A Close Look at the Interview Length for Cell and Landline Telephone Surveys: The Case of the California Health Interview Survey

Jennifer Kali[1] and Ismael Flores Cervantes[1]
[1]Westat, 1600 Research Blvd, Rockville, MD 20850

Abstract

There seems to be a consensus in survey methodology that cell phone interviews are longer than landline interviews within the same study (Brick, et al., 2007; Lynn & Kaminska, 2011; and Vicente, Resis, & Santos, 2009). Several reasons have been hypothesized to explain the differences in length. Among these are satisficing, multitasking, respondent distraction, connectivity issues, or the type of questions related to the telephone service. The available literature is limited and previous studies are based on surveys conducted several years ago. In the meantime, there have been many changes in phone technology, cell phone use, and culture. Additionally, all of these studies were based on short surveys of 15 minutes or less. However, an initial analysis of the 2013-2014 California Health Interview Survey (CHIS) finds little difference in interview length between the two modes.

The CHIS is a dual-frame random digit dial (RDD) telephone survey of the California non-institutionalized population, conducted with both landline and cell phone samples. The CHIS is a large sample of 40,000 respondents, with 30,000 landline- and 10,000 cell phone-completed interviews. The CHIS interview averages 35.8 minutes, much longer in duration than surveys reported in previous literature. The CHIS provides a unique opportunity to explore differences between interviews conducted on the two types of phones as time is recorded by interview section throughout the interview. This study allows a more detailed analysis of the screener and extended interviews and the different sections of the extended interview, such as demographics and health-related questions. Time differences are examined by household size, age, and other relevant variables to examine differences in length of interview by phone types. Results of the analysis show that a part of the differences in duration can be explained by the type of respondents that received a different number of questions. However, some differences still remain after accounting for respondent's characteristics in the model.

Keywords: Cell phone; landline; interview length; CHIS

1. Introduction

Surveys conducted by telephones commonly use a design that includes both landline and cell telephone samples. As the number of cell phone-only households has increased in recent years, the proportion of interviews conducted on cell phones has also increased in dual-frame surveys that include both landline and cell telephone samples. Many studies have found that interviews conducted on cell phones are longer in duration than those conducted on traditional landline telephones.

This difference in duration of the interview has implications on survey management, cost, and labor. The study presented in this article is an effort to understand the differences

between interviews conducted on landline and cell phones. It is not clear from the current literature the elements that drive the differences in duration between the two telephone devices. One hypothesis is that different types of respondents are interviewed on the different devices, and this may lead to differences in duration. For example, Blumberg, Luke, & Cynamon (2006) reported difference in health-related variables between cell phone and landline respondents, and Lynn & Kaminska (2011) reported that cell interviews are typically of people who are younger, higher educated, employed, and a part of larger households. A second hypothesis focuses on the characteristics of the devices themselves and the call quality. Issues related to connectivity, such as static or drop calls, would increase the duration of the call. A third hypothesis regarding the differences in interview duration regards the respondents' behaviors with the devices, such as respondents multitasking during the interview, or the location of respondents when participating in the interview (i.e., at home or somewhere else).

One of the earliest studies to notice the difference in the duration of interviews conducted by landline and cell phone samples was Brick et al. (2007). Their study assessed the feasibility of conducting surveys on cell phones in the United States in 2004 and was part of the practicum of the Joint Program in Survey Methodology at the University of Maryland. The survey collected information on telephone ownership and usage, attitudes toward cell phones, social behaviors, and demographics. They found that for respondents with both a landline and a cell phone, interviews conducted on cell phones were 8.9 minutes on average while those conducted on a landline phone were only 8.2 minutes on average. They did not report interview duration for respondents who did not have both devices. The study found that 44 percent of cell phone interviews were conducted while the respondent was not at home. In contrast, with surveys conducted on landline devices, the respondent is almost certainly at home.

More recently, Vicente, Reis, & Santos (2009) studied the mode effects between surveys conducted on landline phones and those conducted on cell phones in Portugal. The content of the survey focused on Internet usage, attitudes toward the Internet, cultural practices, and demographics. They found that cell phone interviews averaged 11.99 minutes while landline interviews averaged only 10.91 minutes. Additionally, they found that respondents on cell phones were less likely to give "don't know" or "no opinion" responses and were more likely to answer open-ended questions, any of which would lead to longer cell phone surveys. However, landline respondents were less likely to acquiesce (agree to all statements), which could lead to longer landline surveys.

Lynn & Kaminska (2011) analyzed the effects on survey measurement using data from an experiment used to study mixed-mode data collection methods in Hungary as part of the European Social Survey in 2005. The topics covered in the survey were behaviors, such as voting and television, and attitudes toward politics, institutions, immigration, and gender roles, among others. A subset of telephone respondents who had both telephone services were randomized to one type of telephone in the experiment. They found while overall their interview averaged 15.20 minutes; those interviews conducted on cell phones were 2:20 minutes longer than those conducted on landlines. Additionally, they found that cell phone respondents were more likely to report multitasking while participating in the interview, which would likely lead to longer cell interviews.

These are some common characteristics in the surveys analyzed in the previous studies. The length is relatively short and they were conducted several years ago. There have been considerable changes in both phone technology and attitudes toward cell phone use. One

of the goals of our study is to explore the effects in a longer and more current survey. However, due to limitations of the available data, we can only examine the first hypothesis regarding the differences in respondent characteristics between landline and cell phone respondents. Measures for the quality of the call and the interaction and behavior of the respondent and the device were not collected in the survey used in the analysis.

2. The California Health Interview Survey (CHIS)

We examine interview duration between cell and landline phones using data from the 2013-2014 California Health Interview Survey (CHIS). CHIS is a random digit dialing (RDD) dual-frame telephone survey of California's population. CHIS is the largest health survey ever conducted in any state and one of the largest health surveys in the nation. It is a collaborative project of the UCLA Center for Health Policy Research, the California Department of Health Services, and the Public Health Institute. The funding for CHIS includes sources such as the state and federal government agencies and private foundations, among others. Westat is the data collection contractor and develops the weights for analysis of the data collected in CHIS.

CHIS collects extensive information on public health, health status, prevalence of chronic conditions, health-related behaviors, health insurance coverage, and access to health care services. Data from CHIS supports the production of estimates for the state, many counties, and for group's counties in California. The survey also supports the study of the characteristics for the major racial and ethnic groups and a number of smaller ethnic groups within the state. Adults, parents or guardians of children, and adolescents within California households are eligible for sampling. This analysis will focus only on the adult interviews.

CHIS 2013-2014 was a dual-frame random digit dialing (RDD) telephone survey. By design 20 percent of the sample was selected from the cell phone frame and the remaining 80 percent from the landline frame. CHIS was an overlapping design, that is, all sampled phone numbers were eligible regardless of cell phone usage and the frame from which they were sampled. The data of the survey was collected over a two-year period, and the telephones were sampled and released every six months. A total of 38,000 adult interviews were completed in CHIS 2013-2014.

The CHIS 2013-2014 interview has two components: the screener and extended interviews. During the screener interview, the eligibility of the sampled phone number is determined and an eligible respondent among those adults linked to the phone number is selected for the extended interview. However, as explained in detail in Brick, et al. (2012), the content of screener interview differs for landline and cell phones. In general, a landline phone is considered a household appliance that is linked to every person living in that household. The questions in the screener interview for a landline phone focus on determining the eligibility (i.e., residency vs. business) and selecting one respondent from all eligible adults living in the household. In contrast, since a cell phone is typically a personal device and thus the person who answers the phone is the only person linked to the sampled phone number, enumeration and subsampling are unnecessary. Additional eligibility questions are administered to screen out teen cell phones, which are ineligible. The screener questionnaire includes questions to verify that respondents are not driving. Because of these differences between the screener interview for landline and cell phone respondents, data and timing of the screener interviews were not included in this analysis

The extended interview in the CHIS collects demographic and health-related data about the sampled respondent. This information is collected in 12 sections with a total of 572 questions, as shown in Table 1. There are very few differences in the survey questions between surveys conducted on cell and landline phones except for a few question in the last section regarding cell phone usage. The sequence and number of questions asked are a function of the respondent's answers and skip patterns. There are many intricate skip patterns, and thus no respondent is exposed to all 572 questions.

Since CHIS collects health information data, there is a relationship between the demographic and health characteristics of the respondent and the numbers of questions to which respondents are exposed. For example, if a respondent answers "yes" to the question "Have you smoked at least 100 or more cigarettes in your entire lifetime?," the respondent will receive 33 more questions than a respondent who answers "no" to this question. The smoking question has the most direct effect on the number of questions received, but there are more subtle interactions between the characteristics of the respondent and the number of questions to which the respondent is exposed. For example, if the respondent is living in poverty, the respondent will most likely have poorer health and this will lead to more questions. Note that the CHIS is longer and has more intricate skip patterns than the research mentioned previously.

Table 1: Sections of the CHIS Questionnaire

Section	***Question type***	***Number of questions***	***Differ between devices?***	***Skips***
A	Demographic Information	16	No	Yes
B	Health Conditions	47	No	Yes
C	Health Behaviors	66	No	No
D	General Health, Disability, and Sexual Health	28	No	Yes
F	Mental Health	33	No	Yes
G	Demographic Information	36	No	No
H	Health Insurance	107	No	Yes
J	Health Care Utilization and Access	53	No	Yes
K	Employment, Income, Poverty Status, Food Security	30	No	Yes
L	Public Program Participation	15	No	Yes
M	Housing	23	No	Yes
N	Demographic Information	118	Yes	Yes
	Total	572		

3. Descriptive Analysis

The timing of the CHIS extended interview is recorded using time stamps that mark the beginning and ending time of each section in the extended interview. The duration of extended interview is computed by combining all the time stamps for all the sections. Table 2 shows the descriptive statistics for the duration of the extended interview, the number of items to which the respondent was exposed, and the item speed by telephone type. As was found in previous studies, the cell phone interview, at 36.7 minutes, is longer than the landline interview at 33.4 minutes. This can partly be explained by item

exposure. Interviews conducted on cell phones are exposed on average to 201.9 questionnaire items, while those conducted on landline phones are exposed to only 190.4 questionnaire items. This highlights the differences in the types of people who respond to surveys via landline and cell as personal characteristics dictate the number of questions to which a respondent is exposed.

Hox (1994) proposed interview speed as a transformation to interview duration. The speed at which a survey is conducted is the number of questions answered per minute. Looking at speed instead of duration removes the effect of item exposure from the comparison. Surveys conducted on cell phones, with a rate of 5.7 items per minute, are slower than those conducted on landlines, with a rate of 6.0 items per minute. Thus, for the CHIS, surveys conducted on cell phones are longer, in both duration and item exposure, and also slower than those conducted on landlines.

Table 2: Extended interview duration, item exposure, and speed for cell and landline phones; CHIS

Variable	*Device*	*Mean*	*Std. Dev.*	*Minimum*	*Maximum*
Interview duration of extended interview (minutes)	Cell phone	36.7	8.8	18	99
	Landline	33.4	8.7	16	97
Item exposure in extended interview	Cell phone	201.9	24	142	335
	Landline	190.4	23.7	127	329
Number of items/minute (speed)	Cell phone	5.7	1.1	2	10
	Landline	6.0	1.3	2	13

Some records were excluded from the analysis to limit their effect on the results. Only interviews conducted in English were included. Only extended interviews that were completed within the same phone call were included. Ported phone numbers (those that were sampled on the landline frame but were actually cell phone numbers) were excluded. Additionally, extreme outliers in duration were excluded from the analysis.

4. Regression Analysis

In an attempt to explain the factors that influence the differences in durations between phone devices, we implemented two analyses for modeling the duration of the extended interview as a function of the telephone type, respondent's characteristics, and number of items asked. These analyses used two types of regression – proportional hazard regression (also called survival analysis) and linear regression. For each regression analysis, we modeled the duration of the phone call as a function of the device (cell or landline), including various covariates. The coefficient on the device is the outcome of interest as an indicator of a significant difference in interview duration.

There were 19 covariates available for both regression models. These covariates are classified as demographic variables, health-related variables, and paradata. The demographic variables were marital status, race, ethnicity, education attainment, household tenure, household size, presence of children, presence of adolescents, poverty level, employment status, gender, and age. The health variables included were health status (Would you say that in general your health is excellent, very good, good, fair or poor?) and smoking status (Have you smoked at least 100 or more cigarettes in your

entire lifetime?). The paradata include variables for the number of item refusals, number of don't know responses, release group, interviewer, and item exposure. Release group refers to the new samples selected every six months. With each release there were slight modifications to the questionnaire, and including release group serves as a proxy for the questionnaire changes. Interviewer is included as each interviewer has a different cadence, which could influence the speed and duration. All interviewers conducted interviews for both the cell and landline samples.

4.1. Proportional Hazard Regression

The first analysis considered was a type of survival analysis called proportional hazard regression or Cox regression (Hosmer, Lemeshow, & May, 2008) where the time at which an event of interest occurred is modeled. This type of analysis is commonly used in medical studies such as time to disease recovery or death. Because we are analyzing duration of an interview (i.e., time to the end of the interview), this type of regression seemed a fitting model, though no literature was found that used this method for this purpose with survey timing data.

The Cox hazard regression model is

$$H\left(t;\mathbf{X}_i\right)=H_0\left(t\right)\exp\left(\beta_1 X_{i1}+\ldots+\beta_k X_{ik}\right);$$

which is the expected hazard function for an individual i with covariates X_{i1}, …, X_{ik} (i.e., independent or explanatory variables) at time t. The factor $H_0\left(t\right)$ is the baseline hazard function at time t and corresponds to the hazard for a respondent when all explanatory variables are zero. The second factor of the model, which does not depend on time, describes how the hazard changes depending on the values of the explanatory variables.

We fitted two proportional hazard models to the CHIS data. The first model, called the simple model, only has one independent variable – device (cell or landline phone), which is used as our reference. The second model, the full model, included all statistically significant covariates. The coefficients of the regression for the parameters for the simple and full models are shown in Table 3.

The categorical value used as reference in coefficient for the device is landline. This coefficient in both the full and simple models is negative, which indicates that it takes longer for the event to occur (the interview to end) for cell than for landline phones, or interviews on cell phones are longer in duration than interviews on landlines. In both models, the coefficient for device is significant, which confirms that the difference in duration between landline and cell phone interviews is not the result of random variability, even after including covariates in the full model.

Comparing the coefficients of device in the regression model, the full model reduces the risk by two-thirds by accounting for the covariates (-0.3 in the simple model and -0.10 in the full model), which means that although the difference in time between landline and cell phone interviews are small, some difference remains after accounting for covariates. The significant variables in the model are household tenure, poverty level, marital status, release group, smoker status, educational attainment, presence of adolescents, presence of children, and item exposure.

Table 3. Coefficient on device for simple and full proportional hazard regression models

Model	*Parameter*	*Level*	*DF*	*Estimate*	*Error*	*Chi-Square*	*Pr > ChiSq*
Simple	*Device*	Cell phone	1	-0.30	0.02	247.65	<.0001
		Landline	0	0.00			
Full	*Device*	Cell phone	1	-0.10	0.02	23.58	<.0001
		Landline	0	0.00			
	Household tenure	1	1	0.15	0.02	6.16	<.0001
		2	0	0.00	.	.	.
	Poverty level	4	1	0.05	0.03	3.10	0.0781
		6	1	0.35	0.03	138.10	<.0001
		7	1	0.47	0.03	345.28	<.0001
		8	1	0.05	0.03	3.01	0.0827
		1	0	0.00	.	.	.
	Marital status	1	1	0.18	0.02	110.27	<.0001
		0	0	0.00			
	Release group	5	1	-0.12	0.02	32.25	<.0001
		6	1	0.08	0.02	12.03	0.0005
		7	1	0.06	0.02	8.95	0.0028
		8	0	0.00	.	.	.
	Smoking status	0	1	-0.16	0.04	19.53	<.0001
		1	0	0.00	.	.	.
	Education level	2	1	0.28	0.04	47.62	<.0001
		3	1	0.28	0.04	49.72	<.0001
		4	1	0.36	0.04	78.88	<.0001
		1	0	0.00	.	.	.
	Presence of adolescents	1	1	-0.14	0.03	17.88	<.0001
		0	0	0.00	.	.	.
	Presence of children	1	1	-0.20	0.03	55.70	<.0001
		0	0	0.00	.	.	.
	Item exposure		1	-0.01	0.00	1263.44	<.0001

In the second part of the analysis we test the model fit of the proportional hazard regression. Figure 1b shows the survival curves by device. In the diagnostic, we assess the assumption that the ratio of hazards is a constant that does not depend on time. Graphically, when the LogLog plot shows parallel curves, the assumption holds. Figure 1c shows the plots for the CHIS data, and it is apparent the assumption does not hold. The test of hypothesis shows evidence against proportional hazards since the coefficient for interaction term for device and time is statistically different from 0 (χ^2=319.95, p <0.0001). We also examined the residuals of the fitted model. Figure 1d shows deviance residuals, which are often preferred for model checking since they have a more symmetric distribution than other residuals, such as Martingale residuals or Cox residuals. The deviance residuals can be interpreted in a similar manner as the residuals

in linear regression. The funnel pattern of the residuals confirms the lack of fit of the model.

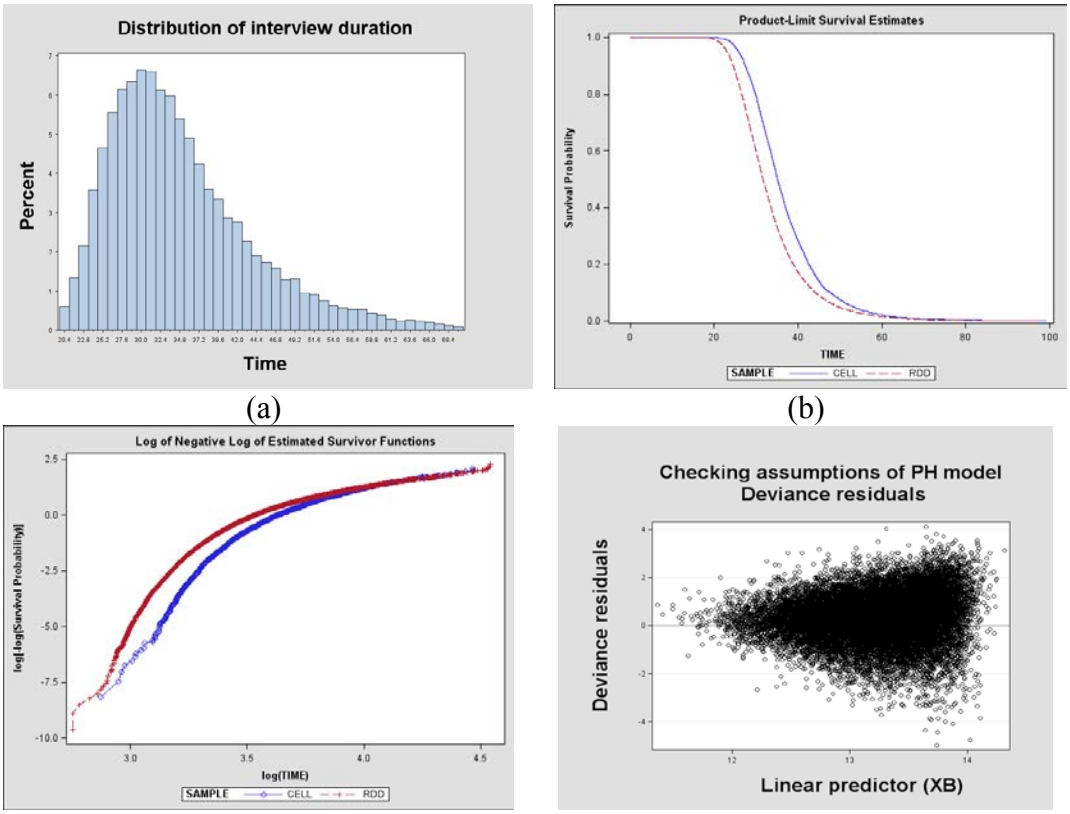

Figure 1: Time distribution, survival curves, and model check for CHIS data

4.2 Linear Regression

The second type of analysis is based on linear regression and modeled the speed of the interview (defined as the ratio of number of items completed in the interview and the total time it took to complete the interview). This is a transformation of the dependent variable for time. The idea of regressing speed instead of time was suggested by Hox (1994) who analyzed interviewer effects by modeling speed as a function of interviewer and explanatory variables. Hox transformed duration into speed because speed follows a normal distribution unlike the distribution of time (see Figure 1a). With the inclusion of the covariates and the dependent variable of speed, the model is explaining both duration and item exposure at the same time. Similar to what Hox found, in CHIS speed follows a normal distribution. Figure 2b shows the relationship between the duration and the number of items respondents were exposed to during the extended interview.

The linear regression model is

$$Y_i = \beta_1 X_{i1} + \ldots + \beta_k X_{ik} + \varepsilon_i ;$$

where the dependent variable Y_i is speed for individual i with the $X_{i1},\ldots,X_{ik}$ are independent or explanatory variables. The interpretation of this model is different from

the proportional hazard model because the linear regression examines factors that explain why the speed of questions answered in cell phones is slower than the speed of landlines independently of the numbers of questions asked (see Table 1).

Similarly to the proportional hazard modeling, we fitted two linear regression models to the CHIS data. The first model, or simple model, had only one independent variable – device – which is used as our reference. The second model, the full model, included all significant covariates. The coefficients of the regression for parameters for the simple and full models are shown in Table 4.

Table 4: Coefficient estimates for device for simple and full linear regression models

Model	*Parameter*	*Level*	*DF*	*Estimate*	*Standard error*	*t value*	*p-value*
Simple	*Intercept*		1	6.62	0.01	591.91	<.0001
	Device	Cell phone	1	-0.25	0.03	-9.66	<.0001
		Landline	0	0.00	.	.	.
Full	*Intercept*		1	6.04	0.07	86.81	<.0001
	Device	Cell phone	1	-0.13	0.03	-4.86	<.0001
		Landline	0	0.00	.	.	.
	Household tenure	1	1	0.15	0.02	6.16	<.0001
		2	0	0.00	.	.	.
	Poverty level	4	1	0.07	0.04	1.79	0.0728
		6	1	0.52	0.04	13.49	<.0001
		7	1	0.73	0.03	22.75	<.0001
		8	1	0.14	0.04	3.74	0.0002
		1	0	0.00	.	.	.
	Release group	5	1	-0.27	0.03	-9.66	<.0001
		6	1	0.09	0.03	3.14	0.0017
		7	1	0.12	0.03	4.46	<.0001
		8	0	0.00	.	.	.
	Race: White alone	1	1	-0.09	0.03	-3.67	0.0002
		2	0	0.00	.	.	.
	Smoking status	0	1	-0.42	0.05	-9.29	<.0001
		1	0	0.00	.	.	.
	Education level	2	1	0.44	0.05	8.09	<.0001
		3	1	0.45	0.05	8.54	<.0001
		4	1	0.59	0.05	11.18	<.0001
		1	0	0.00	.	.	.

The categorical value used as the reference for the device is landline. The coefficient is negative, which indicates that interviews conducted on cell phones are slower than those conducted on landline phones. In both models, the coefficient for device is significant, which confirms that the difference in duration between landline and cell phone interviews is not the result of random variability, even after including covariates in the full model.

The value of the coefficient for the simple model that does not include covariates is -0.25, while for the full model it is -0.12. Including the respondents' characteristics into the

model accounts for close to half of the effect of the device on speed. Although the differences in speed between landline and cell are smaller after accounting for covariates, the difference is not fully explained by the covariates. The significant variables in the model are household tenure, poverty level, release group, smoker status; White alone race indicator, and educational attainment. Note that item exposure was not included in this model as it is not relevant when considering speed.

In the second part of the analysis we test the model fit for the linear regression. Figure 2c, shows the plot the residuals against both the predicted values. As shown, the residuals are randomly scattered about 0 and with equal width (i.e., constant variance assumption) throughout the predicted values. Figure 3d confirms these results graphically with a Q-Q plot of the residuals.

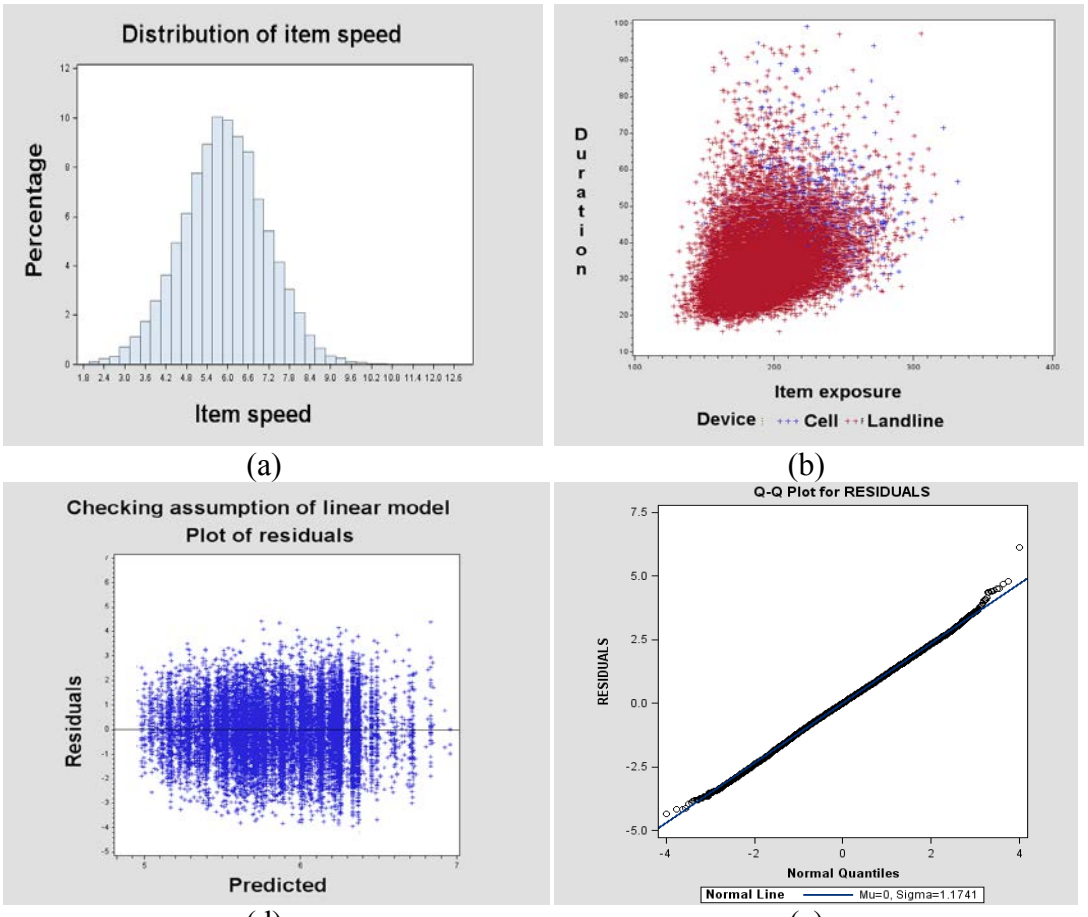

Figure 2: Distribution of interview duration, relationship between duration items exposed, and linear regression diagnostics for CHIS data

5. Discussion

CHIS data supports what other studies have found: Surveys conducted on cell phones are longer in duration than those conducted on landlines phones. People responding on landline phones differ from people responding on cell phones in general, as is seen in the CHIS data and numerous other studies. The results of modeling the CHIS data indicate

that personal characteristics of respondents do contribute to the differences in interview duration as cell phone respondents are exposed to more questions than landline respondents. However, personal characteristics do not account for the all of the differences. Even after accounting for demographic variables, health variables, and available paradata, interviews conducted on cell phones are longer in duration, or slower in speed, than those conducted on landline phones.

Calls conducted on cell phones are becoming more common each year. The impact of these longer interviews on survey management, cost, and labor is unknown. It will be beneficial to continue to study how and why there are differences in duration and speed between cell and landline phones. Some of the previous research has indicated respondents using cell phones are multitasking while responding to the interview and are not in their homes (and thus maybe somewhere distracting). Unfortunately these kinds of paradata were not collected for CHIS. More directed research is needed to understand the interactions affecting interview length.

With the current CHIS data, more research is possible within the 12 sections of the questionnaire. It would be interesting to investigate if the relationships modeled in this analysis hold true for all sections or if there are differences based on section topic or section order (early in interview or toward the end of the interview).

Acknowledgements

We would like to thank Dr. J. Michael Brick for his support to this work.

References

Blumberg, S., Luke, J., & Cynamon, M. (2006). Telephone coverage and health survey estimates: evaluating the need for concern about wireless substitution. *American Journal Public Health, 96*, 926-931.

Brick, J. M., Brick, P. D., Dipko, S., Presser, S., Tucker, C., & Yuan, Y. (2007). Cell phone survey feasibility in the United States: Sampling and calling cell numbers versus landline numbers. *Public Opinion Quarterly, 71*, 23-29.

Brick, J. M., Flores Cervantes, I., Kali, J., Norman, G., & Grant, D. (2012). A review of the cell phone sample component of the California Health Interview Survey. *Joint Statistical Meetings, Survey Research Methods Section*, (pp. 3885-3895).

Hosmer, D. W., Lemeshow, S., & May, S. (2008). *Applied survival analysis: Regression modeling of time to event data* (2nd ed.). New York: John Wiley and Sons.

Hox, J. J. (1994). Hierachical regression models for interviewer and respondent effects. *Sociological Methods and Reserach, 22*, 300-318.

Lynn, P., & Kaminska, O. (2011). The impact of mobile phones on survey measurement error. *Public Opinion Quarterly*, First published online December 8, 2012 .

Vicente, P., Resis, E., & Santos, M. (2009). Using mobile phones for survey research: A comparative analysis between data collection via mobile phones and fixed phones. *International Journal of Market Research, 51*(5), 1-21.