
2015년 FCSM 리서치 컨퍼런스

[Federal Committee on Statistical Methodology Research Conference]

참석 결과보고

2016. 1.



통계개발원
조사연구실



목 차

I. 출장 개요	1
II. 컨퍼런스 주요내용	2
III. 주제별 내용	3
IV. 시사점	19

[붙임] 컨퍼런스 세션별 주제 목록

I 출장 개요

1. 컨퍼런스 개요

- 컨퍼런스명 : 2015 FCSM Research Conference
 - 기 간 : 2015. 12. 1(화)~12. 3(목)(3일간)
 - 개최지 : 미국, 워싱턴 D.C.
 - 관련 사이트 URL
: <http://fcsm.sites.usa.gov/reports/research/2015-research-conference>
- * FCSM(Federal Committee on Statistical Methodology) :
미연방정부 관리예산처(OMB)에 의해 창설되었으며, 연방통계 품질에 영향을 줄 수 있는 방법론과 통계적 이슈에 관해 관리예산처와 통계정책 관계부처에 정보를 제공하는 관계부처합동 위원회임

2. 출장 개요

- 출장자 : 조사연구실 최지형 사무관, 성주현 주무관
- 기 간 : 2015. 11. 30(월)~2015. 12. 5(토)(4박 6일)
- 출장배경 및 목적
 - 최신 조사방법론에 대한 이해 및 연구동향 파악
 - * 'Federal Committee on Statistical Methodology(FCSM) Research Conference'는 데이터 품질 제고를 위한 새로운 조사방법론 소개 및 방법론 관련 다양한 주제에 대해 정보를 공유할 수 있는 장임
 - 국가통계 생산 및 관련기관에서 약 800여명 등록 및 참석

3. 출장 주요 내용

- 최근 조사방법론에 대한 연구 및 동향 파악
 - 총 40개 세션의 145개 연구주제로 구성되어 있으며 조사방법론과 관련된 최근 동향과 다양한 연구주제를 다룸
- 조사방법론 개발·연구를 위한 자료수집
 - 자료연계를 통한 2차 자료생성, 무응답 대체, 측정오차 축소, 혼합모드 자료수집, 모바일조사 설계 등

II 컨퍼런스 주요 내용

- 자료연계, 조사표 설계, 무응답 대체와 관련된 연구가 많이 다뤄졌으며, 이외에 혼합모드 자료수집, 시계열 예측, 베이지안 방법 응용 등도 다루고 있음
 - (자료연계) 데이터의 가치를 높이기 위한 행정자료 연계, 새로운 데이터 생성을 위한 2차 자료생성 기존의 자료연계 뿐만 아니라, 데이터의 품질 향상을 위한 레코드 연계 활용 평가, 측정오차를 줄이기 위한 행정자료 활용 등 데이터 품질 제고를 위한 목적으로도 연구가 진행되고 있음
 - (조사표 설계) 조사환경이 어려워짐에 따라 응답자 편의 및 조사의 정확성 제고를 위해 웹조사, 모바일조사 등에 대한 연구가 활발히 이루어지고 있음
인터넷조사와 함께 모바일기기 중 스마트폰을 이용한 조사에 대한 관심도 증가하고 있으며, 이에 따라 모바일 기기 화면설계 및 구성에 대한 연구가 수행되고 있음
 - (무응답 대체) 핫덱대체를 위한 분류트리 활용, 행정자료의 누락 데이터 처리방법 등 무응답 대체에 관한 여러 주제를 다루고 있음
- 또한, 미국 센서스국을 비롯한 노동통계국, 사법통계국 등에서 사전연구로써 진행중에 있는 연구들(Case Study)이 발표되었음

Ⅲ 주제별 내용

① 자료연계(Data Linkage)

- ❖ Linking BEA Surveys of Foreign Direct Investment in the United States with BLS Register and Survey Data: Twenty-First Century Methods(U.S. Bureau of Labor Statistics)
 - (연구배경) 외국인직접투자에 따른 고용 영향을 측정할 필요가 제기되면서, 신규조사는 많은 비용과 시간을 필요로 하므로 사업체 등록자료와 외국직접투자에 대한 자료를 연계함으로써 필요한 통계 자료의 생산을 제안함
 - BLS(노동통계국)는 사업체 등록자료(business register)로써 QCEW(EINs포함) 자료를 가지고 있고, BEA(경제분석국)는 외국인직접투자(FDI, EINs포함)에 관한 정보를 수집하고 있음
 - QCEW: quarterly census of employment and wages
 - EINs: employer identification numbers
 - (연구방법) 매칭 프로세스 1단계는 자동매칭(사업체 연계를 위해 QCEW EINs와 BEA조사의 각 계열회사 EINs를 이용), 2단계는 매칭분석(자동 매칭된 사업체 내검, 누락 사업체 확인 등), 매칭 우선순위 분석 등으로 진행함
 - (연구내용) 이 연구의 주요 작업은 BEA자료와 사업체기반의 QCEW EINs를 연계하는 것이며, BEA와 BLS의 마이크로데이터를 연계함으로써 FDI와 연관된 고용과 임금 지표를 개발하기 위한 개념적 프레임워크를 제시함
- ❖ An 'Algorithmic Links with Probabilities' Crosswalk for USPC Codes(US Census Bureau)
 - (연구배경) 특허는 혁신과 기술변화 및 보급에 대한 유용한 대용물(proxy)이 될 수 있지만 경제분석과 같은 경우에 특허데이터를 완벽

하게 활용하기 위해서는 이용데이터가 경제활동 측정과 결합된 특허를 필요로 함. 그러나 이것은 어렵다고 증명되었음

- (연구방법) 최근 Lybbert and Zolas(2014)는 국제특허분류(International Patent Classification; IPC)를 몇몇 '빅데이터' 요소와 결합하여 산업 데이터와 연계하는 작업을 하고, 이를 'Algorithmic Links with Probabilities' 접근이라고 명명하였음. 이 연구에서 유사한 접근을 활용하여 새로운 특허분류인 U.S. Patent Classification(USPC) 시스템에 적용
 - (연구결과) 그 결과 U.S. 특허와 북미산업분류(NAICS), 국제표준산업분류(ISIC), 무역코드(HS), 표준국제무역분류(SITC)와 연계하는 USPC-산업 연계 색인을 완성함
 - (기대효과) 이 연계의 결과로서 특허에 대한 USPC-IPC 연계 교차분석이 가능하고, 연구자로 하여금 미국특허를 산업수준에 따른 다양한 경제데이터를 연계하는데 활용할 수 있게 됨
 - 나아가 상호활용에 대한 잠재적 이용가능성과 경제와 특허 트렌드에 대한 다양한 분야에 걸쳐 특허정도를 측정하는데 어떻게 활용되는지 등을 확인할 수 있음
- ❖ Creating Improved Survey Data Products Using Linked Administrative-Survey Data(University of Chicago)
- (연구배경) 중요한 정책 목적으로 활용되는 자기보고(self-reported) 형식의 조사자료에는 심각한 측정오류(무응답, 표본추출, 조사범위, 자료처리, 측정 등의 오류)가 있음
 - (연구방법) 조사오류를 수정하는 방법으로는 조사자료(survey data)를 행정자료(administrative data)와 연계하는 것과, 행정자료 값으로 직접 대체하는 것, 연계된 자료에 대해 모형을 설정하고 더 최근의

조사가 잘된 마이크로데이터로 값을 대체하는 것이 있음

- CPS(Current Population Survey) 조사자료, 의료보장제도(Medicaid) 및 식료품할인구매제도(SNAP)의 행정자료를 이용하여 연구함

○ (연구결과) 조사자료 결과물(microdata, summary files, report tabulations)의 측정오차를 줄이고 정책분석을 위한 조사자료 개선을 위해 결합 모델링(incorporating modelling)을 시작해야 함

- 데이터의 세계는 변화되고 있고, 연방통계 커뮤니티는 데이터의 타당성을 향상시키는 방법들을 찾음으로써 보조를 맞추기를 요구하므로 매우 중요한 조사자료 결과물을 개선시키기 위해 연계자료 모델링 (또는 다른 전략)을 시작하는 것을 늦춰서는 안됨

❖ Using the Census to Evaluate Administrative Records and Vice Versa(U.S. Census Bureau)

○ (연구배경) 센서스 전수 활용을 위해 행정자료의 정확성과 유효 적용범위를 평가하는 이상적인 방법은 2010년 4월 1일 각 가구별 실제거주와 거주자수 비교를 통해서일 것임. 2010 센서스에서 이러한 정보를 제공한다 할지라도 모든 센서스 enumeration이 동일하게 신뢰할 수 있는 것은 아님

- 조사응답데이터의 품질을 평가하는 보편적인 방식은 같은 사람에 대해 행정자료 정보와 비교해보는 것. 예를 들어 Meyer and George (2011)는 ACS(American Community Survey)와 CPS(Current Population Survey) 두 조사의 식료품할인구매 영수증과 행정자료의 식료품 할인권과 비교

- 최근에 연구자들은 Groen(2012)처럼 행정자료원내에서의 시스템 오류에 대해 기록하기 시작. 미 통계국에서는 연구자들이 센서스나 행정자료기록들이 모두 완벽하지 않을 가능성을 열어두고 조사

응답자들을 평가하기 위한 조사툴로서 행정자료를 사용하고 있음.
이번 연구는 이와 맥락을 같이함

- (연구결과) 이번 연구는 일부 센서스 응답은 행정자료 기록보다 더 높은 품질이고 나머지는 비교적 낮은 품질일 것이라고 가정. 우리가 가정한 센서스응답자들의 특성은 정확성과 관련이 있다는 것을 확인함으로써 각각의 품질을 평가하기해 조사자료와 행정자료원을 종합적으로 사용할 것을 제안

② 조사표 설계(Survey questionnaire)

- ❖ Designing Web Surveys for Different Screen Sizes and Operating Systems: What Respondents are Using(ICF International)
- (연구배경) 견고한 설문지 설계의 전반적인 원칙은 측정오류를 가져올 가능성을 최소화하기 위해 응답자에게 동일한 자극을 제시하는 것. 웹 설문을 작성하기 위해 응답자들이 사용할 수 있는 스마트폰, 테블릿, 컴퓨터 및 노트북을 포함하는 다양한 기기 사용 증가와 함께 조사표 설계자들은 동일한 자극을 제시하는 웹 설문지를 설계하는데 새로운 국면을 맞고 있음
- (연구방법) 좋은 디자인 결정을 위해 ICF는 사용된 운영시스템의 형태뿐만 아니라 이러한 형태의 기기사용빈도에 대한 데이터를 수집함. 이러한 데이터는 ICF웹조사를 수행한 사람들을 대상으로 그리고 로그인했지만 조사를 시도하지 않거나 웹조사를 다 작성하기 전에 멈춰버린 사람들에게 대해서 데이터를 수집함
- (연구 결과) 연령별로는 25세 이상의 응답률이 82.6%로 가장 높았으며, 그 형태별로는 스마트폰이 가장 많이 사용되었고, 그다음으로는 노트북을 포함한 컴퓨터가 그리고 테블릿이 가장 낮았음. 또한

사용하는 브라우저의 경우에는 크롬이 가장 많았음. 스크롤바없는 스크린의 높이는 38%만이 중요하다고 응답했고, 연령대가 있는 경우에는 운영시스템이 스크린사이즈보다 더 중요한 것으로 나타났음

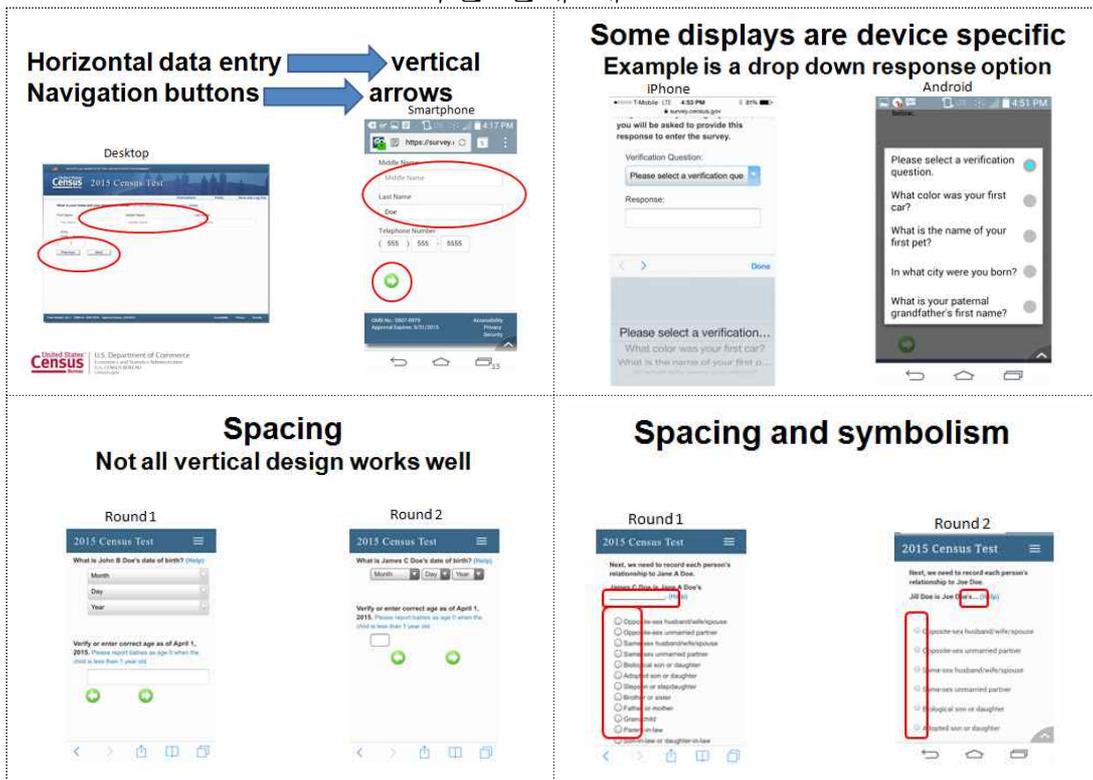
❖ Internet and Smartphone Coverage in a National Health Survey: Implications for Alternative Modes(Survey Research Center, University of Michigan)

- (연구배경) 인터넷이 가능한 스마트폰의 증가는 인터넷 커버리지의 문제와 그것의 커버리지 바이어스 영향을 다시 살펴볼 기회를 주었음
 - 많은 연구에서는 모바일 커버리지와 인터넷 커버리지를 분리하여 연구를 해온 반면, 공동범위의 보완적인 효과는 살펴보지 않아 왔음
- (연구방법) 이 연구에서는 인터넷기반 추적조사의 실행가능성을 탐색하기 위해 NSFG*(National Survey of Family Growth)에 두 가지의 면접자 관찰(인터넷 접근, 스마트폰 소유)을 추가하였음
 - * 15~44세 남녀를 대상으로 2년간('12.9월~'14.8월) 분기조사 실시
 - 전반적으로, 응답자의 82.2%가 인터넷 접근을 하고, 76.1%가 스마트폰을 가지고 있는 것으로 나타났고, 이것은 89.4%가 전통적 기기 또는 스마트폰을 통해 인터넷에 접근한다는 것을 의미함
- (연구결과) 인터넷 커버리지, 스마트폰 커버리지, 인터넷과 스마트폰이 결합된 커버리지의 사회·인구학적 연관성 연구를 제안함
 - 또한, 출산율, 가족형태 등과 관련하여 NSFG에서 작성되는 주요 추정치에 관한 상이한 커버리지의 영향을 살펴볼 것을 제안함
 - 이 연구에서는 선택적 모드와 관련된 무응답 바이어스 문제를 말하는 것이 아니라 웹기반 모드의 자료수집 방식으로 전환하는데 있어서 야기되는 커버리지 바이어스의 영향을 언급하고 있음

❖ Optimizing the Decennial Census for Mobile : a Case Study(U.S. Census Bureau)

- (연구배경) 미국 센서스국은 2020년 센서스를 위해 인터넷 응답을 선택할 수 있도록 하기 위해 노력하고 있음
 - 2014년에 온라인 도구는 데스크탑 또는 랩탑에 적합한 뷰를 위해 디자인 되었는데, 이들 조사가 테블릿 또는 스마트폰에 응답이 되었다고 해도 화면 디스플레이가 매우 작았고, 이용자는 분명하게 질문을 읽고 응답할 수 있도록 화면을 확대하거나 다른 조작들이 가능하기를 요구해 왔음
 - 2014년에 모바일 소유 통계는 스마트폰을 가지고 있는 성인은 64%, 인터넷 접근을 위해 스마트폰에 의존하는 경우는 성인의 7%인 것으로 나타남
 - 모바일 기기 소유의 성장과, 모바일 기기를 통한 인터넷 접근 인구에서의 차이를 통해 센서스국은 2020년 센서스를 위해 모바일에 적합한 설계를 해야 한다고 인식함
- (연구내용) 2015년 센서스 테스트에서 전통적인 데스크탑 및 랩탑 디자인과 함께 모바일에 적합한 온라인 센서스 조사를 개발했음
 - 더 작은 기기에 대한 가이드 마련(문항은 줌 또는 스크롤 없이 읽을 수 있어야 함, 화면을 완전히 사용, 배너/내비게이션/메뉴에 대한 제안 등)
 - 모바일에 적합한 화면 설계, 조사항목 배열, 보기항목 구성, 조사시간, 조사완료여부 등 테스트를 실시
- (연구결과) 모바일 기기에 맞게 조사표를 최적화함으로써 더 유용할 수 있음을 알 수 있었고, 모바일이 PC보다 시간이 더 걸리고, 조사 중단이 증가하는 것으로 나타났음(그러나 조사시간과 조사완료에 대한 테스트가 이용가능성 향상의 가장 좋은 측정방법은 아님)

< 화면 설계 테스트 >



❖ Gridlock : Improving Questionnaire Design for Mobile Devices(U.S. Census Bureau)

- (연구배경) 스마트폰과 태블릿을 포함하여 모바일 기기로 온라인 조사에 참여하는 참가자 비중은 증가하고 있으며, 모바일 기기의 화면구성은 온라인조사에서 성공적인 측정을 위한 결정적인 키임
- 조사의 25~35%가 모바일 기기로 조사를 시작, 조사의 20~30%가 모바일 기기로 조사를 완료함(모바일 기기의 조사중단(break-off)이 더 많이 발생함)
- 온라인에서의 조사항목 제시방법 중에서, 단일 응답을 요구하는 단일 화면에서의 단일 항목들은 온라인조사에서 일반적인 형태이나, 조사항목을 읽고 이해하는 시간, 응답을 위한 손동작, 다음 문항으로의 이동을 위한 다음 버튼 선택 등으로 인한 시간 소요로 효율적이지 않음

- 반면, 그리드형태는 응답 매뉴얼의 표준화, 응답 의미의 표준화로 스크린 전체에 걸쳐 제시되는 단일 항목보다 더 효율적임
- (연구방법) 모바일의 그리드를 향상시키기 위해 두 가지 가능성을 연구하였는데, 그리드를 단순화하고, 응답 범주 수를 축소하여 모바일에 친근하도록 그리드를 만드는 것임
- 테스트를 위해 사용한 기기는 데스크탑/랩탑, 태블릿, 스마트폰으로 3가지이며, 그리드형태는 Traditional, Banked single items, Response on right, Focal element, Focal element으로 4가지임

< 조사항목 그리드 설계 >

Traditional Grid				Banked single Items			
	Not important	Important	Extremely important	It is low in calories			
It is low in calories	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Not important	Important	Extremely important	<input type="radio"/>
It is refreshing	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	It is refreshing			
It is a good value for the money	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Not important	Important	Extremely important	<input type="radio"/>
	Not important	Important	Extremely important	It is a good value for the money			
It gives you quick energy	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Not important	Important	Extremely important	<input type="radio"/>
It has a great taste	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Focal Element			
It is easy to drink on the go	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Item 1 of 7 (0 completed)			
	Not important	Important	Extremely important	It is low in calories			
It has a unique combination of ingredients	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Not important	Important	Extremely important	<input type="radio"/>

Responses on Right		Focal Element	
It is low in calories		It is low in calories	
It is refreshing		It is refreshing	
It is a good value for the money		It is a good value for the money	
It gives you quick energy	Not important	It gives you quick energy	
It has a great taste	Important	It has a great taste	
It is easy to drink on the go	Extremely important	It is easy to drink on the go	
It has a unique combination of ingredients		It has a unique combination of ingredients	

- (연구결과) 조사완료 시간은 Traditional, Banked single items, Response on right, Focal element 순으로 짧게 나타나, Focal element형태가 가장 많은 시간이 소요된 것으로 나타남

- 대부분 새로운 그리드형태는 조사시간이 더 오래 걸리며, Traditional 그리드는 모든 기기에 걸쳐 좋은 기능을 보임
- 응답 항목수가 더 적은 경우(Friendly unipolar, 4점 척도)과 더 많은 경우(UnFriendly bipolar, 7점 척도)로 구분하여 4가지 그리드 형태를 테스트한 결과, 3가지 기기에 대해 Friendly unipolar(응답 항목수가 더 적은) 경우 조사완료시간이 더 짧았음
- 전반적으로, 효율성 측면에서 Traditional 그리드를 능가하지 못했으며, Banked single items 그리드는 모바일 기기로 효율적이고 잘 작동될 수 있음

③ 무응답 대체(Imputation)

- ❖ Treatment of Missing Data in Hierarchically-Structured Administrative Records: A Case Study in the Bakken Region Using FBI's National Incident Based Reporting System(RTI international)
- (연구배경) 국가의 주요관심지표를 추정하기 위해 행정자료에 크게 의존하는 경향에 대응하여 연방 통계대행기관은 행정시스템을 통해 수집되는 데이터의 품질을 평가하고 품질 향상을 위한 활동을 확대 및 개선해 왔음. 행정자료가 종종 다른 대행기관이나 지역 단위 통계가 수집될 경우, 개인이나 집계수준에서 손실자료(missing data)가 나타나는 것은 피할 수 없고 추정시 이는 강조되어야만 함
- * NIBRS(The National Incident-Based Reporting System): 통계나 조사 목적으로 사용되어지기 위해서 행정자료로부터 데이터를 수집하기 위해 고안된 시스템. 그것은 범죄사건에 포괄적인 정보제공을 위해 UCR(Uniform Crime Reporting) 프로그램의 일환으로서 FBI에 의해 개발

- (연구방법) 이번연구에서, 기타 관련 있는 외부데이터를 활용하여 NIBRS에서 손실데이터를 처리하기 위해 개발된 임퓨테이션방법을 소개하고자 함. NIBRS의 계층적 구조하에서 이 특정방법은 다양한 수준에서 발생하는 데이터 손실(a) 사건수준에서 각 사건 내에서의 아이템 손실, b) 특정연도에만 보고하는 몇몇기관 때문에 발생하는 월별 기관단위손실, c) NIBRS에 데이터를 일년 내내 제출하지 않은 몇몇기관 때문에 발생하는 연도별 기관수준에서의 데이터 손실)을 강조. 제안된 방법은 2006년에서 2012년에 오일생산이 증가하면서 얼마나 범죄와 법집행이 변화했는지 검토하기 위해서 NIBRS데이터를 활용하여 미국의 Bakken지역 연구에 적용. 이 임퓨테이션 기술에 의해 소개된 우리의 추정에 있어서 불확실성을 평가하기 위해 분산추정방법 또한 제시될 것임

- (연구결과) 데이터용량, 무응답, 데이터에디팅부분을 처리하는 것이 힘든 과제였으며, 다른 여러 가지 임퓨테이션 방법을 적용하여 실험한 결과, 처음에 사람수준에서 핫덱방식으로 결측치를 대체하고 대형기관수준으로 집계하고 다시 임퓨테이션을 처리하는 것이 가장 적절. 이번 연구는 넓은 관점에서 이 연구는 계층적 구조로된 행정자료에서 손실자료를 어떻게 처리하는지를 보여주는 예시가 될 수 있음

❖ Using Classification Trees to Recommend Hot Deck Imputation Methods: A Case Study(U.S. Census Bureau)

- (연구배경) 미 통계국은 매 5년마다 경제총조사를 실시. 모든 대상사업체로부터 일반적인 정보를 수집하는 것 이외에 경제총조사는 각 사업체의 생산제품에 대한 구체적인 정보를 수집. 2017년에 실시한 경제총조사는 여러분야 데이터 수집으로부터 경제전반과 제품간의 통계표 작성을 위해 제품복미제품분류 시스템을 이용할 예정. 이것은 제품과 산업을 연계하는 현재의

수집방법으로부터의 중요한 출발이며, 거래분야 특정 자료손실 조정관행을 불가능하게 함. 학제간 리서치팀은 이러한 문제를 다루기 위해 설립되었고 그결과는 손실된 제품데이터 처리에 대해 하나의 통합된 방법을 권고

- (연구방법) 이 팀은 2012년과 2017년의 경험적이고 가상의 경제 총조사 자료를 활용하여 다양한 평가연구를 수행. 연구된 모든 산업에서 핫덱 임퓨테이션방식이 고려된 방법들 가운데 최고의 절충안. 그러나 추천된 변화(nearest neighbor or random)는 무역분야에서 나뉘어졌음. 따라서 팀은 사업체나 제품의 어떤 특성이 다른 하위분야를 통해 다른 것 보다 더 통계적 특성을 나타내는 하나의 방법이 예측될 수 있는지를 조사
- (연구내용) 탐색분석은 분류트리분석을 사용. 분류트리는 범주 결과로부터 시작하고 설명적 변수를 나타내는 가지로 늘어남. 트리의 가장꼭대기마디는 결과를 예측하는 것과 가장 관련이 높은 공변량을 나타냄. 연구결과 적용에 있어서 결과는 핫덱 방법의 선택이었으며, 예측변수는 임퓨테이션셀내에서 사업체 데이터의 특성임. 이번 연구에서 다른 것보다 하나의 방법을 더 선호하게 되는 근본적인 원인을 이해하는 것을 개발하기 위해 어떻게 분류트리를 사용했는지를 설명하고 2017년 경제 총조사에 적용될 제한된 가이드라인을 제공함

④ 혼합모드 자료수집(Multimode data collection)

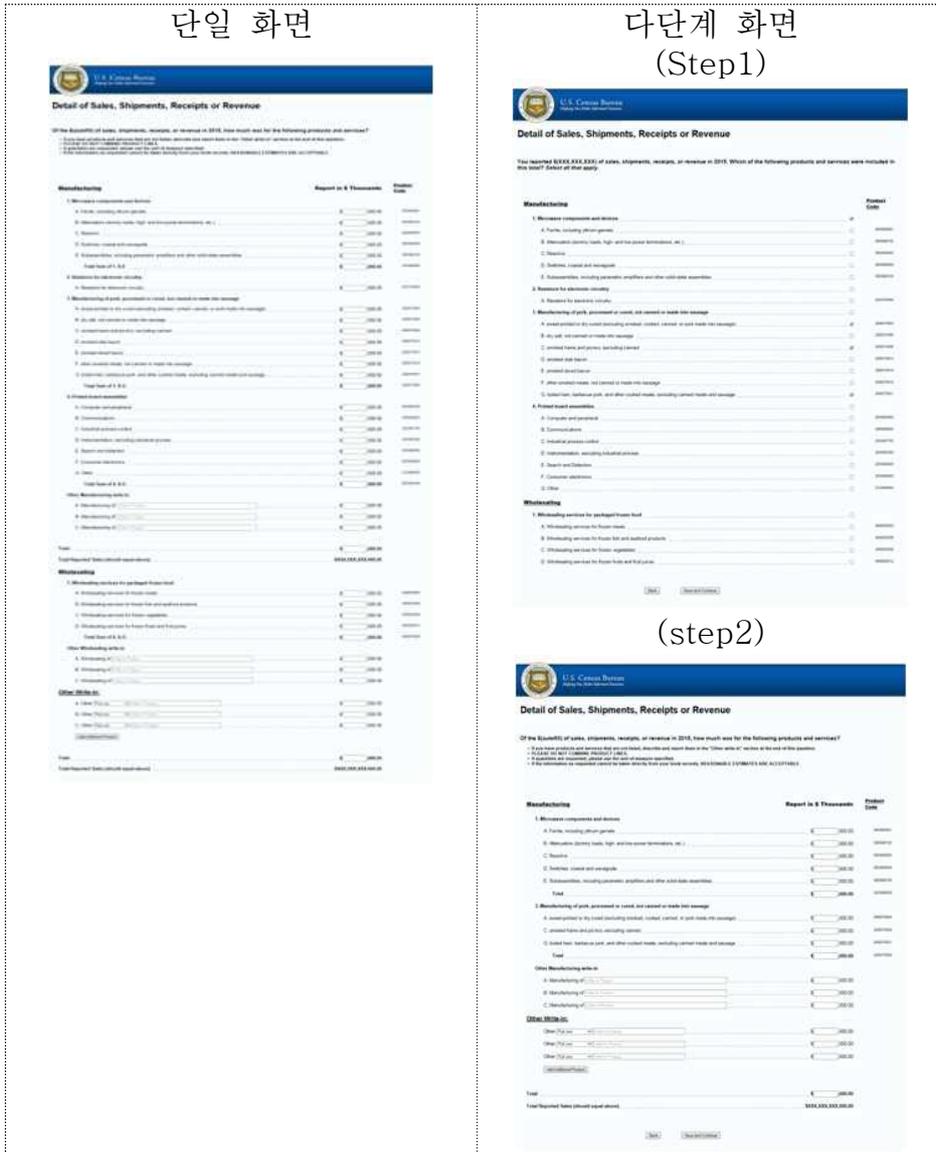
- ❖ Data-Driven Decision Making and the Design of Economic Census Data Collection Instruments(U.S. Census Bureau)
- (연구내용) 경제총조사는 미국 경제와 관련하여 이용가능한 가장 상세한 벤치마크 데이터를 제공하며 GDP 추정을 지원함

- 경제총조사는 5년마다 실시되며, 2012년에는 자기기입 종이조사표와 전자조사표를 통해 조사되었고, 2017년에는 전자적으로 자료수집이 될 것임
- 과거에는 종이조사표 전략에 따랐던 반면, 향후에는 모든 웹 데이터 수집으로 종이조사표 설계의 한계에 얽매이지 않고 자동화된 설계 특성과 더 높은 수준의 조사표 화면 설계(tailoring) 기회들을 가짐
- (연구결과) 화면 설계(질적 방법론)에 있어서, 응답자는 화면에 한번에 제시되는 상품과 서비스 목록이 더 적은 것을 선호함
 - 응답률과 품질 향상을 위해 조사문항과 디자인이 응답자의 관심을 불러일으킬 수 있어야 함. 특히 시작부분에서 응답자가 연구자에게 제공할 것을 가지고 있다고 것을 보여 주어야 함
 - 응답자가 화면에서 보는 정보의 양을 축소, 처음에 리스트된 품목 라인은 응답자의 산업으로 화면이 설계(tailoring)되어 쉽게 확인할 수 있는 것을 원함
 - 단일 화면 설계(스크롤링 화면으로 제품 리스트를 보여주는 것)와 다단계 화면 설계(여러 스크린에 걸쳐 제품 제시하는 것으로 처음 화면에서 제품을 선택하고 이어지는 화면에서 응답값을 입력하도록 함) 중 응답자들은 길게 제시된 제품 리스트를 당황스러움을 느꼈으며, 더 짧고 더 구분된 리스트를 선호하였고 응답자들은 다단계 화면 설계를 선호하였음
 - 제품 제시(양적 방법론)에 있어서, 응답자가 진행해야 하는 정보를 양을 축소시키고, 응답자에게 제시할 제품들을 결정하기 위해 일관성 있고 자동화된 방법론을 개발하도록 함

< 화면 설계 비교 >

단일 화면

다단계 화면
(Step1)



❖ Evaluating Collection Mode Effect with Web Data Collection in the Canadian Labor Force Survey(Statistics Canada)

- (연구배경) 캐나다 노동력조사는 월간 가계조사로 실업과 기타 노동 특성을 측정하는 조사임. 이 조사는 rotating panel조사로서 가구는 6개월 연속으로 표본이 되고 그 패널은 전화와 직접면접방식 자료 수집에 이용되어왔음. 2015년에 노동력조사는 웹방식 데이터 수집 방법을 도입. 전화나 직접 대면조사의 마지막에 전화나 직접 면접관을 만나는 대신 응답자들은 다음달에 전자설문지를 온라인

으로 작성을 선택할 수 있음. 주요 추정에 있어서 새로운 전략의 영향을 평가하기 위해 전자 설문지는 각 지역에 새로운 표본의 절반에 전자설문지를 제공하면서 무작위 램덤실험으로 도입함

- (연구내용) 디자인기반과 모델기반 분석 두 분석 모두 자료수집방법 변화가 가져올 수 있는 오류를 평가하기 위해 수행. 두 전략간 직접 비교가 실업률과 기타 주요 조사추정을 위해 이루어짐. 마르코프 잠재그룹분석은 반복적인 매월 자료수집을 통해 노동력 오분류를 변화를 평가하기 위해 이용. 결국, 성향점수매칭은 자료수집 방법의 자가선택을 조정하는 동안 자료수집모드를 통해 응답을 비교하는데 사용함
- (연구결과) 제공된 전자설문지 절반정도에서 유의한 차이를 보이지 못함. 낮은 비중(20%)으로 인해 전자설문도입으로 인한 유의한 오류나 영향은 무시할 정도임. 향후 전체 전자설문서 작성을 진행하여 지속적으로 모니터링할 계획임. 기타 추정이나 세부그룹에서의 영향 검토할 예정임

5] 기 타

- ❖ Self-reported vs. Administrative Race/Ethnicity Data and Racial Disparities in Criminal Justice System(U.S. Bureau of Justice Statistics)
- (연구배경) 연방통계기관으로서 BJS(사법통계국)는 자료수집에 있어서 인종과 민족성에 대한 OMB(관리예산처)의 1997년 표준을 고수해야 함
- OMB 표준 원칙은 인종 질문에 대한 개인의 응답은 자기신분증명(Self-identification)에 기반을 두어야 한다는 것임. BJS는 교정인구에 대한 행정자료를 주기적으로 받고 있는데 이 자료는 개인의 인종과

히스패닉 분류에 대해 자기신분증명에 기반하지 않음

- (연구방법) 이 연구는 BJS의 시설조사*에서 자기신분증명이 된 인종·히스패닉 범주에 따른 재소자 총계와 주(state)의 행정자료**를 비교함

* Survey of Prison Inmates(교도소 수감자 조사), National Inmate Survey

** National Corrections Reporting Program and National Prisoner Statistics Program

- (연구내용) 자기보고(Self-reported)와 비교하여, 행정자료는 백인 또는 흑인인 재소자 비중을 과대 추정하는 경향이 있는 반면, 히스패닉, 아메리카 원주민, 아시안, 다민족 재소자 비중은 과소 추정되는 경향이 있음

- 측정에서의 차이는 인종, 성, 연령과 자료수집 연도를 포함한 요인들의 결합에 따라 크게 달라지며, 포아송분포를 이용하여 행정자료로부터 자기신분증명이 된 인종으로 재소자의 총계를 예측함
- 이 연구에서 인종의 서로 다른 측정이 범죄재판시스템에서 인종간 차이를 추정하는데에 얼마나 영향을 미치는지를 연구함
- 행정자료로부터 얻어진 것에 비해 자기신분증명에 근거한 측정의 경우 히스패닉 대 백인의 구금율의 비중은 약 68% 증가하고, 흑인 대 백인 구금율의 비중은 단지 7% 더 높았음

❖ Characteristics of Special Purpose Entities in Measures of U.S. Direct Investment Abroad(Bureau of Economic Analysis)

- (연구배경) 국제가이드라인에 따라 국제무역 및 투자에 대한 공식통계는 다국적기업(MNEs)내 거래도 포함. 그러나 공식통계의 이해와 해석에 있어 복잡한 요소는 하나 혹은 그 이상의 특수목적회사(SPEs)을 가진 다국적기업에 대한 거래 및 상태임. 국제가이드라인은 특수목적회사(SPEs)와의 거래 및 특수목적회사(SPEs)내에서의 위치는 공식통계에 포함되어야하고

특수목적회사(SPEs)의 역할을 보다 더 잘 이해하기 위해서 추가적인 정보가 제공되어야만 한다고 권고

* 특수목적회사(SPEs)는 일반적으로 거의 물리적인 형태나 거의 경제적 활동이 없음. 예를 들어 금융 및 지주회사, 특허 및 라이선싱회사, 임대회사 등이 이에 포함

- (연구내용) 최근 연구들은 통계에서 특수목적회사(SPEs)의 효과를 연구하고 있지만, 미국 해외직접투자(USDIA)에 관한 거래나 상태에 포함된 특수목적회사(SPEs)의 근본적인 특성에 대해서는 거의 알려지지 않았음. 따라서 이 연구는 미국 해외직접투자(USDIA)에 관한 거래나 상태에 포함된 특수목적회사(SPEs)의 실험적인 사례를 제공
- (연구방법) 특히, 이 연구는 미상무부 경제분석국에서 수집한 미국모회사나 외국계열에 관한 데이터를 활용하여 미국 다국적기업 기업이 모회사인 비거주 특수목적회사의 금융 및 운영특성을 제시. 이번분석은 비거주 특수목적회사와 비거주, 비특수목적회사 대응과 평균비교검정을 포함. 비용에 대한 대응물로서 보고된 매출액과 생산액에 대한 대응물로서 부가가치를 기반으로 특수목적회사(SPEs)와 비특수목적회사에 대한 듀얼비용기능에 대한 추정을 포함
- (연구결과) 주요 결과는 비거주 특수목적회사, 비거주, 비특수목적회사 대응의 보고된 재무재표와 손익계산서 사이에는 큰 차이를 보였으며, 그 큰차이는 특수목적회사(SPEs)와 비특수목적회사에 따른 부가가치 차이임. 그러나 보고된 매출액에 대한 부가가치효과는 특수목적회사(SPEs)와 비특수목적회사가 혼합되어 있음

Ⅲ 시사점

- 자료연계, 조사표설계, 무응답대체 분야의 연구가 활발히 진행되고 있으며, 특히 자료연계 분야는 행정자료와 조사자료간의 연계를 통한 2차 자료생성 뿐만 아니라 통계자료 정확성 제고를 위한 수단으로써의 활용도 연구되고 있음
 - 행정자료 및 다양한 자료원을 활용 및 연계하여 추가적인 통계자료를 생성하고 이 과정에서 발생하는 무응답이나 자료손실 부분에 대해 다양하게 연구되고 있었음
 - 다양한 조사방법론에 대한 이론들을 실무에 적용하여 국가통계생산의 품질향상개선에 노력하고 있는 점이 인상적이었고, 특히 시대의 흐름에 따른 행정자료 활용이나 연계부분에 대한 내용들이 많았음
- 행정자료의 통계자료로써의 활용이 중요해지면서 이와 관련한 데이터 매칭, 무응답 대체방법 등에 대한 연구도 진행되고 있음
 - ⇒ 또한, 행정자료의 중요도가 높아짐에 따라 행정자료 수집 및 관리 단계에서의 에디팅 기법, 다른 행정자료 또는 조사자료와의 연계가 가능하도록 주요항목(연계변수)의 관리 등 행정자료의 품질 제고 방안 연구가 병행되어야 할 것으로 보임
 - ⇒ 자료연계의 경우 행정자료와 조사자료간, 행정자료간 연계는 활발히 진행되고 있는 것과 같이, 조사자료간 연계 및 이에 따른 가중치 산출방법 등도 활발히 연구될 필요성이 있다고 생각됨

□ 조사표 설계 분야는 인터넷 기반 데스크탑을 비롯하여 모바일 기기 (스마트폰)를 이용한 조사표 설계에 대한 연구가 현장에서의 적용을 위해 실질적으로 진행되고 있었음

⇒ 조사환경의 악화와 우리나라의 ICT 발달 등 급격히 변화하는 환경을 고려할 때 조사품질 확보를 위해 인터넷조사 뿐만 아니라 모바일조사(스마트폰 이용)를 위한 사전적인 연구·분석이 필요함

□ 국가통계기관(미국 통계국)과 수준 높은 학계와의 공동연구가 이루어지고 있으며, 국가통계 개발·개선을 위해 통계학적인 연구가 뒷받침되고 있었음

⇒ 향후 학계 및 통계작성기관과의 연구 및 정보 공유를 통해 깊이 있는 연구가 가능하도록 국내 통계 인프라 구축과 협력이 필요하다고 판단됨

붙임 컨퍼런스 세션별 주제

※ 「2015 FCSM Final Program and Abstracts」(컨퍼런스 프로그램 내역 및 발표자료 요약본)(파일첨부)

날짜	세션별 주제
12. 1. (1일차)	<ul style="list-style-type: none"> • Linking administrative records to enhance data • Designing web surveys • Nonresponse adjustment and variance estimation • Imputation and the impact of non-response • Experiences from the national hospital care survey in using administrative claims and electronic health record data • Studying non-probability studies • Improvements in survey estimation methodology • Economic statistics • Linking data to generate new statistics: jobs, patents, and FDI • Evaluation of surveys through pretesting and other cognitive methods • Application of bayesian methods • Adaptive design(1)
12. 2. (2일차)	<ul style="list-style-type: none"> • Imputation, multiple imputation and administrative records • Demonstrations on data visualization and metadata systems • Innovative recruiting and questionnaire testing methods • Time series and forecasting • Methods for missing data imputation • Innovations in sample design • Increasing response rates • Demographics in large-scale federal surveys • Comparing and understanding differences in estimates across surveys • Linguistic and cultural issues in survey design and evaluation research

(계속)

날짜	세션별 주제
12. 2. (2일차)	<ul style="list-style-type: none"> • Reducing measurement error: editing techniques and use of administrative data • Predictive analytics in a range of circumstances • Geospatial techniques • Reducing measurement error: editing techniques and use of administrative data • Respondent contact techniques and response rates • measuring quality, spending and reform • What's new with longitudinal surveys?
12. 3. (3일차)	<ul style="list-style-type: none"> • Challenges and benefits from linking administrative and survey data • Multimode data collection • Frame development and respondent selection • Health insurance estimates: assessing the impact of a redesign • Experiences and evaluations in using record linkage to improve the quality of survey products and data • Mobile survey design • Issues with Disclosure and data confidentiality • Modeling Preliminary estimates • Adaptive design(2) • Wealth data and wealth data quality • Remote access and analysis tools for working with government data