 1, A = 1)$	$P(S = 1, A = 0)$	$P(S = 1, A = 1)$

Under the negative monotonicity assumption, Aronow et al. (2015) modify the difference-in-means estimator as (7) for using A_i to aid the estimation of S_i.

$$P(S_i = 1) = \left(\sum_{i=1}^{n} \frac{A_i}{n}\right) + \left(\sum_{i=1}^{n} \frac{1 - A_i}{n}\right) \times \left\{ \frac{\sum_{i=1}^{n} Y_i T_i (1 - A_i)}{\sum_{i=1}^{n} T_i (1 - A_i)} - \frac{\sum_{i=1}^{n} Y_i (1 - T_i)(1 - A_i)}{\sum_{i=1}^{n} (1 - T_i)(1 - A_i)} \right\} \tag{7}$$

NOTES: Additional assumptions:

- A_i is independent of T_i.
- A_i is related to S_i in a negatively monotonic manner.

Equation (8) is the counterpart under the positive monotonicity assumption:

$$P(S_i = 1) = \left(\sum_{i=1}^{n} \frac{A_i}{n}\right) \left\{ \frac{\sum_{i=1}^{n} Y_i T_i A_i}{\sum_{i=1}^{n} T_i A_i} - \frac{\sum_{i=1}^{n} Y_i (1 - T_i) A_i}{\sum_{i=1}^{n} (1 - T_i) A_i} \right\} \tag{8}$$

NOTES: Additional assumptions:

- A_i is independent of T_i.
- A_i is related to S_i in a positively monotonic manner.

Eady (2017) uses the same type of auxiliary information to improve Imai's (2011) maximum likelihood estimator. Like Aronow et al. (2015), Eady also assumes that

A_i is monotonically related to S_i. This assumption—though simplifying estimation—severely restricts the choice of auxiliary information. Except for a direct question on the sensitive issue of interest, it is hard to see another source of auxiliary information that could possibly satisfy the monotonicity assumption. This is a major limitation of Aronow et al.'s and Eady's methods because surveys cannot afford to adopt both a direct and an indirect approach for the same question.

Nonetheless, Tsai (2017, in part *Not just unbiased but precise: How auxiliary information can improve modeling based on the item-count technique*) shows that the monotonicity assumption is not essential. In fact, any individual-level information that is predictive of but extraneous to the sensitive issue of interest has the potential to improve the item-count estimation. "Extraneity" means that A_i is not a regressor for modeling S_i. (Eady also implicitly makes this assumption.) "Predictivity" requires A_i to be statistically correlated with S_i; this assumption is weaker than monotonicity and thus places fewer restrictions on the choice of auxiliary information. (Besides, A_i still has to be independent of T_i, but this assumption is statistically testable.)

Consider the 1991 National Race and Politics Survey again. Actually, it did not include a direct question on the sensitive issue of interest, but there were other sources of auxiliary information. For example, a question in that survey asked respondents: "How do you feel about blacks buying houses in white suburbs? Strongly in favor/Somewhat in favor/Somewhat opposed/Strongly opposed." This house-buying question and the key item of the item-count question (that is, "a black family moving in next door [makes you upset]") are largely tautological, so it is reasonable to expect some correlation between respondents' answers to these questions (predictivity). Also because of tautology, it is unnecessary to include that house-buying variable as a regressor to model the key item (extraneity). If the independence assumption holds too, then it is legitimate to dichotomize that variable and use it to improve the item-count estimation. (Note that the monotonicity assumption is not an essential concern.)

Equation (9) shows the likelihood function of Tsai's (2017) estimator, where $\boldsymbol{\psi}_s$, $\boldsymbol{\kappa}_s$, and $\boldsymbol{\delta}$ are coefficients to be estimated:

$$\begin{aligned}\prod_{i=1}^{n}\sum_{s=0}^{1}\{&P(R_i = Y_i - sT_i|A_i = a, S_i = s, \mathbf{X}_i = \mathbf{x}, \boldsymbol{\psi}_s)P(A_i = a|S_i = s, \mathbf{X}_i = \mathbf{x}, \boldsymbol{\kappa}_s)\\ &\times P(S_i = s|\mathbf{X}_i = \mathbf{x}, \boldsymbol{\delta})(\mathbf{1}_{Y_i \neq J+1})^{1-s}(\mathbf{1}_{Y_i \neq 0 \cup T_i = 1})^{s}\}\end{aligned} \tag{9}$$

NOTES:

1. $\mathbf{1}_{\text{Condition}} = 1$ if Condition holds; $\mathbf{1}_{\text{Condition}} = 0$ otherwise.
2. Additional assumptions:
    - A_i is independent of T_i.
    - A_i is predictive of S_i.
    - A_i is extraneous to S_i.
    - Distributional assumptions of R_i.

This estimator models the factorized joint probability of R_i, A_i, and S_i. As always, the coefficient $\boldsymbol{\delta}$ is the primary focus; it estimates the association between respondents' characteristics and their answers to the sensitive issue. Eady's (2017) estimator—which rests on the assumption of either negative monotonicity, $P(A_i = 1|S_i = 0, \mathbf{X}_i, \boldsymbol{\kappa}_s) = 0$, or positive monotonicity, $P(A_i = 1|S_i = 1, \mathbf{X}_i, \boldsymbol{\kappa}_s) = 1$—is a special case of Tsai's (2017) estimator.

Equation (9) is specific for the standard design of the item-count technique, but its idea applies to nonstandard designs too. In the appendix, I extend that estimator to Corstange's (2009) partial item-count technique.

3 Commands

`kict` contains four commands: `kict deff` tests for the assumption of no design effect; `kict ls` performs least-squares estimation described in (1), (2), (3), (5), (7), (8), and (15); `kict ml` performs maximum likelihood estimation described in (4), (6), (9), (16), and (17); `kict pfci` estimates profile-likelihood confidence intervals (CI) for a regression coefficient yielded by `kict ml`. This section provides an overview of these commands.

3.1 Syntax

`kict deff` *depvar* [*if*] [*in*] [*weight*] `,` `condition(`*varname*`)` `nnonkey(`*#*`)`
[[`no`] `prob` `verbose` [`no`] `test` [`no`] `gms` `nsim(`*#*`)`]

`kict ls` *depvar* [*depvar2*] [*indepvars*] [*if*] [*in*] [*weight*] `,`
`condition(`*varname*`)` `nnonkey(`*#* [*#*]`)` `estimator(linear|nonlinear)`
[`duallist` `itemsum` `auxiliary(`*varname*`)` `monotony(positive|negative)`
gmm_options]

`kict ml` *depvar* [*depvar2*] [*indepvars*] [*if*] [*in*] [*weight*] `,`
`condition(`*varname*`)` `nnonkey(`*#* [*#*]`)`
`estimator(imai|tsaieady|corstange)` [`pscale(`*matname*`)` `protect(`*#*`)`
`rseed(`*#*`)` `verbose` `direct(`*varlist*`)`
`distribution(poisbino|binomial|betabino|comabino)` `duallist`
`auxiliary(`*varname* [*varname*]`)` `xauxiliary` `monotony(positive|negative)`
ml_options]

kict pfci *indepvar* [, equation(*eqname*) level(#) upper lower ptolerance(#) itolerance(#) btolerance(#) ciiterate(#) mliterate(#) protect(#) rseed(#)]

indepvar specifies an independent variable of the active kict ml model.

depvar specifies a variable that records respondents' answers to an item-count question (Y_i). For the dual-list design, *depvar* and *depvar2* specify variables that record respondents' answers to the first and second item-count questions, respectively (that is, Y_i^A for Q^A and Y_i^B for Q^B). fweights, iweights, and pweights are allowed; see [U] **11.1.6 weight**. Because kict is programmed using Stata's gmm and ml, many options of these two official commands are also allowed; see [R] **ml**, [R] **gmm**, and the help files of kict in Stata.

3.2 Options for kict deff

condition(*varname*) specifies a dummy variable that identifies the treatment status (T_i). The dummy takes on the value 0, representing the short-list group, or 1, representing the long-list group. condition() is required.

nnonkey(#) specifies a positive integer that represents the number of nonkey items. nnonkey() is required.

[no]prob displays the estimated joint probabilities of a respondent's answer to the key item (S_i) and the number of nonkey items that the respondent would answer in the affirmative (R_i). kict deff estimates the joint probabilities based on Glynn's (2013, 166) formula.

verbose displays the computational log of the joint probabilities.

[no]test performs a test for design effects developed by Blair and Imai (2012, 64–65, 74). The null hypothesis of the test states that all the estimated joint probabilities are positive, whereas the alternative hypothesis states that any of the estimated joint probabilities are negative.

By definition, probabilities cannot be negative. If some of the estimated joint probabilities are negative, kict deff performs the test to check whether those negative estimates have arisen by chance. Rejection of the null hypothesis indicates the presence of design effects. Consequently, the estimates based on the item-count technique are questionable.

kict deff splits the estimated joint probabilities into two sets and tests them separately. The first set includes the probabilities $\Pr(R=0,S=0)$, $\Pr(R=1,S=0),\dots,\Pr(R=J,S=0)$, and the second set includes $\Pr(R=0,S=1),\Pr(R=1,S=1),\dots,\Pr(R=J,S=1)$. (See also Kudo [1963].)

[no]gms performs the test with the method of generalized moment selection. Blair and Imai (2012, 64–65, 74) proposed using this method to improve the power of the test. The basic idea behind this method is to exclude the joint probabilities that

are clearly larger than zero from the test. (See also Andrews and Soares [2010, 129, 135].)

`nsim(#)` performs # Monte Carlo simulations to compute p-values when analytical solutions are not viable. The default is `nsim(1000)`.

3.3 Options for kict ls

`condition(`*varname*`)` specifies a dummy variable that identifies the treatment status. For the dual-list design, the dummy takes on the value 0 for those in the short-list group of the first item-count question and in the long-list group of the second item-count question. However, the dummy takes on the value 1 for those in the long-list group of the first item-count question and in the short-list group of the second item-count question (see table 1). `condition()` is required.

`nnonkey(#` [`#`]`)` specifies a positive integer that represents the number of nonkey items. For the dual-list design, specify two positive integers that represent the numbers of the nonkey items in the first and second item-questions, respectively. (The order of the numbers matters!) `nnonkey()` is required.

`estimator(linear|nonlinear)` specifies the type of least-squares estimator to be used. `linear` performs the linear least-squares estimation [for example, (1) and (2)]; `nonlinear` performs the nonlinear least-squares estimation [for example, (3)]. `estimator()` is required.

`duallist` specifies that `kict ls` perform estimation for the dual-list design of the item-count technique [for example, (5) and (15)].

`itemsum` specifies that `kict ls` perform estimation for the item-sum technique. This option pertains only to `estimator(linear)`.

`auxiliary(`*varname*`)` specifies an auxiliary variable for improving the linear least-squares estimation for the standard item-count technique [that is, (7) and (8)]. This option pertains only to `estimator(linear)` and requires `monotony()`.

`monotony(positive|negative)` specifies that `kict ls` perform estimation under a monotonicity assumption. The positive monotonicity assumes that $P(A_i = 0, S_i = 1) = 0$; the negative monotonicity assumes that $P(A_i = 1, S_i = 0) = 0$ (see table 2).

gmm_options; see `help kict ls` for a list of options, and see [R] **gmm** for option descriptions.

3.4 Options for kict ml

`condition(`*varname*`)` and `nnonkey(# [#])`; see the above descriptions. `condition()` and `nnonkey()` are required.

`estimator(imai|tsaieady|corstange)` specifies the type of maximum likelihood estimator to be used. `imai` performs (4); `tsaieady` performs (9); `corstange` performs (6). `estimator()` is required.

`pscale(`*matname*`)` specifies a row vector of positive numbers to be used as the scale parameters of Cauchy distributions. `kict ml` uses this vector (and fixes location parameters at zero) to construct priors and performs the quasi-Bayesian estimation. The row length of the vector must equal the number of coefficients.

Consider a vector $\mathbf{v} = [10, ., 15]$. Specifying `pscale(v)` imposes a Cauchy$(0, 10)$ prior on the first coefficient, a Cauchy$(0, 15)$ prior on the third coefficient, and no prior on the second coefficient.

`protect(`#`)` specifies that `kict ml` perform # optimizations with random-selected initial values and report the result with the maximum likelihood. Specifying a large number such as `protect(50)` provides reasonable assurance that the result found is global rather than just a local maximum.

`rseed(`#`)` specifies the random-number seed for the option `protect(`#`)`. If the seed is not specified, the random-number generator starts in whatever state it was last in.

`verbose` displays iteration logs of the log likelihood when `protect(`#`)` is specified. (Compare [`no`]`log`, which controls the iteration display for each optimization.)

ml_options; see `help kict ml` for a list of options, and see [R] **maximize** and [R] **estimation options** for option descriptions.

Options specific to estimator(corstange)

`direct(`*varlist*`)` specifies J dummy variables; each records the short-list group's answers to a nonkey item. These dummies may be 0, indicating a negative outcome, or 1, indicating a positive outcome.

`distribution(poisbino|binomial)` specifies whether `kict ml` uses Corstange's approximate likelihood function or Blair and Imai's exact likelihood function in estimation. The default is `distribution(poisbino)`.

The approximate likelihood function assumes that the number of affirmative answers to the key and nonkey items $(S_i + R_i)$ is binomially distributed (`binomial`), whereas the exact likelihood function uses the Poisson-binomial distribution (`poisbino`).

Options specific to estimator(imai)

`duallist` specifies that `kict ml` perform estimation for the dual-list design of the item-count technique [that is, (16)].

`distribution(binomial|betabino|comabino)` specifies the distribution of respondents' answers to nonkey items. The default is `distribution(binomial)`.

Options specific to estimator(tsaieady)

`auxiliary(`*varname* [*varname*]`)` specifies a dummy variable to be the auxiliary variable (A_i). This dummy acts both as a dependent variable in the `kappa0` and `kappa1` equations and as a covariate in the `Psi0p` and `Psi1p` equations (or in the `Thetaj` equation).

If a recoded version of the dummy is also specified after the original dummy, `kict ml` uses the first variable as the dependent variable and uses the second one as the covariate.

`xauxiliary` specifies that `kict ml` add interaction terms between the auxiliary variable (A_i) and every one of the covariates ($\mathbf{X}_i$) into the `Psi0p` and `Psi1p` equations (or the `Thetaj` equation).

`monotony(positive|negative)` specifies that `kict ml` perform estimation under one of the monotony assumptions. The positive monotonicity assumes that $P(A_i = 0, S_i = 1) = 0$; the negative monotonicity assumes that $P(A_i = 1, S_i = 0) = 0$ (see table 2).

3.5 Options for kict pfci

`equation(`*eqname*`)` specifies an equation name of the active `kict ml` model. `kict pfci` estimates the profile-likelihood CI for the coefficient `_b[`*eqname*`:`*indepvar*`]`. The default is `equation(Delta)`.

`level(`#`)` sets the confidence level to be estimated. The default is `level(95)`.

`upper` and `lower` specify `kict pfci` to estimate the upper and the lower limit CI, respectively. The default is estimation of both limits.

`ptolerance(`#`)`, `itolerance(`#`)`, `btolerance(`#`)`, and `ciiterate(`#`)` specify convergence and stopping criteria.

`ptolerance()` specifies the difference between the targeted log likelihood and the log likelihood of the `kict ml` model in the current `kict pfci` iteration. `kict pfci` declares convergence if abs(`ptolerance()`) ≤ #. The default is `ptolerance(0)`.

`itolerance()` specifies the difference in the log likelihood of the `kict ml` model between the previous and the current `kict pfci` iteration. `kict pfci` declares convergence if abs(`itolerance()`) ≤ #. The default is `itolerance(0)`.

`btolerance()` specifies the difference in the coefficient `_b[`*eqname*`:`*indepvar*`]` between the previous and the current `kict pfci` iteration. `kict pfci` declares convergence if abs(`btolerance()`) ≤ #. The default is `btolerance(1e-5)`.

`ciiterate()` specifies the maximum number of `kict pfci` iterations.

`mliterate(`#`)`, `protect(`#`)`, and `rseed(`#`)` change the setup that was used in estimation of the active `kict ml` model. Specifically, `kict pfci` computes CI by refitting the model in each iteration. These three options allow `kict pfci` to execute `kict ml` using a different setup. See `help kict ml` for detailed descriptions of these options. (`mliterate(#)` is the `iterate(#)` in `kict ml`.)

4 Examples and technical details

In this section, I demonstrate the use of `kict`. I also provide further technical details and practical considerations about the applications of the methods reviewed previously.

4.1 The standard item-count technique

I first revisit the item-count question in the 1991 American National Race and Politics Survey. The data contain an indicator variable (`group`) that takes the value 0 or 1 to indicate whether a respondent is in the short-list or long-list group. The item-count response (`itemcount`) takes an integer ranging from 0 to 4 because there is one key item and three nonkey items. There are also several demographic variables, including residential region (southerner versus others), age (divided by 100), gender (male or not), education level (college educated and higher versus others). After casewise deletion, 1,213 white respondents are available for analysis.[4]

```
. use "section4-1-race.dta"
(http://sda.berkeley.edu/cgi-bin/hsda?harcsda+natlrace)
. codebook group itemcount south age_rs male college, compact
Variable    Obs Unique      Mean  Min  Max  Label
-------------------------------------------------------------------------------
group      1213      2   .514427    0    1  treatment status
itemcount  1213      5  2.169002    0    4  response to the item-count question
south      1213      2  .2349547    0    1  white southerner
age_rs     1213     70  .4273207  .18  .88  respondent´s age/100
male       1213      2  .4402308    0    1  male respondent
college    1213      2  .5803792    0    1  having a college degree
-------------------------------------------------------------------------------
```

Diagnostics

The validity of the item-count technique rests on at least three assumptions (footnote 3). A common practice to check the first assumption—treatment randomization—is to test for differences between the short-list and long-list groups' responses to important variables in the survey. No significant difference between groups indicates an effective randomization of treatment.

```
. ttest    age_rs, by(group)
. tabulate south   group, chi2
. tabulate male    group, chi2
. tabulate college group, chi2
. tabulate house   group, chi2
```

4. The complete dataset is available on the website of the Survey Documentation and Analysis at the University of California, Berkeley (http://sda.berkeley.edu/cgi-bin/hsda?harcsda+natlrace).

The second assumption—no liar—requires respondents in the long-list group to answer the key item truthfully. It is not statistically feasible to check this assumption, not only because respondents' answers to the key item are by design unobserved but also because their truthful answers are unknown (otherwise there is no point in using the item-count technique).

The third assumption—no design effect—requires respondents not to change their answers to nonkey items depending on whether the key item appears in the list. Suppose that a respondent in the short-list group answers one nonkey item affirmatively. If he or she were assigned to the long-list group, his or her answer must be either "one" or "two" (that is, he or she either answers one nonkey item affirmatively or answers one nonkey item plus the key item affirmatively). Blair and Imai (2012, 64) propose a statistical test for the no-design-effect assumption. `kict deff` implements that test and summarizes the results in two tables:

```
. kict deff itemcount, nnonkey(3) condition(group)
Joint distributions of the key and non-key items
```

	Coef	Robust SE	z	P>z
Pr(R=0,S=1).	-0.0168664	0.0083700	-2.0151	0.0219
Pr(R=0,S=0)	0.0304487	0.0068782	4.4268	1.0000
Pr(R=1,S=1)	0.0101269	0.0242760	0.4172	0.6617
Pr(R=1,S=0)	0.2139818	0.0174483	12.2638	1.0000
Pr(R=2,S=1)	0.0200497	0.0280796	0.7140	0.7624
Pr(R=2,S=0)	0.3568603	0.0263428	13.5468	1.0000
Pr(R=3,S=1)	0.0544872	0.0090863	5.9966	1.0000
Pr(R=3,S=0)	0.3309118	0.0220162	15.0304	1.0000

```
Test for design effects (with GMS)
```

Ha: Pr<0	K	Lambda	P>Lambda	#P>Lambda
Pr(R ,S=0)	0	0.0000000	1.0000	1.0000
Pr(R ,S=1)	1	4.0606140	0.0219	0.0439

```
# Bonferroni-adjusted p-values
```

The first table lists estimated probabilities of all possible types of item-count responses. For example, `Pr(R=0,S=1)` estimates the joint probability of answering all nonkey items negatively and answering the key item affirmatively. This estimate is a nonsensical value (-0.02), raising doubts about the validity of the no-design-effect assumption. The next step is to check whether such a negative estimate has arisen by chance. The second table shows two hypothesis tests: one tests the null hypothesis that none of `Pr(R=r,S=0)` is smaller than zero; the other tests that none of `Pr(R=r,S=1)` is smaller than zero.[5]

When either of these tests is statistically significant, researchers should conclude that the no-design-effect assumption does not hold.[6]

The difference-in-means estimator

Let us assume for a moment that the 1991 American National Race and Politics Survey satisfies all three fundamental assumptions of the item-count technique. Then a basic analysis is to estimate the proportion of people who answer the key item affirmatively. `kict ls` with the option `estimator(linear)` can perform the difference-in-means estimation (1):

```
. kict ls itemcount, nnonkey(3) condition(group) estimator(linear)
  (output omitted)
Initial weight matrix: Unadjusted                 Number of obs   =      1,213
GMM weight matrix:     Robust

Linear least squares estimator
```

	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf. Interval]	
Delta						
_cons	.0677974	.0495375	1.37	0.171	-.0292943	.1648891
Gamma						
_cons	2.134126	.0331137	64.45	0.000	2.069224	2.199027

```
Instruments for equation Delta: _cons
Instruments for equation Gamma: _cons
```

In this output, the central focus is `Delta_cons`. It means that 6.8% of white Americans would be upset by a black family moving in next door.[7] As for `Gamma_cons`, it gives a predicted number of affirmative answers to the nonkey items—that is, white Americans, on average, feel angry over 2.1 nonkey items. This statistic is of no immediate interest to researchers, though it must be estimated simultaneously.

5. Both tests use the generalized moment selection procedure to improve statistical power (Blair and Imai [2012, 66–68]). The second column of the table shows the number of estimates needed to be tested, (`K`), which ranges between 0 and J. In the first test, all estimates are positive, so $\texttt{K} = 0$. Consequently, the test statistic (`Lambda`) is zero, and the p-value is one. In the second test, one estimate is negative, that is, `Pr(R=0,S=1)`, so $\texttt{K} = 1$. In this case, the test is equivalent to a z test for the null hypothesis that `Pr(R=0,S=1)` ≥ 0. When $\texttt{K} > 1$, all negative estimates are tested simultaneously.
6. Blair and Imai (2012, 66–68) develop a likelihood-based method to adjust for design effects. However, the method requires researchers to make additional assumptions about the pattern of the design effect. Different assumptions require different likelihood functions. `kict` currently does not perform this method.
7. Another way to estimate `Delta_cons` is to use two built-in commands: `mean itemcount, over(group)` and `lincom _b[`*long_list*`]-_b[`*short_list*`]`.

Note that, regardless of estimators, `kict`'s output always uses `Delta` to label the coefficients that are directly related to the key item. The remaining coefficients in the output are labeled according to estimator developers' notations to suit the convenience of interested readers for referring back to the formula in section 2 and in the literature. For space considerations, I concentrate on the `Delta` coefficients in the rest of this demonstration.

The linear least-squares estimator

Kuklinski, Cobb, and Gilens (1997) analyze whether whites in the Southern United States have a higher level of racial prejudice than other white Americans do. They answer this question by applying the differences-in-means estimator to the subsample of Southern whites and the other whites in the survey and then testing for the difference between two estimates.

Taking a more sophistic approach, Imai (2011, 412) estimates the regional difference in racial prejudice while controlling for other sociodemographic characteristics. The output below replicates his analysis by a linear least-squares regression [similar to (2)].[8] The coefficient of interest is `Delta_south`, which is 0.20; this indicates that, ceteris paribus, racial prejudice is more prevalent among Southern whites than among other whites by 20 percentage points, but the difference does not reach statistical significance at the conventional level of 0.05.

8. This output replicates columns 2 and 3 of Imai's table 1, except that I rescale the variable age by dividing it by 100 to make all regressors have a similar range. This difference applies for the following outputs that replicate Imai's table 1. Another way to produce this output is to type `regress itemcount group##south group##c.age group##male group##college, vce(r)`.

```
. kict ls itemcount south age_rs male college, nnonkey(3) condition(group)
> estimator(linear)
```

(*output omitted*)

```
Initial weight matrix: Unadjusted                Number of obs   =      1,213
GMM weight matrix:     Robust

Linear least squares estimator
```

	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
Delta						
south	.2019751	.116308	1.74	0.082	-.0259844	.4299346
age_rs	.7309383	.3016095	2.42	0.015	.1397944	1.322082
male	.1802289	.0976275	1.85	0.065	-.0111174	.3715752
college	.1144616	.0969011	1.18	0.238	-.0754611	.3043843
_cons	-.434299	.1586645	-2.74	0.006	-.7452758	-.1233222
Gamma						
south	-.1802073	.0736831	-2.45	0.014	-.3246235	-.0357911
age_rs	.2047344	.1974631	1.04	0.300	-.1822862	.591755
male	-.2017671	.0646587	-3.12	0.002	-.3284958	-.0750385
college	-.3940769	.0634893	-6.21	0.000	-.5185137	-.2696402
_cons	2.406064	.1039832	23.14	0.000	2.202261	2.609867

```
Instruments for equation Delta: south age_rs male college _cons
Instruments for equation Gamma: south age_rs male college _cons
```

One disadvantage of the linear least-squares estimation is unreasonable prediction. For example, a 21-year-old college-educated female non-Southern white's probability of feeling anger over a black family moving in next door is estimated at $-0.17(= 0.20 \times 0 + 0.73 \times 0.21 + 0.18 \times 0 + 0.11 \times 1 - 0.43)$. This is nonsensical because a probability cannot be a negative value. We can avoid this problem by using the nonlinear least-squares estimator.[9]

9. The predicted number of nonkey items that upset this individual is $2.05(= -0.18 \times 0 + 0.20 \times 0.21 - 0.20 \times 0 - 0.39 \times 1 + 2.41)$, which is 0.18 items more than the predicted number for her Southern white counterpart (as `Gamma_south` $= -0.18$).

The nonlinear least-squares estimator

`kict ls` with the option `estimator(nonlinear)` performs nonlinear least-squares estimation [similar to (3)]. It uses a logistic function to restrict predicted values within the unit interval. Based on the output below,[10] the 21-year-old college-educated female non-Southern white's probability of feeling anger over a black family moving in next door is now estimated at a positive value: $\text{logit}^{-1}(2.49\times0+2.6\times0.21+3.10\times0+0.61\times1-7.08) = 0.003$.

```
. kict ls itemcount south age_rs male college, nnonkey(3) condition(group)
> estimator(nonlinear) vce(robust)
  (output omitted)
Initial weight matrix: Unadjusted                Number of obs   =      1,213
GMM weight matrix:     Robust

Nonlinear least squares estimator
(Logistic function)
```

	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
Delta						
south	2.486014	1.265948	1.96	0.050	.0048011	4.967227
age_rs	2.630872	3.145072	0.84	0.403	-3.533355	8.7951
male	3.073732	2.772236	1.11	0.268	-2.35975	8.507215
college	.6131016	1.029562	0.60	0.552	-1.404803	2.631007
_cons	-7.068127	3.629267	-1.95	0.051	-14.18136	.0451053
Gamma						
south	-.2765303	.1161733	-2.38	0.017	-.5042258	-.0488348
age_rs	.3306478	.350308	0.94	0.345	-.3559432	1.017239
male	-.3322204	.1070148	-3.10	0.002	-.5419656	-.1224753
college	-.6617426	.1131421	-5.85	0.000	-.8834969	-.4399882
_cons	1.388113	.1868327	7.43	0.000	1.021928	1.754298

```
Instruments for equation Delta: south age_rs male college _cons
Instruments for equation Gamma: south age_rs male college _cons
```

Because of the logistic parameterization, `Delta` coefficients can be interpreted based on predicted values or odds ratios similarly to logistic regression coefficients. For example, `Delta_south` is 2.49, suggesting that, holding all other variables constant, Southern whites' odds of being upset by a black family moving in next door are higher than the other whites' odds by a factor of $e^{2.49}$ (see Long [1997, 68–82]).[11]

10. This output replicates columns 4 and 5 of table 1 in Imai (2011, 412).
11. Odds ratios do not directly apply to Gamma coefficients, but predicated values do. For example, the expected number of nonkey items that upset the 21-year-old college-educated female non-Southern whites is $\text{logit}^{-1}(-0.28\times0+0.33\times0.21-0.33\times0-0.66\times1+1.39)\times3 = 2.07$ (where "3" is the number of nonkey items).

Imai's maximum likelihood estimator

Imai's maximum likelihood estimator (4) is an alternative to the nonlinear least-squares estimator. By default, `kict ml` with the option `estimator(imai)` implements this estimator using the following setups:

$$
\begin{aligned}
S_i|\mathbf{X}_i = \mathbf{x} &\overset{\text{indep}}{\sim} \text{Bernoulli}\left\{\text{logit}^{-1}(\mathbf{x}'\boldsymbol{\delta})\right\} \\
\text{logit}^{-1}(\mathbf{x}'\boldsymbol{\delta}) &= P(S_i = 1|\mathbf{X}_i = \mathbf{x}) \\
R_i|S_i = s, \mathbf{X}_i = \mathbf{x} &\overset{\text{indep}}{\sim} \text{Binomial}\left\{J, \text{logit}^{-1}(\mathbf{x}'\boldsymbol{\psi}_s)\right\} \\
\text{logit}^{-1}(\mathbf{x}'\boldsymbol{\psi}_s) &= P(R_{i,j} = 1|S_i = s, \mathbf{X}_i = \mathbf{x})
\end{aligned}
$$

`kict ml` also produces a set of `Delta` coefficients that estimate the associations between respondents' characteristics ($\mathbf{X}_i$) and their potential answers to the key item (S_i). The interpretation of these coefficients is the same as explained previously because `kict ml` still uses the logistic function for parameterization. For example, `Delta_age` is 6.55, meaning that being one year older is expected to increase the odds of feeling angry over a black family moving in next door by a factor of $e^{6.55\times 0.01}$. (Remember that the variable age is rescaled.)[12]

```
. kict ml itemcount south age_rs male college, nnonkey(3) condition(group)
> estimator(imai) iterate(100)
```

(*output omitted*)

```
Imai MLE                                        Number of obs     =      1,213
Log likelihood = -1441.0198
```

itemcount	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
Delta						
south	1.379264	.8195656	1.68	0.092	-.2270549	2.985583
age_rs	6.550166	2.072924	3.16	0.002	2.487309	10.61302
male	1.365503	.6120502	2.23	0.026	.1659071	2.5651
college	-.1821738	.5686168	-0.32	0.749	-1.296642	.9322947
_cons	-6.225903	1.045659	-5.95	0.000	-8.275357	-4.176448
Psi0						
south	-.2986788	.1068365	-2.80	0.005	-.5080746	-.0892831
age_rs	.3142651	.3173355	0.99	0.322	-.3077011	.9362313
male	-.2177434	.0858093	-2.54	0.011	-.3859265	-.0495603
college	-.4881385	.0866453	-5.63	0.000	-.6579601	-.3183169
_cons	1.156234	.1564676	7.39	0.000	.8495628	1.462905
Psi1						
south	-.2698412	.589824	-0.46	0.647	-1.425875	.8861925
age_rs	-1.281746	1.571162	-0.82	0.415	-4.361168	1.797675
male	-1.688119	1.629799	-1.04	0.300	-4.882467	1.506228
college	-.9536414	.7143459	-1.33	0.182	-2.353734	.4464509
_cons	3.779721	2.154579	1.75	0.079	-.4431775	8.002619

```
*Sum(nonkey items) ~ Binomial
```

12. This output replicates columns 8 and 9 of table 1 in Imai (2011, 412).

Furthermore, the `Psi0` and `Psi1` equations model respondents' counts of affirmative answers to nonkey items (R_i). These equations, though of no immediate interest to researchers, are crucial to the validity of the `Delta` coefficients. Note that `kict ml` by default assumes that R_i conditionally follows a binomial distribution. Substantively, this implies two additional assumptions: 1) $(R_{i,j} \perp R_{i,k})|S_i, \mathbf{X}_i$—each respondent's answer to a nonkey item is independent of his or her answer to any other nonkey item; 2) $P(R_{i,j} = 1|S_i, \mathbf{X}_i) = P(R_{i,k} = 1|S_i, \mathbf{X}_i)$—the probability that a respondent answers a nonkey item affirmatively is identical to the probability that he or she answers any other nonkey item affirmatively.

These assumptions do not always hold true. For example, a commonly used strategy to prevent respondents from answering all items affirmatively or negatively is to include an extremely low or high prevalent nonkey item into the list. This strategy jeopardizes assumption 2. Another strategy is to use negatively correlated nonkey items (Glynn 2013, 163). However, this risks violating assumption 1.

A way to relax these two assumptions is to use other distributions to model R_i. `kict ml` currently provides two alternative distributions: the beta-binomial distribution, which is particularly useful when nonkey items are positively correlated, and the Conway–Maxwell binomial distribution, which allows either a positive or negative correlation among nonkey items.[13]

$$
\begin{aligned}
R_i|S_i = s, \mathbf{X}_i = \mathbf{x} &\overset{\text{indep}}{\sim} \text{Beta-Binomial}\left\{J, \frac{\mu_s(1-\varrho_s)}{\varrho_s}, \frac{(1-\mu_s)(1-\varrho_s)}{\varrho_s}\right\} \\
\mu_s &= \text{logit}^{-1}(\mathbf{x}'\boldsymbol{\psi}_s) \\
\varrho_s &= \text{logit}^{-1}(\mathbf{x}'\boldsymbol{\rho}_s) \\
R_i|S_i = s, \mathbf{X}_i = \mathbf{x} &\overset{\text{indep}}{\sim} \text{Conway–Maxwell-binomial}(J, \mu_s, \varrho_s) \\
\mu_s &= \text{logit}^{-1}(\mathbf{x}'\boldsymbol{\psi}_s) \\
\varrho_s &= \mathbf{x}'\boldsymbol{\nu}_s
\end{aligned}
$$

Nonetheless, both alternative distributions require `kict ml` to estimate additional coefficients, adding considerable complexity to optimization. Moreover, neither alternative guarantees to fit R_i well, especially when the relationships among nonkey items are highly complicated (for example, correlations vary from one pair of nonkey items to another, and the probability of answering an item affirmatively is different from that of another item). In fact, because respondents do not report their answers to each nonkey

13. The interpretation of `Psi0` and `Psi1` coefficients ($\boldsymbol{\psi}_s$) varies according to the distribution specification. In the binomial model, $\text{logit}^{-1}(\mathbf{x}'\boldsymbol{\psi}_s)$ is the probability of answering a nonkey item affirmatively, conditional on the key item and regressors, and this probability applies to every nonkey item. In the Beta–Binomial model, the probability is random over nonkey items and follows a beta distribution; $\text{logit}^{-1}(\mathbf{x}'\boldsymbol{\psi}_s)$ is the average of them, and $\text{logit}^{-1}(\mathbf{x}'\boldsymbol{\rho}_s)$ is the correlation among nonkey items (Imai 2011, 410–411). In the Conway–Maxwell-binomial model, $\text{logit}^{-1}(\mathbf{x}'\boldsymbol{\psi}_s)$ does not have a substantive meaning except when $\mathbf{x}'\boldsymbol{\nu}_s = 1$, which means no correlation among nonkey items. In that case, the model degenerates to a binomial model. In contrast, $\mathbf{x}'\boldsymbol{\nu}_s < 1$ indicates a positive correlation among nonkey items, whereas $\mathbf{x}'\boldsymbol{\nu}_s > 1$ indicates a negative correlation. However, note that $\mathbf{x}'\boldsymbol{\nu}_s$ itself is not a usual correlation coefficient (Kadane 2016).

item, there is insufficient information to identify an appropriate distribution for R_i. This is an inherent limitation of the standard item-count technique. Some studies have attempted to address this limitation (for example, Corstange [2009]; Tsai [2018]), but more work is still needed.

Another significant aspect of Imai's estimator is regarding the relationship between responses to the key and nonkey items. By default, `kict ml` uses the most flexible specification to model that relationship: the `Psi0` coefficients are free to differ from their counterparts in the `Psi1` equation. This specification, as illustrated by (10), not only allows $R_{i,j}$ to be correlated with S_i but also allows the correlation to be moderated by $\mathbf{X}_i$.[14]

$$\begin{aligned} P(R_{i,j}=1|S_i,\mathbf{X}_i) &= \operatorname{logit}^{-1}\{\psi_{\alpha 0}+\mathbf{X}_i'\boldsymbol{\psi}_{\beta 0}+S_i(\psi_{\alpha 1}-\psi_{\alpha 0}) \\ &\quad +(S_i\mathbf{X}_i)'(\boldsymbol{\psi}_{\beta 1}-\boldsymbol{\psi}_{\beta 0})\} \qquad (10) \\ P(R_{i,j}=1|S_i,\mathbf{X}_i) &= \operatorname{logit}^{-1}\{\psi_{\alpha 0}+\mathbf{X}_i'\boldsymbol{\psi}_{\beta 0}\} \qquad (11) \\ P(R_{i,j}=1|S_i,\mathbf{X}_i) &= \operatorname{logit}^{-1}\{\psi_{\alpha 0}+\mathbf{X}_i'\boldsymbol{\psi}_{\beta 0}+S_i(\psi_{\alpha 1}-\psi_{\alpha 0})\} \qquad (12) \\ P(R_{i,j}=1|S_i,\mathbf{X}_i) &= \operatorname{logit}^{-1}\{\psi_{\alpha 0}+S_i(\psi_{\alpha 1}-\psi_{\alpha 0})\} \qquad (13) \\ P(R_{i,j}=1|S_i,\mathbf{X}_i) &= \operatorname{logit}^{-1}(\psi_{\alpha 0}) \qquad (14) \end{aligned}$$

Imai (2011, 412) proposes a simpler specification. He assumes that $R_{i,j}$ and S_i are independent after controlling $\mathbf{X}_i$ and accordingly restricts all `Psi0` coefficients to match their `Psi1` counterparts (11). `kict ml` can perform this specification by imposing two constraints: `[Psi0=Psi1]` and `[Psi0=Psi1]:_cons`. There are other specifications. For example, (12) allows for a correlation between $R_{i,j}$ and S_i while assuming no moderate effect of $\mathbf{X}_i$ (use the constraint `[Psi0=Psi1]`). Equation (13) assumes no correlation between $R_{i,j}$ and $\mathbf{X}_i$ (`[Psi0p=Psi1p]`, `[Psi0p]:X1=0`, `[Psi0p]:X2=0`, ...). Equation (14) assumes that neither S_i nor $\mathbf{X}_i$ correlates with $R_{i,j}$ (`[Psi0p=Psi1p]`, `[Psi0p=Psi1p]:_cons`, `[Psi0p]:X1=0`, `[Psi0p]:X2=0`, ...). All these specifications are nested in (10). The likelihood-ratio test is suitable for comparing and choosing among specifications.

Advanced issues for maximum likelihood estimators

Imai's maximum likelihood estimator is complex and hence potentially difficult to optimize (so are the other maximum likelihood estimators demonstrated later). None of Stata's built-in optimization algorithms guarantee finding the global maximum. A trick to handle this issue is to optimize a model from different random initial values. `kict ml` with the option `protect(#)` does this trick. Many trials provide reasonable assurance that the result is not a local but a global maximum.

For example, the syntax below requires `kict ml` to perform 51 optimizations (each with 100 iterations at the most). Optimization 0 uses Stata's default initial values;

14. In (10)–(14), $\mathbf{X}_i$ is defined as a set of explanatory variables without a constant term. Elsewhere in this article, $\mathbf{X}_i$ is a set of explanatory variables with a constant term.

optimizations 1–50 use random-initial values. `rseed()` specifies the starting random-number seed. `kict ml` reports the result with the largest log-likelihood value among the trials, but users may add the option `verbose` to the syntax to monitor the result of every trial.

```
. kict ml itemcount south age_rs male college, nnonkey(3) condition(group)
> estimator(imai) protect(50) iterate(100) rseed(123)
Likelihood verification 0   = -1441.019764032303
Likelihood verification 1   = -1443.973781847681 .
Likelihood verification 2   = -1441.019764032303
Likelihood verification 3   = -1441.019764032303
  (output omitted)
```

Furthermore, because of its complexity, Imai's estimator is also prone to the problems of complete separation and boundary values. The quasi-Bayesian approach is a possible solution to these problems. Following Gelman et al. (2008), `kict ml` with the option `pscale()` uses independent Cauchy distributions as weakly informative priors, each centered at zero and with user-specified scales. Before fitting a model by this approach, Gelman et al. suggest centering and rescaling regressors in the first instance. In our example, that is,

```
. quietly summarize    south       // center the binary variable
. generate c_south   = south  -r(mean)
. quietly summarize    age_rs      // center and rescale the continuous variable
. generate cr_age_rs =(age_rs -r(mean))/(r(sd)*2)
. quietly summarize    male
. generate c_male    = male   -r(mean)
. quietly summarize    college
. generate c_college = college-r(mean)
```

The next step is to decide which coefficients need priors and how informative their priors should be. For illustrative purposes, I impose a prior with a scale parameter 2.5 on every slope coefficient in the `Delta` and `Psi0` equations. For intercepts, I use a prior with a scale parameter 10. No prior is for the `Psi1` coefficients because I intend to constrain them and their `Psi0` counterparts to be equal. Accordingly, I generate a matrix of scale parameters, define two constraints, and fit the model by the quasi-Bayesian approach as follows:

```
. matrix cauchyscale = (2.5, 2.5, 2.5, 2.5, 10), /// // for Delta coefficients
>                      (2.5, 2.5, 2.5, 2.5, 10), /// // for Psi0  coefficients
>                      (  .,   .,   .,   .,  .)      // for Psi1  coefficients
. constraint 1 [Psi0=Psi1]:_cons
. constraint 2 [Psi0=Psi1]
. kict ml itemcount c_south cr_age_rs c_male c_college, nnonkey(3)
> condition(group) estimator(imai) iterate(100) constraints(1 2)
> pscale(cauchyscale)
  (output omitted)
```

No matter whether the quasi-Bayesian approach is adopted, `kict ml` always constructs CI for regression coefficients based on the assumption that the sampling distributions of the coefficient estimates are asymptotically normal. `kict pfci` is a postestimation command for producing CI based on the profile likelihood. The profile-based CI does not rest on the normality assumption (Royston 2007) and is hence more robust than the normal-based CI when the sample size is small or when nonnormal priors are used.

`kict pfci` works for one coefficient at a time. For example, the output below is specifically for coefficient `c_South` in equation `Delta`. The normal-based results are identical to those reported in the output of `kict ml` (shown previously). The calculation of the profile-based CI involves reoptimizing the model many times (the option `mliterate(100)` sets that each optimization perform a maximum of 100 iterations).

```
. kict pfci c_south, equation(Delta) mliterate(100)
Finding the upper bound:
Profile iteration 1   [ p= -.244799 i= -1.67593 b=  1.02679 ]
Profile iteration 2   [ p=  .233177 i= -.477976 b=  .149981 ]
Profile iteration 3   [ p= -.006101 i=  .239278 b= -.073167 ]
Profile iteration 4   [ p= -.000142 i= -.005959 b=  .001866 ]
Profile iteration 5   [ p=  9.2e-08 i= -.000143 b=  .000045 ]
Profile iteration 6   [ p= -1.4e-12 i=  9.2e-08 b= -2.9e-08 ]
Finding the lower bound:
Profile iteration 1   [ p=  .173514 i= -2.09424 b= -1.02679 ]
Profile iteration 2   [ p= -.167539 i=  .341053 b=  .085073 ]
Profile iteration 3   [ p= -.003881 i= -.163658 b= -.041791 ]
Profile iteration 4   [ p=  .000091 i= -.003972 b= -.000991 ]
Profile iteration 5   [ p= -4.7e-08 i=  .000091 b=  .000023 ]
Profile iteration 6   [ p= -6.8e-13 i= -4.7e-08 b= -1.2e-08 ]
ICT Model   : Imai MLE
Coefficient : [Delta]:c_south =  1.51914
Original LL = -1468.83335013
Targeted LL = -1470.75407954
```

	Normal-based	Profile-based	Difference
95% CI			
Upper bound	2.54592570	2.62464986	-0.07872416
lower bound	0.49234671	0.53465991	-0.04231320
Interval	2.05357899	2.08998996	-0.03641097
Std. Err.	0.52388182	.	.

All the commands, options, and discussions above apply to the other maximum likelihood estimators demonstrated later in this section.

4.2 The partial item-count technique

To gauge Lebanese attitudes toward voting rights for illiterates, Corstange (2009) uses the partial item-count technique in a Lebanese nationwide face-to-face survey. He requires a quarter of randomly selected respondents to answer each of the following items separately and directly:

> "There has been some debate recently over who should have the right to vote in Lebanese elections. I'll read you some different groups of people: please tell me if they should be allowed to vote or not.
>
> (1) Young people between the ages of 18 to 21;
>
> (2) Lebanese expatriates living abroad;
>
> (3) Palestinians without Lebanese citizenship."

The rest of respondents answer to the question below in the item-count format, where the third item in the list is the key item:

> "There has been some debate recently over who should have the right to vote in Lebanese elections. I'm going to read you the whole list, and then I want you to tell me how many of the different groups you think should be allowed to vote. Don't tell me which ones, just tell me how many.
>
> (1) Young people between the ages of 18 to 21;
>
> (2) Lebanese expatriates living abroad;
>
> (3) Illiterate people;
>
> (4) Palestinians without Lebanese citizenship."

After casewise deletion, the sample size for analysis is 909 (see the `codebook` output). The dataset contains an indicator variable (`group`) that takes on the value 0, representing the short-list group, or 1, representing the long-list group. Because there are three nonkey items and a key item, the item-count response is a variable taking on integers from 0 to 4. The sample size of this variable is 714 because only the long-list group answers the item-count question. The sample size of the variables `direct_young`, `direct_expat`, and `direct_pal` is 195, because only the short-list group answers these nonkey items in the direct questioning format.[15] There are also variables for multivariate analysis. `christian`, `shia`, `sunni`, and `muslinmin` are four mutually exclusive binary variables representing respondents' religions (the following example uses `christian` as the baseline category). `electricity`, `deconfess`, and `education` measure respondents' material well-being, attitudes on fuller democratization, and levels of educational attainment, respectively (Corstange 2009, 58).

15. I thank Daniel Corstange for providing me with this dataset.

```
. use "section4-2-suffrage.dta", clear
(https://doi.org/10.1093/pan/mpn013)
. codebook group itemcount direct_young direct_expat direct_pal
> christian shia sunni muslinmin electricity deconfess education, compact
Variable        Obs Unique      Mean  Min       Max  Label
--------------------------------------------------------------------------------
group           909      2  .7854785    0         1  treatment status
itemcount       714      5  2.536415    0         4  count of items on the long...
direct_young    195      2  .9487179    0         1  suffrage for young people ...
direct_expat    195      2        .8    0         1  suffrage for Lebanese expa...
direct_pal      195      2   .025641    0         1  suffrage for Palestinians ...
christian       909      2  .3223322    0         1  christian
shia            909      2  .3333333    0         1  shia
sunni           909      2  .3146315    0         1  sunni
muslinmin       909      2   .029703    0         1  muslin minority
electricity     909     17  2.544857    0  4.898979  average hour/day the elect...
deconfess       909      2  .6963696    0         1  support for removing the s...
education       909      5  .4642464    0         1  education level (rescaled ...
--------------------------------------------------------------------------------
```

Corstange (2009) conducts an analysis by using a maximum likelihood estimator (6) that assumes that each respondent's item count approximately follows a binomial distribution. `kict ml` with the `estimator(corstange)` and `distribution(binomial)` options performs that estimator with the following setups:

$$
\begin{aligned}
S_i|\mathbf{X}_i = \mathbf{x} &\overset{\text{indep}}{\sim} \text{Bernoulli}\left\{\text{logit}^{-1}(\mathbf{x}'\boldsymbol{\delta})\right\} \\
\text{logit}^{-1}(\mathbf{x}'\boldsymbol{\delta}) &= P(S_i = 1|\mathbf{X}_i = \mathbf{x}) \\
R_{i,j}|\mathbf{X}_i = \mathbf{x} &\overset{\text{indep}}{\sim} \text{Bernoulli}\left\{\text{logit}^{-1}(\mathbf{x}'\boldsymbol{\theta}_j)\right\} \\
\text{logit}^{-1}(\mathbf{x}'\boldsymbol{\theta}_j) &= P(R_{i,j} = 1|\mathbf{X}_i = \mathbf{x}) \\
(S_i + R_i)|\mathbf{X}_i = \mathbf{x} &\overset{\text{indep}}{\sim} \text{Binomial}(J+1, \pi_i) \\
\pi_i &= \left\{\text{logit}^{-1}(\mathbf{x}'\boldsymbol{\delta}) + \sum_{j=1}^{J} \text{logit}^{-1}(\mathbf{x}'\boldsymbol{\theta}_j)\right\} / (J+1)
\end{aligned}
$$

The output below replicates table 1 of Corstange (2009, 59). The `Delta` equation estimates the probability of supporting the enfranchisement of illiterates. The `Theta` equations from 1 to 3 estimate the probabilities of supporting the enfranchisement of three different groups described in nonkey items. Corstange does not use any variable to model the first and third nonkey items. This is equivalent to restricting slope coefficients in the `Theta1` and `Theta3` equations to be zero.

```
. kict ml itemcount shia sunni muslinmin electricity deconfess education,
> cond(group) nn(3) est(cors) direct(direct_young direct_expat direct_pal)
> distribution(binomial) constraint(1 2 3 4 5 6 7) noomitted
  (output omitted)
Corstange MLE                                   Number of obs     =        909
Log likelihood = -951.91779
 ( 1)  [Theta1]shia = 0
 ( 2)  [Theta1]sunni = 0
 ( 3)  [Theta1]muslinmin = 0
 ( 4)  [Theta1]electricity = 0
 ( 5)  [Theta1]deconfess = 0
 ( 6)  [Theta1]education = 0
 ( 7)  [Theta1]shia - [Theta3]shia = 0
 ( 8)  [Theta1]sunni - [Theta3]sunni = 0
 ( 9)  [Theta1]muslinmin - [Theta3]muslinmin = 0
 (10)  [Theta1]electricity - [Theta3]electricity = 0
 (11)  [Theta1]deconfess - [Theta3]deconfess = 0
 (12)  [Theta1]education - [Theta3]education = 0
```

itemcount	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
Delta						
shia	-.3531789	.9537647	-0.37	0.711	-2.222523	1.516166
sunni	-.517838	.9495531	-0.55	0.586	-2.378928	1.343252
muslinmin	-.4424527	1.364906	-0.32	0.746	-3.117619	2.232713
electricity	.8797809	.3104834	2.83	0.005	.2712447	1.488317
deconfess	1.618598	.9684688	1.67	0.095	-.2795657	3.516762
education	-.3511779	1.331978	-0.26	0.792	-2.961806	2.25945
_cons	-1.33745	1.035527	-1.29	0.197	-3.367045	.692145
Theta1						
_cons	2.895317	.3254694	8.90	0.000	2.257408	3.533225
Theta2						
shia	-.0065264	.4387895	-0.01	0.988	-.8665379	.8534852
sunni	.4297074	.4631838	0.93	0.354	-.4781161	1.337531
muslinmin	-.0034925	.8365783	-0.00	0.997	-1.643156	1.636171
electricity	-.411367	.1543246	-2.67	0.008	-.7138376	-.1088964
deconfess	-1.379712	.5458077	-2.53	0.011	-2.449475	-.3099483
education	-1.104436	.8292863	-1.33	0.183	-2.729808	.520935
_cons	3.893472	.7971586	4.88	0.000	2.33107	5.455874
Theta3						
_cons	-3.659585	.4535932	-8.07	0.000	-4.548611	-2.770558

```
*Sum(all of items) ~ Binomial
```

Blair and Imai (2012) improve Corstange's estimator by using a Poisson-binomial distribution to model respondents' item counts. `kict ml` with the `estimator(corstange)` and `distribution(poisbino)` options performs that modified estimator with the setups below. (The output is omitted because its arrangement and interpretation are the same as the previous.)

$$
\begin{aligned}
S_i|\mathbf{X}_i = \mathbf{x} &\overset{\text{indep}}{\sim} \text{Bernoulli}\left\{\text{logit}^{-1}(\mathbf{x}'\boldsymbol{\delta})\right\} \\
\text{logit}^{-1}(\mathbf{x}'\boldsymbol{\delta}) &= P(S_i = 1|\mathbf{X}_i = \mathbf{x}) \\
R_{i,j}|\mathbf{X}_i = \mathbf{x} &\overset{\text{indep}}{\sim} \text{Bernoulli}\left\{\text{logit}^{-1}(\mathbf{x}'\boldsymbol{\theta}_j)\right\} \\
\text{logit}^{-1}(\mathbf{x}'\boldsymbol{\theta}_j) &= P(R_{i,j} = 1|\mathbf{X}_i = x) \\
(S_i + R_i)|\mathbf{X}_i = \mathbf{x} &\overset{\text{indep}}{\sim} \text{Poisson-Binomial}\,(J+1, \\
&\qquad \left[\text{logit}^{-1}(\mathbf{x}'\boldsymbol{\delta}), \left\{\text{logit}^{-1}(\mathbf{x}'\boldsymbol{\theta}_j)\right\}_{j=1}^{J}\right])
\end{aligned}
$$

Blair and Imai's estimator generally outperforms Corstange's when the probability of answering an item in the affirmative varies greatly from one item to another. Nonetheless, note that both estimators assume that items are conditionally independent. It remains unclear about which estimator is more robust to the violation of that assumption.

4.3 Item-count techniques with auxiliary information

Eady (2017, 252–257) investigates Canadians' attitudes toward gender equality through an Internet survey by using both the standard item-count technique and a direct question. A random half of respondents (the long-list group) are prompted with the following item-count question:

> "How many of the following do you agree with?
>
> - There should be more funding for the arts
> - The government should raise taxes on gasoline
> - Corporations are taxed too much
> - Women are as competent as men in politics
> - Unions have too much power"

The other half of respondents (the short-list group) are prompted by the same question without the sensitive item (the fourth one). Later in the same survey, all respondents are required to answer the sensitive question directly. In other words, those in the long-list group answer the key item twice:

> "Do you agree or disagree with the following statement? 'Women are as competent as men in politics.' "

The sample size for analysis is 22,372 (see the `codebook` output below). The dataset contains an item-count variable (`itemcount`) and a group indicator (`group`). A more noteworthy variable is `direct_key`, which recodes respondents' answers to the direct question on the sensitive item. This variable provides auxiliary information for the item-count estimation. There are also five variables for modeling the key item (`ideology`, `gender`, `education`, `mothertongue`, `region`).[16]

```
. use "section4-3-gender.dta", clear
(https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/PZKBUX)
. codebook group itemcount direct_key ideology gender age education mothertongue
> region, compact
Variable        Obs Unique      Mean  Min  Max  Label
-------------------------------------------------------------------------------
group         22372      2  .4974969    0    1  treatment status: 0=short, 1=...
itemcount     22372      6  2.046129    0    5  count of items
direct_key    22372      2  .9773824    0    1  direct question on the key it...
ideology      22372     11  5.800107    0   10  0=right, 10=left
gender        22372      2  .3815931    0    1  0=male, 1=female
age           22372      5   2.22926    0    4  0=18-29, 1=30-39, 2=40-49, 3=...
education     22372      3   1.47993    0    2  0=high school-, 1=college, 2=...
mothertongue  22372      3  .6557304    0    2  0=English, 1=French, 2=other
region        22372      4   1.62766    0    3  0=Ontario, 1=Atlantic, 2=Queb...
-------------------------------------------------------------------------------
```

Aronow et al.'s (2015) difference-in-means estimator

Eady (2017) argues that, while respondents who deny gender equality may misreport their attitudes for avoiding social disapproval, those who agree with the norm should have no incentive to give a false answer. In this regard, Aronow et al.'s difference-in-means estimator with the positive monotonicity assumption (8) is suitable to estimate the prevalence of the belief in gender equality in Canada. `kict ls` with the options `auxiliary()` and `monotony(positive)` can perform this estimation.

16. The author thanks Gregory Eady for releasing this dataset for replication. The dataset can be downloaded from: https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/PZKBUX.

```
. kict ls itemcount, estimator(linear) condition(group) nnonkey(4)
> auxiliary(direct_key) monotony(positive)
```

(*output omitted*)

```
Initial weight matrix: Unadjusted                 Number of obs   =     22,372
GMM weight matrix:     Robust

Linear least squares estimator
monotonic auxiliary information)
```

	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
Delta						
_cons	.8712999	.0113876	76.51	0.000	.8489806	.8936191
Zeta						
_cons	.9773824	.000994	983.25	0.000	.9754342	.9793307
DeltaA0						
_cons	.1516317	.0829693	1.83	0.068	-.0109853	.3142486
GammaA0						
_cons	1.360465	.0597298	22.78	0.000	1.243397	1.477533
DeltaA1						
_cons	.8914626	.0116158	76.75	0.000	.8686961	.9142291
GammaA1						
_cons	1.616624	.0079985	202.12	0.000	1.600947	1.632301

(*output omitted*)

In the output above, `Delta_cons` is the estimated prevalence of the belief in gender equality. `kict ls` produces this estimate based on both the item-count variable `itemcount` and the auxiliary variable `direct_key`. The standard error (0.0113) is almost the same with that of the estimate based only on `itemcount` (0.0115—not displayed in the output). This suggests that `direct_key` does not provide sufficient information for a substantial reduction in estimation uncertainty.

`Zeta_cons` is the estimated prevalence of the belief in gender equality based only on `direct_key`. This estimate is often subject to the misreporting bias. In this example, `Zeta_cons` is higher than `Delta_cons`, suggesting that many people who deny gender equality misreport their attitudes. As for `DeltaA0_cons`, `DeltaA1_cons`, `GammaA0_cons`, and `GammaA1_cons`, these are the estimates of $P(S_i = 1|A_i = 0)$, $P(S_i = 1|A_i = 1)$, $P(R_i = 0|A_i = 0)$, and $P(R_i = 0|A_i = 1)$, respectively.

This difference-in-means estimator rests on two assumptions. First, A_i is independent of T_i. We may test this assumption with `ttest direct_key, by(group)`. The second assumption is that A_i is monotonically related to S_i. If this assumption and the three fundamental assumptions for the item-count technique (footnote 3) all hold true, $\widehat{P}(S_i = 1|A_i = 0)$ must be close to 0 (Aronow et al. 2015, 8–9). Based on this idea, we can jointly test these four assumptions with `test _b[DeltaA0:_cons]=0`.[17]

Eady–Tsai's maximum likelihood estimator

Besides the overall prevalence of the belief in gender equality, Eady is also interested in whether that belief varies along ideological lines. He conducts a regression analysis using the maximum likelihood estimator shown in (9) (hereafter "Eady–Tsai's estimator"). `kict ml` with the options `estimator(tsaieady)` and `auxiliary()` performs that estimator by the setups below:

$$
\begin{aligned}
S_i|\mathbf{X}_i = \mathbf{x} &\overset{\text{indep}}{\sim} \text{Bernoulli}\left\{\text{logit}^{-1}(\mathbf{x}'\boldsymbol{\delta})\right\} \\
\text{logit}^{-1}(\mathbf{x}'\boldsymbol{\delta}) &= P(S_i = 1|\mathbf{X}_i = \mathbf{x}) \\
A_i|S_i = s, \mathbf{X}_i = \mathbf{x} &\overset{\text{indep}}{\sim} \text{Bernoulli}\left\{\text{logit}^{-1}(\mathbf{x}'\boldsymbol{\kappa}_s)\right\} \\
\text{logit}^{-1}(\mathbf{x}'\boldsymbol{\kappa}_s) &= P(A_i = 1|S_i = s, \mathbf{X}_i = \mathbf{x}) \\
R_i|A_i = a, S_i = s, \mathbf{X}_i = \mathbf{x} &\overset{\text{indep}}{\sim} \text{Binomial}\left\{J, \text{logit}^{-1}(\mathbf{x}'\boldsymbol{\psi}_s)\right\} \\
\text{logit}^{-1}(\mathbf{x}'\boldsymbol{\psi}_s) &= P(R_{i,j} = 1|A_i = a, S_i = s, \mathbf{X}_i = \mathbf{x})
\end{aligned}
$$

The option `monotony()` can be specified, if necessary, to impose the monotonicity assumption on $P(A_i = 1|S_i = s, \mathbf{X}_i = \mathbf{x})$.

The output below replicates the second model in table 4 of Eady (2017, 256). Two coefficients are particularly of interest. First, `Delta_ideology` suggests that the more politically left individuals are, the more likely they are to believe in gender equality. The standard error (0.032) is slightly smaller than that of the estimate without auxiliary information (0.041—not displayed in the output). Second, `Kappa0_ideology` suggests that, for those who deny gender equality, the more individuals are to the left politically, the more likely they are to misreport their belief [remember that the `kappa0` equation represents $\widehat{P}(A_i = 1|S_i = 0, \mathbf{X}_i = \mathbf{x})$].

17. Use test `_b[DeltaA1:_cons]=1` to jointly test the negative monotonicity and the three fundamental assumptions.

```
. kict ml itemcount ideology female age30_39 age40_49 age50_64 age65
> edu_college edu_university tongue_french tongue_other
> region_atlantic region_quebec region_west group,
> estimator(tsaieady) condition(group) nnonkey(4) auxiliary(direct_key)
> monotony(positive) constraints(1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16)
> noomitted
  (output omitted)
Imai-Eady-Tsai MLE                               Number of obs     =     22,372
Log likelihood = -30870.813
 ( 1)  [Psi0]ideology - [Psi1]ideology = 0
 ( 2)  [Psi0]female - [Psi1]female = 0
 ( 3)  [Psi0]age30_39 - [Psi1]age30_39 = 0
 ( 4)  [Psi0]age40_49 - [Psi1]age40_49 = 0
 ( 5)  [Psi0]age50_64 - [Psi1]age50_64 = 0
 ( 6)  [Psi0]age65 - [Psi1]age65 = 0
 ( 7)  [Psi0]edu_college - [Psi1]edu_college = 0
 ( 8)  [Psi0]edu_university - [Psi1]edu_university = 0
 ( 9)  [Psi0]tongue_french - [Psi1]tongue_french = 0
 (10)  [Psi0]tongue_other - [Psi1]tongue_other = 0
 (11)  [Psi0]region_atlantic - [Psi1]region_atlantic = 0
 (12)  [Psi0]region_quebec - [Psi1]region_quebec = 0
 (13)  [Psi0]region_west - [Psi1]region_west = 0
 (14)  [Psi0]group = 0
 (15)  [Psi1]group = 0
 (16)  [Delta]group = 0
------------------------------------------------------------------------------------
       itemcount |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----------------+------------------------------------------------------------------
Delta            |
        ideology |   .2111368   .0323968     6.52   0.000     .1476402    .2746334
          female |   .6171109   .2130978     2.90   0.004     .1994468    1.034775
        age30_39 |  -.2401679   .3927619    -0.61   0.541    -1.009967    .5296313
        age40_49 |  -.5068702    .392248    -1.29   0.196    -1.275662    .2619217
        age50_64 |  -.5119114   .3513178    -1.46   0.145    -1.200482    .1766587
           age65 |  -.9533061   .3572732    -2.67   0.008    -1.653549   -.2530636
     edu_college |   .0966809   .2425528     0.40   0.690    -.3787139    .5720756
  edu_university |   .7246482   .2406512     3.01   0.003     .2529806    1.196316
   tongue_french |   .3620602    .298036     1.21   0.224    -.2220796       .9462
    tongue_other |   .0890361   .2524217     0.35   0.724    -.4057013    .5837735
 region_atlantic |    .960454   .7531654     1.28   0.202     -.515723    2.436631
   region_quebec |     .13694   .3159932     0.43   0.665    -.4823952    .7562752
     region_west |   .0281941   .2306993     0.12   0.903    -.4239681    .4803564
                 |
           _cons |   .8550994   .5192563     1.65   0.100    -.1626243    1.872823
------------------------------------------------------------------------------------
```

Kappa0						
ideology	.1817083	.0450091	4.04	0.000	.0934921	.2699244
female	.3807	.2878155	1.32	0.186	-.183408	.9448081
age30_39	.2852318	.6576717	0.43	0.665	-1.003781	1.574245
age40_49	.8709722	.6244763	1.39	0.163	-.3529789	2.094923
age50_64	1.111552	.5759109	1.93	0.054	-.0172123	2.240317
age65	1.372159	.579951	2.37	0.018	.2354757	2.508842
edu_college	.0089005	.312138	0.03	0.977	-.6028786	.6206797
edu_university	-.2811058	.3197367	-0.88	0.379	-.9077783	.3455667
tongue_french	.0315154	.3987122	0.08	0.937	-.7499461	.8129769
tongue_other	-.2810905	.3282465	-0.86	0.392	-.9244418	.3622608
region_atlantic	-.9609013	1.307092	-0.74	0.462	-3.522754	1.600951
region_quebec	.1628963	.4181187	0.39	0.697	-.6566013	.9823939
region_west	-.5366885	.3199466	-1.68	0.093	-1.163772	.0903952
group	.0169536	.1289836	0.13	0.895	-.2358496	.2697568
_cons	-.4611224	.8003032	-0.58	0.564	-2.029688	1.107443
Psi0						
ideology	.0192199	.0032666	5.88	0.000	.0128175	.0256222
female	-.0352134	.0147167	-2.39	0.017	-.0640576	-.0063693
age30_39	-.0591878	.0238349	-2.48	0.013	-.1059033	-.0124722
age40_49	-.0776767	.0256279	-3.03	0.002	-.1279065	-.027447
age50_64	-.1755678	.0218041	-8.05	0.000	-.2183031	-.1328325
age65	-.1111554	.0233784	-4.75	0.000	-.1569761	-.0653346
edu_college	-.0314031	.0245304	-1.28	0.200	-.0794818	.0166756
edu_university	.1686017	.0215607	7.82	0.000	.1263435	.2108598
tongue_french	-.0150518	.0243219	-0.62	0.536	-.0627218	.0326183
tongue_other	.0088287	.0216414	0.41	0.683	-.0335876	.0512451
region_atlantic	-.1037261	.0346632	-2.99	0.003	-.1716648	-.0357874
region_quebec	-.032388	.0249119	-1.30	0.194	-.0812144	.0164384
region_west	-.0521116	.0186764	-2.79	0.005	-.0887166	-.0155065
direct_key	.4063125	.0760064	5.35	0.000	.2573428	.5552823
_cons	-.5868745	.0575188	-10.20	0.000	-.6996092	-.4741397
Psi1						
ideology	.0192199	.0032666	5.88	0.000	.0128175	.0256222
female	-.0352134	.0147167	-2.39	0.017	-.0640576	-.0063693
age30_39	-.0591878	.0238349	-2.48	0.013	-.1059033	-.0124722
age40_49	-.0776767	.0256279	-3.03	0.002	-.1279065	-.027447
age50_64	-.1755678	.0218041	-8.05	0.000	-.2183031	-.1328325
age65	-.1111554	.0233784	-4.75	0.000	-.1569761	-.0653346
edu_college	-.0314031	.0245304	-1.28	0.200	-.0794818	.0166756
edu_university	.1686017	.0215607	7.82	0.000	.1263435	.2108598
tongue_french	-.0150518	.0243219	-0.62	0.536	-.0627218	.0326183
tongue_other	.0088287	.0216414	0.41	0.683	-.0335876	.0512451
region_atlantic	-.1037261	.0346632	-2.99	0.003	-.1716648	-.0357874
region_quebec	-.032388	.0249119	-1.30	0.194	-.0812144	.0164384
region_west	-.0521116	.0186764	-2.79	0.005	-.0887166	-.0155065
_cons	-.4939727	.0388722	-12.71	0.000	-.5701609	-.4177845

```
*Sum(nonkey items) ~ Binomial
*Positive monotony : Pr(Auxiliary=1 | key item=1) = 1
```

Several technical points are noteworthy. First, the model above does not estimate the `kappa1` equation because of the positive monotonicity assumption. (`kappa1` is relating to

$\widehat{P}(A_i = 1|S_i = 1, \mathbf{X}_i = \mathbf{x})$, which is fixed at 1 in this example.)[18] Second, Eady includes `group` into the `kappa0` equation as a covariate to release the assumption that A_i and T_i are independent. Eady, however, does not include `group` into the other equations, so the model constrains the respect coefficients to be zero. Third, when `monotony(positive)` is specified, `kict ml` automatically drops A_i (`direct_key`) from the `Psi1` equation to prevent perfect collinearity. Likewise, under the negative monotonicity assumption, A_i is dropped from the `Psi0` equation. Fourth, `kict ml` by default assumes that R_i follows a binomial distribution (see *Imai's maximum likelihood estimator* on page 411 for the implications of this assumption). Moreover, Eady assumes that R_i and S_i are correlated (that is, allowing `Psi0_cons` $\neq$ `Psi1_cons`), but there is no moderate effect of $\mathbf{X}_i$ on that correlation (that is, restricting `[Psi0=Psi1]` for all slopes except those for group and `direct_key`).

4.4 The dual-list item-count technique

Consider the dual-list design presented in table 1: two item-count questions having the same key item but different nonkey items—three nonkey items for Q^A and two for Q^B. To demonstrate this design, I construct an artificial dataset as follows:

```
. clear all
. quietly set obs 1000
. set seed 123
. generate x1 = rnormal()                 // 1st independent variable (observed)
. generate x2 = rbinomial(1, 0.5)         // 2nd independent variable (observed)
. generate x3 = rbeta(5, 10)              // 3rd independent variable (observed)
. generate group = cond(_n<=_N*0.5, 1, 0) // group indicator
. label define group 0 "group0" 1 "group1"
. label values group group
. generate groupA = group
. generate groupB = 1 - group
. generate key = cond(1+2*x1-3*x2+0*x3+logit(runiform())>0,1,0)
> // key item (unobserved)
. generate nonkeyA1 = rbinomial(1, 0.2) // 1st nonkey item of list A (unobserved)
. generate nonkeyA2 = rbinomial(1, 0.5) // 2nd nonkey item of list A (unobserved)
. generate nonkeyA3 = rbinomial(1, 0.8) // 3rd nonkey item of list A (unobserved)
. generate nonkeyB1 = rbinomial(1, 0.3) // 1st nonkey item of list B (unobserved)
. generate nonkeyB2 = rbinomial(1, 0.7) // 2nd nonkey item of list B (unobserved)
. // answers to the item-count question A (observed)
. generate itemcountA = cond(group==1, nonkeyA1+nonkeyA2+nonkeyA3+key,
>                                      nonkeyA1+nonkeyA2+nonkeyA3)
. // answers to the item-count question B (observed)
. generate itemcountB = cond(group==0, nonkeyB1+nonkeyB2+key,
>                                      nonkeyB1+nonkeyB2)
```

18. By default, `kict ml` does not make any monotonicity assumption and estimates both the `kappa0` and `kappa1` equations. In that case, discussions of (10)–(14) apply here for the specification of `kappa0` and `kappa1`, except that now $\kappa_{0\alpha}$, $\kappa_{0\beta}$, $\kappa_{1\alpha}$, $\kappa_{0\beta}$, and A_i substitute for $\psi_{0\alpha}$, $\psi_{0\beta}$, $\psi_{1\alpha}$, $\psi_{1\beta}$, and R_i, respectively.

The output below demonstrates that each item-count variable is sufficient to produce an estimate of the population average response to the key item (`Est_QA` and `Est_QB`). The mean of these two estimates produces a more precise estimate: the standard error of `Est_Dual` is smaller. Note that the third and fourth lines in the `nlcom` syntax are the dual-list version of the difference-in-means estimator (5). (The first and second lines in the `nlcom` syntax are unnecessary if you are not interested in the separate estimates.)

```
. quietly mean itemcountA itemcountB, over(group)
. nlcom (Est_QA: _b[itemcountA:group1]-_b[itemcountA:group0])
>       (Est_QB: _b[itemcountB:group0]-_b[itemcountB:group1])
>       (Est_Dual:(_b[itemcountA:group1]-_b[itemcountA:group0] +
>                _b[itemcountB:group0]-_b[itemcountB:group1])/2)
```

(*output omitted*)

Mean	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
Est_QA	.486	.0512634	9.48	0.000	.3855256	.5864744
Est_QB	.46	.0470178	9.78	0.000	.3678467	.5521533
Est_Dual	.473	.0346619	13.65	0.000	.4050639	.5409361

`kict ls` with the options `estimator(linear)` and `duallist` is a shortcut to performing the dual-list version of the difference-in-means estimator, as demonstrated below. The numbers 3 and 2 in `nnonkey()` are the numbers of the nonkey items of `itemcountA` and `itemcountB`, respectively. `Delta_cons` in the output below is the same as `Est_Dual` shown above. `GammaA_cons` is the predicted number of affirmative answers to the nonkey items of Q^A, and `GammaB_cons` is the predicted number of affirmative answers to the nonkey items of Q^B.

```
. kict ls itemcountA itemcountB, nnonkey(3 2) condition(group) estimator(linear)
> duallist vce(robust)
```

(*output omitted*)

```
Initial weight matrix: Unadjusted                Number of obs   =      1,000
GMM weight matrix:     Robust

Linear least squares estimator
(Dual-list design)
```

	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
Delta						
_cons	.473	.0346297	13.66	0.000	.4051271	.5408729
GammaA						
_cons	1.416	.0337918	41.90	0.000	1.349769	1.482231
GammaB						
_cons	1.014	.0284887	35.59	0.000	.9581631	1.069837

(*output omitted*)

`kict ls` can also help to conduct multivariate analysis for the dual-list design. For example, the following syntax performs linear and nonlinear least-squares regressions (15) to model the key item by three covariates, `x1`, `x2`, and `x3`:

```
. kict ls itemcountA itemcountB x1 x2 x3, nnonkey(3 2) condition(group)
> estimator(linear) duallist vce(robust)
  (output omitted)
. kict ls itemcountA itemcountB x1 x2 x3, nnonkey(3 2) condition(group)
> estimator(nonlinear) duallist vce(robust)
  (output omitted)
```

Regarding maximum likelihood estimation, `kict lm` with the options `duallist` and `estimator(imai)` performs the estimator derived in (16). The first part of the following syntax performs that estimator; the second and third parts of the syntax perform the single-list version of Imai's maximum likelihood estimator for Q^A and Q^B, respectively. The variances of the `Delta` coefficients in `IML_Dual` are roughly half as large as those in `IML_A` and `IML_B`. This example clearly illustrates how the dual-list design can help to reduce estimation uncertainty.

```
. quietly kict ml itemcountA itemcountB  x1 x2 x3, nnonkey(3 2)
> condition(group) estimator(imai) dual
. estimate store IML_Dual
. quietly kict ml itemcountA x1 x2 x3, nnonkey(3) condition(groupA)
> estimator(imai)
. estimate store IML_QA
. quietly kict ml itemcountB x1 x2 x3, nnonkey(2) condition(groupB)
> estimator(imai)
. estimate store IML_QB
. quietly estimates table IML_QA IML_QB IML_Dual, se
. matlist r(coef)
```

	IML_QA		IML_QB		IML_Dual	
	b	var	b	var	b	var
Delta						
x1	1.596935	.2382511	2.439722	.2071857	2.058573	.1160627
x2	-2.712104	.5736409	-4.084287	.6506392	-3.386827	.3206916
x3	.7659092	7.468236	6.666645	7.089496	3.86412	3.457439
_cons	.979539	.8184387	-.9694352	.8468057	-.0054919	.3895544
Psi0						
x1	.0879528	.0053847	-.0963225	.0072097	.z	.z
x2	-.1514808	.0222804	.1209527	.0312425	.z	.z
x3	.8357167	.2342122	-.1528504	.3109436	.z	.z
_cons	-.0887079	.0489448	-.0064979	.0524877	.z	.z
Psi1						
x1	.2032532	.0061563	.1184679	.0155941	.z	.z
x2	-.2188636	.026593	.0256345	.0695691	.z	.z
x3	.0770539	.2844814	.2218973	.4887539	.z	.z
_cons	-.3239008	.039845	-.1008481	.0822806	.z	.z

PsiA0						
x1	.z	.z	.z	.z	.0897869	.0049914
x2	.z	.z	.z	.z	-.1547374	.0200325
x3	.z	.z	.z	.z	.818051	.2108259
_cons	.z	.z	.z	.z	-.0763265	.0395926
PsiA1						
x1	.z	.z	.z	.z	.1882043	.0068843
x2	.z	.z	.z	.z	-.1782136	.0278689
x3	.z	.z	.z	.z	-.0536808	.2770275
_cons	.z	.z	.z	.z	-.2681234	.0425379
PsiB0						
x1	.z	.z	.z	.z	-.1012927	.0076713
x2	.z	.z	.z	.z	.1233779	.0326852
x3	.z	.z	.z	.z	-.0632922	.320587
_cons	.z	.z	.z	.z	-.0487254	.0575907
PsiB1						
x1	.z	.z	.z	.z	.170599	.0128146
x2	.z	.z	.z	.z	-.1012491	.0508176
x3	.z	.z	.z	.z	.5220651	.4272538
_cons	.z	.z	.z	.z	-.2429835	.0663069

4.5 The item-sum technique

To demonstrate the item-sum design illustrated in *The item-sum technique* on page 397, I construct an artificial dataset as follows:

```
. clear all
. quietly set obs 1000
. set seed 123
. generate x1 = rnormal()         // 1st independent variable (observed)
. generate x2 = rbinomial(1, 0.5) // 2nd independent variable (observed)
. generate x3 = rbeta(5, 10)      // 3rd independent variable (observed)
. generate group = cond(_n<=_N*0.5, 1, 0)       // group indicator
. label define group 0 "short_list" 1 "long_list"
. label value group group
. generate key = 1+50*x1-100*x2+rnormal(200,100) // key item (unobserved):
. replace  key = 0 if key<0                      // earnings from undeclared work
(119 real changes made)
. generate nonkey1 = rnormal(500,50)             // 1st nonkey item (unobserved)
. replace  nonkey1 = 0 if nonkey1<0              // expenditure on housing
(0 real changes made)
. generate nonkey2 = rnormal(15,10)              // 2nd nonkey item (unobserved)
. replace  nonkey2 = 0 if nonkey2<0              // expenditure on house
(61 real changes made)
. // answers to the item-sum question (observed)
. generate sumofitems = cond(group==1, nonkey1+nonkey2+key, nonkey1+nonkey2)
```

`kict ls` with the options `estimator(linear)` and `itemsum` performs linear least-squares estimation for the key item. In the artificial example, `Delta_cons` is 147.9, suggesting that the population average earnings from undeclared work is around €147.9 per month. `Gamma_cons` suggests that, on average, an individual's monthly expenditure on housing plus weekly hours of TV watching is 516.4.

```
. kict ls sumofitems, nnonkey(2) condition(group) estimator(linear) itemsum
  (output omitted)
Initial weight matrix: Unadjusted                  Number of obs   =      1,000
GMM weight matrix:     Robust

Linear least squares estimator
------------------------------------------------------------------------------
             |               Robust
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Delta        |
       _cons |   147.9285   5.650742    26.18   0.000     136.8532    159.0037
-------------+----------------------------------------------------------------
Gamma        |
       _cons |   516.3799    2.32784   221.83   0.000     511.8174    520.9424
------------------------------------------------------------------------------
Instruments for equation Delta: _cons
Instruments for equation Gamma: _cons
```

On multivariate analysis, the interpretation of the `Delta` coefficients is the same as the interpretation of the linear regression coefficients. Suppose that `x2` is a variable of gender taking on the value 1 for men or 0 for women. `Delta_x2` in the output below indicates that, ceteris paribus, men's average monthly earnings from the undeclared work are €74.9 less than women's.

```
. kict ls sumofitems x1 x2 x3, nnonkey(2) condition(group) estimator(linear)
> itemsum
  (output omitted)
Initial weight matrix: Unadjusted                  Number of obs   =      1,000
GMM weight matrix:     Robust

Linear least squares estimator
------------------------------------------------------------------------------
             |               Robust
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Delta        |
          x1 |   36.94052   5.190272     7.12   0.000     26.76777    47.11327
          x2 |  -74.92421   10.09597    -7.42   0.000    -94.71194   -55.13648
          x3 |  -54.58159   40.84395    -1.34   0.181    -134.6343    25.47108
       _cons |   206.4933   15.65327    13.19   0.000     175.8135    237.1732
-------------+----------------------------------------------------------------
Gamma        |
          x1 |    4.23793   2.388061     1.77   0.076    -.4425832    8.918443
          x2 |  -6.113935   4.631708    -1.32   0.187    -15.19192    2.964045
          x3 |   22.23792   18.58975     1.20   0.232    -14.19732    58.67316
       _cons |   511.7117    7.05401    72.54   0.000     497.8861    525.5373
------------------------------------------------------------------------------
Instruments for equation Delta: x1 x2 x3 _cons
Instruments for equation Gamma: x1 x2 x3 _cons
```

5 Concluding remarks

In this article, I reviewed recent developments of the item-count technique and introduced a package, `kict`, for statistical analysis of item-count data. The review provides a general guide to the applications of the technique, and the package offers every convenience for data analysis. However, note that every item-count technique and their respective statistical methods rest on certain assumptions and have their own limitations. Although I gave an overview of those issues, applied researchers should still refer to and familiarize themselves with the relevant literature to avoid any misuse.

The `kict` package has provided many commonly used methods for researchers to analyze several popular types of item-count data. Undeniably, there are methods that have not been built in `kict`, especially those that are overly individualized such as the estimators with corrections for certain bias (for example, Blair and Imai [2012, 66–68]; Blair, Chou, and Imai [2018]). Those estimators are difficult to program because their formulas are highly contingent upon individual researchers' assumptions about the bias to be corrected. Of course, I could have programmed those estimators based on one or two assumptions, but the resulting commands would have been severely limited and of little practical use. Abuse is also a matter of concern because careless users may misperceive those commands as "one-size-fits-all" corrections. I decided to postpone programming overly individualized methods but will resume doing it after formulating an appropriate strategy for programming.

The development of the item-count technique is still ongoing. There are more and more interesting variant designs and new methods for data analysis (for example, Chaudhuri and Christofides [2013, 115–150]; Blair, Imai, and Lyall [2014]; Chou, Imai, and Rosenfeld [Forthcoming]; Ibrahim [2016]; Tsai [2018]). I will continue to update the `kict` package accordingly.

6 References

Andrews, D. W. K., and G. Soares. 2010. Inference for parameters defined by moment inequalities using generalized moment selection. *Econometrica* 78: 119–157.

Aronow, P. M., A. Coppock, F. W. Crawford, and D. P. Green. 2015. Combining list experiment and direct question estimates of sensitive behavior prevalence. *Journal of Survey Statistics and Methodology* 3: 43–66.

Blair, G., W. Chou, and K. Imai. 2018. List experiments with measurement error. Technical report. Princeton University. https://imai.fas.harvard.edu/research/files/listerror.pdf.

Blair, G., and K. Imai. 2012. Statistical analysis of list experiments. *Political Analysis* 20: 47–77.

Blair, G., K. Imai, and J. Lyall. 2014. Comparing and combining list and endorsement experiments: Evidence from Afghanistan. *American Journal of Political Science* 58: 1043–1063.

Chaudhuri, A., and T. C. Christofides. 2013. *Indirect Questioning in Sample Surveys*. Berlin: Springer.

Chou, W., K. Imai, and B. Rosenfeld. Forthcoming. Sensitive survey questions with auxiliary information. *Sociological Methods and Research*.

Corstange, D. 2009. Sensitive questions, truthful answers? Modeling the list experiment with LISTIT. *Political Analysis* 17: 45–63.

Droitcour, J., R. A. Caspar, M. L. Hubbard, T. L. Parsley, W. Visscher, and T. M. Ezzati. 1991. The item count technique as a method of indirect questioning: A review of its development and a case study application. In *Measurement Errors in Surveys*, ed. P. P. Biemer, R. M. Groves, L. E. Lyberg, N. A. Mathiowetz, and S. Sudman, 185–210. Hoboken, NJ: Wiley.

Drukker, D. 2014. Using gmm to solve two-step estimation problems. The Stata Blog: Not Elsewhere Classified. http://blog.stata.com/2014/12/08/using-gmm-to-solve-two-step-estimation-problems/.

Eady, G. 2017. The statistical analysis of misreporting on sensitive survey questions. *Political Analysis* 25: 241–259.

Gelman, A. 2014. Thinking of doing a list experiment? Here's a list of reasons why you should think again. Statistical Modeling, Causal Inference, and Social Science. https://andrewgelman.com/2014/04/23/thinking-list-experiment-heres-list-reasons-think/.

Gelman, A., A. Jakulin, M. G. Pittau, and Y.-S. Su. 2008. A weakly informative default prior distribution for logistic and other regression models. *Annals of Applied Statistics* 2: 1360–1383.

Glynn, A. N. 2013. What can we learn with statistical truth serum? Design and analysis of the list experiment. *Public Opinion Quarterly* 77: 159–172.

Holbrook, A. L., and J. A. Krosnick. 2010. Social desirability bias in voter turnout reports: Tests using the item count technique. *Public Opinion Quarterly* 74: 37–67.

Ibrahim, F. 2016. An alternative modified item count technique in sampling survey. *International Journal of Statistics and Applications* 6: 177–187.

Imai, K. 2011. Multivariate regression analysis for the item count technique. *Journal of the American Statistical Association* 106: 407–416.

Jann, B. 2005. rrlogit: Stata module to estimate logistic regression for randomized response data. Statistical Software Components S456203, Department of Economics, Boston College. https://ideas.repec.org/c/boc/bocode/s456203.html.

Kadane, J. B. 2016. Sums of possibly associated Bernoulli variables: The Conway–Maxwell-binomial distribution. *Bayesian Analysis* 11: 403–420.

Kudo, A. 1963. A multivariate analogue of the one-sided test. *Biometrika* 50: 403–418.

Kuklinski, J. H., M. D. Cobb, and M. Gilens. 1997. Racial attitudes and the "new south". *Journal of Politics* 59: 323–349.

Long, J. S. 1997. *Regression Models for Categorical and Limited Dependent Variables.* Thousand Oaks, CA: Sage.

Royston, P. 2007. Profile likelihood for estimation and confidence intervals. *Stata Journal* 7: 376–387.

Sniderman, P. M., P. E. Tetlock, and T. Piazza. 1991. Data of 1991 race and politics survey. Survey Research Centre, University of California, Berkeley. http://sda.berkeley.edu/cgi-bin/hsda?harcsda+natlrace.

Trappmann, M., I. Krumpal, A. Kirchner, and B. Jann. 2014. Item Sum: A New Technique for Asking Quantitative Sensitive Questions. *Journal of Survey Statistics and Methodology* 2: 58–77.

Tsai, C.-l. 2017. Solutions to turnout over-reporting: What is out there, what works, and can we do better? Doctoral dissertation, University of Essex.

———. 2018. "A random-wording design for improving the item-count technique". Paper presented at the 2018 Annual Meeting of the American Political Science, August 30–September 2, 2018, Boston, MA.

Von Hermanni, H. 2016. "List experiment: Past and present". Paper presented at the Conference of European Statistics Stakeholders, October 20–21, 2016, Budapest, Hungary.

Wolter, F., and B. Laier. 2014. The effectiveness of the item count technique in eliciting valid answers to sensitive questions. An evaluation in the context of self-reported delinquency. *Survey Research Methods* 8: 153–168.

About the author

Chi-lin Tsai is a postdoctoral research fellow at the Department of Political Science at National Cheng-chi University, Taiwan. He received his PhD in political science from the University of Essex in the UK in 2017. His research interests lie in the intersection of survey methodology, electoral studies, and comparative politics.

Appendix

I first modify Imai's least-squares estimators [(2) and (3)] to accommodate to the dual-list design of the item-count technique. I continue to use the notations introduced in *The dual-list item-count technique* on page 394 and additionally define $R^A_{i,j}$ and $R^B_{i,k}$ as respondent i's potential answers to the jth and kth nonkey items of Q^A and Q^B, respectively, where $j = 1, \ldots, J$, and $k = 1, \ldots, K$. Accordingly, the respondent's answers to Q^A and Q^B are $Y_i^A = R_i^A + T_i S_i$ and $Y_i^B = R_i^B + (1 - T_i)S_i$, respectively, where $R_i^A = \sum_{j=1}^{J} R_{i,j}^A$ and $R_i^B = \sum_{k=1}^{K} R_{i,k}^B$. Moreover, let $Y_i^L = T_i Y_i^A + (1 - T_i)$ Y_i^B. I consider the following simultaneous equation model for analyzing the dual-list item-count data:

$$
\begin{aligned}
Y_i^A &= f(\mathbf{X}_i, \boldsymbol{\gamma}^A) + T_i g(\mathbf{X}_i, \boldsymbol{\delta}) + \epsilon_i^A \\
Y_i^B &= h(\mathbf{X}_i, \boldsymbol{\gamma}^B) + (1 - T_i) g(\mathbf{X}_i, \boldsymbol{\delta}) + \epsilon_i^B
\end{aligned}
\tag{15}
$$

NOTES:

a) $g(\mathbf{X}_i, \boldsymbol{\delta}) = E(S_i|\mathbf{X}_i) = P(S_i = 1|\mathbf{X}_i)$.
$f(\mathbf{X}_i, \boldsymbol{\gamma}^A) = E(R_i^A|\mathbf{X}_i)$.
$h(\mathbf{X}_i, \boldsymbol{\gamma}^B) = E(R_i^B|\mathbf{X}_i)$.

b) Additional assumptions: $E(\epsilon_i^A|\mathbf{X}_i, T_i) = 0$ and $E(\epsilon_i^B|\mathbf{X}_i, T_i) = 0$.

In the linear least-squares estimation, $g(\cdot)$, $f(\cdot)$, and $h(\cdot)$ are identity-link functions. For nonlinear estimation, `kict ls` defines $g(\mathbf{X}_i, \boldsymbol{\delta}) = \operatorname{logit}^{-1}(\mathbf{x}'\boldsymbol{\delta})$, $f(\mathbf{X}_i, \boldsymbol{\gamma}^A) = J \operatorname{logit}^{-1}(\mathbf{x}'\boldsymbol{\gamma}^A)$, and $h(\mathbf{X}_i, \boldsymbol{\gamma}^B) = K \operatorname{logit}^{-1}(\mathbf{x}'\boldsymbol{\gamma}^B)$. Conceptually, the coefficients are estimated through the following three-step procedure:

$$
\text{Step 1} : \widehat{\boldsymbol{\gamma}}^A = \arg\min_{\boldsymbol{\gamma}^A} \sum_{i \in (T_i = 0)} \left\{ Y_i^A - f(\mathbf{X}_i, \boldsymbol{\gamma}^A) \right\}^2
$$

$$
\text{Step 2} : \widehat{\boldsymbol{\gamma}}^B = \arg\min_{\boldsymbol{\gamma}^B} \sum_{i \in (T_i = 1)} \left\{ Y_i^B - h\left(\mathbf{X}_i, \boldsymbol{\gamma}^B\right) \right\}^2
$$

$$
\text{Step 3} : \widehat{\boldsymbol{\delta}} = \arg\min_{\boldsymbol{\delta}} \sum_{i=1}^{N} \left\{ Y_i^L - T_i f\left(\mathbf{X}_i, \widehat{\boldsymbol{\gamma}}^A\right) - (1 - T_i)\, h\left(\mathbf{X}_i, \widehat{\boldsymbol{\gamma}}^B\right) - g\left(\mathbf{X}_i, \boldsymbol{\delta}\right) \right\}^2
$$

Computationally, `kict ls` converts these steps into moment conditions and uses the generalized method of moments (`gmm`) for estimation (see also Drukker [2014]; Imai [2011, 415]).

Second, I attempt to modify Imai's maximum likelihood estimator (4) for the dual-list item-count technique. Equation (16) shows the likelihood function of one possible modification:

$$
\begin{aligned}
&\left[\prod_{i=1}^{n}\sum_{s=0}^{1}\left\{P\left(R_i^A=Y_i^A-sT_i|S_i=s,\mathbf{X}_i=\mathbf{x},\boldsymbol{\psi}_s^A\right)P\left(S_i=s|\mathbf{X}_i=\mathbf{x},\boldsymbol{\delta}\right)\right.\right.\\
&\qquad\left.\left.\times\left(\mathbf{1}_{Y_i^A\neq J+1}\right)^{1-s}\left(\mathbf{1}_{Y_i^A\neq 0\cup T_i=1}\right)^{s}\right\}\right]\\
&\qquad\times\left(\prod_{i=1}^{n}\sum_{s=0}^{1}\left[P\left\{R_i^B=Y_i^B-s\left(1-T_i\right)|S_i=s,\mathbf{X}_i=\mathbf{x},\boldsymbol{\psi}_s^B\right\}P\left(S_i=s|\mathbf{X}_i=\mathbf{x},\boldsymbol{\delta}\right)\right.\right.\\
&\qquad\left.\left.\times\left(\mathbf{1}_{Y_i^B\neq K+1}\right)^{1-s}\left(\mathbf{1}_{Y_i^B\neq 0\cup T_i=0}\right)^{s}\right]\right) \qquad (16)
\end{aligned}
$$

NOTES:

a) $\mathbf{1}_{\text{Condition}}=1$ if Condition holds; $\mathbf{1}_{\text{Condition}}=0$ otherwise.
b) Additional assumptions: the distributions of R_i^A and R_i^B.

Third, I apply Eady and Tsai's idea about auxiliary information to Corstange's partial item-count technique. Equation (17) is the likelihood function:

$$
\begin{aligned}
&\prod_{i\in(T_i=0)}\left\{\prod_{j=1}^{J}P\left(R_{i,j}=r_j|A_i=a,\mathbf{X}_i=\mathbf{x},\boldsymbol{\theta}_j\right)\right.\\
&\qquad\left.\times\sum_{s=0}^{1}P(A_i=a|S_i=s,\mathbf{X}_i=\mathbf{x},\boldsymbol{\kappa}_s)P(S_i=s|\mathbf{X}_i=\mathbf{x})\right\}\\
&\qquad\times\prod_{i\in(T_i=1)}\sum_{s=0}^{1}\left\{\left(\mathbf{1}_{Y_i\neq J+1}\right)^{1-s}\left(\mathbf{1}_{Y_i\neq 0}\right)^{s}\right.\\
&\qquad\times\left\{\sum_{u\in U^{Y_i-s}}\prod_{j=1}^{J}P\left(R_{i,j}=r_{u,j}|A_i=a,\mathbf{X}_i=\mathbf{x},\boldsymbol{\theta}_j\right)\right\}\\
&\qquad\left.\times P\left(A_i=a|S_i=s,\mathbf{X}_i=\mathbf{x},\boldsymbol{\kappa}_s\right)P\left(S_i=s|\mathbf{X}_i=\mathbf{x},\boldsymbol{\delta}\right)\right\} \qquad (17)
\end{aligned}
$$

NOTES:

a) $\mathbf{1}_{\text{Condition}}=1$ if Condition holds; $\mathbf{1}_{\text{Condition}}=0$ otherwise.
b) U^{Y_i-s} is a set of combinations of $R_{i,j}$ that satisfy $R_i=Y_i-s$.
c) u is one of the combinations in U^{Y_i-s}.
$r_{u,j}=1$ if $R_{i,j}=1$ in u; $r_{u,j}=0$ otherwise.
d) Additional assumptions:
   - A_i is independent of T_i.
   - A_i is predictive of S_i.
   - A_i is extraneous to S_i.
   - S_i, $R_{i,j}$, and $R_{i,k} (j\neq k)$ are independent after controlling for $\mathbf{X}_i$.