The impact of COVID-19 lockdowns on physical domestic violence: Evidence from a list randomization experiment

Catherine Porter [a,*], Marta Favara [b], Alan Sánchez [c], Douglas Scott [b]

[a] *University of Lancaster, LA1 4YX, Lancaster, United Kingdom*
[b] *University of Oxford, United Kingdom*
[c] *Grupo de Análisis para El Desarrollo (GRADE), Peru*

ARTICLE INFO

ABSTRACT

We quantify the increase in physical domestic violence (family or intimate partner violence) experienced by young people aged 18–26 during the 2020 COVID-19 lockdowns in Peru. To do this we use an indirect methodology, the double list randomization experiment. The list experiment was embedded in a telephone survey to participants of the Young Lives study, a long-standing cohort survey. We find that 8.3% of the sample experienced an increase in physical violence within their households during the lockdown period. Those who had already reported experiencing domestic violence in the last round of (in-person) data collection in 2016 are more likely to have experienced increased physical violence during the COVID-19 lockdown, with 23.6% reporting an increase during this time. The reported increase in violence does not differ significantly by gender. List experiments, if carefully conducted, may be a relatively cheap and feasible way to elicit information about sensitive issues during a phone survey.

1. Introduction

Domestic violence was a global problem before 2020, and government policies enacted by countries around the world to reduce the spread of the COVID-19 virus, including stay-at-home requirements or lockdowns, may have exacerbated the problem. Our intended aim is to generate timely evidence about the impact of the COVID-19 lockdowns on young people's experiences of physical domestic violence in Peru,[1] a country which has been one of the hardest hit by the pandemic, in terms of both cases and fatalities per capita, while also operating extremely strict lockdown policies. We address a number of methodological and ethical concerns related to measuring violence within the household, through the use of a List Randomization approach (List Experiment, or LE), a technique originally used to correct for biases in surveys where respondents are asked questions on sensitive topics (Miller, 1984; Raghavarao & Federer, 1979). This experiment is included in a phone survey administered to a longitudinal cohort sample, for which we have

rich background data, contact information recently collected, and where enumerators know the participants from previous visits. To our knowledge, this is the first study of physical domestic violence which employs an LE implemented during a phone survey.

Even before the pandemic, the prevalence of physical violence perpetrated by household members against children and young people was as high as 60% in some countries (Devries et al., 2018), while it is estimated that approximately 1 in 3 women globally are subjected to physical and/or sexual violence by an intimate partner in their lifetime (World Health Organization, 2013). The Council of Europe on 20th April (2020) noted "emerging data are showing an alarming increase in the number of reported cases of certain types of such [violence against women and girls, as well as domestic] violence worldwide". It is therefore an urgent priority to measure the effect of the pandemic and associated lockdown policies on this aspect of young people's lives (United Nations, 2020).[2]

Unfortunately, the steps taken to control the pandemic also make the

---

measurement of domestic violence more difficult. Most of the recent evidence comes from developed countries and relies on administrative data (such as police reports and calls to helplines) or internet surveys. While informative, for many reasons administrative data tends to underestimate the true prevalence of violence. Many people are unable to report their cases to the authorities, or choose not to report to avoid retaliation. Incentives to underreport are likely to increase during lockdowns, as victims might feel they have nowhere else to go. Furthermore, this data often includes limited information about victims' characteristics, making the identification of vulnerable groups problematic. Reports from helplines face similar limitations, and internet surveys (while convenient) are less representative in a developing country context, due to limited internet access (especially among the poorest).

A priori, potential increases in domestic violence during lockdowns in developing countries could be measured through phone surveys. However, this invokes multiple challenges. First, there are limitations in the ability to adequately sample from a given population for a phone survey. Second, people are known to underreport when directly asked about sensitive topics (Tourangeau & Yan, 2007). Third, individuals might not have enough privacy to answer questions, especially in overcrowded homes. Fourth, the inability over the telephone to observe visual clues of distress. Finally, the strain that asking questions about domestic violence may pose for already stressed individuals requires the implementation of a response protocol if a case of domestic violence is uncovered and a participant asks for help (particularly if there is not enough support available locally).

One recent study in Peru uses administrative data on phone calls to the helpline for domestic violence (Línea 100) to show that violence against women increased by 48% between April and July 2020 (Agüero, 2020). As noted above, however, many cases of violence may remain unreported, and those individuals that do report may be significantly different from those who do not.[3] While focused on a specific age group, our analysis complements this study in three ways. First, we elicit a measure of violence in a way that maximizes the chances of a truthful response. Second, we are able to examine the heterogeneity of impacts based on past and present characteristics. Third, we expand on the sparse literature measuring the prevalence of domestic violence against men and boys. A limitation of our approach, however, is that constraints on our survey methodology allow only one general question on violence. We therefore do not capture the different domains of violence, the perpetrator, or the intensity of the situation. Nevertheless, we show that the LE method, if designed carefully, can provide urgent policy-relevant information on who is most at risk of violence, without compromising the anonymity of the respondent. Furthermore, it is relatively cheap to administer, adding no more than 5 min to a phone survey, and can be used to measure a variety of sensitive issues, not only violence within the household.

Our findings show that 8.3% of the sample experienced an increase in physical domestic violence since the beginning of the COVID-19 stay-at-home requirements. Our results indicate that there is no significant difference in the probability of experiencing an increase in violence between males and females. Instead, the most significant predictor of such an increase is a previous history of physical violence (reported in an earlier survey round in 2016). Among this group, representing around one-sixth of the sample, 23.6% experienced an increase in violence during the lockdowns, while the estimated proportion among those who did not experience violence previously is much lower, at 5.4%. Extending our analysis, we do not find evidence that geographical variation in the duration of the lockdown in Peru affected the probability of more violence.

The rest of the paper is structured as follows. Section 2 describes the situation in Peru, within the context of the COVID-19 pandemic. Section 3 provides a brief literature review and a discussion of the different approaches to measuring domestic violence. Section 4 describes the data used for the empirical analysis, as well as the methodology and approach we followed to design the LE. Section 5 presents our empirical strategy, while Section 6 presents our main results. Section 7 discusses our findings and concludes.

2. Country context

Peru is an upper-middle income country with a population of approximately 32 million. It occupied the highest position in the global rankings of COVID-19 cases and deaths per capita during most of 2020, despite extremely lengthy and restrictive stay-at-home requirements. Soon after the first cases were detected, the government declared a state of national emergency, closed borders, and implemented mandatory social isolation from March 16th, 2020 (which, from here onwards, we refer to as "lockdown"). This period of lockdown was initially planned to last 15 days (*Decreto supremo 044-2020-PCM*). During this period, the right to free movement was limited to essential activities, including buying food and medicines and participating in essential economic activities (but excluding physical attendance at schools, universities and other educational institutions). A few days later, additional restrictions on movement for children and senior citizens, bans on going out on Sundays and an evening curfew (from 4pm to 5am) were instituted. The lockdown was extended on five consecutive occasions, finally ending on the 30th of June (a total of 107 days), although the length of the daily curfew reduced as time went by.
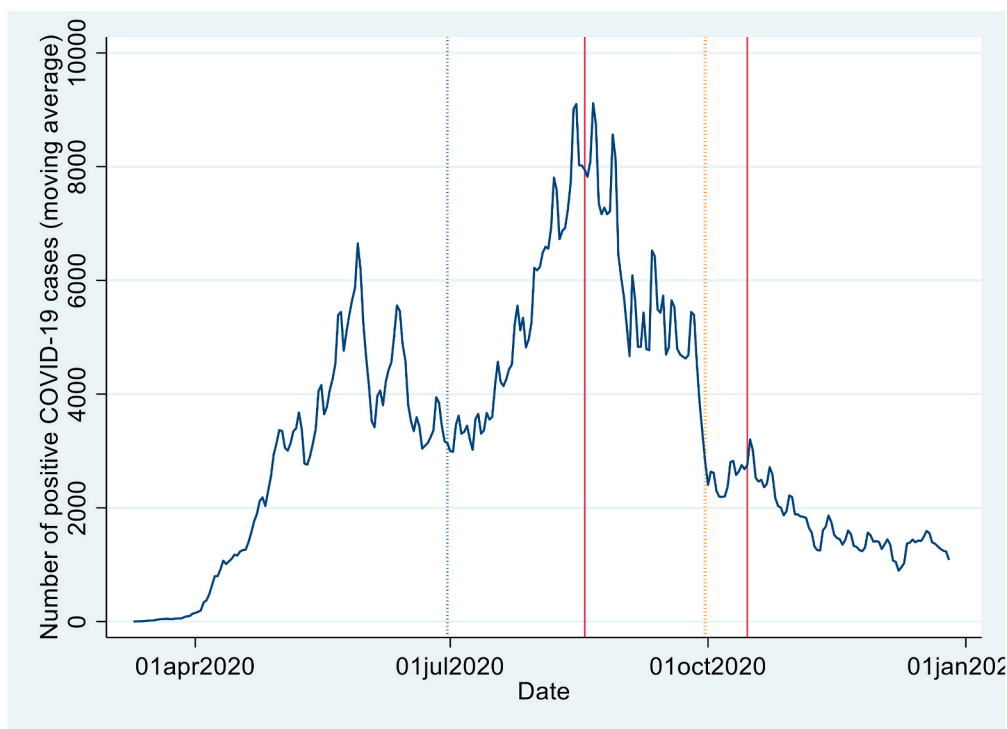
In July, the government moved to a phase of local lockdowns (at both the Regional and Provincial level) in areas where there was a rapid increase in the number of COVID-19 cases, and this phase continued until September 30th. As a result, certain Regions/Provinces remained in lockdown for the entire months of July, August and/or September. Except for some schools in rural areas, all educational institutions remained closed for the entire year, and social isolation remained mandatory for senior citizens and children until October and November, respectively. As of January 1st, 2021, the evening curfew remains in place from 11pm to 4am. Peru's response has, therefore, been extremely stringent, relative to other countries, peaking at 96.3 out of a maximum of 100 (during the month of May) in the COVID-19 Government Stringency Index (Roser et al., 2020).

Fig. 1 reports the evolution of COVID-19 cases since early March. The blue dotted line marks the end of the national lockdown, whereas the orange dotted line indicates the end of the local lockdowns phase. As a reference, the two solid red lines represent the period during which the Young Lives phone survey was administered (from August 18th to October 15th).

According to pre-pandemic data, Peru is also a country with historically high levels of intimate partner and family violence. In a global study by the World Health Organizaton, Peru was ranked 22nd out of 154 countries for lifetime experiences of intimate partner violence among women aged 15–49 (WHO, 2018). The 2019 Peruvian Demographic and Health Survey reported that 30% of women in this age group (who have had a partner) experienced physical violence from their current or last partner. While still high, this prevalence represented an important decline over the past decade (from an earlier rate of 38% in 2009). However, COVID-19 threatens to undo some of the past progress made in reducing the levels of household violence within the country.

3. Literature review

Peterman, O'Donnell, and Palermo (2020) and Peterman and O'Donnell (2020a,b) summarize papers which aim to quantify household violence or intimate partner violence (IPV) since the start of the pandemic. Innovations using administrative data have allowed many

---

[3] As a reference, according to the 2019 Demographic and Health Survey, less than 1 out 3 women in Peru, that are physically abused by their partners, look for help from official institutions.

Note: The number of cases is a daily moving average of the last 5 days. The blue dotted line marks the end of the national lockdown, whereas the orange dotted line the end of the local lockdowns. The red solid lines represent the period during which the Young Lives list experiment took place. *Source: Government of Peru.*

Fig. 1. Evolution of COVID-19 positive cases and length of lockdowns in Peru
Note: The number of cases is a daily moving average of the last 5 days. The blue dotted line marks the end of the national lockdown, whereas the orange dotted line the end of the local lockdowns. The red solid lines represent the period during which the Young Lives list experiment took place. *Source: Government of Peru.*

studies to capture changes in incidence at the population level, including clinical assessments, phone calls to helplines and police dispatch/crime data (Agüero, 2020; Hsu & Henke, 2020; Leslie & Wilson, 2020; Silverio-Murillo, Balmori de la Miyar, & Hoehn-Velasco, 2020). Of the few studies utilizing primary data, Béland, Brodeur, Haddad, and Mikola (2020) use an internet survey to measure domestic violence in Canada and find evidence of links between reported violence, unmet financial obligations and social isolation. Similarly, Arenas-Arroyo, Fernandez-Kranz, & Nollenberger (2021) use an online survey to estimate a 23% increase in intimate partner violence during the lockdown in Spain, driven primarily by psychological violence (with little evidence of increased physical violence). Some studies have also used perceptions of general (for example, village-level) frequency of domestic violence (Halim, Can, & Perova, 2020; Mahmud & Riley, 2020).

Only three surveys in low- and middle-income countries, to our knowledge, have used direct questions in a quantitative phone survey to measure domestic violence during the COVID-19 lockdowns. Hamadani et al. (2020) find that 6.5% of mothers enrolled in a child iron supplement program report an increase in physical violence in Bangladesh, while Abrahams, Boisits, Schneider, Prince, and Lund (2020) find that levels of domestic violence (physical, psychological and sexual) increased among pregnant women in Cape Town (by approximately 3%). In rural Kenya, Egger et al. (2021) also estimate a 4% increase in domestic violence (defined similarly) experienced by women during the pandemic, although this result was not statistically significant. To the best of our knowledge, no study has yet included an LE in a phone survey during the pandemic, and we do not know of any phone survey LE on domestic violence.

The LE, also known as item response count, was introduced by Raghavarao and Federer (1979) and Miller (1984), to correct for biases in surveys where people tended to provide socially desirable responses when asked questions relating to sensitive topics (Corstange, 2009; Holbrook & Krosnick, 2010; Imai, 2011). The method was originally used to investigate social attitudes, such as racism or homophobia (Kuklinski, Cobb, & Gilens, 1997; Lax, Phillips, & Stollwerk, 2016), but has now been applied to various topics in developing countries, including misallocation of loans for small firms in Peru and the Philippines (Karlan & Zinman, 2011), illegal migration in Mexico, Morocco, Ethiopia, and the United States (McKenzie & Siegel, 2013), sexual behaviour in Uganda (Jamison, Karlan, & Raffler, 2013), Cameroon and Côte d'Ivoire (Chuang, Dupas, Huillery, & Seban, 2021), and Senegal (Lépine, Treibich, & D'Exelle, 2020,b), as well as female genital mutilation in Ethiopia (De Cao & Lutz, 2018).

More recently LEs have also been used to estimate intimate partner violence and domestic violence (Agüero & Frisancho, 2021; Gibson, Gurmu, Cobo, Rueda, & Scott, 2020; Peterman, Palermo, Handa, Seidenfeld, & Zambia Child Grant Program Evaluation Team, 2018). Recent studies show that, while the WHO's multiple question approach provides a more comprehensive measure, LEs tend to outperform many standard survey methods used to assess the prevalence of domestic violence. For example, Bulte and Lensink (2019) and Cullen (2020) show evidence from Vietnam, Nigeria and Rwanda that LEs substantially increase reports of intimate partner violence. Similarly, Asadullah, De Cao, Khatoon, and Siddique (2020) and Gibson et al. (2020) show that,

compared to direct survey questions, LEs are more likely to elicit truthful attitudes towards intimate partner violence. In Peru, Agüero and Frisancho (2021) find that there are no significant differences in direct versus indirect reporting of physical violence on average, although they do find that more educated respondents underreport when using direct questions.

LEs have been included in other multi-topic surveys in the past. For example, a LE on vote-buying was embedded in the Mexico 2012 Panel Study, designed and fielded by Greene, Domınguez, Lawson, and Moreno (2012), while Peterman et al. (2018) conclude from a study of intimate partner violence in Zambia that it is possible to implement this type of experiment in a large, multi-topic survey.

We found one study which embedded a LE into a phone survey in a developing country, to examine the prevalence of vote buying in Tanzania (Croke, 2017). Since the onset of the pandemic several studies have embedded an LE into an online survey, to test for truthful reporting of hygenic/social distancing behaviours during the pandemic (Bowles, Larreguy, & Liu, 2020; Larsen, Nyrup, & Petersen, 2020; Timmons, McGinnity, Belton, Barjaková, & Lunn, 2020; Vandormael, Adam, Greuel, & Bärnighausen, 2020)), but none thus far that have been incorporated into a phone survey.

4. Data and methodology

4.1. The Young Lives study

The Young Lives (YL) study tracks the livelihoods of 12,000 children (now adolescents and young adults) in Ethiopia, India (Telangana and Andhra Pradesh), Peru, and Vietnam. Participants have been visited in-person on five occasions, approximately once every three years, initially in 2002 and most recently in 2016. Over time, the study has demonstrated relatively low attrition rates compared to other longitudinal studies in developing countries (Sánchez and Escobal, 2020). In Peru, the initial survey collected information on 2766 participants; 2052 children who were born around the millennium (Younger Cohort, YC) and another 714 who were around eight years old at the time (Older Cohort, OC). The sample was randomly selected from the universe of districts in the country in 2001, excluding the wealthiest 5%. Although it was not intended to be a nationally representative study, statistical analysis with data from the Peruvian Demographic and Health Survey indicates that the sample contains households across the entire wealth spectrum (Escobal & Flores, 2008). Peru is the only YL country in which experiences of domestic violence were collected in previous rounds through a self-administered questionnaire.

The COVID-19 outbreak began when fieldwork was about to start for the sixth round of (in-person) data collection, with the two cohorts ranging in age between 18-19 and 25–26 years old. In late 2019, in preparation for the sixth round, the sample interviewed in 2016 was tracked, first by phone and then with a small in-person tracking operation. This combined effort allowed the survey team to track and obtain up-to-date contact information for 90% of the 2016 sample (2215 out of the 2468 participants). In recognition of the health crisis, fieldwork was postponed, and instead, a three-part phone survey was implemented, aimed at measuring the short-term impacts of the COVID-19 crisis on the survey participants: the 'Listening to Young Lives at Work: COVID-19 Phone Survey'. The LE took place during the second call of the phone survey, between 18th August and 15th October of 2020, by which point the national lockdown had ended (although some regions and provinces were still in lockdown due to an increase in COVID-19 cases. See Fig. 1).[4] Only the 2215 participants found in the tracking were contacted for the phone survey.

The survey included 1992 complete interviews, representing 90% of the tracking sample and 81% of the 2016 sample. Of those participants that were not found, 154 did not have a phone number, whereas 58 had one but did not answer. There were 20 refusals, 6 incomplete interviews and 1 participant passed away. In addition, 16 participants who were not included in the tracking sample were later found and interviewed. In Table A1, columns (2 and 3), we analyze the nature of attrition in relation to the sample of respondents for which we collected information in the most recent (fifth) round in 2016. Overall, attrited participants are less likely to be from the top wealth quintile and more likely to come from households where the mother has completed fewer years of formal education. There are no discernible differences by sex, age, native tongue, or area of residence between attrited and non-attrited respondents. Conditional on pre-pandemic characteristics, those who were exposed to violence in the past are less likely to have answered the phone survey. We come back to this aspect, and its implications for the interpretation of our main results, in Section 7.

Table 1 presents an overview of the characteristics of our analytical sample, after omitting 151 (7.6%) observations with missing information on key variables. Unlike the majority of studies on domestic violence, which focus on female respondents, our sample is evenly split by gender. The phone survey took place when the national lockdown was over, but the local lockdowns were on-going (see Fig. 1). For this reason, exposure to lockdown days depends on the place of residence and date of interview. On average, individuals had experienced 132 days

Table 1
Descriptive statistics of the young lives sample.

	All	Younger Cohort		Older Cohort	
		Males	Females	Males	Females
Age	20.46	18.90	18.88	25.93	25.91
Urban	0.81	0.81	0.79	0.86	0.87
First language is Spanish	0.87	0.87	0.85	0.91	0.87
Mother incomplete primary education	0.33	0.33	0.33	0.36	0.33
Mother complete primary education (only)	0.30	0.30	0.30	0.29	0.36
Mother complete secondary education (only)	0.21	0.21	0.21	0.24	0.18
Mother incomplete/ complete higher education	0.15	0.16	0.16	0.12	0.13
Days in lockdown	131.65	132.23	130.44	132.71	132.93
More than 107 days in lockdown	0.55	0.57	0.55	0.52	0.55
Household job loss and/or non-payment of wages	0.68	0.63	0.69	0.78	0.71
Household member receives Juntos (2016)	0.19	0.23	0.24	0.05	0.03
AWSA (2016)	0.62	0.59	0.63	0.63	0.66
Physical violence by family/partner (2016)	0.15	0.10	0.20	0.09	0.23
Highest education grade achieved	12.15	11.95	12.14	12.55	12.41
Wealth Index Quintile 1 (2016)	0.17	0.17	0.21	0.12	0.12
Wealth Index Quintile 2 (2016)	0.20	0.21	0.21	0.16	0.21
Wealth Index Quintile 3 (2016)	0.19	0.19	0.15	0.31	0.23
Wealth Index Quintile 4 (2016)	0.21	0.20	0.22	0.18	0.24
Wealth Index Quintile 5 (2016)	0.23	0.23	0.22	0.23	0.21
Proportion Female	*0.50*	*0.51*		*0.46*	
Observations	1841	705	726	221	189

Note: Physical violence is defined as physical violence perpetrated by family members, spouse and/or boyfriend/girlfriend. AWSA refers to the Attitudes to Women Scale for Adolescents and takes values between 0 and 1, with higher scores representing views more consistent with gender equality.

---

[4] Given that this second call took place only a few weeks after the national lockdown ended in Peru, any measurement error related to the recall period is likely to be small.

of lockdown at the time of the interview, while 55% of respondents came from provinces where the lockdown had been extended beyond the initial 107 days (imposed at the national level). Most respondents were located in urban areas, in-line with the 2017 national census.

Under the assumption that increases in household violence during the pandemic are not independent of past experiences, we also report an indicator variable based on the information recorded in the 2016 self-administered questionnaires (SAQs).[5] More specifically, this indicator variable is defined by the following question, asked to both the OC (aged 22 at the time) and to the YC (aged 15): "*Have you ever been beaten up or physically hurt in other ways by the following people?*" At the time, the prevalence of physical domestic violence (defined as perpetrated either by the spouse/partner or another family member) was 15% in the sample overall.[6] However, females were more likely to experience physical domestic violence in both cohorts, and the gender gap was larger in the OC.

Domestic and intimate partner violence is also commonly linked to the presence of social norms around gender and gender roles (Heise & Kotsadam, 2015; Gibbs et al., 2020). In recognition of this, Table 1 summarises scores from the Attitudes to Women Scale for Adolescents (AWSA) (Galambos, Petersen, Richards, & Gitelson, 1985), again measured in the 2016 survey round. The AWSA is a 12-item scale including statements referring to the rights, freedoms, and role of males and females, girls in education, sports, dating and families, and to adult roles in parenting and housework. Individual scores range between 0 and 1, with higher scores indicating views more consistent with gender equality.

Economic shocks (job losses in particular) are also commonly cited as having links to violence within the household (Anderberg, Rainer, Wadsworth, & Wilson, 2016; Béland et al., 2020; Bhalotra, Kambham-pati, Rawlings, & Siddique, 2020; Schneider, Harknett, & McLanahan, 2016), as are general levels of wealth and poverty, which are likely to be correlated with a wide variety of risk factors (Gibbs et al., 2020; Vyas & Watts, 2009). Therefore, we include in our analysis a variable indicating if a member of the respondent's household had experienced a job loss and/or non-payment of (full) wages during the pandemic (68% of sample households had) and split our sample according to five quintiles of the YL wealth index (Briones, 2017). Previous studies have also found a mitigating effect on levels of domestic violence from involvement in conditional cash transfer (CCT) programs (see Buller et al., 2018). Therefore, we also consider a variable indicating if a member of the respondent's household was a recipient of the *Juntos* CCT program in the last survey round (19% of households were).

4.2. Design of the list experiment

In a standard LE, sample participants are randomly split into two groups. One group is asked about a number of non-sensitive statements (items) unrelated to the sensitive topic, and another group receives the same list of items, alongside an extra sensitive item (here, related to physical violence). In each case, the respondent is asked to confirm how many of the items apply to them, without disclosing which specific items they agreed with. The difference between the mean count of applicable items between the sensitive versus non-sensitive list recipients provides an estimate of the prevalence of the sensitive behavior the researcher aims to measure. We follow this approach in our experiment. However, to improve the precision of our estimates, the roles of treatment and control groups are reversed in a second LE, with four new control items

introduced and those previously assigned to the control group now receiving the additional sensitive item (see Droitcour et al., 1991).

In order to maximize statistical precision (minimize the variance of the responses) we are mindful of two design issues related to the correlation between control items and the types of control items (themes). One problem when designing a LE is the potential for *ceiling effects* (Kuklinski et al., 1997), which can occur when a respondent would honestly respond "yes" to all of the non-sensitive (control) items. This would imply that a treatment group respondent would no longer have the desired level of protection to honestly report their response to the sensitive item (if they agree with all statements asked, then they clearly agree with the sensitive item embedded within the list). The analogous problem occurs if a respondent would answer "no" to all of the non-sensitive items (*floor effects*). In this case, a respondent may be concerned that responding "yes" only once would risk revealing their response to the targetted statement.

Glynn (2013) notes three generally accepted pieces of design advice. The first two are to avoid the use of many high (low)-prevalence, non-sensitive items. For example, while it is important to minimize ceiling effects, if a respondent suspects that all non-sensitive items have a low prevalence, she/he may also become concerned about the level of privacy protection and underreport her/his answers (Tsuchiya, Hirai, & Ono, 2007). Third, lists should not be too short, as short lists may also increase the likelihood of ceiling effects (Kuklinski et al., 1997). Glynn (2013) argues that these three pieces of design advice tend to lead to increased variance in the list (Corstange, 2009; Tsuchiya et al., 2007) and an unfortunate trade-off between ceiling effects and variable results (an apparent bias-variance trade-off). He proposes that it may be possible to reduce both bias and variance by constructing the non-sensitive (control) list from items that are likely to be negatively correlated. We follow this principle when selecting our control items.

The subject matter of the non-sensitive items is also important. Many experiments in the literature use baseline statements that are largely innocuous, aside from the singular sensitive statement, potentially making the targeted statement embedded in the list obvious. Kuklinski et al. (1997) warn against *contrast effects* (where the resonance of the sensitive item overwhelms the control items), and Droitcour et al. (1991) advise that the non-sensitive items should be on the same subject as the sensitive item. Chuang et al. (2021) investigate whether the topic and sensitivity of the baseline items influence LE results. They randomly vary whether the baseline items are innocuous, or whether they include items that are related to the targeted statements (either in topic and/or by being somewhat sensitive statements themselves). They find that sensitive baseline items produce more precise prevalence estimates than non-sensitive baseline items. We therefore, include statements relating to feelings and behaviour during the lockdown period, such that the statement on increased violence should not stand out.

By design, the LE introduces noise into the estimate of the prevalence of the sensitive trait (in this case having experienced an increase in physical domestic violence). As noted above, in our experiment, we follow the double list design (Droitcour et al., 1991), which uses two different lists of non-sensitive control items. The main advantage is that this method can substantially improve the precision of our estimates (since we have two LEs using the same sensitive question). Glynn (2013) notes that to improve the precision of a double LE, there needs to be a positive correlation between responses to items on the two lists. The disadvantage is that there are two lists to complete during the survey and this may increase participant burden or cognitive load during the interview.

We conducted a pilot study by phone during the week of July 25th-31st, to assess the prevalence and correlation between potential control items. Our pilot survey used a purposive sample, comprised of 202 participants, aged between 17 and 27 years of age, representing males and females, as well as urban and rural areas (in 13 of the 24 regions of the country).

Based on the results from our pilot survey, we choose lists with a

---

[5] See Sánchez and Hidalgo (2019) for further details on the self-administered questionnaire (SAQ) data.

[6] There are no comparable official statistics for this definition and for young people, but as a reference, according to the 2019 Peruvian Demographic and Health Survey, 30% of women aged 15 to 49 have been victims of Intimate Partner Violence.

relatively low prevalence of ceiling or floor effects, and items with a high correlation between the number of agreed answers, when asked either as part of a list or the sum of individual items (a full piloting analysis is shown in Appendix B). The final statements are listed below.

List 1

Group 1: Control	Group 2: Treatment
• During the lockdown, I started drinking more alcohol than usual. • During the lockdown, I was able to spend more free time than previously doing exercise. • I felt much closer to my family during the lockdown than I did before. • During the lockdown, there were more arguments in our household than there were before.	• During the lockdown, I started drinking more alcohol than usual. • During the lockdown, I was able to spend more free time than previously doing exercise. • I felt much closer to my family during the lockdown than I did before. • During the lockdown, there were more arguments in our household than there were before. • I was physically hurt more often by someone in my household during the lockdown.

List 2

Group 1: Treatment	Group 2: Control
• I had more caring responsibilities when the schools were closed during the lockdown. • I spent much more time working during the lockdown than I normally would have. • I talked on the phone to my friends much more than usual during the lockdown. • During the lockdown, I ran out of the money I needed for basic things more often than before. • I was physically hurt more often by someone in my household during the lockdown.	• I had more caring responsibilities when the schools were closed during the lockdown. • I spent much more time working during the lockdown than I normally would have. • I talked on the phone to my friends much more than usual during the lockdown. • During the lockdown, I ran out of the money I needed for basic things more often than before.

Unfortunately, there was no space in the phone survey to include multiple statements on different aspects of violence (Agüero & Frisancho, 2021), and little time to pilot potential statements, especially those of a sensitive nature, and over the phone. We therefore, simplify a sensitive item validated by Peterman et al. (2018) that covers physical violence only.[7] Our sensitive statement, "*I was physically hurt more often by someone in my household during the lockdown",* is intended to measure the *increase in* domestic violence (perpetrated by any household member) that affected young people during the COVID-19 lockdown periods in Peru.

Finally, we randomize the survey participants into two groups, to receive the sensitive item in only one of the two lists (Kennedy and Mann, 2015). As noted by Blair and Imai (2012) and Glynn (2013), using blocking rather than simple randomization is preferable to achieve balance across participant attributes that we are later interested in comparing. Therefore, we block by cohort, gender and urban/rural location, since the analysis of previous data collected on our sample suggests that increases in domestic violence might vary according to these characteristics. We also automatically select the randomization which has the best balance according to the YL Home Environment for Protection (Basic YL-HEP) index (Brown, Ravallion, & Van de Walle, 2020) and highest parental education. We do not randomize the order of the questions, for reasons of power, and to keep the instructions as simple as possible for enumerators.

Aside from the design of the experiment, mentioning violence at all, even in a confidential setting, and even indirectly, still raises an ethical concern. In recognition of this, we prepared a fieldworker manual which described the protocol the enumerator should follow if the subject of violence came up in the discussion, and a consultation guide with information about the resources available (such as helplines and psychological support). This consultation guide was provided to all respondents at the end of the survey (by email or through WhatsApp), irrespective of their responses in the LE.[8]

A further practical concern is the issue of numeracy in areas with poor education and the ability of participants to recall answers to a list of statements during a phone survey. Lépine, Treibich, Ndour, Gueye, and Vickerman (2020) asked respondents to pass marbles from one hand to the other (behind their backs) to keep count of the number of statements they agree with, when conducting an in-person survey. We adapt this idea to the context of our own experiment, by asking respondents to count on their fingers. Enumerators reported this worked well for participants to keep track of their responses.

5. Empirical strategy

The estimate of the overall proportion of the sample who experienced an increase in domestic violence during the lockdowns can be derived from the simple system of linear equations shown in (1.1) and (1.2). The variable Y_{il} indicates the number of positive responses by individual i to the statements in list l and the indicator variable T_{il} takes value 1 for inclusion within the treatment group for that list. The constant term γ_{0l} indicates the average number of positive responses to the four control questions in each list, while the estimated coefficient δ measures the increase in the number of affirmative responses resulting from receiving the additional sensitive statement (for those in the treatment group). The term ε_{il} represents the error in each equation. Given a set of testable assumptions (discussed in detail in Section 6.3), δ should give an unbiased estimate of the prevalence of the sensitive trait within the sample overall. The two-stage approach used to estimate δ from (1.1) and (1.2) is discussed below.

$$Y_{i1} = \gamma_{01} + \delta T_{i1} + \varepsilon_{i1} \tag{1.1}$$

$$Y_{i2} = \gamma_{02} + \delta T_{i2} + \varepsilon_{i2} \tag{1.2}$$

In an extension to the model, we include characteristics which may be correlated with increased violence. In this second model, a vector of individual and household characteristics **X** is included in the previous system of equations. Here the influence of the additional covariates on the answers to the control and the sensitive statements is measured by their associated coefficients **γ** and **δ**.

$$Y_{i1} = \gamma_{01} + \boldsymbol{\gamma_1 X_i} + T_{i1}*(\delta_0 + \boldsymbol{\delta X_i}) + \varepsilon_{i1} \tag{2.1}$$

$$Y_{i2} = \gamma_{02} + \boldsymbol{\gamma_2 X_i} + T_{i2}*(\delta_0 + \boldsymbol{\delta X_i}) + \varepsilon_{i2} \tag{2.2}$$

To estimate the parameters of the model described by equations (2.1) and (2.2), we rely on a two-stage approach, adapted from Imai (2011) and Blair and Imai (2012). In the first stage, predictions of responses to the control statements, conditional on individual characteristics, are obtained for the first and second lists separately, as specified in equations (3.1) and (3).2):

$$Y_{i1}(0) = \gamma_{01} + \boldsymbol{\gamma_1 X_i} + \varepsilon_{i1}, \quad \forall T_{i1} = 0 \tag{3.1}$$

$$Y_{i2}(0) = \gamma_{02} + \boldsymbol{\gamma_2 X_i} + \varepsilon_{i2}, \quad \forall T_{i2} = 0 \tag{3.2}$$

---

[7] The original statement was "*I have been slapped, punched, kicked, or physically harmed by my partner*" (Peterman et al., 2018), which we simplified, and also generalised, given that young people are predominantly living with parents and/or other adults who may also harm them as well as spouses/partners.

[8] The consultation guide was multi-purpose, not specific to violence issues. It included information on how to find work, scholarships, eligibility for benefits etc. It is available at www.ninosdelmilenio.org/2020/11/11/guia-para-consultas.

This yields the conditional estimates of the mean number of affirmative responses to the control statements $E[Y_{il}(0) \mid X_i = x]$. The predicted response to the sensitive item $\widehat{Z}_i$ for those in either treatment group can then be generated as the estimated number of affirmative responses to the four control statements, deducted from the observed number of affirmative responses to all five statements $Y_{il}(1)$.

In the second stage, $\widehat{Z}_i$ can then be regressed on the individual characteristics in X to estimate the δ parameters of the model.

$$\widehat{Z}_i = \delta_0 + \boldsymbol{\delta} X_i + \varepsilon_i \tag{4}$$

Where: $\widehat{Z}_i = Y_{i1}(1) - E[Y_{i1}(0) \mid X_i = x]$, $\forall \ T_{i1} = 1$ and: $\widehat{Z}_i = Y_{i2}(1) - E[Y_{i2}(0) \mid X_i = x]$, $\forall \ T_{i2} = 1$

As estimates of the predicted response to the sensitive item $\widehat{Z}_i$ are derived (in part) from predictions of the number of responses to the control statements (the first stage), standard errors are bootstrapped in all reported regressions.[9] To maintain consistency in our approach, the overall prevalence of increased violence (without covariates), described by equations (1.1) and (1.2), is calculated using an analogous two-stage approach, whereby $E[Y_{il}(0) \mid X_i = x]$ is replaced with the (unconditional) mean value of affirmative responses $E[Y_{il}(0)]$ and $\widehat{Z}_i$ in equation (4) is regressed on an intercept only.

The vector of individual and household characteristics X includes the lockdown duration, either defined as a dummy variable (indicating if a respondent experienced more than the initial 107 days of national lockdown) or as a continuous measure (the number of days extension beyond the 107 days), controls for the gender and age (cohort) of the individual, as well as their years of education and their AWSA score. The vector of covariates also includes the indicator variable for past experiences of physical violence within the household. In addition, at the household level, we control for pre-pandemic wealth status (in wealth index quintiles), urban location, whether a household member experienced a job loss or wage cut during the pandemic, and whether the household was previously a recipient of the Juntos conditional cash transfer programme.

6. Results

6.1. Summary of the list experiment

Before reporting the results of our analysis, Table 2 provides a summary of the responses for both the treatment and control groups in both lists. This breakdown indicates that only a small number of individuals in the control groups reported agreeing with either zero statements (3.2% and 1.7%, in lists 1 and 2, respectively) or all four control statements (1.7% and 3.1%). Instead, control respondents were most likely to agree with two of the four control items, in line with the design of the experiment and a negative correlation between pairs of control statements (see Section 4.2). This pattern is consistent across both male and female respondents.

6.2. Regression results

Table 3 reports the results of our analysis, based on the empirical approach outlined in Section 5. Column 1 reports the estimated coefficient representing the prevalence of increased violence for the sample overall (the intercept represents δ in equations (1.1) and (1.2)). The coefficient indicates that 8.3% of our sample of young people experienced an increase in physical domestic violence during the lockdown.

Having established that physical violence increased during the

---

[9] Our results remain largely unchanged with alternative treatments of the standard errors, namely clustering at the individual or the survey cluster (sentinel site) level. These results are available from the authors upon request.

Table 2
Summary results of the list experiment.

List 1	Number of statements						
	0	1	2	3	4	5	Mean number of responses
All (%, treatment group)	3.7	26.4	52.1	15.8	1.6	0.3	1.862
All (%, control group)	3.2	32.4	52.9	9.8	1.7		1.744
Females (%, treatment group)	3.1	29.0	49.0	16.9	1.8	0.2	1.860
Females (%, control group)	1.5	32.4	55.2	9.2	1.7		1.773
Males (%, treatment group)	4.3	24.0	55.1	14.7	1.5	0.4	1.864
Males (%, control group)	5.0	32.4	50.5	10.4	1.7		1.715

List 2	Number of statements						
	0	1	2	3	4	5	Mean number of responses
All (%, treatment group)	4.3	28.6	40.8	20.7	5.2	0.4	1.950
All (%, control group)	3.7	29.6	42.8	20.8	3.1		1.899
Females (%, treatment group)	4.3	29.0	39.7	21.9	4.7	0.4	1.951
Females (%, control group)	4.0	30.1	42.1	21.8	2.0		1.878
Males (%, treatment group)	4.3	28.3	41.9	19.4	5.6	0.4	1.950
Males (%, control group)	3.5	29.2	43.4	19.9	4.1		1.920

lockdown for a significant proportion of the sample, columns 2 and 3 report the effect of the additional covariates on the predicted (linear) probability of experiencing increased violence. Based on these results, for the overall sample of 1,841, there is no evidence that a relatively longer lockdown period, either expressed as a dichotomous variable (column 2) or as a continuous variable (column 3), significantly altered this probability. However, two key results are indicated in columns 2 and 3. First, the change in the probability of increased violence is not significantly different between males and females. While the underlying prevalence of violence is almost certainly higher for females (see Table 1), an increase in experiences of violence during the lockdowns in Peru appears no more likely for either gender. Second, the most significant and influential variable is the past experience of violence, reported by approximately 15% of the sample in the previous survey round (2016). This indicator increases the probability of more domestic violence during the lockdown by more than 18%.

Fig. 2 shows the estimated proportion experiencing increased violence for specific groups (calculated at the sample mean of all covariates). When considering the effect of gender, only an increase in violence for males (predicted at 10.1%) is significantly different from zero, although there is no significant difference between this prediction and the prediction for females (6.1%). Perhaps the clearest result in Fig. 2 is that 23.6% of those who previously reported violence in 2016 gave responses in the LE which imply that they had also experienced an increase in violence during the lockdown. For those with no history of previous violence, the probability of an increase is much smaller (estimated at 5.4%).

It should be noted that the use of the double list design substantially increases the precision of our estimates, relative to the use of either of the two lists independently. We find that the estimated standard errors around the prevalence of increased violence are consistently smaller than those estimated from either list separately. This is true for the

Table 3
The prevalence of increased violence during the lockdown.

Dependent Variable:	(1)	(2)	(3)
Predicted response to sensitive item ($\hat{Z}$)			
Intercept	0.083***	−0.002	−0.013
	(0.025)	(0.183)	(0.183)
Female		−0.040	−0.040
		(0.053)	(0.053)
Younger Cohort		0.081	0.081
		(0.064)	(0.064)
Years of Education		−0.008	−0.008
		(0.018)	(0.018)
Lockdown Extension (=1)		0.004	
		(0.053)	
Lockdown Extension (Days)			0.001
			(0.001)
Job Loss/Non-Payment (2020)		0.037	0.036
		(0.053)	(0.053)
Juntos recipient (2016)		−0.101	−0.100
		(0.086)	(0.086)
AWSA (2016)		−0.073	−0.079
		(0.259)	(0.258)
Past Violence (2016)		0.181**	0.182**
		(0.074)	(0.074)
Living in urban areas (2020)		−0.018	−0.013
		(0.081)	(0.081)
Wealth Index Quintile 1 (2016)		0.148	0.151
		(0.099)	(0.099)
Wealth Index Quintile 2 (2016)		0.082	0.081
		(0.085)	(0.085)
Wealth Index Quintile 4 (2016)		0.019	0.021
		(0.077)	(0.077)
Wealth Index Quintile 5 (2016)		0.083	0.082
		(0.083)	(0.083)
Observations	1841	1841	1841
F-test: Wealth quintiles (*p*-value)		0.571	0.573

Notes: Bootstrapped standard errors are reported in parentheses * p < 0.1, **p < 0.05, ***p < 0.01. AWSA refers to the Attitudes to Women Scale for Adolescents and takes values between 0 and 1, with higher scores representing views more consistent with gender equality.

sample overall, as well as sub-groups defined by gender, age, location (urban/rural) and previous experiences of physical violence. A summary of these efficiency gains can be found in Appendix Table D1.

6.3. Robustness checks

The validity of the LE methodology is based on three assumptions: successful randomization of the treatment, no design effects, and the absence of ceiling and floor effects (Lépine, Treibich, & D'Exelle, 2020). Successful randomization is required such that individuals allocated to each group are, on average, likely to agree with the same number of non-sensitive statements in any given list. The no design effects assumption is necessary so that the inclusion of the sensitive item does not change the number of positive answers to the non-sensitive items. Finally, as noted earlier, the absence of ceiling and floor effects is required, as individuals may be reluctant to provide truthful answers if

they believe they no longer benefit from the privacy of their responses (Kuklinski et al., 1997).

Table C1 in Appendix C suggests that the randomization of the LE was successful, given that individual and household characteristics do not significantly differ between the two treatment groups. To assess whether the inclusion of the sensitive statement modified the answers to the non-sensitive statements, we implemented two statistical tests developed by Blair and Imai (2012). Table C2 in Appendix C reports the results of these tests for both LEs. The Bonferroni-corrected minimum *p*-values indicate that we fail to reject the null hypothesis of no design effect in both LEs. Finally, any bias resulting from ceiling or floor effects should be minimal, as the proportion of respondents in the control group who agreed with all or none of the non-sensitive items is never more than 5% in either list (see Table 2). The lack of ceiling and floor effects is, in part, due to the time taken to pilot potential non-sensitive control items, which gave us the opportunity to select control statements that minimised the likelihood of such responses.

7. Conclusions

Understanding the effect of the global COVID-19 pandemic on domestic violence perpetrated against young people is of paramount importance. We employ a double list randomization approach to estimate the increases in physical domestic violence experienced by young people in Peru during the severe lockdowns implemented throughout much of 2020. This LE is embedded, for the first time, in a phone survey conducted among a long-standing cohort study. The extra questions are incorporated without problems into the survey and lengthen it by less than 5 min. It would, therefore, be relatively easy to replicate this method in similar phone surveys, to measure a wide variety of sensitive topics, not just violence within the household, though with appropriate ethical considerations, including the availability of resources for participants experiencing the sensitive issue.

This study demonstrates that it is not only possible to measure increases in domestic violence during the COVID-19 pandemic, but that it is possible to do so discretely and with the appropriate level of ethical consideration for the well-being of respondents. This is evidenced by only around 1% (15 individuals), of those interviewed, not providing answers to the questions posed in our LE, and none of our enumerators reporting any instances of discomfort or distress associated with this part of the survey. We also demonstrate that it is possible to adapt the list randomization approach to be administered via a multitopic phone survey, and with no requirement that participants possess even basic levels of numeracy.

While the list randomization method is viable in this context, implementing the experiment over the phone, and as part of a larger survey, provided specific challenges, especially while our respondents were experiencing the obvious stress associated with the pandemic. Providing consultation guides including information on helplines and psychological support to all participants in the LE, and the careful selection and piloting of control items was a crucial part of the experiment. The fact that the study was known to participants likely improved their trust, and we had pre-pandemic information that we could use in the analysis. Particularly, Peru was the only country of the Young Lives Survey to have a pre-covid measure of domestic violence, which has proven to be a key correlate in our analysis.

There are three findings emerging from our study: first, we show an increase in domestic violence since the beginning of the COVID-19 stay-at-home requirements in Peru. We find that 8.3% of young people experienced an increase in physical domestic violence during the lockdown. Due to the nature of attrition in the sample, and the relatively higher probability of those with a previous history of violence being omitted from the final sample (see Appendix A), this is likely to reflect a lower bound of the increase in violence. This estimate would represent a greater increase in violence than that reported in phone surveys conducted in South Africa, Bangladesh and Kenya during the pandemic (see
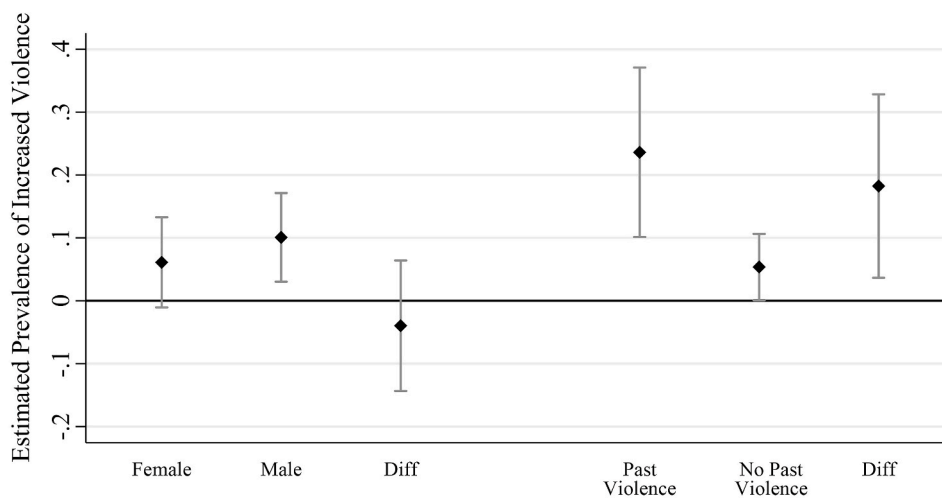
Notes: Predictions for each group are calculated at the sample mean of all other covariates. Vertical bars indicate a 95% confidence interval around predictions.

Fig. 2. The predicted prevalence of increased violence by sub-groups
Notes: Predictions for each group are calculated at the sample mean of all other covariates. Vertical bars indicate a 95% confidence interval around predictions.

Abrahams et al., 2020; Egger et al., 2021; Hamadani et al., 2020), However, due to differences in methodology (LE, as opposed to direct questions), our focus on physical violence only (as opposed to a broader definition) and differences in the population of interest (males and female, as opposed to women only), results may not be directly comparable. Second, while the pre-pandemic levels of violence are shown to be higher for females in our sample, our results indicate no significant difference in the probability of increased violence occurring during the pandemic between females and males. Third, the increase in domestic violence appears to be largely determined by those who have experienced violence in the past, about 15% of the sample, whose probability of experiencing more domestic violence during the lockdown is 23.6% (relative to 5.4% for those who did not previously report violence).

From a policy perspective, our results provide useful insights in characterizing which sub-groups of the population should be closely monitored as more vulnerable to being victims of violence during these challenging times. Given the constraints of time and cognitive load for participants, the results are necessarily relatively crude, with only a single question on physical violence and no measure of the intensity of the violence (or who the perpetrator was). However, we consider that this approach provides a viable, pragmatic, and relatively cheap solution to the problem of missing information on sensitive topics during a global pandemic.

Ethical approval

Ethical approval for the Young Lives COVID-19 phone survey was obtained from the University of Oxford (UK) and the Instituto de Investigación Nutricional (Peru).

CRediT authorship contribution statement

Catherine Porter: Conceptualization, Methodology, Investigation, Writing – original draft, Writing – review & editing, Funding acquisition. **Marta Favara:** Conceptualization, Methodology, Investigation, Writing – review & editing, Supervision, Project administration, Funding acquisition. **Alan Sánchez:** Conceptualization, Investigation, Writing – review & editing. **Douglas Scott:** Methodology, Investigation, Validation, Formal analysis, Writing – original draft, Writing – review & editing.

Declaration of competing interest

None.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ssmph.2021.100792.

References

Abrahams, Z., Boisits, S., Schneider, M., Prince, M., & Lund, C. (2020). Domestic violence, food insecurity and mental health of pregnant women in the COVID-19 lockdown in Cape Town, South Africa. Preprint. https://doi.org/10.21203/rs.3.rs-10 2205/v1.

Agüero, J. M. (2020). Covid-19 and the rise of intimate partner violence. *World Development, 137*, 105217.

Agüero, J. M., & Frisancho, V. (2021). Measuring violence against women with experimental methods. Economic development and cultural change. Preprint. https://www.journals.uchicago.edu/doi/abs/10.1086/714008.

Anderberg, D., Rainer, H., Wadsworth, J., & Wilson, T. (2016). Unemployment and domestic violence: Theory and evidence. *The Economic Journal, 126*(597), 1947–1979.

Arenas-Arroyo, E., Fernandez-Kranz, D., & Nollenberger, N. (2021). Intimate partner violence under forced cohabitation and economic stress: Evidence from the COVID-19 pandemic. *Journal of Public Economics, 194*, 104350.

Asadullah, M. N., De Cao, E., Khatoon, F. Z., & Siddique, Z. (2020). Measuring gender attitudes using list experiments. *Journal of Population Economics.* https://doi.org/10.1007/s00148-020-00805-2

Béland, L. P., Brodeur, A., Haddad, J., & Mikola, D. (2020). *COVID-19, family stress and domestic violence: Remote work, isolation and bargaining power (No. 13332).* Institute of Labor Economics (IZA).

Bhalotra, S., Kambhampati, U., Rawlings, S., & Siddique, Z. (2020). *Intimate partner violence: The influence of job opportunities for men and women. Policy research working paper 9118*. Washington DC: World Bank.

Blair, G., & Imai, K. (2012). Statistical analysis of list experiments. *Political Analysis, 20* (1), 47–77.

Bowles, J., Larreguy, H., & Liu, S. (2020). Countering misinformation via WhatsApp: Preliminary evidence from the COVID-19 pandemic in Zimbabwe. *PloS One, 15*(10), Article e0240005.

Briones, K. (2017). *'How many rooms are there in your house?' Constructing the Young Lives wealth index. Young Lives Technical Note* (Vol. 43). Oxford: Young Lives.

Brown, C. S., Ravallion, M., & Van de Walle, D. (2020). *Can the world's poor protect themselves from the new coronavirus? NBER working paper No. 27200*. May 2020.

Buller, A. M., Peterman, A., Ranganathan, M., Bleile, A., Hidrobo, M., & Heise, L. (2018). A mixed-method review of cash transfers and intimate partner violence in low-and middle-income countries. *The World Bank Research Observer, 33*(2), 218–258.

Bulte, E., & Lensink, R. (2019). Women's empowerment and domestic abuse: Experimental evidence from Vietnam. *European Economic Review, 115*, 172–191.

Chuang, E., Dupas, P., Huillery, E., & Seban, J. (2021). Sex, lies, and measurement: Consistency tests for indirect response survey methods. *Journal of Development Economics, 148*, 102582.

Corstange, D. (2009). Sensitive questions, truthful answers? Modeling the list experiment with LISTIT. *Political Analysis, 17*(1), 45–63.

Council of Europe. (2020). Declaration of the committee of the parties to the Council of Europe convention on preventing and combating violence against women and domestic violence (istanbul convention) on the implementation of the convention during the COVID-19 pandemic Accessed on 2nd January 2021 from https://rm.coe. int/declaration-committee-of-the-parties-to-ic-covid-/16809e33c6n-cases-of-violenc e-against-women.

Croke, K. (2017). Tools of single party hegemony in Tanzania: Evidence from surveys and survey experiments. *Democratization, 24*(2), 189–208.

Cullen, C. (2020). *Method matters: Underreporting of intimate partner violence in Nigeria and Rwanda. Policy research working paper 9274*. Washington DC: World Bank.

De Cao, E., & Lutz, C. (2018). Sensitive survey questions: Measuring attitudes regarding female genital cutting through a list experiment. *Oxford Bulletin of Economics & Statistics, 80*(5), 871–892.

Devries, K., Knight, L., Petzold, M., Merrill, K. G., Maxwell, L., Williams, A., et al. (2018). Who perpetrates violence against children? A systematic analysis of age-specific and sex-specific data. *BMJ paediatrics open, 2*(1).

Droitcour, J., Caspar, R. A., Hubbard, M. L., Parsley, T. L., Visscher, W., & Ezzati, T. M. (1991). The item-count technique as a method of indirect questioning: A review of its development and a case study application. In P. P. Biemer, R. M. Groves, L. E. Lyberg, N. A. Mathiowetz, & S. Sudman (Eds.), *In measurement errors in surveys* (pp. 185–210). New York: John Wiley & Sons.

Egger, D., Miguel, E., Warren, S. S., Shenoy, A., Collins, E., Karlan, D., et al. (2021). Falling living standards during the COVID-19 crisis: Quantitative evidence from nine developing countries. *Science Advances, 7*(6).

Escobal, J., & Flores, E. (2008). *An assessment of the young lives sampling approach in Peru, technical note* (Vol. 3). Oxford: Young Lives.

Galambos, N. L., Petersen, A. C., Richards, M., & Gitelson, I. B. (1985). The attitudes toward women scale for adolescents (AWSA): A study of reliability and validity. *Sex Roles, 13*(5–6), 343–356.

Gibbs, A., Dunkle, K., Ramsoomar, L., Willan, S., Jama Shai, N., Chatterji, S., et al. (2020). New learnings on drivers of men's physical and/or sexual violence against their female partners, and women's experiences of this, and the implications for prevention interventions. *Global Health Action, 13*(1), 1739845.

Gibson, M. A., Gurmu, E., Cobo, B., Rueda, M. M., & Scott, I. M. (2020). Measuring hidden support for physical intimate partner violence: A list randomization experiment in South-Central Ethiopia. *Journal of Interpersonal Violence*. https://doi. org/10.1177/0886260520914546

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.

Greene, K. F., Domınguez, J., Lawson, C., & Moreno, A. (2012). *The Mexico 2012 Panel study*.

Halim, D., Can, E. R., & Perova, E. (2020). *What factors exacerbate and mitigate the risk of gender-based violence during COVID-19? East asia and pacific gender innovation lab*.

Hamadani, J. D., Hasan, M. I., Baldi, A. J., Hossain, S. J., Shiraji, S., Bhuiyan, M. S., et al. (2020). Immediate impact of stay-at-home orders to control COVID-19 transmission on socioeconomic conditions, food insecurity, mental health, and intimate partner violence in Bangladeshi women and their families: An interrupted time series. *The Lancet Global Health*. https://doi.org/10.1016/S2214-109X(20)30366-1

Heise, L. L., & Kotsadam, A. (2015). Cross-national and multilevel correlates of partner violence: An analysis of data from population-based surveys. *The Lancet Global Health, 3*(6), e332–e340.

Holbrook, A. L., & Krosnick, J. A. (2010). Social desirability bias in voter turnout reports: Tests using the item count technique. *Public Opinion Quarterly, 74*(1), 37–67.

Hsu, L. C., & Henke, A. (2020). COVID-19, staying at home, and domestic violence. *Review of Economics of the Household*, 1–11.

Imai, K. (2011). Multivariate regression analysis for the item count technique. *Journal of the American Statistical Association, 106*(494), 407–416.

Jamison, J. C., Karlan, D., & Raffler, P. (2013). Mixed-method evaluation of a passive mHealth sexual information texting service in Uganda. *Information Technologies and International Development, 9*(3), 1–28.

Karlan, D., & Zinman, J. (2011). *List randomization for sensitive behavior: An application for measuring use of loan proceeds. NBER working paper No. 17475*.

Kennedy, C., & Mann, C. B. (2015). "RANDOMIZE: Stata module to create random assignments for experimental trials, including blocking, balance checking, and automated rerandomization," Statistical Software Components S458028. Boston College Department of Economics. revised 01 Jun 2017.

Kuklinski, J., Cobb, M., & Gilens, M. (1997). Racial attitudes and the "new South". *The Journal of Politics, 59*(2), 323–349.

Larsen, M., Nyrup, J., & Petersen, M. B. (2020). Do survey estimates of the public's compliance with COVID-19 regulations suffer from social desirability bias? *Journal of Behavioral Public Administration, 3*(2).

Lax, J. R., Phillips, J. H., & Stollwerk, A. F. (2016). Are survey respondents lying about their support for same-sex marriage? Lessons from a list experiment. *Public Opinion Quarterly, 80*(2), 510–533.

Lépine, A., Treibich, C., & D'Exelle, B. (2020a). Nothing but the truth: Consistency and efficiency of the list experiment method for the measurement of sensitive health behaviours. *Social Science & Medicine, 266*, 113326.

Lépine, A., Treibich, C., Ndour, C. T., Gueye, K., & Vickerman, P. (2020b). HIV infection risk and condom use among sex workers in Senegal: Evidence from the list experiment method. *Health Policy and Planning, 35*(4), 408–415.

Leslie, E., & Wilson, R. (2020). Sheltering in place and domestic violence: Evidence from calls for service during COVID-19. *Journal of Public Economics, 189*(2020), 104241.

Mahmud, M., & Riley, E. (2020). Household response to an extreme shock: Evidence on the immediate impact of the Covid-19 lockdown on economic outcomes and well-being in rural Uganda. *World Development, 140*(2021), 105318.

McKenzie, D., & Siegel, M. (2013). Eliciting illegal migration rates through list randomization. *Migration Studies, 1*(3), 276–291.

Miller, J. D. (1984). *A new survey technique for studying deviant behavior. PhD diss*. The George Washington University.

Peterman, A., & O'Donnell, M. (2020a). *COVID-19 and violence against women and children: A second research round up*. Washington (DC): Center for Global Development.

Peterman, A., & O'Donnell, M. (2020b). *COVID-19 and violence against women and children. A third research round up for the 16 days of activism*. Washington (DC): Center for Global Development.

Peterman, A., O'Donnell, M., & Palermo, T. (2020b). *COVID-19 and violence against women and children. what have we learned so far*. Washington (DC): Center for Global Development.

Peterman, A., Palermo, T. M., Handa, S., Seidenfeld, D., & Zambia Child Grant Program Evaluation Team. (2018). List randomization for soliciting experience of intimate partner violence: Application to the evaluation of Zambia's unconditional child grant program. *Health Economics, 27*(3), 622–628.

Peterman, A., Potts, A., O'Donnell, M., Thompson, K., Shah, N., Oertelt-Prigione, S., et al. (2020a). *Pandemics and violence against women and children*. CGD Working Paper, 2020-4 (528).

Raghavarao, D., & Federer, W. T. (1979). Block total response as an alternative to the randomized response method in surveys. *Journal of the Royal Statistical Society, Series B, Methodological, 41*(1), 40–45.

Roesch, E., Amin, A., Gupta, J., & García-Moreno, C. (2020). Violence against women during covid-19 pandemic restrictions. https://doi.org/10.1136/bmj.m1712.

Sánchez, A., & Hidalgo, A. (2019). *Medición de la prevalencia de la violencia física y psicológica hacia niñas, niños y adolescentes, y sus factores asociados en el Perú. Evidencia de Niños del Milenio. Avance de Investigación N° 38*. Lima: GRADE.

Schneider, D., Harknett, K., & McLanahan, S. (2016). Intimate partner violence in the great recession. *Demography, 53*(2), 471–505.

Silverio-Murillo, A., Balmori de la Miyar, J. R., & Hoehn-Velasco, L. (2020). *Families under confinement: COVID-19, domestic violence, and alcohol consumption*. Andrew Young School of Policy Studies Research Paper Series Forthcoming.

Timmons, S., McGinnity, F., Belton, C., Barjaková, M., & Lunn, P. (2020). It depends on how you ask: Measuring bias in population surveys of compliance with COVID-19 public health guidance. *Journal of Epidemiology & Community Health*.

Tourangeau, R., & Yan, T. (2007). Sensitive questions in surveys. *Psychological Bulletin, 133*(5), 859.

Tsuchiya, T., Hirai, Y., & Ono, S. (2007). A study of the properties of the item count technique. *Public Opinion Quarterly, 71*, 253–272.

United Nations. (2020). *The Impact of COVID-19 on children*. Policy Brief.

Vandormael, A., Adam, M., Greuel, M., & Bärnighausen, T. (2020). A short, animated video to improve good COVID-19 hygiene practices: A structured summary of a study protocol for a randomized controlled trial. *Trials, 21*(1), 1–3.

Vyas, S., & Watts, C. (2009). How does economic empowerment affect women's risk of intimate partner violence in low and middle income countries? A systematic review of published evidence. *Journal of International Development: Journal of Developments in Sustainable Agriculture, 21*(5), 577–602.

World Health Organization. (2018). *Violence against women prevalence estimates, 2018*. Geneva: World Health Organization. Retrieved from www.who.int/publications/i /item/violence-against-women-prevalence-estimates.

World Health Organization, London School of Hygiene and Tropical Medicine, and South African Medical Research Council. (2013). *Global and regional estimates of violence against women: Prevalence and health effects of intimate partner violence and non-partner sexual violence*.