

국제조사방법론 심포지엄 참가
결 과 보 고 서

2008. 11.



조 사 관 리 국

Survey Management Bureau

국제조사방법론 심포지엄 참가 결과보고

□ 회의 명칭 및 목적

- 24회 국제 조사방법론 심포지엄
 - 캐나다 통계청 주관으로 매년 개최되며 조사방법에 관한 세계적 동향, 효율적 조사표 구성, 자료수집 방법, 무응답 대책 및 처리 방법 등 다양한 최신 조사방법론 및 분석기법 소개
- 최근 조사방법 동향, 자료수집방법, 무응답 대처방안 및 처리방법, 가중값 조정 등 다양한 조사방법론에 대한 공유
- 동일조사에 다양한 자료수집 방법 적용, 무응답 가중치 처리방안 등은 최근 열악한 환경에서 자료수집 방법 및 무응답 대처방안 등을 실제로 적용할 수 있는 방안 모색
- 자료수집 계획 및 설계, 조사표 설계 및 평가, 자료수집에 최신 기법 적용, 추론시 보조자료 적용 등을 전문가 그룹 토의 및 실무경험을 교환함으로써 향후 우리청의 조사방법 및 추정기법 검토시 반영

□ 회의 기간 및 장소

- 기간 : 2008. 10. 28.(화)~10. 31.(금)
- 장소 : 캐나다 오타와

□ 참가자 및 출장기간

- 참가자 : 조사관리국 조사기획과 신명철, 표본관리과 정구현
- 출장기간 : 2008. 10. 27.(월)~11. 4.(화)

□ 상세 출장일정

일정		업무수행내용	참가과정
10.27 (월)	-	인천 → 밴쿠버 이동	
		밴쿠버 도착	
		오타와 출발	
10.28 (화)	00:10	오타와 도착	
	02:00	호텔도착(Four Points by Sheraton 호텔)	
	09:00	워크숍 참가	
	09:00~17:00	Mixed-Mode Surveys : Innovation & Challenge	workshop
10.29 (수)	09:00~10:00	Waksberg Award Address International surveys : Motives and Methodologies	session 1
	10:30~12:00	Bias and Estimation issues	session 4
		Selection Models for evaluating assumption of methods that compensate for missing values in sample surveys	
		The effect of non-response follow-up in a survey on living among immigrants in Norway	
		Non-response bias analysis using abbreviated interviews Data collection challenges in dual frame surveys	
	13:30~15:00	Monitoring collection	session 5
		Statistics canada's computer assisted personal interview monitoring program	
		CARI as a tool for monitoring field interviews and improving field data collection CARI : a tool for improving data quality now and next time	
	15:30~17:00	Blaise : New development	session 8
		Offline questionnaires at statistics netherlands	
Experience with Multi-mode surveys New usage of blaise : biometrics inputs			
17:00~19:00	Cocktail reception		
10.30 (목)	09:00~10:00	Keynote Address Dynamic survey design managed by modeled praradata	session 10
	10:00~10:30	Poster session	
		Coverage Evaluation of the linkage between the canadian community health survey and hospital data New strategies for dealing with non-response followup in tobacco cessation studies	
	10:30~12:00	Census Data collection	session 12
Building a business case for census internet data collection Mailing census questionnaires : result from 2006 and plans for 2011			

		Changes in the management of data collection operations for the canadian census in 2011	
		A census of population based on an administrative register	
	13:30~15:00	Collection via the internet	session 14
		The internet : a new data collection method for the census	
		Going web-only in a complex enterprise survey	
		Accuracy of travel data samples utilizing online vs mail methodologies	
		Can we make official statistics with self-selection web surveys?	
	15:30~17:00	Analysing the collection process using paradata1	session 16
		using paradata to estimate the average time it would take respondents to complete a CATI questionnaires	
		Coverage and nonresponse in a random digit dialing survey :	
		Using paradata to manage nonresponse in the survey of labor and income dynamics	
	8:45~10:15	Analysing the collection process using paradata2	session 17
		Data collection research using paradata at statistics canada	
		Identification of optimal call patterns for intensive followup in business surveys using operations research methods	
		Effects of number of call attempts on nonresponse rate and nonresponse bias	
	10:45~12:15	Response Burden	session 19
		Measure and followup of the response burden from enterprises and organizations at statistics sweden	
		Developing statistics new zealand's respondent load strategy	
		Measuring respondent burden in the uk office for national statistics	
	13:45~15:15	Questionnaire Design and Testing2	session 20
		Reducing the Number of cognitive by adding other cognitive method of testing	
		Presenting 'don't know' in web survey	
		How to Develop business surveys continuously in cyclic model	
11.2-4	-	캐나다 → 인천	

※ 11월 1일에 귀국하는 비행기 편이 없음

□ 소요예산 : 9,564,900원

□ 주요 내용

1. 자료수집과정에서 복합방법 사용

○ 복합방법의 정의

- 복합방법은 조사원과 응답자간의 의사소통을 위해 두 가지 이상의 방법을 사용
 - 의사소통(communication)과정은 초기 접촉단계와 자료수집 과정에서 발생

○ 복합방법을 사용하는 이유

- 주어진 조건(시간, 예산, 질문내용 등)하에 최선의 자료를 수집
- 전체 오차를 감소하기 위해

○ 조사시 발생하는 오차의 종류

- 포함오차(coverage error)는 관심변수에 대한 모집단의 모든 단위를 포함하지 않는 경우
 - ※ 인터넷 보급률
한국 : 70.5%, 북미 73.6%, 전 세계 21.9%
- 표본오차(sampling error)는 전수조사를 하지 않고 표본조사를 함으로써 발생하는 오차
 - 확률표본에서만 계산이 가능
- 무응답오차(nonresponse error)는 만나지 못하거나(noncontact)거나 거부(refusal)시 발생되며, 이 때 응답자와 무응답자는 관심변수에서 차이를 보일 때 발생하는 오차
 - 응답률이 최근 20년 동안 지속적으로 감소하고 있으며, 이로 인해 자료수집의 복합방법의 사용이 제기됨
- 측정오차(measurement error)는 복합방법에서 가장 문제가 됨
 - 자료수집과정에서 발생하는 오차로 조사원의 태도, 응답자 기억력, 방

법별 차이 등에 의해 발생

○ 복합방법의 가장 적합한 사용 예

- American Community Survey의 경우 처음에 우편조사(mail survey), 무응답 추적 전화조사 실시하며 마지막에 무응답 면접조사를 실시

○ 복합방법 사용시 발생 문제점

- 실제 값(true value)이 방법별로 차이가 있을 수 있음
- 서로 다르게 조사된 자료를 하나로 통합이 가능한지?
 - 횡단면 조사 : 분산이 커짐
 - 종단면 조사 : 시간의 교락효과가 발생 가능
- 다른 방법으로 조사된 자료가 국가간에 비교가 가능한지?

○ 의사소통의 접촉단계

- 사전에 고지(알림)의 효과
 - 표본틀 보완, 응답율 증가, 적법성과 신뢰도 증가
- 사전에 알림 단계에서 여러 가지 방법을 혼용하여도 측정오차에 영향 없음

○ 의사소통의 독촉단계

- 단순한 알림은 비용부담이 적고 응답율 제고가 가능

○ 의사소통의 자료수집단계

- 복합방법은 측정오차를 가져올 가능성이 높음
- 몇 개의 질문내용은 하나의 주된방법, 몇 개는 다른 방법을 사용하되 모든 응답자에게 동일한 절차를 사용(GOOD)
예) 일상적인 내용은 CATI 조사, 민감한 항목은 CASI로 전환

- 종단면/패널조사 첫조사 수집방법 달리 적용(So-So)
- 여러 가지 방법을 동시에 사용(BAD)

○ 수집방법이 왜 다른가?

- 조사자의 영향(Interviewer Impact)
- 대중매체 영향(Media Related Factors)
- 정보의 전달(Information Transmission)

○ 자료수집을 위해 복합방법 접근

- 한 가지 수집방법을 주로 사용, 다른 방법은 보조
- 다중방법을 동시에 사용

○ 복합방법이 성공하기 위한 조건

- 더욱 더 많은 현장보고서 필요
- 과거, 국제적, 소프트웨어 개발자로부터 경험공유
- 전문가 필요 등

○ 복합방법에 대한 검토의견

- 과학, 기술 및 사회적 여건에 의해 수집방법의 다양화 불가피
- 시험조사 등을 통해 조사별로 가장 적합한 복합방법 선택
- 수집방법별 차이에 대한 분석작업을 통해 측정오차를 감소하는 노력 필요

※ 자세한 내용은 붙임의 파워포인트 자료 참조

2. 자료수집의 모니터링

○ 캐나다의 CAPI 모니터링 프로그램

- 캐나다에서 CATI는 이미 중앙 전화조사에 공식적인 절차를 지님
 - 전화조사는 잘 정비된 기준으로 조사이행 여부를 평가
 - 작업의 질적인 요소 평가(좋은 점과 개선이 필요한 점)

- 피드백 제공

- 목적 :

- 객관적이고 연속적인 측정을 제공
- 조사에서 자료수집의 일관성 유지
- 조사원의 잘못조사를 발견하기 위해

- 모니터링 전략

- 조사내용의 일부분을 발췌하여 녹음
- 모니터링
- 보고서 및 피드백 제공

- 샘플링

- 1주일에 적정량의 표본을 선택
- 발췌하여 녹음한 것 중 20개 질문을 선택
- 일자와 시간에 맞추어 오디오 파일 소팅
- 계통추출

- 모니터링

- 표준을 지키지 않는 레코딩이 있는지 확인
- 보고서에 오류통계표 작성 : 조사자, 감독자, 매니저에게 송부

- 피드백

- 문제를 방지하고 긍정적인 역량을 강화
- 조사에 크게 영향을 미치는 심각한 오류 : 즉시 통보
- 간단한 오류 : 일주일 단위로 통보

- 향후 보완책

- 레코드의 질적인 문제

- 거부율의 차이에 대한 분석
- 레코드 협조율에 대한 분석

3. 센서스 자료수집

○ 사업체 센서스에서 인터넷 자료수집방법

- 소개

- 센서스 개발자와 사회관련 연구자들은 조사방법으로 인터넷 고려
- 초기 하드웨어 및 소프트웨어의 비용이 큼
- 컴퓨터 지식 및 접속율 확대되어 조사가능
- 당장은 모든 종이 조사표를 대체하지는 않을 것임

- 국제적 동향(가구부문)

Country	Internet Experience	Future Plans
Canada	Started in 2006 with a take up rate of 18 percent	• 2011: increase take-up rate to 35-40 percent, target internet users, and reuse system where possible
Australia	Started in 2006 with a take up rate of 9 percent	• 2011: promote the internet on the doorstep (if internet accepted, the paper form will not be provided)
New Zealand	Started in 2006 with a take up rate of 7 percent	• 2011: increase take-up rate to 20-30 percent, increase awareness via publicity, provide PIN to all households
England and Wales	Completed feasibility study	• 2011: plan to implement for the first time, estimate a 20 percent take-up rate
USA	Tested on 4 occasions 2005 test take up rate of 7.2 percent	• 2010: will not be offered to every dwelling

- 인터넷 자료수집방법의 장점 및 한계

- 응답률 제고
- 사회적/ 공적인 기대 부응
- 데이터 질적 내용 개선
- 장기간으로 보면 경제적 이익

- 다양한 자료수집방법에 맞게 조사표 설계
- 인터넷의 안정성 및 신뢰성 확보 필요
- 다양한 대중매체에 접근

- 사업체에서 인터넷 도입
- 가구에 대한 인터넷조사는 일반화 경향
- 사업체에 대한 인터넷조사 도입은 시대적 요구임

○ 우편조사 조사표 : 2006년 결과 및 2011년 계획

- 2006년 이전 센서스 자료수집

- 국가의 면적을 센서스 조사구로 분할
- 조사원에 의해 조사표 배부
- 막대한 노동력 필요
- 조사원이 살고 있는 지역이나 근처 지역은 사생활보호 제기
- 조사원이 모든 업무수행 : 리스트 작성, 조사표 배부, 무응답자 추적조사, 거주지 분류 및 내용점검

- 2006년 우편조사 계획

- 노동력 감소 : 우편조사로 조사방법 변경
- 사생활 내용 보완 : 우편이나 이메일로 조사표 송부
- 우편조사 실시 기본조건
 - 캐나다 우체국에서 주소 기반으로 배달이 가능
 - 캐나다 통계청의 주소록(address register)이 전지역을 포함
 - 캐나다 통계청은 지역코드에 맞도록 공간데이터 인프라 구축

- 캐나다 우체국에서 주소 기반으로 배달이 가능
 - 캐나다 우편은 주소(civic sort)와 점유자 성명(knowledge sort)으로 전달
 - 센서스는 주소단위로 송부하되 점유자 이름을 사용하지 않음

- 캐나다 통계청의 주소록(AR)은 전지역을 포함
 - 전지역을 커버하기 위해 이전 센서스를 벤치마킹

- 행정자료를 이용하여 주소록 update
- 우편배달 지역의 모든 주소는 현장에서 확인

- 캐나다 통계청은 지역코드에 맞도록 공간데이터 인프라 구축
 - 행정자료로부터 얻어진 주소는 우편블록별로 연결
 - 필드에서 update나 수집과정을 감소하기위해 행정자료 사용

- 면접용 리스트(Block canvass)는 센서스 7~9개월 전부터 update
- 시간이 지남에 따라 커버리지가 이슈가 됨 : 면접용 리스트 구비가 늦어짐
- 신축 거주지의 리스트 정비: 이는 누락될 가능성이 잠재됨
- 2006년에는 캐나다 통계청은 우편조사의 목표치는 가구의 64%였음

- 2006년 우편조사 설계

- 우편조사는 아래 사항을 기초로 설계됨
 - 주소록 커버리지
 - 우편 조사 가능한 주소의 결정
 - 운영가능한 요구사항

- 우편조사의 운영체제 개발
- 2001년부터 공표지역(Dissemination Area)을 도입
- 만약 주소록에 의해 공표지역을 커버한다면 공표지역을 우편조사로 선택
- 주소록의 분포는 매우 치우침
- 우편조사가 가능한 주소록의 커버리지가 95%이상인 모든 DA를 선택
- 2006년 총조사 거주지역의 70%가 우편조사 지역으로 가능
- 2006년 우편조사 지역은 캐나다 전체 면적의 0.3%만이 대표(도시지역)
- 캐나다 통계청은 9.4백만 개의 조사표를 보냄
- 반송된 조사표는 존재하지 않거나 비어있는 거주지가 대부분임

- 2011년 우편조사 계획

- 2006년 보다 우편조사 확대
- 2011년에는 주거지역의 80~85%에 우편발송 예정
- 2006년 우편조사 리스트 활용

- 2011년 우편조사 설계

- 우편조사 지역에 대한 설계는 큰 변동이 없음
- 2006년 센서스 주소록 활용
- DA 보다는 먼저 조사단위를 기초로 설계
- 주소록의 커버리지를 2006년 88.8%에서 95.7%로 확대
- 2006년 도시지역에서 군부지역으로 보다 확대
- 캐나다 우체국의 civic sort를 이용하여 공조
- 리스트 구축
 - .. 안정된 지역을 확인
 - .. 행정자료로부터 수집단위의 주소를 효율적으로 할당하기 위해 지리적 코딩을 중요
 - .. 안정적인 리스트를 구축하기 위해 새로운 규칙을 개발 필요

- 2009년 센서스 테스트

- 리스트 확인 작업
- 확장된 우편조사 지역으로부터 장소 추출
- 테스트 목적은 확장된 우편조사지역이 우편조사 가능여부를 확인
- 테스트는 안정된 지역과 불안정된 지역을 예측하기 위한 작업

- 결론

- 우편조사 조사표
 - .. 노동력을 감소하기 위해 실시
 - .. 응답자의 비밀보호를 개선
 - .. 주소록을 수정 및 보완 필요
 - .. 응답자들을 정확한 블록에 할당하기 위해 공간적 인프라 구축 필요
 - .. 캐나다 우체국에 의해 주소기반으로 배달이 가능하도록 구축

○ 2011년 캐나다 센서스에서 자료수집의 운영시스템 변화

- 2006년 이전 센서스

- 1971 : 자계식 조사표와 조사원 조사표 배부 및 회수 도입

- 1986 : 도시지역은 조사표 배부/우편회수, 군부는 조사원 배부/회수
- 1991 : 군부지역 조사표배부/우편회수(국가적 조사방법)
- 1996 : 98% 자계식, 2% 조사원 타계식 조사
- 2001 : 인터넷 조사 시험조사 도입, 하나의 조사 블록으로 형성

- 2006년 센서스

- 대부분의 조사표는 우편배부하기 위해 주소록(AR) 사용
- 우편 또는 인터넷을 통한 응답 가능
- 특별한 계층과 무응답자를 위한 현장 방문조사
- 위험을 줄이기 위해 과거방법을 보완

- 2006년 센서스 평가

- 기대보다 높은 인터넷 응답
- 질문서의 70%가 우편조사를 취함으로 현장 방문조사 감소
- 조사방법을 관리하는 통합시스템 운영
- 데이터 처리과정의 합리화

- 2006년 센서스 도전

- 현장조사 담당자와 의사소통
- 관리정보를 위한 적시성
- 현장 방문조사원 모집 및 보유

- 2011년 센서스 운영의 특징

- Master control system 운영
- 현장조사 운영시스템 운영
- 통합 시스템 운영(인력관리, 모집)
- 순차적인 조사방법(Wave) 접근

- 개선된 특징

- 질문지 회수 결과에 즉각적인 운영

- .. 매일 마스터 콘트롤 시스템에 등록
- .. 캐나다 우체국에서 우편조사 수령 등록
- .. 현장관리 시스템을 통한 현장조사 등록
- .. 인터넷조사 확인

- 순차적인 조사방법
- .. 응답자를 고무하기 위한 접근방법
- .. 초기 인터넷조사/우편조사를 위한 우편송부, 무응답 추적조사를 위한 면접조사
- .. 현장조사의 업무 경감

- 현장조사 운영 시스템

- 응답된 질문지를 빠른 시간내 알림
 - .. 우편, 인터넷, 중앙 전화 등을 통해 접수된 조사표를 MCS로부터 1일 2회 update

- 매일 조사현황 파악(Daily capture of productivity)
 - .. 조사완료된 리스트를 조사원과 감독자에게 보고
 - .. 조사완료 상황(조사표, 거부율, 빈집 등)
 - .. 거주지 접촉회수
 - .. 주된 자료수집방법

- 매일 투입비용 파악(Daily capture of pay claims)
 - .. 1일 노동시간 및 이행된 조사현황
 - .. 1일 소요비용(이동, 기타)
 - .. 소요 예산과 조사현황(가치)에 관한 관리정보의 신속한 이용

- 관리정보 시스템
 - .. 현장조사 조사원의 지출, 인원, 위치, 진행정도를 신속히 보고
 - .. 보고서를 해석할 수 있는 도구(지표)
 - .. 비용과 노력에 대한 관계 측정
 - .. 애로사항 및 완화전략 등 개선사항

- 응답자 요구에 신속한 대응

- Help Line에 받은 전화로부터 요청의 대응
- 지리정보 시스템과 MCL와 연결하여 현장에서 적절히 처리할 수 있도록 알림
- 요청 받은 내용에 대해 조사원과 감독관에게 알림

- 자료처리 자동화
 - 지역별로 응답과 목표량의 모니터링
 - 원하는 지역수준의 결과공표 여부 확인
 - 업무 종료를 신속히 알림

- 방문기록(visitation record) 확보
 - 2011년 거주지역의 20%는 list/leave 방법을 통한 수집
 - master control system의 리스트에 없는 거주지역
 - 감독자에 의한 현장에서 VR 확보
 - 모집단의 신속한 조사완료
 - 자료수집의 신속한 관리

- 결론

- 2011년 새로운 방법과 시스템을 통하여 운영방법을 대폭 개선
- 2016년에는 접촉시도에 따른 기록을 이용하여 노력도를 측정하는 등 보다 합리적으로 추진할 계획

○ 센서스의 새로운 자료수집 방법 : The Internet

- 캐나다 2006년 센서스 특징

- 매 5년 주기로 실시(every 5 years)
- 법적인 의무사항(legal obligation)
- 자계식 조사(self-enumeration)
- 12.4백만 가구(31.6백만명)
- 2가지 질문 형태
 - 80% short form, 20% long form
- 2가지 공식언어(영어, 프랑스어)
- 조사표

- .. 종이조사표
우편발송(mail-out) 2/3, 조사원배부/면접(drop-off/interview) 1/3
- .. 웹 조사표(센서스 웹 접속, 접속코드 사용)

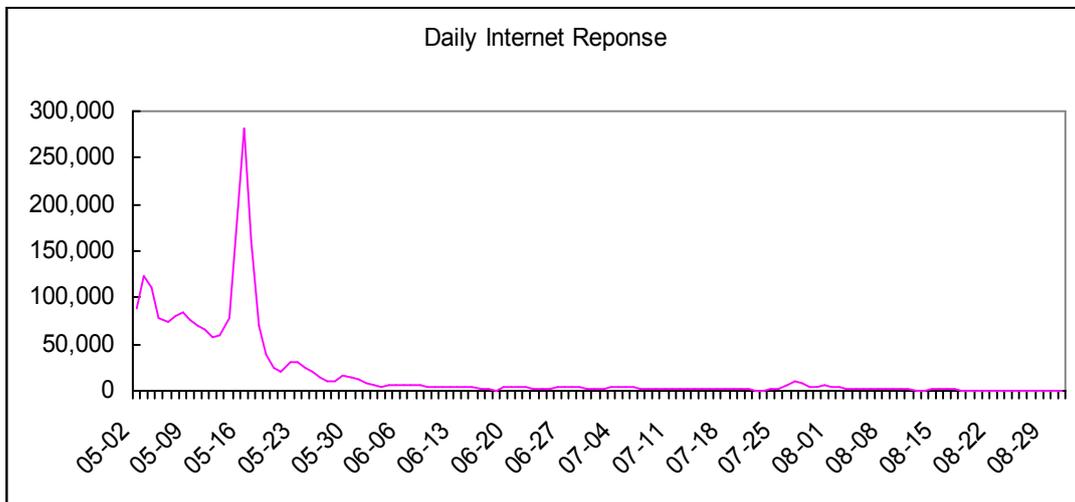
· 자료수집 방법

- .. 인터넷
- .. 종이조사표
- .. 전화(census help line)
- .. 무응답 추적조사(nonresponse follow-up)

- 자료수집 결과

수집방법	구성비(%)
인터넷	18.3
종이	63.1
CHL	2.3
NRFU	16.3

인터넷 조사 결과



· 조사방법별 거부율

조사방법	Short(%)	Long(%)
인터넷	2.5	5.7
종 이	5.6	39.1

- mode effects
 - .. 우편주소
 - .. 숫자변수
 - .. drop-down 메뉴

- 인터넷 조사의 장점

- 조사표의 빠른 응답
- 조사표를 우편으로 보낼 필요 없음
- 별도의 자료를 입력할 필요 없음
- 자료내검 가능, 조사표를 완성도 높음, 보완조사가 덜 필요
- 온라인으로 도움기능 가능
- 문제 발견시 수정이 쉬움
- 인터넷 사용(접속률 : 70%, 2006 센서스 : 18.3%)
 - .. 인터넷조사는 충분히 홍보가 되지 않음, 아직 종이조사표 선호

- 인터넷 조사 홍보에 관한 연구

- 종이 조사표를 제공하지 않고, 인터넷을 통한 조사표를 완성하도록 요구
- 종이 조사표 대신하여 가구에 아래의 사항을 편지로 보냄
 - .. 센서스 웹사이트 정보(URL)
 - .. 접속코드(access code)
 - .. 무료 전화번호
- 인터넷으로 응답이 가능할 지역을 미리 결정
- 4만 가구의 잠정표본을 선정
 - .. IRP(internet response promotion) 표본 - 2만 가구(편지를 받음)
 - .. control 표본 - 2만 가구(종이조사표 받음)

· IRP 결과

조사방법	센서스	control 표본	IRP 표본
인터넷	18.3	23.8	61.8
우편	63.1	61.1	7.6
CHL	2.3	2.1	7.7
NRFU	16.3	13.0	22.9

- 추적조사(follow-up surveys)
 - 우편에 대한 부정적 반응이 적음
 - 21%가 인터넷에 대해 기술적인 어려움을 가짐
 - 18%가 유효한 전달을 받음
 - 인터넷을 통하지 않는 40%는 CHL에 전화
 - 인터넷조사를 하지 않는 이유
 - ... 종이조사표 선호
 - ... 기술적인 어려움
 - ... 우편에 어려움 가짐(문맹 등)

- 2011년 센서스 계획

- 인터넷조사를 위한 거대한 규모로 홍보 실시
 - 60% 캐나다 가구에 대해 실시
- 목표 인터넷조사 구성비 : 40%
- 사전에 편지를 받은 가구가 인터넷조사에 응할 가능성 높음
 - 2006년 자료를 로지스틱 회귀분석한 결과
- 2009년에 센서스 테스트에서 가설의 검정
- 순차적인 자료수집 방법 활용
 1. 편지 보냄
 2. 독촉 편지
 3. 무응답자에게 종이 조사표 우편배부
 4. 무응답 추적조사(면접조사)