

A Service of



Leibniz-Informationszentrum Wirtschaft Leibniz Information Centre

Rao, Lakshman Nagraj; Gentile, Elisabetta; Pipon, Dave; Roque, Jude David; Thuy, Vu Thi Thu

Working Paper

The impact of computer-assisted personal interviewing on survey duration, quality, and cost: Evidence from the Viet Nam Labor Force Survey

GLO Discussion Paper, No. 605

Provided in Cooperation with:

Global Labor Organization (GLO)

Suggested Citation: Rao, Lakshman Nagraj; Gentile, Elisabetta; Pipon, Dave; Roque, Jude David; Thuy, Vu Thi Thu (2020): The impact of computer-assisted personal interviewing on survey duration, quality, and cost: Evidence from the Viet Nam Labor Force Survey, GLO Discussion Paper, No. 605, Global Labor Organization (GLO), Essen

This Version is available at: https://hdl.handle.net/10419/222414

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



The impact of computer-assisted personal interviewing on survey duration, quality, and cost: Evidence from the Viet Nam Labor Force Survey*

Lakshman Nagraj Rao,^a Elisabetta Gentile,^{ab} Dave Pipon,^a Jude David Roque,^a and Vu Thi Thu Thuy^c

^a Economic Research and Regional Cooperation Department, Asian Development Bank (ADB); ^b Global Labor Organization; ^c Population and Labor Statistics Department, General Statistics Office of Viet Nam

Abstract

We use a randomized field experiment to estimate the effect of computer-assisted personal interviewing (CAPI) on interview duration, number of errors, respondent perceptions, and cost. During Quarter 3 of the 2017 Labor Force Survey data collection for Ho Chi Minh City, Viet Nam, 15 households were randomly selected and interviewed using pencil-and-paper interviewing (PAPI), while another 15 households were randomly selected and interviewed using CAPI within each of a total of 180 sample enumeration areas. On average, CAPI interviews lasted 9.4 minutes less and had 0.8 less errors per questionnaire relative to PAPI. Respondents were more likely to perceive interview duration as long or very long when the enumerator was female or educated to college level or above, which is contrary to our experimental findings. Finally, the break-even number of interviews that make CAPI cost-effective is 1,769, which is lower than prior estimates and reflects the rapidly decreasing cost of technology.

Keywords: computer-assisted personal interviewing; data quality; randomized experiment; survey; labor statistics

IEL codes: C81, J21

^{*} We thank the General Statistics Office, Ministry of Planning and Investment, Viet Nam for their close collaboration in this project and implementing the fieldwork. Anna Christine Durante and Rea Jean Tabaco provided excellent research assistance. This paper benefited greatly from helpful comments received from Elaine Tan. We are grateful to all colleagues who offered help and advice, including Rana Hasan, Matteo Lanzafame, Pamela Lapitan, Yasuyuki Sawada, and Joseph E. Zveglich. Finally, we thank seminar attendees at ADB and conference participants at the 2018 Asian and Australasian Society of Labour Economics (AASLE) Conference and the 62nd ISI World Statistics Congress. The views expressed in this paper are those of the authors and do not necessarily reflect the views and policies of the Asian Development Bank (ADB) or its Board of Governors or the governments they represent. The standard disclaimer applies.

1. Introduction

Labor force surveys (LFS) are used to calculate the official measures of employment and unemployment in a country; they are the basis for labor market policy analysis and planning and provide essential data required for economic and social research. Yet, most LFS users do not concern themselves with the underlying data collection process, associated data errors, and potential implications for economic analysis.

In many developing economies, data collection continues to rely on the Pen-and-Paper Personal Interviewing (PAPI) technique, which involves a printed survey questionnaire coupled with a survey manual as means of gathering information in the field by an enumerator. Once questionnaires have been filled up and compiled. they would still need to be verified by a field supervisor before being manually entered into a data entry application. The best practice for data entry protocol involves two independent inputs of data and then a field by field comparison of the two entries by the data entry software. Inconsistent values between the two data entries are subsequently resolved either by pulling out the questionnaire and revisiting the field, or in the worst-case scenario, making a callback to the household. This is followed by a final set of data cleaning activities to produce a freshly minted dataset. It is only after a successful completion of these back-end processes that analysis is possible. Therefore, not only does PAPI require the printing and transportation of paper-based questionnaires to the field and then back to headquarters, but the process of data coding is extremely time-consuming and error-prone if the data ought to be entered manually.

The types of errors that can potentially affect PAPI surveys are grouped into four categories. Skip errors arise when the enumerator fails to skip the next question, or a series of questions based on the interviewee's response. For example, if the respondent answers 'No' when asked if they have children, then the following question(s) on the number, age, and sex of the children should be skipped. Data validation errors happen when an input to a question is restricted to a range (or list) of choices, but the response is outside that range. For example, if the respondent answers a question on total hours worked per week with a number that is higher than the total hours available in a week. Logic errors are identified through the implementation of cross/logical checks across modules of a questionnaire; for example, the same respondent answered 'Male' in one section of the questionnaire and 'Currently pregnant' in another section of the questionnaire. Finally, missing errors refer to unanswered items that are required to be filled based on the condition from the previous item. For example, if the respondent answers 'Yes' when asked if they have children, then the following questions on the number, age, and sex of children should not be blank.

Computer-Assisted Personal Interviewing (CAPI) involves transforming a paper-based questionnaire into an electronic questionnaire using one of the many platforms available¹. The electronic questionnaire is then downloaded to handheld devices like tablets and smartphones, allowing enumerators to record and store the responses directly into the devices. Algorithms alerting the enumerator of any of the

4

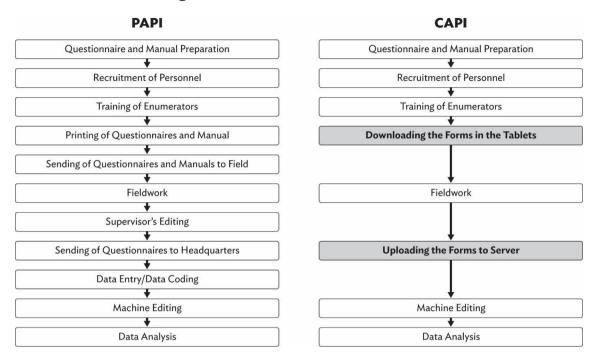
¹ For an overview of available CAPI platforms, see ADB (2019).

four types of aforementioned errors can be built into the questionnaire. The prompts provide the enumerators with the chance to confirm in real time whether responses are indeed valid. The data collected from CAPI surveys is then uploaded to either a physical or cloud server and is immediately ready for analysis. Figure 1 provides a comparison of the PAPI and CAPI workflow.

Far from being a new development, CAPI emerged three decades ago as a result of rapid advances in computing technology. However, it has remained the prerogative of developed economies (see, for example, Sebestik et al. 1988, Baker 1992, Banks and Laurie 2000), while in most developing countries major nationally representative surveys are still PAPI-based. Personal data assistants (PDAs) and rudimentary survey software were first used for implementing face-to-face surveys in developing economies only in the late 2000s (Byass et al. 2008; Caeyers et. al 2012). Lately, the emergence of low-cost smartphones, tablets, and other handheld devices, accompanied by a growth in the number of off-the-shelf survey software options, have made CAPI a more viable option. These new software support multilanguage implementation, GPS, audio, video, barcode, and other forms of multimedia data, among other novel features. However, the relative benefits of each of these novel sources of data and the cost of collecting them remain at the forefront of concerns for National Statistics Offices (NSOs).

From an NSO perspective, a transition from PAPI to CAPI makes operational sense if it delivers improved data quality and faster data availability, even though acquiring the necessary equipment may increase costs during the initial implementation stage, and respondents may need time to adjust to the new survey mode. Tabletbased surveys allow for built-in validation checks to reduce the occurrence of errors and improve data accuracy (Sebestik et al. 1988; Nicholls et al. 1997). For example, imposing a range of acceptable values for a question would lead the survey to reject input values outside that range. In addition, CAPI supports functionalities such as GPS tracking and the collection of photos and other multimedia files which can be used as paradata to further validate the quality of data (Choumert-Nkolo et al. 2019). Second, it does not require second data entry, which means that data is available for analysis much faster (de Leeuw 2008). Next, CAPI surveys require far less or no paper printouts, thus reducing transportation and enumerator costs, and making it possible to increase sample sizes. Finally, it is likely that in today's day and age, with technology becoming ubiquitous, survey respondents would be comfortable with being interviewed using modern technology. Limited evidence on respondent perceptions from developed economies points in favor of CAPI on this domain. However, there are likely to be differences in perception of technology between rural and urban households (Sun et al. 2020; ADB 2019). The earlier literature considered CAPI a viable option only for large surveys, due to the fixed front-end cost of hardware and software development, as opposed to the variable back-end cost of logistics, data entry and data cleaning (Weeks 1992). However, with the cost of technology decreasing at a relatively fast pace and the availability of a variety of CAPI platforms that are low-cost and easy to use, CAPI has become a cost-effective option with smaller and smaller sample sizes (Rahija 2016; ADB 2019). An important point to note is that the cost advantage of CAPI materializes as the relevant knowledge capital is accumulated and economies of scale are achieved, so that each subsequent survey becomes less and less expensive.

Figure 1. PAPI versus CAPI workflow



Source: ADB 2019.

In this paper, we use a randomized rollout of CAPI and PAPI surveys during the LFS data collection for Ho Chi Minh City, Viet Nam, in Quarter 3 of 2017 to determine the effect of CAPI surveys on interview time, number of errors, respondent perceptions, and cost implications. We find that CAPI reduced interview duration by 9.4 minutes per questionnaire and reduced the average number of errors per questionnaire by 0.8 errors, which amounts to roughly 46,000 errors for Quarter 3 of the 2017 LFS for Viet Nam. We also find that respondents are more likely to perceive interview duration as long or very long rather than just right when the enumerator is female or educated to college level or above, which is contrary to our experimental findings. Finally, the break-even number of interviews that make CAPI cost-effective is 1,769, which is lower than prior estimates, consistent with the fact that the cost of technology keeps decreasing at a relatively fast pace.

Previous studies in developing countries have focused on surveys that are not conducted regularly by NSOs, with only one study conducted by Caeyers et al. (2012) attempting to implement a randomized experiment in Tanzania. This study was conducted in 2009, when the landscape for CAPI was drastically different. Furthermore, the same enumerators were used for both CAPI and PAPI, which could lead to contamination, a potential threat to the identification strategy. That is because CAPI interviewers, having learned how the checks work on this new technology, may change the way they conduct PAPI surveys. Our study contributes to the literature by providing experimental evidence for a mainstream survey conducted regularly in a developing Asian economy. Furthermore, it accounts for contamination more rigorously by randomly assigning enumerators to CAPI or PAPI. Finally, it provides updated cost estimates factoring in the availability of newer tablets and free CAPI software and attempts to explain factors that affect respondent perception on interview duration.

The rest of the paper is organized as follows: Section 2 provides a brief overview of the existing literature; Section 3 introduces the setting and experimental design. Section 4 dives into the results from the empirical analysis, and Section 5 concludes with some policy implications and directions for future research.

2. Evidence from the Literature

The earlier literature on CAPI has largely been conceptual, calling for more evidence-based and systematic studies, especially in a developing country context. For ease of exposition, we present the findings from the literature across four dimensions: interview duration and timeliness, survey cost, improved data quality, and respondent perceptions.

Interview length has implications for questionnaire completion rates, as well as response burden. The evidence on the impact of CAPI in this area is mixed: Caeyers et al (2012) conducted a randomized control trial of households in Tanzania and found that interview times were reduced by 10% with CAPI relative to PAPI, mainly attributable to CAPI's automated routing and skipping features. Some of the earlier research corroborates this and finds that interview lengths would be 16 to 20 percent shorter with CAPI (Lynn 1998; Baker et al. 1995). Zhang et al. (2012), on the other hand, compare smartphones and PAPI in data collection of infant feeding practices in rural China and find that the average duration for the smartphone method is almost a minute shorter than pen-and-paper, but the difference is not statistically significant. Similarly, examining over 12,000 households in Ethiopia to compare Android-based electronic data collection and the traditional paper method, King et al. (2013) do not find a significant difference in interview durations between methods.

There is also evidence that CAPI interviews might take longer than PAPI ones (Martin et al. 1993; Fuchs et al. 2000; Watson and Wilkins 2011). Nonetheless, this may be attributable to the adjustment to new technology, as well as the comparison with experienced interviewers already proficient in administering PAPI surveys. In his study of household surveys in Africa, Leisher (2014) finds that the average interview duration for CAPI only drops after three full days of using tablets. This might reflect the time it takes for interviewers to familiarize themselves with the technology and suggests a lag before the benefits of CAPI versus PAPI materialize in terms of interview duration.

Next, we analyze the evidence on CAPI's ability to generate and deliver data in a timely and more efficient manner. Martin et al. (1993) acknowledge CAPI's ability to reduce the time interval between fieldwork and the availability of data, and King et al. (2013), studying Android-based data collection in Ethiopia, find that electronic data collection saves time by providing results immediately. In practice, because CAPI software has faster back-end processing, much of the data processing and cleaning is done by the end of an interview. This also implies the elimination of data entry, and therefore a reduction in data delivery time. Therefore, whether time is

gained or lost with CAPI interviews, substantial time savings are still realized through more timely and efficient data delivery.

A similar argument can be made for cost. CAPI requires a sizable investment in hardware, more front-end design and development work, and more training on how to use the software and devices than PAPI (Schrapler et al. 2010). However, the front-end fixed costs must be balanced against the potential savings from the elimination of back-end costs related to producing and dealing with paper forms, data entry, and data cleaning (Martin et al. 1993).

Caeyers et al. (2012) find that CAPI would be most cost-effective for samples larger than 4,000, since in this case the reduction of variable costs associated with PAPI would exceed the fixed costs related to CAPI. Leisher (2014) finds that the cost per interview of tablet-based surveys was 74% lower than the paper-based average and that the savings originate from the reduced need for data cleaning and processing, as well as lower enumerator fees. This is further corroborated by King et al. (2013) and Zhang et al. (2012), who find that savings from CAPI are greater than the initial front-end expenses, such as hardware and software.

As discussed in Section 1, the earlier literature established that CAPI has tremendous potential for improving data quality. However, few studies have attempted to systematically and empirically study the effect of CAPI on data quality and analysis. Caeyers et al. (2012) conducted a randomized control trial on three groups of households from Pemba Island in Tanzania. The first group had a PAPI survey administered to them. The second group had a 'restricted CAPI' survey, i.e. a CAPI survey without consistency checks. The third group had the 'full CAPI' survey, i.e. with consistency checks. The purpose of comparing restricted CAPI to full CAPI was to isolate the effect of consistency checks, which are believed to be important for data quality. They found the share of questionnaires with at least one missing or illogical value to be 83% with PAPI, 40% with restricted CAPI, and only 2% with full CAPI. They also found that the CAPI study groups resulted in more refined consumption estimates. For instance, had the authors done the survey using PAPI, they would have concluded that 21% of households were malnourished, i.e. living on less than 1,500 kcals a day. The full CAPI survey, on the other hand, would have concluded that only 8% of households were malnourished. The difference in estimates was found to be statistically significant at 1%. The study attributed the 13% difference between full CAPI and PAPI as 6% due to CAPI's consistency checks mechanism, and 7% due to CAPI's other features, such as the automated routing mechanism. The study noted that 94% of the errors in PAPI that were because of routing or skip errors were avoided with the CAPI software.

Fafchamps et al. (2012) sought to measure sales and profits in microenterprises in Ghana through four rounds of quarterly data collection from 2008 to 2009. Such data tend to be noisy, with high standard deviations and low autocorrelation over time of sales and profits for a given firm. The authors implemented two types of consistency checks in their survey: first, a cross-sectional comparison of profits and sales for a given firm, which was implemented in all rounds; and second, a panel check that alerted for large differences in sales and profits compared to the earlier round. The panel checks occurred in round 2 onwards for sales, and in round 3

onwards for profits. The results showed that the checks minimized variations in responses and increased the autocorrelation of sales and profits for given firms through data collection rounds. However, only about 3 to 13% of the observations in a survey round were revised, amounting to a modest effect. The study concluded that more than mismeasurement, it was the intrinsic volatility and variation of the microenterprise sector that played a bigger part in the noisiness of the data. The study then flagged this as a limitation of CAPI's consistency checks in improving data quality for firm profit and sales.

The difference in findings between the two studies suggests that there is room for further research to investigate the impact of CAPI on different types of data and different contexts. In particular, as both studies took place in Africa, it would be interesting to study the transition to CAPI in different regions of the world.

In the early days of CAPI, there was concern that respondents would be reluctant to be interviewed using modern technology due to unfamiliarity with the technology or privacy concerns. However, studies mostly found no significant differences in nonresponse or refusal rates between PAPI and CAPI (see, for example, Nicholls et al. 1997; Baker et al. 1995; Tourangeau et al. 1997). With the turn of the 21^{st} century, mobile phone ownership and internet usage have been climbing steadily in the developing world, especially in emerging economies. In Viet Nam, for example, 97 percent of adults own a mobile phone, 67 percent of which are smartphones; and 64 percent of adults use at least one social media platform or messaging app (Silver et al. 2019). Therefore, distrust of digital technologies is less likely to be an issue today than 30 years ago.

Enumerator characteristics, on the other hand, may have an impact on nonresponse rates as well as response accuracy. Di Maio and Fiala (2019) conducted a large-scale experiment in Uganda in which enumerators and respondents were randomly paired to study the 'enumerator effect.' They found that the experience and capacity of enumerators to elicit the correct answers from respondents have a large bearing on survey results. Case study evidence from four African countries as presented by Randall et al. (2013) corroborates these findings. The study shows that enumerator persuasion skills and the ability to follow precise wordings in interview manuals are crucial for a successful interview, especially in Africa, where survey jargon varies throughout regions and territories, and oftentimes interviewers are socioeconomically distant from respondents.

Di Maio and Fiala (2019) also found a strong and robust effect of enumerator gender on responses: respondents were less likely to report strong voting preferences when the enumerator was male. The authors hypothesized that female enumerators may be perceived as less frightening or intimidating. Axinn (1989) used data from a study in Nepal that randomly assigned enumerators to households to test the influence of enumerator gender (among other things) on data accuracy and found that gender affects responses to sensitive questions, observing under-reporting in male-led interviews relative to female-led ones.

The enumerator effect could be a factor in the transition from PAPI to CAPI as well. Schrapler et al. (2010), for example, used a multivariate logit model to determine

the factors surrounding nonresponse within the context of the transition from PAPI to CAPI in the German Socioeconomic Panel. They found that the first three survey waves exhibited significantly higher probability of refusals than PAPI, specifically in questions that had to do with gross and net income. While privacy and confidentiality concerns related to new technology cannot be excluded, the authors suggested that the enumerator's lack of confidence in CAPI, given its novelty, was conveyed to the respondents.

In sum, the literature on the effect of CAPI has been mostly focused on advanced economies, where front-end costs and skill requirements were less of a barrier to adoption in the early stages. That resulted in a lack of rigorous studies set in developing countries, and the little evidence available is overwhelmingly from Sub-Saharan Africa. Nonetheless, a key finding emerges that the effects of CAPI in areas such as data quality or respondent perceptions may depend on the type of survey. A second consideration is that enumerators play a key role in the successful rollout of CAPI. Finally, the declining cost of handheld devices and the availability of CAPI platforms that are free of charge and easy to learn is further shifting costs in favor of CAPI, requiring updated estimates of the threshold for cost effectiveness.

This paper makes several contributions to the literature. We focus on CAPI implementation on the actual LFS of Viet Nam, a major nationally representative survey in a developing economy. We estimate the effect of CAPI on survey duration, data quality, and respondent perceptions. Finally, we provide updated calculations of the break-even point for CAPI to become viable.

3. Setting and Experimental Design

The Labor Force Survey (LFS) is a major household survey conducted on a quarterly basis by the General Statistics Office (GSO) of Viet Nam. The survey instrument comprises 71 questions grouped into three sections, as shown in appendix table A1.²

The sampling frame for the LFS is based on the 2014 Intercensal Population and Housing Survey. The sample was drawn from a two-stage stratified sample, while the enumeration areas (EAs) were selected by probability proportional to the size of the two independent sub-sample frames (urban and rural). This is to ensure that the sample is representative nationally and annually for 63 provinces/cities. Further, the nationally representative sample can be disaggregated quarterly down to 6 economic regions, the cities of Hanoi and Ho Chi Minh City, and rural and urban areas. A total of 3,828 EAs were selected for the 2017 round of LFS, as shown in appendix table A2.

The randomized experiment was conducted by the GSO with technical assistance

٠

² The full questionnaire is available in the online Appendix.

from the Asian Development Bank³ and targeted Quarter 3 of the data collection in Ho Chi Minh City. Field operations were implemented during the first 15 days of July, August, and September 2017, respectively. Within each EA, 30 households were randomly selected, 15 randomly assigned to CAPI, and the remainder assigned to PAPI as part of the regular GSO LFS data collection efforts. Data collection within each EA was implemented such that CAPI and PAPI households were surveyed simultaneously. This ensured that households in the treated group did not learn of the differences in survey conduct for the control group (and vice versa), which could have led them to change their behavior.

Enumerators were separate for CAPI and PAPI surveys, but they all came from a general pool of enumerators in the GSO that have been contracted to conduct the LFS for several years. Both sets of enumerators were trained on the technical concepts of the questionnaire and the checks that were required of them to verify alongside field practice. In the case of the CAPI enumerators, an additional day of training was provided on how to use the tablets to navigate through the questionnaire and send the data file to their supervisors.

The difference in the total number of households interviewed is due to a small number of CAPI households refusing to participate in the survey, whereas there were no such refusals for PAPI households.⁴ Household weights were appropriately adjusted for the CAPI sample to account for attrition, since the sample design did not permit random replacement of households. This was done to adjust for minor changes in selection probabilities for the EAs with less than 15 CAPI households.

[Insert Table 1 about here]

The difference in the total number of household members interviewed between CAPI and PAPI is due to the difference in household composition, in particular the number of adult members of the household. While the main respondent for the LFS is the household head, when other members of the household are present, they are requested to provide feedback on targeted questions.

The CAPI survey was based on a CSPro ⁵ Android platform. For an effective comparison of the number of errors in CAPI surveys with respect to PAPI, the same checks that PAPI enumerators were trained to perform on the paper questionnaire were programmed into the CAPI questionnaire. In total, we implemented 28 checks for skip errors, 76 checks for data validation errors, 26 checks for logic errors, and 69 checks for missing errors.⁶

[Insert Table 2 about here]

³ Regional Capacity Development Technical Assistance (R CDTA) 9018 "Implementing Information and Communication Technology (ICT) Tools to Improve Data Collection and Management of National Surveys supporting the Sustainable Development Goals."

⁴ Dropping from our sample the enumerator areas where there were refusals to participate in CAPI surveys does not change the results of or analysis.

⁵ CSPro is a free software package developed by the US Census Bureau, primarily for the editing, processing, and dissemination of census and survey data, most widely used by National Statistics Offices (NSOs) in Asia and the Pacific.

⁶ The full list of implemented checks by type of error is available in the online Appendix.

Table 2 presents weighted summary statistics for both independent and dependent variables in our analysis: first, a series of household characteristics (panel A), followed by enumerator characteristics (panel B), and outcomes of interest (panel C). In the last column, we report a paired t-test for the null hypothesis that the group means are different for CAPI and CAPI surveys⁷.

Panel A in table 2 suggests that CAPI and PAPI households have comparable characteristics overall, indicating that the randomization was well implemented. Considering that the randomization took place within the context of a major nationally representative survey, the achievement of balanced samples is nontrivial. That said, CAPI households appear to have a slightly higher monthly wage, but the standard deviation is also higher than PAPI households.

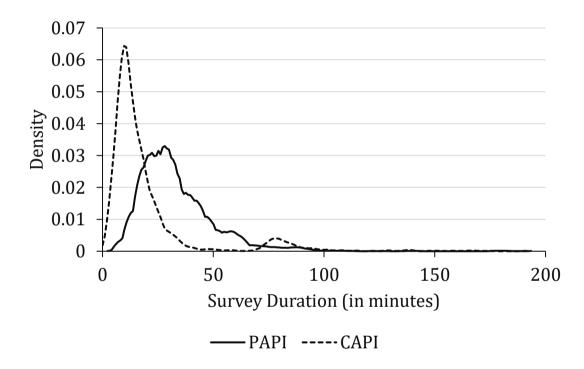


Figure 2. Density plot of survey duration

Source: Authors

Panel B in table 2 shows that although the enumerators were randomly assigned to either CAPI or PAPI households, a larger proportion of enumerators were assigned to PAPI relative to CAPI. This was done to ensure that there would be no delays on PAPI households, which form part of the national estimates for labor force statistics. A slightly higher percentage of CAPI enumerators (73.77%) were female relative to PAPI enumerators (69.47%), while the average age, experience, and proportion with college degrees are close for both cohorts. Although there is a statistically significant difference in the number of adults in the household between CAPI and PAPI households, the mean values are close in magnitude (2.80 for PAPI versus 2.86 for

⁷ Note that enumerator characteristics remain unweighted in the analysis, since the weights being employed are at the household level.

CAPI). Overall, we fail to reject the null hypothesis that the group means are equal for CAPI and PAPI enumerators.

Finally, Panel C in table 2 shows that average interview duration is significantly higher for PAPI than CAPI surveys. Interestingly, the average number of errors is slightly higher for CAPI surveys, but the difference in means is statistically insignificant. Figures 2 and 3 display the density plots of our outcomes of interest. The distribution of survey duration for both CAPI and PAPI is skewed to the right, but skewness for CAPI is larger. Furthermore, the distribution for CAPI is smoother than PAPI. The distribution of survey errors is also skewed to the right and remarkably similar for CAPI and PAPI.

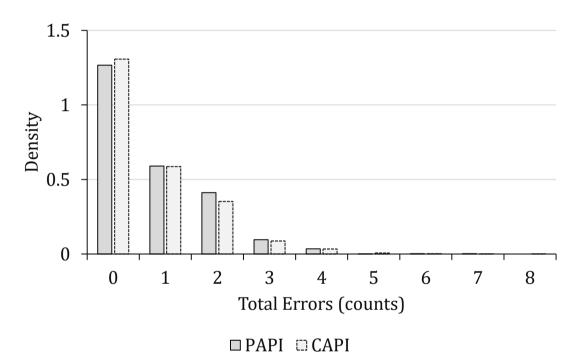


Figure 3. Density plot of total errors

Source: Authors

4. Results

A. Effect of CAPI on Total Survey Duration and Total Survey Errors

We define total survey duration as the number of minutes it took for a questionnaire to be finished by the household. In the case of a PAPI survey, enumerators accurately recorded the start and end time on the questionnaire, whereas a CAPI survey automatically recorded each interview's start time, end time, and duration by question, which could be added up to get a total survey duration per household.

The following simple OLS model computes the difference between the mean duration of CAPI versus PAPI interviews:

$$Y_{ijc} = \alpha + \beta \cdot CAPI_i + HH_i\Gamma + ENUM_i\Delta + \zeta_c + \eta_i + \epsilon_{ijc}$$
 (1)

 Y_{ijc} represents total survey duration (in minutes) for household i surveyed by enumerator j in cluster (i.e. enumeration area) c. $CAPI_i$ is the binary variable indicating treatment status (=1 for being randomized into the CAPI households, =0 otherwise), HH_i is a set of household head characteristics, and $ENUM_j$ is a set of enumerator characteristics. The coefficient α is the mean duration for PAPI interviews, and the coefficient β gives the difference between the mean duration of CAPI and PAPI interviews. Γ and Δ are vectors of coefficients for household head characteristics and enumerator characteristics, respectively. ζ_c represents cluster fixed effects, η_j enumerator fixed effects, and ϵ_{ijc} is the error term. Finally, we apply household weights to address survey design.

Column (1) in table 3 displays the results from the specification in equation (1). It shows that CAPI reduces interview duration by 9.5 minutes, and the coefficient is significant at the 1 percent level. The number of adults in the household increases interview duration by roughly 3 minutes. That may be because the enumerator needs to determine whether other members of the household have to provide feedback on targeted questions.

[Insert Table 3 about here]

Enumerator characteristics are significant determinants of interview duration: on average, interviews conducted by female enumerators last 7 minutes less than male enumerators, and the coefficient is significant at the 1 percent level. Interviews are one minute shorter for every additional year of experience of the enumerator. Finally, interviews conducted by college-educated enumerators, on average, are shorter by 5.7 minutes with respect to enumerators with education below college level, and the coefficient is significant at the 5 percent level.

In figure 4, we plot the coefficients with their standard error bands from a simple regression of the total duration of a CAPI interview on its place in the full sequence of interviews carried out by each CAPI enumerator. In other words, the explanatory variables of interest are whether the survey is the n^{th} conducted by each enumerator, where $n = \{2, 3, ..., 50 \text{ or more}\}$ and n = 1 is the excluded category. We end the sequence with '50 or more' because from roughly 50 interviews onward there are fewer and fewer observations. The controls in the regression include the full set of household characteristics in table 3 (i.e. sex, age, and wage of household head, and number of adults in the household), enumerator fixed effects, and enumeration area fixed effects.

Figure 4 shows evidence of a learning curve for CAPI enumerators. Despite some noise in the data, the overall trend is that the more interviews they conduct, the shorter the duration compared to their first interview.

We now turn our attention to total survey errors, which we define as the sum of all

types of errors (i.e. skip errors, data validation errors, logical errors, and missing errors) found within the questionnaire. We use the model in equation (1) once again, where the outcome of interest is now total survey errors. The results are in Column (2) of table 3.

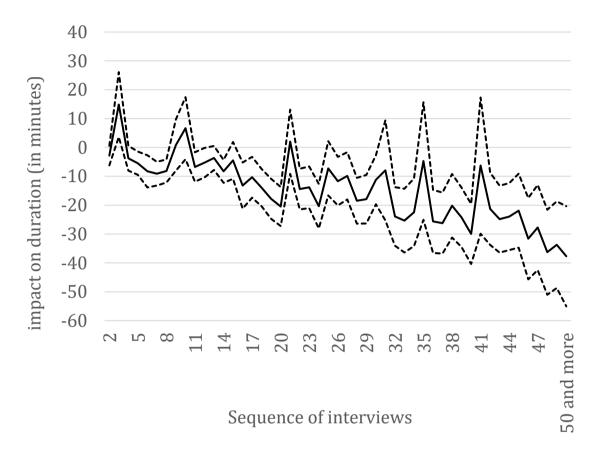


Figure 4. Impact of enumerator learning on CAPI interview duration

Source: Authors

We find that CAPI decreased the average number of errors per questionnaire by roughly 0.8 errors. In order to properly assess the magnitude of this effect, we must consider that the questionnaire comprised only 71 questions, making it short relative to, for example, enterprise surveys, which can have more than three times that number of questions. If we multiply 0.8 errors per questionnaire by 15 households per EA by 3,828 EAs for Quarter 3 of 2017 LFS, it amounts to a total number of errors just shy of 46,000. As mentioned in Section 3, both PAPI and CAPI surveys were conducted by well-trained enumerators who had no statistically significant differences in observable characteristics such as age, years of experience, gender breakdown, and education. Furthermore, the two sets of enumerators were randomly assigned to each group, further bolstering the validity of the statistically significant difference in average number of errors per questionnaire.

Household characteristics also affect the total number of errors. Older household heads are associated with a small but highly significant reduction in the average number of errors. Wage of household head and number of adults in the households

are associated with significant increases in the number of errors at the 5 percent and 1 percent level, respectively. Interestingly, enumerator characteristics do not have a significant effect on the incidence of errors in the questionnaires.

B. Respondent perceptions

The survey had a module at the end to assess household heads' perceptions of different aspects of the interview process. Appendix table A3 compares the responses to this module between CAPI and PAPI households. The first question is about perceptions of interview duration. Roughly a quarter of household heads found the interview long or very long regardless of survey method. Interestingly, household heads that were interviewed with the PAPI method were more likely to rate interview duration as short or very short than those interviewed with CAPI⁸.

We use a multinomial logistic regression model to investigate what factors affect household heads' perceptions of interview duration. We grouped the five possible ratings in appendix table A3 into three categories: "long or very long," "length is just right," and "short or very short." "Length is just right" (*just right*) is the reference category against which we compare the other two alternatives— "long or very long" (*long*) and "short or very short" (*short*)—as follows:

$$\ln \frac{\Pr(Z_{ijc} = long)}{\Pr(Z_{ijc} = just \ right)} = \alpha_1 + \beta_1 \cdot CAPI_i + PERC_i\Theta_1 + HH_i\Gamma_1 + ENUM_j\Delta_1$$
(2)

$$\ln \frac{\Pr(Z_{ijc} = short)}{\Pr(Z_{ijc} = just \ right)} = \alpha_2 + \beta_2 \cdot CAPI_i + PERC_i\Theta_2 + HH_i\Gamma_2 + ENUM_j\Delta_2$$
(3)

 Z_{ijc} represents household heads' perceptions of interview duration for household i surveyed by enumerator j in enumeration area c. It is a categorical variable that can take the values of 'long,' 'just right,' or 'short' as explained above. $PERC_i$ is a set of two variables related to household heads' perceptions (questions 2 and 4 in appendix table A3). The first is a binary indicator of whether the respondent enjoyed the interview process. The second is a categorical variable that takes the value of 0 if the respondent found the interview process easy, 1 if neither difficult nor easy, and 2 if difficult. Finally, Θ represents the vector of coefficients for $PERC_i$.

[Insert Table 4 about here]

Column (1) in table 4 displays the results from the specification in equation (2). The estimated coefficients represent the relative log odds of rating survey duration as

⁸ This is likely due to a learning curve for enumerators when it comes to navigating a CAPI survey, as shown in figure 4. In other words, the first few households take longer, but over time, CAPI enumerators catch up with their PAPI counterparts. Hence the number of cases falling in the 'just right' category is close between CAPI and PAPI. As mentioned in Section 1, the cost advantage of CAPI materializes as the relevant knowledge capital is accumulated.

long with respect to just right associated with each variable. For example, the relative log odds decrease by 0.19 for CAPI households compared to PAPI households, and the coefficient is significant at the 10 percent level. Similarly, the relative log odds decrease by 0.92 for household heads who enjoyed the interview process with respect to those who did not. The relative log odds of rating survey duration as long with respect to just right increase by 1.97 if moving from the lowest level of difficulty in answering questions (i.e. 'easy') to the highest (i.e. 'difficult'). The coefficients for household characteristics are all insignificant, except for the number of adults in the household, which is positive and significant. That is, again, because the survey contains questions such as employment status or sources of income for all adults in the household, thus increasing the perceived length of the interview process.

Perhaps the most interesting finding from Column (1) in table 4 is that enumerator characteristics have significant associations with household heads' perceptions of survey duration. The multinomial logit for female relative to male enumerators is 0.28 unit higher for rating survey duration as long with respect to just right, and the coefficient is significant at the 1 percent level. A one-year increase in enumerator age is associated with a 0.02 decrease in the relative log odds of rating survey duration as *long* with respect to *just right*. A one-year increase in enumerator experience is associated with a 0.05 decrease in the relative log odds of rating survey duration as *long* with respect to *just right*. Finally, the relative log odds increase by 0.36 for enumerators with education at college level or above, with the coefficient significant at the 1 percent level. While the results in table 3 show that interviews conducted by female enumerators or enumerators with college education and above lasted roughly 7 and 6 minutes less, respectively, table 4 suggests that household heads associated these same enumerator characteristics with a perception of longer interview duration. This is likely due to heterogeneity in frames of reference for household heads in terms of average survey durations from past experiences with participating in a survey, which may lead to differences in subjective perception of survey duration at any given level of objectively measured survey duration. 9 Furthermore, this could also be due to the socio-economic distance between interviewers and respondents discussed in Randall et al. (2013), despite the fact that the Ho Chi Minh City households were prevalently urban. 10

Column (2) in table 4 is mostly the mirror image of column (1), with a few exceptions. First, the relative log odds of rating survey duration as *short* with respect to *just right* decrease by 0.90 for CAPI households compared to PAPI households, and the coefficient is significant at the 1 percent level. A one dollar increase in the wage of the household head is associated with a 0.001 decrease in the relative log odds of rating survey duration as *short* with respect to *just right*. While the coefficient is relatively small, it is highly significant at the 1 percent level. The

⁹ While formally testing this hypothesis may be difficult due to data limitations, comparisons across subjective and objective measures for other variables such as welfare have been explored in depth in the economics literature and highlighted the role of frame-of-reference-bias (Beegle, Himelein, Ravallion 2016).

¹⁰ We explored this possibility further by adding an interaction of household head sex and enumerator sex to the models in equations (1), (2), and (3). The interaction is negative for total survey duration and positive for total survey errors, but statistically insignificant.

relative log odds associated with enumerator sex, age, and education are statistically insignificant.

C. Cost comparison

As mentioned in Section 1, both CAPI and PAPI surveys have fixed and variable costs, but they differ in terms of where the costs are concentrated. The cost comparison in table 5 shows that while CAPI has higher fixed costs related to front-end design and development work, PAPI has higher variable costs related to logistics of producing and dealing with paper forms, data entry, and data cleaning.

[Insert Table 5 about here]

The information in table 5 allows us to calculate the break-even sample size at which CAPI becomes more cost-effective than PAPI (N_{BE}):

$$N_{BE} = -\frac{FC_{CAPI} - FC_{PAPI}}{VC_{CAPI} - VC_{PAPI}} \tag{4}$$

where FC is fixed cost and VC is variable cost. We find that N_{BE} = 1,769. This is substantially lower than the threshold of 4,000 calculated by Caeyers et al. (2012), demonstrating that CAPI is becoming cost-effective for increasingly small sample sizes.

5. Conclusion

In this paper, we use the randomized rollout of CAPI and PAPI surveys during the LFS data collection for Ho Chi Minh City, Viet Nam, in Quarter 3 of 2017 to determine the effect of CAPI surveys on interview time, number of errors, cost implications, and respondent perceptions. Our findings confirm the benefits of CAPI in terms of interview duration and data quality: CAPI reduced interview duration by 9.4 minutes per questionnaire and reduced the average number of errors per questionnaire by 0.8 errors, which when scaled up amounts to a total number of errors just shy of 46,000 for the whole LFS data collection of Quarter 3 of 2017.

While our experimental results found that interviews conducted by female enumerators and enumerators with college education and above were roughly 7 and 6 minutes shorter, respectively, household heads' perception of interview duration went in the opposite direction. Household heads were more likely to rate interview duration as long or very long rather than just right when the enumerator was female or educated to college level or above. These findings suggest that socio-economic dynamics might be at play between enumerators and responders, even though the Ho Chi Minh City households were prevalently urban.

One key feature distinguishing CAPI from PAPI is the distribution of survey costs. PAPI has higher variable costs associated with data entry, printing, storage, logistics,

etc. CAPI, on the other hand, is traditionally associated with higher fixed costs of software development and hardware such as tablets and power banks. These costs, however, have been steadily decreasing over time. Consequently, the break-even sample size at which CAPI becomes more cost-effective than PAPI has been decreasing, too, and we calculate it at 1,769 interviews.

While our findings are encouraging, it is worth noting that, in order to reap the full benefits of CAPI, enough time must be allocated for programming and piloting the questionnaire and for enumerator training. Handheld devices now come in a wide range of prices and features, which calls for careful consideration of the performance-price ratio, especially for surveys with specific requirements such as GPS location or sound recording. While power banks have eased concerns over intermittent access to electricity in the field, poor data connectivity can be crippling in developing countries and may require alternative 'off-line' arrangements for data collection. Also, as of 2020, 107 countries in the world (of which 66 are developing or transition economies) have put in place legislation to secure the protection of data and privacy (UNCTAD 2020) and more countries have draft legislations in the works. Understanding how these laws may impact the implementation of CAPI surveys—e.g. those that rely on cloud-based servers to store the data uploaded by the enumerators—and securing compliance could be a challenge, especially in the early stages of implementation.

Our paper contributes to the growing literature on the impact of survey design choices and methods on survey measurement (see, for example, McKenzie and Rosenwig 2012; Godlonton et al. 2017; Dillon and Rao 2018) by focusing on one kind of digital technology, namely CAPI. The findings from our randomized experiment led the General Statistics Office (GSO) in Viet Nam to scale up CAPI operations for LFS data collection to the national level starting in 2018, showing that this kind of methodological research has significant policy implications.

With technology permeating the survey world in developing economies in the form of web surveys, SMS surveys, and interactive voice responses (Lau et al. 2019; Dabalen et al. 2016), particularly in light of global pandemics such as COVID-19, our results provide useful insights on the careful adoption of new technologies. There is still much to be learned about the impact of survey technology on measurement errors in labor force surveys, particularly in developing economies. Further exploration of the sources and determinants of these biases will improve causal estimates of these relationships by researchers and data-driven policy making.

References

- Asian Development Bank (ADB). 2019. The CAPI Effect: Boosting Survey Data Through Mobile Technology. A Special Supplement of the Key Indicators for Asia and the Pacific 2019. Manila: Asian Development Bank, April 2019.
- Axinn, William G. 1989. "Interviewer and data quality in less developed setting." *Journal of Official Statistics* 5: 265-280
- Baker, Reginald P. 1992. "New Technology in Survey Research: Computer-assisted Personal Interviewing (CAPI)". *Social Science Computer Review* 10: 145-157.
- Baker, Reginald P. and Norman N. Bradburn and Robert A. Johnson. 1995. "Computer-assisted Personal Interviewing: An Experimental Evaluation of Data Quality and Costs." *Journal of Official Statistics* 11(4). pp. 413-431.
- Banks, Randy and Heather Laurie. 2000. "From PAPI to CAPI: The Case of the British Household Panel Survey." *Social Science Computer Review* 18(4): 397-406.
- Byass, Peter and Sennen Hounton and Ouédraogo Moctar and Henri Somé and Ibrahima Diallo and Edward Fottrell and Axel Emmelin and Nicolas Meda. 2008. "Direct data capture using hand-held computers in rural Burkina Faso: Experiences, benefits and lessons learnt." *Tropical Medicine & International Health* 13: 25–30.
- Caeyers, Bet and Neil Chalmers and Joachim De Weerdt. 2012. "Improving Consumption Measurement and Other Survey Data Through CAPI: Evidence from a Randomized Experiment." *Journal of Development Economics* 98(2): 19-33.
- Choumert-Nkolo, Johanna and Henry Cust and Callum Taylor. 2019. "Using paradata to collect better survey data: Evidence from a household survey in Tanzania." *Review of Development Economics* 23(2): 598-618.
- Couper, Mick P. 2005. "Technology Trends in Survey Data Collection." *Social Science Computer Review* 23: 486-501.
- Dabalen, Andrew and Alvin Etang and Johannes Hoogeveen and Elvis Mushi and Youdi Schipper and Johannes von Engelhardt. 2016. *Mobile Phone Panel Surveys in Developing Countries: A Practical Guide for Microdata Collection*. Washington D.C.: World Bank.
- De Leeuw, Edith D. and Joop J. Hox and Ger Snijkers. 1995. "The Effect of Computer-Assisted Interviewing on Data Quality. A Review." *International Journal of Market Research* 37: 325-344.
- De Leeuw, Edith D. 2008. "The Effect of Computer-Assisted Interviewing on Data Quality: A Review of the Evidence." MethodikA: Utrecht University.
- Dillon, Andrew and Lakshman Nagraj Rao. 2018. "Land Measurement Bias: Comparisons from Global Positioning System, Self-Reports, and Satellite Data." ADB Economics Working Paper No. 540. Manila: Asian Development Bank.
- Di Maio, Michele and Nathan Fiala. 2019. "Be Wary of Those Who Ask: A Randomized Experiment on the Size and Determinants of the Enumerator

- Effect." The World Bank Economic Review lhy024.
- Fafchamps, Marcel and David McKenzie and Simon Quinn and Christopher Woodruff. 2012. "Using PDA Consistency Checks to Increase the Precision of Profits and Sales Measurement in Panels." *Journal of Development Economics* 98. pp. 51-57.
- Fuchs, Marek and Mick P. Couper and Sue Ellen Hansen. 2000. "Technology Effects: Do CAPI or PAPI Interviews Take Longer?" *Journal of Official Statistics* 16(3):273-286.
- Godlonton, Susan and Manuel A. Hernandez and Mike Murphy. 2018. "Anchoring Bias in Recall Data: Evidence from Central America." *American Journal of Agricultural Economics* 100(2): 479-501.
- King, Jonathan D. and Joy Buolamwini and Elizabeth A. Cromwell and Andrew Panfel and Tesfaye Teferi and Mulat Zerihun and Berhanu Melak and Jessica Watson and Zerihun Tadesse and Danielle Vienneau and Jeremiah Ngondi and Jürg Utzinger and Peter Odermatt and Paul M. Emerson. "A Novel Electronic Data Collection System for Large-Scale Surveys of Neglected Tropical Diseases." PLoS ONE 8(9).
- Lau, Charles Q. and Ansie Lombaard and Melissa Baker and Joe Eyerman and Lisa Thalji. 2018. "How Representative Are SMS Surveys in Africa? Experimental Evidence from Four Countries." *International Journal of Public Opinion Research* 31(2): 309–330.
- Leisher, Craig. 2014. "A Comparison of Tablet-Based and Paper-Based Survey Data Collection in Conservation Projects." *Social Sciences* 3(2): 264-271.
- Lynn, Peter. 1998. "Data Collection Mode Effects on Responses to Attitudinal Questions." *Journal of Official Statistics* 14(2): 1-14.
- McKenzie, David and Mark Rosenzweig. 2012. "Preface for symposium on measurement and survey design." *Journal of Development Economics* 98(1): 1-2
- Martin, Jean and Colm O'Muircheartaigh and John Curtice. 1993. "The Use of CAPI for Attitude Surveys: An Experimental Comparison with Traditional Methods." *Journal of Official Statistics* 9(3): 641-661.
- Nicholls, William L. and Edith D. de Leeuw. 1996. "Factors in Acceptance of Computer-Assisted Interviewing methods: A Conceptual and Historical Review." *Proceedings of the Section of Survey Research Methods*, American Statistical Association 758-763.
- Nicholls, William L. and Reginald P. Baker and Jean Martin. 1997. "The Effect of New Data Collection Technologies on Survey Data Quality," in *Survey Measurement and Process Quality*, edited by Lars Lyberg and Paul Biemer and Martin Collins and Edith De Leeuw and Cathryn Dippo and Norbert Schwarz and Dennis Trewin, 221-248. New York: Wiley, February 10, 1997.
- Olsen, Randy. 1992. "The Effects of Computer Assisted Interviewing on Data Quality." In 4th Social Science Methodology Conference. Trento, Italy.
- Rahija, Michael and Titus Mwisomba and Mlemba Abassy Kamwe and James Muwonge and Ugo Pica-Ciamarra. 2016. "Are CAPI Based Surveys a Cost-

- effective and Viable Alternative to PAPI Surveys? Evidence from Agricultural Surveys in Tanzania and Uganda." Proceedings of the ICAS VII Seventh International Conference on Agricultural Statistics, Rome, October 24-26, 2016.
- Randall, Sara and Ernestina Coast and Natacha Compaore and Philippe Antoine. 2013. "The power of the interviewer: a qualitative perspective on African survey data collection." *Demographic research* 28 (27): 763-792.
- Ravallion, Martin and Kristen Himelein and Kathleen Beegle, "Can Subjective Questions on Economic Welfare Be Trusted?" *Economic Development and Cultural Change* 64(4): 697-726.
- Schrapler, Jörg-Peter and Jürgen Schupp and Gert G. Wagner. 2010. "Changing from PAPI to CAPI: Introducing CAPI in a Longitudinal Study." *Journal of Official Statistics* 26(2): 239-269.
- Silver, Laura and Aaron Smith and Courtney Johnson and Kyle Taylor and Jingjing Jiang and Monica Anderson and Lee Rainie. 2019. *Mobile Connectivity in Emerging Economies*. Washington D.C.: Pew Research Center.
- Sun, Hanyu and Frederick G Conrad and Frauke Kreuter. 2020. "The Relationship between Interviewer-Respondent Rapport and Data Quality." *Journal of Survey Statistics and Methodology* smz043.
- Tourangeau, Roger and Kenneth Rasinski and Jared B. Jobe and Tom W. Smith and William F. Pratt. 1997. "Sources of Error in a Survey of Sexual Behavior." *Journal of Official Statistics* 13: 341-365.
- Watson, Nicole and Roger Wilkins. 2012. "Experimental Change from Paper-Based Interviewing to Computer-Assisted Interviewing in the HILDA Survey." Melbourne Institute Working Paper No. 6/12. The Melbourne Institute of Applied Economic and Social Research.
- Weeks, Michael F. 1992. "Computer-Assisted Survey Information Collection: A Review of CASIC Methods and their Implications for Survey Operations." *Journal of Official Statistics* 8(4): 445-466.
- Sebestik, Jutta and Harvey Zelon and Dale Dewitt and Jim O. O'Reilly and Kevin McGowan. 1988. "Initial Experiences with CAPI." Proceedings of the Bureau of the Census Fourth Annual Research Conference, Washington, D.C.: U.S. Department of Commernce, Bureau of the Census, 357-365.
- United Nations Conference on Trade and Development (UNCTAD). 2020. "Data Protection and Privacy Legislation Worldwide." https://unctad.org/en/Pages/DTL/STI and ICTs/ICT4D-Legislation/eCom-Data-Protection-Laws.aspx, accessed on February 11, 2020.
- Zhang, Shuyi and Qiong Wu and Michelle HMMT van Velthoven and Li Chen snd Josip Car and Igor Rudan and Yanfeng Zhang and Ye Li and Robert W Scherpbier. 2012. "Smartphone Versus Pen-and-Paper Data Collection of Infant Feeding Practices in Rural China." *Journal of Medical Internet Research* 14(5).

Table 1. Sampling characteristics

	CAPI	PAPI
Total number of enumeration areas	180	180
Total number of households interviewed	2,633	2,700
Average number of households per enumeration area	15	15
Total number of household members interviewed	9,502	9,533
Total number of household heads interviewed	2,633	2,700

Table 2. Summary statistics

	CAPI	PAPI	Total	t-statistic
		A. Household	characteristics	
Fraction of female household heads	41.72 (49.05)	41.07 (49.47)	41.40 (49.26)	-0.48
Age of household head	49.64 (14.68)	49.41 (15.09)	49.53 (14.89)	-0.67
Monthly wage (USD) of household head	249.09 (471.26)	245.74 (314.58)	247.41 (401.37)	-0.26
Number of adults in the household	2.86 (1.30)	2.80 (1.27)	2.83 (1.29)	-1.77*
Number of observations	2,633	2,700	5,333	
		B. Enumerator	Characteristics	
Fraction of female enumerators	73.77 (44.35)	69.47 (46.30)	71.15 (45.45)	-0.58
Age	34.52 (12.54)	35.91 (12.67)	35.37 (12.59)	0.67
Experience	4.02 (4.46)	4.46 (5.60)	4.29 (5.17)	0.53
Fraction with college education and above	75.41 (43.42)	73.68 (44.27)	74.36 (43.81)	-0.24
Number of observations	61	95	156	
		C. Outcome	es of interest	
Total interview duration	21.71 (41.69)	34.53 (30.03)	28.13 (36.94)	9.28***
Number of errors per household	0.85 (1.91)	0.78 (1.01)	0.81 (1.53)	-1.56
Number of observations	2,633	2,700	5,333	

Notes: Standard deviations in parentheses. The t-statistics represent a paired samples statistic with standard errors clustered by country for the null hypothesis that the group means are equal for CAPI and CAPI questionnaires. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 3. The effect of CAPI on total survey duration and total survey errors

Table 3. The effect of CAFT on total survey	Total Survey	-
	Duration	Errors
	(1)	(2)
	(-/	(-/
Method of data collection	0.4.50 dadada	0.505.66
CAPI (base = $PAPI$)	-9.463***	-0.797**
TT 1111 2 2 2 2	(2.690)	(0.337)
Household characteristics		
Sex of household head (1 if female)	0.634	-0.030
	(1.224)	(0.044)
Age of household head	-0.072**	-0.008***
	(0.032)	(0.001)
Wage (per month in USD)	1.44e-05	0.000247**
	(0.00137)	(0.000101)
Number of adults in the household	2.917***	0.266***
	(0.371)	(0.020)
Enumerator characteristics		
Sex (1 if female)	-7.022***	0.063
	(2.474)	(0.312)
Age	-0.194**	0.015*
	(0.092)	(0.008)
Experience (in years)	-0.987**	0.0333
	(0.380)	(0.036)
Squared Experience (in years)	0.023**	-0.002**
	(0.010)	(0.001)
Education (base = below college level)	-5.745**	0.060
,	(2.313)	(0.102)
Constant	53.44***	0.526
	(6.189)	(0.535)
	, ,	, ,
Observations	5,333	5,333
R-squared	0.190	0.153
Enumeration Area Fixed Effects	Yes	Yes
Enumerator Fixed Effects	Yes	Yes
Sampling weights	Yes	Yes

Note: Standard errors of the means in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, 1% levels, respectively.

Table 4. Multinomial logit model of household heads' perceptions of survey duration

Table 4. Multinomial logit model of household	Long	Short
	(1)	(2)
Method of data collection		
CAPI (1 if CAPI)	-0.1917**	-0.9046***
Chi i (i ii Chi i)	(0.0942)	(0.1367)
Respondents perception	(0.0) 12)	(0.1307)
Enjoyment (1 if Yes)	-0.9181***	0.7561***
Zinjoyinoni (Tir Tes)	(0.0958)	(0.1899)
Difficulty to answer (0 if easy)	(0.0500)	(0.1055)
Neither difficult nor easy	0.5546***	-1.1674***
•	(0.1173)	(0.1411)
Difficult	1.9717***	-1.2481***
	(0.1751)	(0.4696)
Household characteristics	,	,
Sex of household head (1 if female)	0.0882	0.1397
	(0.0954)	(0.1342)
Age of household head	-0.0033	0.0059
-	(0.0034)	(0.0048)
Wage (per month in USD)	-0.00004	-0.00116***
	(0.00014)	(0.00032)
Number of adults in the household	0.4044***	-0.5675***
	(0.0406)	(0.0758)
Enumerator characteristics		
Sex (1 if female)	0.2750***	0.2132
	(0.1028)	(0.1607)
Age	-0.0239***	0.0029
	(0.0046)	(0.0051)
Experience (in years)	0.0512***	-0.0094
	(0.0090)	(0.0093)
Education (1 if college level or above)	0.3576***	-0.1089
	(0.1150)	(0.1361)
Constant	-1.8140***	-0.5630
	(0.3011)	(0.4398)
Observations	3,393	3,393
Pseudo R-squared	0.159	0.159
Enumeration area fixed effects	No	No
Enumerator fixed effects	No	No
Sampling weights	Yes	Yes

Note: The reference category for the multinomial logistic regression is "Length is just right." Standard errors of the means in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, 1% levels, respectively.

Table 5. PAPI-CAPI cost comparison

	CAPI	PAPI
	A. Fixed	costs
Software	-	-
Programming	$USD 400 \times 29 \text{ days} =$ = $USD 11,500$	USD 400 × 15 days = = USD 6,000
Total fixed costs (FC $_i$)	USD 11,500	USD 6,000
	B. Variabl	le costs
Tablets (including data and power bank)	USD 195 ÷ 912 days in the field ÷ 5 interviews per day = = USD 0.04 per interview	-
Enumerators Data entry	· -	USD 0.65 per interview
Printing Miscellanea (storage, bags, logistics, etc.)	- USD 0.50 per interview	USD 1.50 per interview USD 1.50 per interview
Total variable costs (VC_i)	USD 0.54 per interview	USD 3.65 per inverview

Note: Enumerator costs are the same regardless of survey mode; therefore they are not included in the cost analysis.

Table A1. Components of survey instrument

Section and su	bsection	Count
Part 1: Inform	nation on Usual Residents of Household	8
Part 2: Some 1	Basic Characteristics of Respondents	7
Part 3: Questi	ons for Classifying the Economically Active Status	
A:	Information on Employment	8
	I. The Main Job	15
	II. Income and Working Hours	4
	III. Underemployment	3
	IV. Self-Evaluation	6
Item B	: Information on the Unemployment of Outside the Labor Force	16
C:	Specific Information About Field of Study/Training	4
Total		71

Source: Authors based on the 2017 Labor Force Survey Questionnaire of Viet Nam

Table A2. Total number of sampled enumeration areas per quarter, LFS 2017

Administrative Uni	Urban	Rural	Total
Entire Country	1,668	2,160	3,828
Ha Noi	126	90	216
Ho Chi Minh	150	30	180
Other Provinces	-	-	54-60

Source: General Statistics Office, Viet Nam

Table A3. Respondent perceptions of interview process by method

Table A3. Respondent perceptions of intervio	-	
	CAPI (%)	PAPI (%)
Q1: What did you think of the duration of t	he interview	?
Very long	3.55	3.72
Long	21.17	20.82
Length is just right	69.58	62.01
Short	5.53	12.84
Very short	0.17	0.60
Q2: Did you enjoy participating in the inter	view?	
Yes	66.30	66.70
No	33.70	33.30
Q3: How satisfied or dissatisfied are you wi	th the interv	riew
based on the flow of questions?		
Very dissatisfied	1.03	0.43
Dissatisfied	5.78	4.97
Neither dissatisfied nor satisfied	60.42	54.67
Satisfied	27.68	33.29
Very satisfied	5.10	6.64
Q4: Did you find the questions easy or diffic	cult to answe	er?
Very difficult	0.13	1.22
Difficult	6.95	9.09
Neither difficult nor easy	60.92	48.34
Easy	30.05	37.04
Very easy	1.95	4.30
Q5: If we went through the survey again, do	you think a	ny
answers would change?		
Yes	2.03	7.07
No	97.97	92.93
Q6: Did you feel comfortable talking to the	interviewer	?
Not comfortable	1.66	0.85
A little comfortable	10.42	10.34
Comfortable	29.12	33.46
Very comfortable	58.80	55.34
Q7: If we're not recording the answers (just	t talking to v	ou), how
much would your answer have changed?	, turning to j	04), 110 ;;
Totally different	0.39	2.05
A bit different	12.19	12.56
No change	87.42	85.38

ONLINE APPENDIX The impact of computer-assisted personal interviewing on survey duration, quality, and cost: Evidence from the Viet Nam Labor Force Survey

Lakshman Nagraj Rao, Elisabetta Gentile, Dave Pipon, Jude David Roque, and Vu Thi Thu Thuy

28 April 2020

MINISTRY OF PLANNING AND INVESTMENT

GENERAL STATISTIC OFFICE

THE 2017 LABOR FORCE SURVEY QUESTIONNAIRE ON LABOR AND EMPLOYMENT INFORMATION

Г

(The information collected from the Labor Force Survey will be used and kept confidentially in accordance with Statistical Law)

SAMPLE DIGITS USED TO FILL IN LARGE BOX 0 1 2 3 4 5 6 7 8 9	1 NUMBER OF ACCTUAL	LY USUAL RESIDENTS OF HOU	ISEHOLD:
CROSS "X" INTO SMALL BOX TO INDICATE THE CORESSPONDING ANSWER X	1. NOMBER OF AGOTORE	ET GOOKE REGIDEIVIO OF FIOO	JOET IOLD
PROVINCE/CITY:	OF WHICH, FEMALES:		
DISTRICT/QUARTER/TOWN/PROVINCIAL GOVERNED CITY:	2. NUMBER OF RESPOND	PENTS AGED 15*:	
COMMUNE/WARD/DISTRICT TOWN:	OF WHICH, FEMALE RES	PONDENTS AGED 15+:	
ENUMERATION AREA NUMBER			
ENUMERATION AREA'S NAME:	SIG	NATURE AND VERIFICATIO	ON
HOUSEHOLD NUMBER:		FULL NAME	SIGNATURE
FULL NAME OF HOUSEHOLD HEAD.			
FULL NAME OF HOUSEHOLD HEAD:	INFORMATION PROVIDER		
ADDRESS OF HOUSEHOLD:	INFORMATION PROVIDER		
	INFORMATION PROVIDER INTERVIEWER		

PART 1: INFORMATION ON USUAL RESIDENTS OF HOUSEHOLD

SERIAL NUMBER	SERIAL Nº					
Pls, let me know the full name of each person usually residing in the household, starting with the household head? (USE THE INTERVIEW PROCEDURE TO DEFINE ACTUALLY USUAL RESIDENT)						
What is [NAME]'s relationship to the household head?	HH. HEAD					
3. Is [NAME] male or female?	MALE1 FEMALE2					
4. In what solar month and year was [NAME] born?	MONTH YEAR Q6 NOT STATED 9998					
5. At present, what is [NAME]'s age as of last birthday? WRITE '95' IF AGE IS 95 AND ABOVE	COMPLETED AGE.	COMPLETED AGE.	COMPLETED AGE	COMPLETED AGE.	COMPLETED AGE.	COMPLETED AGE.
CHECK: Q.4 or Q.5: ASK	(PERSONS AGED 15+.					

HOUSEHOLD:		

SERIAL NUMBER OUESTION SERIAL NO							
NAME PART 2 ABROAD 2	SERIAL NUMBER QUESTION	SERIAL NO					
7. What was the main reason why [NAME] moved abroad? MORK	[NAME] residing in	PART 2					
reason why [NAME] moved abroad? STUDY		NAME OF COUNTRY					
[NAME] been resident in that country? YEAR	reason why [NAME]	STUDY	STUDY	STUDY	STUDY2	STUDY2	STUDY2
IF THERE ARE OTHER RESPONDENTS, ASK THE NEXT; OTHERWISE MOVE TO PART 2.	[NAME] been resident in that						
	IF THERE ARE OTHER RESPONDENTS, ASK THE NEXT; OTHERWISE MOVE TO PART 2.						

PART 2: SOME BASIC CHARACTERISTICS OF RESPONDENT

Г

(Only asking those persons aged 15^+ and currently residing in Vietnam – If Q.6 = 1)

NAME AND SERIAL NO QUESTION				
9. What is your current marital status?	NEVER MARRIED 1 MARRIED 2 WIDOWED 3 DIVORCED 4 SEPARATED 5	NEVER MARRIED 1 MARRIED 2 WIDOWED 3 DIVORCED 4 SEPARATED 5	NEVER MARRIED 1 MARRIED 2 WIDOWED 3 DIVORCED 4 SEPARATED 5	NEVER MARRIED 1 MARRIED 2 WIDOWED 3 DIVORCED 4 SEPARATED 5
10. For how long have you been usually residing in the current ward, town or commune?	UNDER 1 MONTH	UNDER 1 MONTH	UNDER 1 MONTH	UNDER 1 MONTH
11. Is your previous place of usual residence a ward, town or commune?	WARD/TOWN	WARD/TOWN	WARD/TOWN1 COMMUNE	WARD/TOWN 1 COMMUNE 2
12. From which province/city did you move?	IN VIET NAM	IN VIET NAM	IN VIET NAM	IN VIET NAM

HOUSEHOLD:		

NAME AND SERIAL NO QUESTION				
13. What was the main reason that you moved to this household?	TO FIND JOB	TO FIND JOB	TO FIND JOB	TO FIND JOB
L	FAMILY/RETIRED	FAMILY/RETIRED	FAMILY/RETIRED	FAMILY/RETIRED
14. What is the highest educational level that you have attained? **ABBREVIATION:** PRI - PRIMARY PROF PROFESSIONAL	(SPECIFY) NEVER ATTENDED SCHOOL1 NOT COMPLETED PR. SCHOOL2 PRIMARY SCHOOL3 LOWER SECONDARY SCHOOL4 UPPER SECONDARY SCHOOL5 MID-TERM PROF. SCHOOL6 PROFESSIONAL COLLEGE7 UNIVERSITY8 ABOVE UNIVERSITY	(SPECIFY) NEVER ATTENDED SCHOOL 1 NOT COMPLETED PR. SCHOOL 2 PRIMARY SCHOOL	(SPECIFY) NEVER ATTENDED SCHOOL 1 NOT COMPLETED PR. SCHOOL 2 PRIMARY SCHOOL	(SPECIFY) NEVER ATTENDED SCHOOL
15. What is the highest vocational skill/ certificate/ diploma that you have attained?	NO TECH. QUALIFICATION OR NO VOCATIONAL SKILL 1 TECH. WORKER WITHOUT DIPL/CERT 2 UNDER 3 MONTH VOCATIONAL SKILL 3	NO TECH. QUALIFICATION OR NO VOCATIONAL SKILL TECH. WORKER WITHOUT DIPL./CERT. UNDER 3 MONTH VOCATIONAL SKILL	NO TECH. QUALIFICATION OR NO VOCATIONAL SKILL 1 TECH. WORKER WITHOUT DIPL./CERT. 2 UNDER 3 MONTH VOCATIONAL SKILL	NO TECH. QUALIFICATION OR NO VOCATIONAL SKILL 1 TECH. WORKER WITHOUT DIPL./CERT. 2 UNDER 3 MONTH VOCATIONAL SKILL
ABBREVIATION: CERT CERTIFICATE TECH TECHNICAL DIPL DIPLOMA	UNDER 3 MONTH VOCATIONAL CERT	UNDER 3 MONTH VOCATIONAL CERT	UNDER 3 MONTH VOCATIONAL CERT	UNDER 3 MONTH VOCATIONAL CERT4 PRIMARY VOCATIONAL CERT5 MID-TERM VOCATIONAL CERT6 VOCATIONAL COLLEGE DIPL7

PART 3: QUESTIONS FOR CLASIFYING THE ECONOMICALLY ACTIVE STATUS

NAME AND SERIAL NO QUESTION				
16. In the last 7 days, did you do any work for at least 1 hour to get salary/wage?	YES1 ITEM A NO2			
17. In the last 7 days, did you engage into or do any productive or business activities for at least 1 hour to generate income?	YES1 ITEM A NO2			
18. In the last 7 days, did you do any unpaid work for at least 1 hour to generate your family income?	YES1 ITEM A NO2			
19. What was the reason that you did not work in the last 7 days?	TEMPORARILY ABSENT	TEMPORARILY ABSENT	TEMPORARILY ABSENT	TEMPORARILY ABSENT

HOUSEHOLD		L	

DUSEHOLD		1				
ハルシヒロいカリン						

NAME AND SERIAL NO QUESTION				
20. You are on the temporary absence from work, so what was your main reason of absence? **ABBREVIATION:** **PROD PRODUCTION**	WORK BY SHIFT	WORK BY SHIFT	WORK BY SHIFT	WORK BY SHIFT
21. During the temporary absence, did you receive salary/wage or get profit from any productive or business activities?	YES1 → ITEM A NO2	YES1 → ITEM A NO	YES1 — ITEM A NO	YES1 → ITEM A NO
22. After the temporary absence, are you going to be recalled for that work?	YES1 NO	YES1 ☐ NO	YES1	YES1 □ NO
23. Hence, how long will you return work?	UNDER 1 MONTH	UNDER 1 MONTH	UNDER 1 MONTH	UNDER 1 MONTH

<u>Check</u>: Q16 = 1 or Q17 = 1 or Q18 = 1 or Q21 = 1 or Q23 = 1 skip to Item A to ask for information on employment; Q19 = (2/3/4/5/6/7) or Q22 = 2 or Q23 = (2/3/4) skip to Item B to ask for information on the unemployment and outside the labour force.

ITEM A: INFORMATION ON EMPLOYMENT

(Asking for the job worked in the last 7 days/or the job worked before the temporary absence)

NAME AND SERIAL NO QUESTION				
I. THE MAIN JOB				
24. In your main job/business, what kind of work do you usually do in last 7 days or temporarily absent? WRITE CLEARLY:				
-MAIN TASKS AND DUTIES - OCCUPATION TITLE (IF ANY) Example:	OCCUPATION TITLE, IF ANY			
Cattle farmer: breed, raise and sell cattle.				
Policeman: patrol the streets Cook: plan and prepare meal Primary school teacher: teach children how to write and write	MAIN TASKS AND DUTIES (ISCO CODE)			
25. What is the main activity of the business or establishment where you work?				
WRITE CLEARLY: - NAME (IF ANY)	NAME OF ESTABLISHMENT (IF ANY)	NAME OF ESTABLISHMENT (IF ANY)	NAME OF ESTABLISHMENT (IF ANY)	NAME OF ESTABLISHMENT (IF ANY)
- MAIN ACTIVITY, GOODS OR SERVICES Example:				
Cong Vi ward police: ensure security and order of the ward.	MAIN ACTIVITY, GOODS OR SERIVES			
Hoa Lan Restaurant: food and drink business.	(ISIC CODE)	(ISIC CODE)	(ISIC CODE)	(ISIC CODE)
Bien Dong Shipping Company: long- distance transport of goods	(13.13.3322)	(.5.5 3322)	(1212 3 322)	(:-:: -)

HOUSEHOLD		
HOUSEHOLD		

SEHOLD				L
NAME AND SERIAL NO QUESTION				
26. Does the establishment that you worked belong to one kind of economic unit as follows? (Read out answer choices) ABBREVIATION: BUS. – BUSINESS ORG. – ORGANIZATION ASS. – ASSOCIATION IN INDIVIDUAL	FARM HOUSEHOLD	FARM HOUSEHOLD	FARM HOUSEHOLD	FARM HOUSEHOLD
27. Are the main products or services that you produced or provided for sales/trade or for own use of household?	ONLY FOR SALE/TRADE	ONLY FOR SALE/TRADE	ONLY FOR SALE/TRADE	ONLY FOR SALE/TRADE
28. Did the establishment that you worked have the business registration?	YES	YES1 NO2	YES1 NO2	YES1 NO2
29. Did the establishment that you worked have the trade union?	YES1 NO2	YES1 NO2	YES1 NO2	YES1 NO2

 Γ NAME AND SERIAL NO QUESTION 30. With the above job, are you? EMPLOYER (USUALY HIRE LABORS) 1 EMPLOYER (USUALY HIRE LABORS).... 1 EMPLOYER (USUALY HIRE LABORS) 1 EMPLOYER (USUALY HIRE LABORS).... 1 OWN ACCOUNT WORKER 2 OWN ACCOUNT WORKER 2 OWN ACCOUNT WORKER... OWN ACCOUNT WORKER 2 (Read out answer choices) FAMILY CONTR. WORKER...... 3 FAMILY CONTR. WORKER...... 3 FAMILY CONTR. WORKER......3 FAMILY CONTR. WORKER 3 Q34 Q34 Q34 Q34 ABBREVIATION: MEMBER OF COOPERATIVE ... 4 MEMBER OF COOPERATIVE .. 4 MEMBER OF COOPERATIVE .. 4 MEMBER OF COOPERATIVE .. 4 CONTR. - CONTRIBUTING EMPLOYEE 5 EMPLOYEE......5 EMPLOYEE 5 EMPLOYEE5 INDEFINITE TERM CONT......1 INDEFINITE TERM CONT......1 INDEFINITE TERM CONT......1 INDEFINITE TERM CONT......1 31. What type of labor contract did you hold, with the above Q33 Q33 Q33 < Q33 iob? 1 TO <3 YEAR CONT. 2 1 TO <3 YEAR CONT. 2 1 TO <3 YEAR CONT.....2 1 TO <3 YEAR CONT.....2 (Read out answer choices) 3 MONTH TO < 1 YEAR CONT... 3 3 MONTH TO < 1 YEAR CONT... 3 3 MONTH TO < 1 YEAR CONT ... 3 3 MONTH TO < 1 YEAR CONT...3 UNDER 3 MONTH CONT 4 UNDER 3 MONTH CONT...... 4 UNDER 3 MONTH CONT4 UNDER 3 MONTH CONT.....4 ABBREVIATION: LUMP SUM CONT 5 LUMP SUM CONT 5 LUMP SUM CONT.....5 LUMP SUM CONT......5 CONT. - LABOR CONTRACT VERBAL AGREEMENT 6 NO LABOR CONTRACT......7 NO LABOR CONTRACT......7 NO LABOR CONTRACT.....7 NO LABOR CONTRACT7 32. What was the reason that ON JOB TRAINING/APPRENTICE. 1 ON JOB TRAINING/APPRENTICE. 1 ON JOB TRAINING/APPRENTICE..1 ON JOB TRAINING/APPRENTICE . 1 vou held such kind of labor ON PROBATION.....2 ON PROBATION.....2 ON PROBATION.....2 ON PROBATION.....2 contract above or no labor contract? SEASONAL JOB...... 3 SEASONAL JOB...... 3 SEASONAL JOB......3 SEASONAL JOB3 DAILY/CASUAL JOB..... 4 DAILY/CASUAL JOB..... 4 DAILY/CASUAL JOB4 DAILY/CASUAL JOB4 DO A FAVOUR 5 DO A FAVOUR 5 DO A FAVOUR 5 DO A FAVOUR 5 AS REGULATED BY THE AS REGULATED BY THE AS REGULATED BY THE AS REGULATED BY THE ESTABLISHMENT 6 ESTABLISHMENT 6 ESTABLISHMENT......6 ESTABLISHMENT......6 OTHERS OTHERS OTHERS OTHERS (SPECIFY) (SPECIFY) (SPECIFY) (SPECIFY)

HOUSEHOLD: .	١		

NAME AND SERIAL NO QUESTION				
33. By what manner is for the payment that you received, with the above job?	FIXED SALARY	FIXED SALARY1 PER WORKING DAY/HOUR2	FIXED SALARY1 PER WORKING DAY/HOUR2	FIXED SALARY1 PER WORKING DAY/HOUR2
with the above job!	PER PRODUCT/PIECE3	PER PRODUCT/PIECE3	PER PRODUCT/PIECE 3	PER PRODUCT/PIECE3
	PER REVENUE4 FIXED SALARY AND REVENUE5	PER REVENUE4 FIXED SALARY AND REVENUE5	PER REVENUE	PER REVENUE4 FIXED SALARY AND REVENUE5
_	REVENUE AND COST	REVENUE AND COST6 OTHERS 7	REVENUE AND COST	REVENUE AND COST6 U
I	OTHERS77	(SPECIFY)	(SPECIFY)	(SPECIFY)
34. Have you paid for social insurance, with the above job?	YES	YES1 NO	YES1 NO	YES1 NO
35. What kind of social insurance have you paid for?	MANDATORY 1	MANDATORY 1	MANDATORY1 VOLUNTARY2	MANDATORY1 VOLUNTARY2
36. For how long have you worked, with the above job?	UNDER 6 MONTHS	UNDER 6 MONTHS	UNDER 6 MONTHS1 6 TO UNDER 12 MONTHS2 1 TO UNDER 5 YEARS	UNDER 6 MONTHS 1 6 TO UNDER 12 MONTHS 2 1 TO UNDER 5 YEARS 3
L	5 TO UNDER 10 YEARS4	5 TO UNDER 10 YEARS	5 TO UNDER 10 YEARS	5 TO UNDER 10 YEARS 4
	10 YEARS AND MORE5 Q39	10 YEARS AND MORE5 Q39	10 YEARS AND MORE5 Q39	10 YEARS AND MORE 5 Q39
37. Before taking in the above job, what did you do?	IN ANOTHER JOB1 WAITING for JOB/OFF SEASON .2	IN ANOTHER JOB1 WAITING for JOB/OFF SEASON .2	IN ANOTHER JOB1 WAITING for JOB/OFF SEASON.2	IN ANOTHER JOB1 WAITING for JOB/OFF SEASON. 2
(Read out answer choices)	LOOKING FOR JOB	LOOKING FOR JOB	LOOKING FOR JOB3	LOOKING FOR JOB 3 OTHERS 4
L	(SPECIFY)	(SPECIFY) Q39 ←	(SPECIFY)	(SPECIFY) Q39 ←

 \Box

NAME AND SERIAL NO **QUESTION** 38. What was the reason that WORKFORCE CUT DOWN/ WORKFORCE CUT DOWN/ WORKFORCE CUT DOWN/ WORKFORCE CUT DOWN/ you left from the previous ORGAN, RESTRUCTURED...... 01 ORGAN, RESTRUCTURED01 ORGAN, RESTRUCTURED01 ORGAN, RESTRUCTURED01 iob? DISSOLVED/IN RUIN 02 DISSOLVED/IN RUIN02 DISSOLVED/IN RUIN......02 DISSOLVED/IN RUIN02 PRODUCTION STOPPED 03 PRODUCTION STOPPED.......03 PRODUCTION STOPPED.......03 PRODUCTION STOPPED.......03 DISMISSED.......04 DISMISSED......04 DISMISSED......04 DISMISSED......04 ABBREVIATION: CONTRACT ENDED05 LOW INCOME 06 LOW INCOME06 LOW INCOME06 LOW INCOME06 ORG. - ORGANIZATION FARMING LAND LOST...... 07 FARMING LAND LOST......07 FARMING LAND LOST......07 FARMING LAND LOST.....07 UNSUITABLE JOB...... 08 UNSUITABLE JOB08 UNSUITABLE JOB08 UNSUITABLE JOB08 RESETTLEMENT 09 RESETTLEMENT......09 RESETTLEMENT......09 RESETTLEMENT......09 EFFECT OF ENVIRONMENT... 10 **EFFECT OF ENVIRONMENT....10** EFFECT OF ENVIRONMENT....10 **EFFECT OF ENVIRONMENT....10** OTHERS OTHERS OTHERS 11 OTHERS (SPECIFY) (SPECIFY) (SPECIFY) (SPECIFY) II. QUESTIONS RELATING TO INCOME AND NUMBER OF WORKING HOURS THAT THE RESPONDENT SPENT IN THE MAIN JOB IN THE LAST 7 DAYS/BEFORE THE **TEMPORARY ABSENCE TOTAL AMOUNT TOTAL AMOUNT TOTAL AMOUNT TOTAL AMOUNT** 39. Last month, how much money did you receive, including overtime remuneration. bonus, occupational (THOUSAND VND) (THOUSAND VND) (THOUSAND VND) (THOUSAND VND) allowances and other welfare payments for all jobs? IN WHICH: IN WHICH: IN WHICH: IN WHICH: In which: A. MAIN JOB (THOUSAND VND) (THOUSAND VND) (THOUSAND VND) (THOUSAND VND) B. Other jobs CHECK: IF THE TOTAL AMOUNT OF ALL JOBS = THE AMOUNT OF MAIN JOB + (THOUSAND VND) (THOUSAND VND) (THOUSAND VND) (THOUSAND VND) THE AMOUT OF OTHER JOBS, IF NOT.

ASK THE RESPONDENT AGAIN.

USEHOLD:				
NAME AND SERIAL NO QUESTION				
40. Actually, how many hours did you spend in all jobs in the last 7 days (including the main job and other jobs)?	TOTAL OF ACTUAL HOURS	TOTAL OF ACTUAL HOURS	TOTAL OF ACTUAL HOURS	TOTAL OF ACTUAL HOURS
In which:	A. MAIN JOB	A. MAIN JOB	A. MAIN JOB	A. MAIN JOB
A. Main job				
B. Other jobs CHECK: IF ACTUALLY HOURS OF ALL JOBS = HOURS OF MAIN JOB + HOURS OF OTHER JOBS. IF NOT, ASK QUESTION AGAIN.	B. OTHER JOBS	B. OTHER JOBS	B. OTHER JOBS	B. OTHER JOBS
41. <u>Usually</u> , how many hours did you spend in all jobs in the last 7 days (including the	TOTAL OF USUAL HOURS .	TOTAL OF USUAL HOURS .	TOTAL OF USUAL HOURS .	TOTAL OF USUAL HOURS .
main job and other jobs)? In which:	A. MAIN JOB	A. MAIN JOB	A. MAIN JOB	A. MAIN JOB
A. Main job				
B. Other jobs	B. OTHER JOBS	B. OTHER JOBS	B. OTHER JOBS	B. OTHER JOBS
CHECK: TOTAL OF ACTUAL HOUF	RS IN Q.40 DIFFER WITH TOTAL	OF USUAL HOURS IN Q.41 ASK	Q.42; OTHERWISE ASK Q.43	
42. Last 7 days, why did you actually work for less/more hours than usual?	INJURIED/ILL	INJURIED/ILL	INJURIED/ILL	INJURIED/ILL
	(SPECIFY)	(SPECIFY)	(SPECIFY)	(SPECIFY)

NAME AND SERIAL NO QUESTION				
III. QUESTIONS RELATING TO 1	THE UNDER-EMPLOYMENT			
43. For such total number of hours worked above, would you like to work more time?	YES1 ☐ NO2 ☐ → Q.46	YES1 ☐ NO2 → Q.46	YES1 ☐ NO2 ☐→ Q.46	YES1 ☐ NO2 ☐→ Q.46
44. You would like to work more time but are you available to work immediately?	YES1 NO	YES1 ☐ NO2 → Q.46	YES1 ☐ NO2 ☐→ Q.46	YES1 NO
45. How many additional hours would you like to work per week?	HOURS/WEEK	HOURS/WEEK	HOURS/WEEK	HOURS/WEEK
IV. QUESTION THAT RESPOND	ENTS SELF EVALUATE			
46. Do you think that your job is suitable with your training?	YES	YES	YES	YES
47. The job that you are taking is a temporary one to wait for or seek another job, isn't it?	YES	YES	YES	YES
48. In the last 30 days, did you intent to look for another job?	YES	YES1 NO	YES	YES
49. For the next 2 weeks, if other job is available, will you ready to work?	YES1 NO	YES1 NO2	YES	YES

HOUSEHOLD:		

NAME AND SERIAL NO QUESTION								
50. Even you are working, which	STUDYING OR TRAINING1	STUDYING OR TRAINING1	STUDYING OR TRAINING1	STUDYING OR TRAINING 1				
of the following best describes your main activity	WORKING FOR PAY OR PROFIT 2	WORKING FOR PAY OR PROFIT 2	WORKING FOR PAY OR PROFIT2	WORKING FOR PAY OR PROFIT 2				
or situation at present?	FARMING OR FISHING MAINLY FOR FAMILY CÓNUMPTION 3	FARMING OR FISHING MAINLY FOR FAMILY CÓNUMPTION3	FARMING OR FISHING MAINLY FOR FAMILY CÓNUMPTION3	FARMING OR FISHING MAINLY FOR FAMILY CÓNUMPTION 3				
(Read out answer choices) (MAIN ACTIVITY BY SELF	LOOKING FOR A JOB4	LOOKING FOR A JOB4	LOOKING FOR A JOB4	LOOKING FOR A JOB4				
RESPONDENT MAY BE	HOUSEHOLD RESPONSIBILITIES 5	HOUSEHOLD RESPONSIBILITIES 5	HOUSEHOLD RESPONSIBILITIES5	HOUSEHOLD RESPONSIBILITIES 5				
ACTIVITY WHICH IS THE MOST TIME OR IMPORTANT)	DISABILITY, INJURY, LONG-TERM ILLNESS6	DISABILITY, INJURY, LONG-TERM ILLNESS6	DISABILITY, INJURY, LONG-TERM ILLNESS6	DISABILITY, INJURY, LONG-TERM ILLNESS 6				
	RETIRED OR PENSIONER7	RETIRED OR PENSIONER7	RETIRED OR PENSIONER7	RETIRED OR PENSIONER 7				
	OTHER 8	OTHER 8	OTHER 8	OTHER 8				
	(SPECIFY)	(SPECIFY)	(SPECIFY)	(SPECIFY)				
51. Even you are	STUDYING OR TRAINING1	STUDYING OR TRAINING1	STUDYING OR TRAINING1	STUDYING OR TRAINING 1				
working/absence from work, which of the following best	WORKING FOR PAY OR PROFIT 2	WORKING FOR PAY OR PROFIT 2	WORKING FOR PAY OR PROFIT2	WORKING FOR PAY OR PROFIT 2				
describes your main activity or situation in the last 12	FARMING OR FISHING MAINLY FOR FAMILY CÓNUMPTION3	FARMING OR FISHING MAINLY FOR FAMILY CÓNUMPTION3	FARMING OR FISHING MAINLY FOR FAMILY CÓNUMPTION3	FARMING OR FISHING MAINLY FOR FAMILY CÓNUMPTION 3				
months?	LOOKING FOR A JOB4	LOOKING FOR A JOB4	LOOKING FOR A JOB4	LOOKING FOR A JOB4				
(Read out answer choices)	HOUSEHOLD RESPONSIBILITIES 5	HOUSEHOLD RESPONSIBILITIES 5	HOUSEHOLD RESPONSIBILITIES5	HOUSEHOLD RESPONSIBILITIES 5				
(MAIN ACTIVITY BY SELF RESPONDENT MAY BE	DISABILITY, INJURY, LONG-TERM ILLNESS6	DISABILITY, INJURY, LONG-TERM ILLNESS	DISABILITY, INJURY, LONG-TERM ILLNESS6	DISABILITY, INJURY, LONG-TERM ILLNESS				
ACTIVITY WHICH IS THE MOST TIME OR IMPORTANT)	RETIRED OR PENSIONER7	RETIRED OR PENSIONER7	RETIRED OR PENSIONER7	RETIRED OR PENSIONER 7				
MOST TIME OR IMPORTANT)	OTHER 8	OTHER 8	OTHER 8	OTHER 8				
	(SPECIFY)	(SPECIFY)	(SPECIFY)	(SPECIFY)				
CHECK Q.4 OR Q.5: IF RESPONDE	CHECK Q.4 OR Q.5: IF RESPONDENT AGE 15 TO 59 AND Q14=6/7/8/9 OR Q15=5/6/7 AND Q47 = 1; GO TO ITEM C. OTHERS, ASK THE NEXT RESPONDENT.							

ITEM B: INFORMATION ON THE UNEMPLOYMENT OR OUTSIDE THE LABOUR FORCE

NAME AND SERIAL NO QUESTION				
52. In the last 30 days, did you intent to look for job or prepare to start any productive or business activites?	YES1 NO	YES1 NO	YES1 ☐ NO2 ☐→Q.55	YES
53. How did you look for job or start any productive or business activities? ABBREVIATION: PRO./BUS PRODUCTIVE/BUSINESS ADVER ADVERTISEMENT INTERV. – INTERVIEW	JOB APPLYING	JOB APPLYING	JOB APPLYING	JOB APPLYING
54. How long have you been looking for job?	UNDER 1 MONTH	UNDER 1 MONTH	UNDER 1 MONTH	UNDER 1 MONTH

HOUSEHOLD:		

NAME AND SERIAL NO QUESTION				
QUEUTION				
55. What was the reason that you	TOO OLD/YOUNG, DISABLED01	TOO OLD/YOUNG, DISABLED01	TOO OLD/YOUNG, DISABLED 01	TOO OLD/YOUNG, DISABLED01
did not look for job in the last	TAKING DOMESTIC WORKS02	TAKING DOMESTIC WORKS 02	TAKING DOMESTIC WORKS 02	TAKING DOMESTIC WORKS02
30 days?	SCHOOLING (ST/PU/AP)03	SCHOOLING (ST/PU/AP)03	SCHOOLING (ST/PU/AP)03	SCHOOLING (ST/PU/AP)03
,	NO DEMAND OF WORK/ NO NEED TO WORK04	NO DEMAND OF WORK/ NO NEED TO WORK04	NO DEMAND OF WORK/ NO NEED TO WORK04	NO DEMAND OF WORK/ NO NEED TO WORK04
ABBREVIATION:	BELIEVING NO JOB AVAILABLE/	BELIEVING NO JOB AVAILABLE/	BELIEVING NO JOB AVAILABLE/	BELIEVING NO JOB AVAILABLE/
PRO./BUS -PRODUCTIVE/BUSINESS	NO JOB SUITABLE05	NO JOB SUITABLE05	NO JOB SUITABLE05	NO JOB SUITABLE05
ST - STUDENT	DID NOT KNOW WHERE/HOW	DID NOT KNOW WHERE/HOW	DID NOT KNOW WHERE/HOW	DID NOT KNOW WHERE/HOW
PU - PUPIL	TO FIND JOB06	TO FIND JOB06	TO FIND JOB06	TO FIND JOB
AP – APPENTICE	TEMPORARILY LAID OFF 07	TEMPORARILY LAID OFF07	TEMPORARILY LAID OFF07	TEMPORARILY LAID OFF 07
AI - AIT ENTICE	WAITING JOB APPLICATION/TO	WAITING JOB APPLICATION/TO	WAITING JOB APPLICATION/TO	WAITING JOB APPLICATION/TO
_	START PRO/BUS ACTIVITIES 08	START PRO/BUS ACTIVITIES 08	START PRO/BUS ACTIVITIES 08	START PRO/BUS ACTIVITIES08
'	OFF SEASON09	OFF SEASON09	OFF SEASON09	OFF SEASON09
	BAD WEATHER10	BAD WEATHER10	BAD WEATHER10	BAD WEATHER10
	EFFECT OF ENVIRONEMNT 11	EFFECT OF ENVIRONEMNT11	EFFECT OF ENVIRONEMNT11	EFFECT OF ENVIRONEMNT 11
	FAMILY OBLIGATIONS/	FAMILY OBLIGATIONS/	FAMILY OBLIGATIONS/	FAMILY OBLIGATIONS/
	RELAXING12	RELAXING12	RELAXING12	RELAXING12
	TEMPORARY ILLNESS/INJURY 13	TEMPORARY ILLNESS/INJURY 13	TEMPORARY ILLNESS/INJURY 13	TEMPORARY ILLNESS/INJURY.13
	OTHERS14	OTHERS14	OTHERS14	OTHERS14
	(SPECIFY)	(SPECIFY)	(SPECIFY)	(SPECIFY)
56. For the next 2 weeks, if it is available for a job or a productive/business activity,	YES1 Q.58	YES1 → Q.58	YES1 Q.58	YES1 Q.58
will you be ready to work immediately?	NO2	NO2	NO2	NO2
57. What is the main reason that you	SCHOOLING/TRAINING1	SCHOOLING/TRAINING1	SCHOOLING/TRAINING1	SCHOOLING/TRAINING 1
are unavailable for work	HOUSEWIFE2	HOUSEWIFE 2	HOUSEWIFE2	HOUSEWIFE2
immediately?	FAMILY OBLIGATIONS 3	FAMILY OBLIGATIONS 3	FAMILY OBLIGATIONS3	FAMILY OBLIGATIONS 3
	BAD WEATHER4	BAD WEATHER4	BAD WEATHER4	BAD WEATHER4
	TEMPORARY ILLNESS/INJURY 5	TEMPORARY ILLNESS/INJURY 5	TEMPORARY ILLNESS/INJURY .5	TEMPORARY ILLNESS/INJURY.5
	TOO OLD/YOUNG, DISABLED 6	TOO OLD/YOUNG, DISABLED 6	TOO OLD/YOUNG, DISABLED6	TOO OLD/YOUNG, DISABLED 6
	OTHERS7	OTHERS7	OTHERS 7	OTHERS7
	(SPECIFY)	(SPECIFY)	(SPECIFY)	(SPECIFY)

NAME AND SERIAL NO QUESTION				
58. Have you ever worked before?	YES1 ☐ NOT YET2 ☐ → Q.65	YES1 NOT YET2 Q.65	YES1 ☐ NOT YET2 ☐ → Q.65	YES1
59. For how long have you left the last job?	UNDER 1 MONTH1	UNDER 1 MONTH1	UNDER 1 MONTH1	UNDER 1 MONTH 1
ue.1901	1 TO < 3 MONTHS2			
	3 MONTHS TO < 1 YEAR	3 MONTHS TO < 1 YEAR3	3 MONTHS TO < 1 YEAR3	3 MONTHS TO < 1 YEAR 3
	1 TO < 5 YEARS4	1 TO < 5 YEARS4	1 TO < 5 YEARS4	1 TO < 5 YEARS 4
	5 YEARS AND MORE5	5 YEARS AND MORE5	5 YEARS AND MORE5	5 YEARS AND MORE 5
	Q.65	Q.65	Q.65	Q.65
60. What reason did you quit/leave from the last job?	WORKFORCE CUT DOWN/ ORG. RESTRUCTURED01			
	DISSOLVED/IN RUIN02	DISSOLVED/IN RUIN 02	DISSOLVED/IN RUIN02	DISSOLVED/IN RUIN02
ABBREVIATION:	PRODUCTION STOPPED 03	PRODUCTION STOPPED 03	PRODUCTION STOPPED03	PRODUCTION STOPPED 03
ORG. –ORGANIZATION	DISMISSED04	DISMISSED04	DISMISSED04	DISMISSED04
	CONTRACT ENDED05	CONTRACT ENDED 05	CONTRACT ENDED05	CONTRACT ENDED 05
	LOW INCOME06	LOW INCOME06	LOW INCOME06	LOW INCOME06
	FARMING-LAND LOST07	FARMING-LAND LOST07	FARMING-LAND LOST07	FARMING-LAND LOST07
	EFFECT OF ENVIRONMENT 08	EFFECT OF ENVIRONMENT 08	EFFECT OF ENVIRONMENT08	EFFECT OF ENVIRONMENT 08
	UNSUITABLE JOB 09	UNSUITABLE JOB09	UNSUITABLE JOB09	UNSUITABLE JOB09
	RESETTLEMENT10	RESETTLEMENT 10	RESETTLEMENT10	RESETTLEMENT10
	RETIRED/DUE TO DISABILITY 11	RETIRED/DUE TO DISABILITY 11	RETIRED/DUE TO DISABILITY.11	RETIRED/DUE TO DISABILITY 11
	OTHERS12	OTHERS12	OTHERS12	OTHERS1212
	(OI LOII I)	(OI LOII I)	(OI LOII I)	(OI LOII I)

HOUSEHOLD:				L
NAME AND SERIAL NO QUESTION				
61. In your main job/business, what kind of work did you do before leave the last job?				
WRITE CLEARLY: - MAIN TASKS AND DUTIES				
- OCCUPATION TITLE (IF ANY) Example:	OCCUPATION TITLE, IF ANY			
Cattle farmer: breed, raise and sell cattle.				
Policeman: patrol the streets				
Cook: plan and prepare meal	MAIN TASKS AND DUTIES			
Primary school teacher: teach children how to read, writ and calculate				
	(ISCO CODE)	(ISCO CODE)	(ISCO CODE)	(ISCO CODE)
62. What is the main activity of the business or establishment where did you work in the last job?				
WRITE CLEARLY:	NAME OF ESTABLISHMENT (IF ANY)			

- NAME (IF ANY)
- MAIN ACTIVITY, GOODS OR SERVICES

Example:

Cong Vi ward police: ensure security and order of the ward.

Hoa Lan Restaurant: food and drink business.

Bien Dong Shipping Company: longdistance transport of goods

MAIN ACTIVITY, GOODS OR SERIVES

(ISIC CODE)

MAIN ACTIVITY, GOODS OR SERIVES

(ISIC CODE)

(ISIC CODE)

MAIN ACTIVITY, GOODS OR SERIVES

MAIN ACTIVITY, GOODS OR SERIVES

(ISIC CODE)

NAME AND SERIAL NO QUESTION				
63. With the above-mentioned job, were you? (Read out answer choices) ABBREVIATION: CONTR. – CONTRIBUTING	EMPLOYER	EMPLOYER	EMPLOYER	EMPLOYER
64. Does the last establishment that you worked belong to one kind of economic unit as follows?	FARM HOUSEHOLD	FARM HOUSEHOLD	FARM HOUSEHOLD	FARM HOUSEHOLD
(Read out answer choices) ABBREVIATION:	NON-STATE SERVICE UNIT06 STATE:	NON-STATE SERVICE UNIT 06 STATE:	NON-STATE SERVICE UNIT 06 STATE:	NON-STATE SERVICE UNIT 06 STATE:
BUS. – BUSINESS ORG. – ORGANIZATION ASS. – ASSOCIATION IN INDIVIDUAL	+ LEGISLATIVE, EXECUTIVE, JUDICAL AGENCY			

HOUSEHOLD:		

NAME AND SERIAL NO QUESTION				
65. Even you are not working, which of the following best describes your main activity or situation at present? (Read out answer choices) (MAIN ACTIVITY BY SELF RESPONDENT MAY BE ACTIVITY WHICH IS THE MOST TIME OR IMPORTANT)	STUDYING OR TRAINING	ĐI HỌC/ĐÀO TẠO	ĐI HỌC/ĐÀO TẠO	ĐI HỌC/ĐÀO TẠO
66. Even you are not working, which of the following best describes your main activity or situation in the last 12 months? (Read out answer choices) (MAIN ACTIVITY BY SELF RESPONDENT MAY BE ACTIVITY WHICH IS THE MOST TIME OR IMPORTANT)	STUDYING OR TRAINING	ĐI HỌC/ĐÀO TẠO	ĐI HỌC/ĐÀO TẠO	ĐI HỌC/ĐÀO TẠO
67. CHECK Q 4 OR Q 5, AGED FR	OM 15 TO 59 AND Q14=6/7/8/9	OR Q15=5/6/7; ASK ITEM C. OTI	HERWISE ASK NEXT RESPOND	DENT.

ITEM C. SPECIFY INFORMATION ABOUT FIELD OF STUDY/TRAINNING

(Ask qualified persons aged 15 to 59 who are currently do not having job or having temporary job)

NAME AND SERIAL NO QUESTION				
The next questions, I would like	to ask you some more inforn	nation about the type of highe	st degree you have achieved.	
68. What is field of the highest level of training you have been trained, and the year you graduated from that level?				
SPECIFIED:				
- DEGREE TRAINING				
- TRAINING FIELDS				
- GRADUATION YEAR.	(EDUCATION AND TRAINNING CODE)			
 Example: Bachelor of Business Administration. Vocational training: operating 	(GRADUATION YEAR)	(GRADUATION YEAR)	(GRADUATION YEAR)	(GRADUATION YEAR)
excavator. 69. In addition to the highest degree you mentioned, do you have any other qualification from vocational training?	YES	YES1 ☐ NO2 ☐ → Q72	YES	YES1 ☐ NO2 ☐ → Q72
70. What are the degree?	PRIMARY VOCATIONAL CERTA	PRIMARY VOCATIONAL CERT. A	PRIMARY VOCATIONAL CERT . A	PRIMARY VOCATIONAL CERTA
(MARK ALL ANSWER CODES)	MID-TERM VOCATIONAL CERB	MID-TERM VOCATIONAL CER . B	MID-TERM VOCATIONAL CER B	MID-TERM VOCATIONAL CERB
ABBREVIATION:	MID-TERM PROF.SCHOOLC	MID-TERM PROF.SCHOOL C	MID-TERM PROF.SCHOOL C	MID-TERM PROF.SCHOOLC
CERT CERTIFICATE	VOCATIONAL COLLEGE DIPLD	VOCATIONAL COLLEGE DIPL D	VOCATIONAL COLLEGE DIPL D	VOCATIONAL COLLEGE DIPLD
PROF PROFESSIONAL	PROFESSIONAL COLLEGEE	PROFESSIONAL COLLEGE E	PROFESSIONAL COLLEGE E	PROFESSIONAL COLLEGEE
DIPL DIPLOMA	UNIVERSITYF	UNIVERSITY F	UNIVERSITYF	UNIVERSITYF
	ABOVE UNIVERSITYG	ABOVE UNIVERSITY G	ABOVE UNIVERSITYG	ABOVE UNIVERSITYG

A. SECOND DEGREE	A. SECOND DEGREE	A. SECOND DEGREE	A. SECOND DEGREE
(EDLICATION AND TRAINING CODE)	(EDLICATION AND TRAINING CODE)	(EDLICATION AND TRAINING CODE)	(EDLICATION AND TRAINING CODE)
(EDUCATION AND TRAINNING CODE)	(EDUCATION AND TRAINNING CODE)	(EDUCATION AND TRAINNING CODE)	(EDUCATION AND TRAINNING CODE)
(GRADUATION YEAR)	(GRADUATION YEAR)	(GRADUATION YEAR)	(GRADUATION YEAR)
B. THIRD DEGREE (IF ANY)	B. THIRD DEGREE (IF ANY)	B. THIRD DEGREE (IF ANY)	B. THIRD DEGREE (IF ANY)
(EDUCATION AND TRAINNING CODE) (GRADUATION YEAR)	(EDUCATION AND TRAINNING CODE) (GRADUATION YEAR)	(EDUCATION AND TRAINNING CODE) (GRADUATION YEAR)	(EDUCATION AND TRAINNING CODE) (GRADUATION YEAR)
((EDUCATION AND TRAINNING CODE) (GRADUATION YEAR) THIRD DEGREE (IF ANY) (EDUCATION AND TRAINNING CODE) (GRADUATION YEAR)	(EDUCATION AND TRAINNING CODE) (GRADUATION YEAR) (GRADUATION YEAR) (B. THIRD DEGREE (IF ANY) (EDUCATION AND TRAINNING CODE) (EDUCATION AND TRAINNING CODE) (EDUCATION AND TRAINNING CODE) (GRADUATION YEAR)	(EDUCATION AND TRAINNING CODE) (GRADUATION YEAR) (GRADUATION YEAR) (GRADUATION YEAR) B. THIRD DEGREE (IF ANY) (EDUCATION AND TRAINNING CODE)

SOME ISSUES NEED TO BE CONSIDERED

- Part 1: Interviewer collects information from all members who are currently usual resident in the household, including those who are studying, working or traveling abroad within the time allowed by the agency authorization.
- Part 3: Interviewer collects information to classify the status of economic activity of household members aged 15 and over and current living in Vietnam. This section is designed to be divided into three major categories:
- Item A: For employed people.
- Item B: For those who do not have a job.
- Item C: For those aged 15-59 who are technically professionally qualified, currently not having job or temporarily employed.

Items A and B of Part 3 have additional questions designed to determine what the main activity that the respondent is currently doing / in the last 12 months (questions 50, 51, 65, 66) to support the construction of classify the main activities on the basis of self-awareness / declaration of the respondent. It could be the activity that the respondent spent most of the tim

e or activity that the respondent felt most important but was generally required to be declared as self-aware by the respondent. The answers in this section may differ from the answers in determining the state of economic activity.

Variable Name	Condition
skip_q6	if Q04_1 == 98 & Q05 is not missing
skip_q10	if Q10 == 5 & at least one from Q11 to Q13 are not missing
skip_q12_1_a	if $Q12_1 == 1 & Q12_2b$ is missing
skip_q12_1_b	if $Q12_1 == 2 \& Q12_3b$ is missing
skip_q16	if Q16 == 1 & at least one from Q17 to 23 are not missing
skip_q17	if Q17 == 1 & at least one from Q18 to Q23 are not missing
skip_q18	if Q18 == 1 & at least one from Q19 to Q23 are not missing
skip_q19	if Q19 != 1 & at least one from Q20 to Q51 are not missing
skip_q21	if Q21 == 1 & at least one from Q22 to Q23 are not missing
skip_q22	if Q22 == 2 & at least one from Q23 to Q51 are not missing
skip_q23	if Q23 != 1 & at least one from Q24_1 to Q51 are not missing
skip_q26	if not inlist(Q26,1,2) & Q27 is not missing
skip_q27	if inlist(Q27,3,4) & at least one from Q28 to Q29 are not missing
skip_q30	if inlist(Q30,1,2,3) & at least one from Q31 to 33 are not missing
skip_q31	if $Q31 == 1 & Q32$ is not missing
skip_q34	if $Q34 == 2 & Q35$ is not missing
skip_q36	if inlist(Q36,3,4,5) & at least one from Q37 to Q38 are not missing
skip_q37	if inlist(Q37,2,3,4) & Q38 is not missing
skip_q39_3	if Q39_1 is not missing but Q39_2 or Q39_3 is missing
skip_q40_3	if Q40_1 is not missing but Q40_2 or Q40_3 is missing
skip_q41_1	if Q41_1 is not missing but Q41_2 or Q41_3 is missing
= =	if $inrange(Q05,15,59)$ & $inlist(Q14,6,7,8,9)$ or $inlist(Q15,5,6,7)$ & $Q47 == 1$ & at least one from Q52 to Q66 are not missing
skip_q52	if $Q52 == 2$ & at least one from Q53 to Q54 are not missing
skip_q54	if inlist(Q54,1,2,3,4) & Q55 is not missing
skip_q56	if $Q56 == 1 & Q57$ is not missing
skip_q58	if $Q58 == 2$ & at least one from Q59 to Q64 are not missing
	if inrange(Q05,15,59) & inlist(Q14,6,7,8,9) or inlist(Q15,5,6,7) & at least one from Q68_1 to Q71b_3 are not missing
skip_q69	if $Q69 == 2$ & at least one from Q70 to Q71b_3

Source: Authors

Variable Name	Condition
val_female	if total number of member of hh <= 0
val_hh_15above	if total number of female in hh <= 0
val_15abovefemale	
val_q02	if Q02 is not 1,2,3,4,5 or 6
val_q03	if Q03 is not 1 or 2
val_q04_1	if Q04_1 < 1 or Q04_1 > 12 & Q04_1 == 98
val_q04_2	if Q04_2 < 1900 or Q04_2 > 2018 & Q04_2 != 9998
val_q05	if Q05 < 0 & Q05 > 95
val_q06_1	if Q06_1 is not 1 or 2
val_q07	if Q07 is not 1,2,3,4,5,6 or 9
val_q08_1	if Q08_1 < 0 & Q08_1 > 12 & Q08_1 != 98
val_q08_2	if Q08_2 < 0 & Q08_2 > 12 & Q08_2 != 98
val_q09	if Q09 is not 1,2,3,4,5 or 9
val_q10	if Q10 is not 1,2,3,4,5 or 9
val_q11	if Q11 is not 1,2 or 9
val_q12_1	if Q12 is not 1 or 2
val_q13	if Q13 is not 1,2,3,4,5,6,7,8,9,10 or 98
val_q14	if Q14 is not 1,2,3,4,5,6,7,8,9 or 98
val_q15	if Q15 is not 1,2,3,4,5,6,7 or 98
val_q16	if Q16 is not 1,2 or 9
val_q17	if Q17 is not 1,2 or 9
val_q18	if Q18 is not 1,2 or 9
val_q19	if Q19 is not 1,2,3,4,5,6,7,8,9 or 98
val_q20	if Q20 is not 1,2,3,4,5,6,7,8,9,10,11,12 or 98
val_q21	if Q21 is not 1,2 or 9
val_q22	if Q22 is not 1,2 or 9
val_q23	if Q23 is not 1,2,3,4 or 9
val_q26	if Q26 is not 1,2,3,4,5,6,7,8,9,10,11,12 or 98
val_q27	if Q27 is not 1,2,3,4 or 9
val_q28	if Q28 is not 1,2 or 9
val_q29	if Q29 is not 1,2 or 9
val_q30	if Q30 is not 1,2,3,4,5 or 9 if Q31 is not 1,2,3,4,5,6,7 or 9
val_q31	if Q31 is not 1,2,3,4,5,6,7 or 9
val_q32	if Q32 is not 1,2,3,4,5,6,7 or 9
val_q33	if Q33 is not 1,2,3,4,5,6,7 or 9
val_q34	if Q34 is not 1,2 or 9 if Q35 is not 1,2 or 9
val_q35	
val_q36	if Q36 is not 1,2,3,4,5 or 9
val_q37	if Q37 is not 1,2,3,4 or 9
val_q38 val_39_1	if Q38 is not 1,2,3,4,5,6,7,8,9,10,11 or 98 if Q39_1 < 0 or Q39_1 > 900,000
val_39_1 val_39_3	if Q39_3 < 0 or Q39_3 > 900,000
	if Q40_1 < 0 or Q40_1 >99 & Q40_1 != -1
val_q40_1 val_q40_2	if Q40_1 < 0 or Q40_1 >99 & Q40_1 := -1 if Q40_2 < 0 or Q40_2 > 99 & Q40_2 != -1
vai_q40_2 val_q40_3	if Q40_3 < 0 or Q40_3 > 99 & Q40_3 != -1
*	if Q41_1 < 0 or Q40_1 > 99 & Q40_1 != -1
val_q41_1	$11 \text{ Q+1}_1 \text{ \ } 1 \text{ \ } 0 \text{ \ } 1 \text{ \ } 0 \text{ \ } 1 \text{ \ } 2 \text{ \ } 2 \text{ \ } 3 \text{ \ } \infty \text{ \ } 2 \text{ \ } 40_1 \text{ \ } 1 $

Online Appendix Table 2: Validation error conditions

Variable Name	Condition
val_q41_2	if Q41_2 < 0 or Q40_2 > 99 & Q40_2 != -1
val_q41_3	if $Q41_3 < 0$ or $Q40_3 > 99 & Q40_3 != -1$
val_q42	if Q42 is not 1,2,3,4,5,6,7,8,9,10,11,12 or 98
val_q43	if Q43 is not 1 or 2
val_q44	if Q44 is not 1 or 2
val_q45	if Q45 < 0 & Q45 > 99
val_q46	if Q46 is not 1,2,3 or 4
val_q47	if Q47 is not 1,2 or 9
val_q48	if Q48 is not 1,2 or 9
val_q49	if Q49 is not 1 or 2
val_q50	if Q50 is not 1,2,3,4,5,6,7,8 or 9
val_q51	if Q51 is not 1,2,3,4,5,6,7,8 or 9
val_q52	if Q52 is not 1,2 or 9
val_q53	if Q53 is not 1,2,3,4,5,6,7,8,9 or 98
val_q54	if Q54 is not 1,2,3,4,5,6,7,8,9 or 98
val_q55	if Q55 is not 1,2,3,4,5,6,7,8,9,11,12,13,14 or 98
val_q56	if Q56 is not 1 or 2
val_q57	if Q57 is not 1,2,3,4,5,6,7,8 or 9
val_q58	if Q58 is not 1 or 2
val_q59	if Q59 is not 1,2,3,4,5 or 9
val_q60	if Q60 is not 1,2,3,4,5,6,7,8,9,10,11,12 or 98
val_q63	if Q63 is not 1,2,3,4,5 or 9
val_q64	if Q64 us not 1,2,3,4,5,6,7,8,9,10,11,12 or 98
val_q65	if Q65 is not 1,2,3,4,5,6,7,8 or 9
val_q66	if Q66 is not 1,2,3,4,5,6,7,8 or 9
val_q68_3	if $Q68_3 \le 1973$ or $Q68_3 \ge 2018$
val_q69	if Q69 is not 1,2 or 9
val_q70	if Q70 is not A,B,C,D,E,F,G or -1
val_q71a_3	if Q71a_3 < 1973 or Q71a_3 > 2017
val_q71b_3	if Q71b_3 < 1973 or Q71b_3 > 2017

Source: Authors

Variable Nam	e Condition
logic_q02	if head of hh is more than 2
logic_q05_a	if hh head and spouse age difference is greater than 30
logic_q05_b	if hh head and childrens age difference is greater than 60
logic_q05_c	if spouse and childrens age difference is greater than 50
logic_q09_a	if $Q02 == 2 & Q04$ is 1,2,3,4 or 5
logic_q09_b	if Q05 <= 18 & Q09 is 2,3,4 or 5
logic_q13	if $Q05 \le 18 \& Q13 == 5$
logic_q14	if Q05 <= 18 & Q14 >=6
logic_q15	if Q05 <= 18 & Q15 is 6 or 7
logic_q26	if $Q25_3 > 325 \& Q26 == 1$
logic_q28	if Q26 is 5,6,7,8,9,10 & 11 & Q28 == 2 or Q26 is 1 or 2 & Q28 == 1
logic_q29	if Q27 is 3 or 4 & Q29 is 1 or 2
logic_q30	if (Q26 is 1 or 2 & Q30 is 1,4, or 5) or (Q26 is 7,8,9 or 10 & Q30 is 2 or 3) or (Q26 == 4 & Q30 != 4) or (Q30 == 4 & Q26 != 4) or
	$(Q26 \text{ is } 11 \text{ or } 12 \& Q30 == 4) \text{ or } (Q30 == 3 \& Q25_3 \le 325 \& Q26 != 1)$
logic_q39	if (Q30 == 3 & Q39_1 != 0) or (!inrange(Q39_1,0,100000) or !inrange(Q39_2,0,100000) or !inrange(Q39_3,0,100000)) or (Q39_2
	< Q39_1 & Q39_3 < Q39_1) or (the sum of Q39_2 and Q39_3 != Q39_1)
logic_q40	if $(Q40_1) = 84$ or $Q40_2 = 84$ or $Q40_3 = 84$) or $(Q40_2) = 241_1$ or $Q40_3 = 240_1$) or (the sum of $Q40_2$ and $Q40_3 = 240_1$) or (the sum of $Q40_3$
	$Q40_1$) or $(Q19 == 1 & Q40_2 > 0)$
logic_q41	if $(Q41_1) = 84$ or $Q41_2 = 84$ or $Q41_3 = 84$) or $(Q41_2) = 241_1$ or $Q41_3 = 241_1$ or $Q41_1 = 241_1$ or $Q41_1 = 241_1$ or $Q41_1 = 241_1$ or $Q41_1 = 241_$
	Q41_1)
logic_q45	if $Q45 >= 84$
logic_q46	if $(Q14 \ge 6 \& Q45 == 3)$ or $(Q15 \ge 4 \& Q46 == 3)$
logic_q48	if $Q47 == 9 & (Q48 == 1 \text{ or } Q48 == 2)$
logic_q49	if $Q47 == 9 & (Q49 == 1 \text{ or } Q49 == 2)$
logic_q53	if Q52 == 9 & Q54 is 1,2,3,4,5,6,7,8 or 9
logic_q54	if $Q52 == 9 \& Q54$ is 1,2,3 or 4
logic_q55	if Q52 == 9 & Q55 is 1,2,3,4,5,6,7,8,9,10,11,12,13 or 14
logic_q64	if (Q64 is 1 or 2 & Q63 is 1,4 or 5) or (Q64 is 7,8,9 or 10 & Q63 is 2 or 3) or (Q64 == 4 & Q63 != 4) or (Q63 == 4 & Q64 != 4) or
	(Q64 is 11 or 12 & Q63 == 4) or (Q63 == 3 & Q62 <= 325 & Q64 != 1)
logic_q68	if Q68_3 is not missing & Q14 is 7,8 or 9 & Q05 <= 18
logic_q70	if (Q70 == G & Q14 <= 8) or (Q70 == F & Q14 <= 7) or (Q70 == E & Q14 <= 6) or (Q70 == D & Q15 <= 5) or (Q70 == C & Q14 <= 6) or (Q70 == D & Q15 <= 6) or (Q70 == C & Q14 <= 6) or (Q70 == D & Q15 <= 6) or (Q70 == C & Q14 <= 6) or (Q70 == D & Q15 <= 6) or (Q70 == C & Q14 <= 6) or (Q70 == D & Q15 <= 6) or (Q70 == C & Q14 <= 6) or (Q70 == D & Q15 <= 6) or (Q70 == C & Q14 <= 6) or (Q70 == D & Q15 <= 6) or (Q70 == C & Q14 <= 6) or (Q70 == D & Q15 <= 6) or (Q70 == C & Q14 <= 6) or (Q70 == D & Q15 <= 6) or (Q70 == C & Q14
	$Q15 \le 5$) or $(Q70 == B \& Q15 \le 4)$ or $(Q70 == A \& Q15 \le 3)$

Online Appendix Table 4: Missing Error Conditions

Variable Nam	e Condition
miss_q02	if Q02 is missing & Q03 is not missing
miss_q03	if Q03 is missing & Q02 or Q04 is not missing
miss_q04_1	if Q04 is missing & Q03 or Q06 is not missing
miss_q05	if Q05 is missing & Q04_1 != 98 & Q04_2 != 9998 & Q04 is not missing
miss_q06_1	if Q06 is missing & Q05 \geq 15 & Q07 or Q04 is not missing
miss_q07	if Q07 is missing & Q06_1 == 2 & Q08 is not missing
miss_q08	if Q08 is missing & Q07 is not missing
miss_q09	if Q09 is missing & Q05 < 15 & Q06_1 != 1 & Q10 is not missing
miss_q10	if Q10 is missing & Q05 < 15 & Q06_1 != 1 & Q11 or Q09 is not missing
miss_q11	if Q11 is missing & Q10 != 5 & Q05 < 15 & Q06_1 != 1 & Q10 or Q12 is not missing
miss_q12_1	if Q12 is missing & Q10 != 5 & Q05 < 15 & Q06_1 != 1 & Q13 or Q11 is not missing
miss_q13	if Q13 is missing & Q10 != 5 & Q05 < 15 & Q06_1 != 1 & Q12 or Q14 is not missing
miss_q14	if Q14 is missing & Q13 or Q15 is not missing
miss_q15	if Q15 is missing & Q14 or Q16 is not missing
miss_q16	if Q16 is missing & Q15 is not missing
miss_q17	if Q17 is missing & Q16 != 1 & Q05 < 15 & Q06_1 != 1 & Q16 or Q18 is not missing
miss_q18	if Q18 is missing & Q16 != 1 & Q05 < 15 & Q06_1 != 1 & Q17 or Q19 is not missing
miss_q19	if Q19 is not missing & Q05 < 15 & Q06_1 != 1 & Q18 or Q20 is not missing
miss_q20	if Q20 is missing & Q19 == 1 & Q05 < 15 & Q06_1 != 1 & Q19 or Q21 is not missing
miss_q21	if Q21 is missing & Q19 == 1 & Q05 < 15 & Q06_1 != 1 & Q20 or Q22 is not missing
miss_q22	if Q22 is missing & Q21 != 1 & Q19 == 1 & Q05 < 15 & Q06_1 != 1 & Q21 or Q23 is not missing
miss_q23	if Q23 is missing & Q22 $== 1$
miss_q24	if Q24 is missing & Q25 is not missing
miss_q25	if Q25 is missing & Q24 or Q26 is not missing
miss_q26	if Q26 is missing & Q25 or Q27 is not missing
miss_q27	if Q27 is missing & inlist(Q26,1,2)
miss_q28	if Q28 is missing & inlist(Q27,1,2) & Q29 is not missing
miss_q29	if Q29 is missing & inlist(Q27,1,2) & Q30 is not missing
miss_q30	if Q30 is missing & inlist(Q27,3,4) & Q31 is not missing
miss_q31	if Q31 is missing & inlist(Q30,4,5)

Variable Name Condition	
miss_q32	if Q32 is missing & inlist(Q31,2,3,4,5,6,7)
miss_q33	if Q33 is missing & (Q32 is not missing or Q31 $== 1$)
miss_q34	if Q34 is missing & (inlist(Q30,1,2,3) or Q33 or Q35 is not missing)
miss_q35	if Q35 is missing & Q34 == 1 & Q36 is not missing
miss_q36	if Q36 is missing & Q35 or Q37 is not missing
miss_q37	if Q37 is missing & inlist(Q36,1,2) & Q38 is not missing
miss_q38	if Q38 is missing & inlist(Q36,1,2) & Q37 == 1 & Q39 is not missing
miss_q39	if Q39 is missing & Q40 is missing
miss_q40	if Q40 is missing & Q41 is not missing
miss_q41	if Q41 is missing & Q40 or Q42 is not missing
miss_q42	if Q42 is missing & Q40_1 != Q41_1 & Q42 is not missing
miss_q43	if Q43 is missing & (Q41 $_1$) > 0 & Q41 is not missing) or Q44 is not missing))
miss_q44	if Q44 is missing & Q43 $== 1$ & Q45 is not missing
miss_q45	if Q45 is missing & Q43 == 1 & Q44 == 1 & Q46 is not missing
miss_q46	if Q46 is missing & Q43, Q44 or Q47 is not missing
miss_q47	if Q47 is missing & Q46 or Q48 is not missing
miss_q48	if Q48 is missing & Q47 $== 1$ & Q49 is not missing
miss_q49	if Q49 is missing & Q47 $== 1$ & Q48 or Q50 is missing
miss_q50	if Q50 is missing & Q49or Q51 is not missing
miss_q51	if Q51 is missing & Q50 is not missing
miss_q52	if Q52 is missing & (inlist(Q19,2,3,4,5,6,7) or Q22 == 2 or inlist(Q23,2,3,4))
miss_q53	if Q53 is missing & Q52 $== 1$ & Q54 is not missing
miss_q54	if Q54 is missing & Q52 $== 1$ & Q53 or Q56 is not missing
miss_q55	if Q55 is missing & Q52 == 2 & Q56 is not missing or Q52 == 2
miss_q56	if Q56 is missing & Q55, Q54 or 57 is not missing
miss_q57	if Q57 is missing & Q56 $== 2$ & Q58 is not missing
miss_q58	if Q58 is missing & Q57 or Q59 is not missing
miss_q59	if Q59 is missing & Q58 == 1 or Q60 not missing
miss_q60	if Q60 is missing & (inlist(q59,1,2,3,4) or Q61 is not missing)
miss_q61	if Q61 is missing & Q60, Q62 is not missing

Online Appendix Table 4: Missing Error Conditions

Variable Nan	ne Condition
miss_q62	if Q62 is missing & Q61 or Q63 is not missing
miss_q63	if Q63 is missing & Q62 or Q64 is not missing
miss_q64	if Q64 is missing & Q63 is not missing
miss_q65	if Q65 is missing & Q59 == 5 or Q66 is not missing
miss_q66	if Q66 is missing & Q65 is not misssing
miss_q68	if Q68 is missing & inrange(Q05,15,59) & (inlist(Q14,6,7,8,9) inlist(Q15,5,6,7)) & Q66 is not missing
miss_q69	if Q69 is missing & Q68 or Q70 is not missing
miss_q70	if Q70 is missing & Q69 == 1 & inrange(Q05,15,59) & (inlist(Q14,6,7,8,9) inlist(Q15,5,6,7)) & Q71 is not missing
miss_q71	if Q71 is missing & Q70 is not missing

Source: Authors