

연구논문

코로나19 전후 조사방식의 변화와 모드효과: KICCE 소비실태조사를 중심으로*

최효미**

이 연구는 코로나19의 확산에 따라 대면 면접조사가 극히 어려워진 상황에서 조사방식의 변경이 데이터의 품질에 어떠한 영향을 미쳤는지를 살펴보는 연구로, ‘KICCE 소비실태조사’ 2차년도(2019년)와 3차년도(2020년) 자료를 활용하여 분석하였다. 다수 조사방식 간의 선택효과를 측정하기 위해서 패널 확률효과 다항로짓 분석을 실시하였으며, 조사방식이 데이터 품질에 미치는 영향을 알아보기 위해 미관측 이질성과 시차효과 통제가 가능한 패널 이원고정효과 모형을 적용하였다.

분석 결과, 혼합조사방식은 응답률 측면에서 긍정적 효과가 있음을 확인할 수 있었다. 한편, 총 자녀 수가 많은 경우, 가구소득이 299만 원 이하인 저소득 가구, 그리고 중소도시에 거주하는 경우에 방문조사 대비 방문+유치조사 참여 확률이 높은 것으로 분석됐다. 마지막으로, 조사방식에 따라 주요 측정변수(총 양육비용)가 체계적 차이를 보이는지 검증해 본 결과에서는 방문조사 대비 방문+유치조사를 한 경우에만 유의미한 양의 영향이 포착되었다. 단 이같은 모드효과는 측정 오차라기보다는 조사방식의 선택효과에 기인한 것으로 보인다. 따라서 총 자녀 수가 많거나 대면 접촉이 어려운 가구의 경우 조사방식을 선택하도록 함으로써 응답편의성을 제고할 수 있을 것으로 판단된다.

주제어: 모드효과, 조사방식 선택, 코로나19와 조사방식 변경

* 이 논문은 2020년 한국 조사연구학회 추계학술대회에서 발표한 논문을 수정, 보완한 것으로, ‘영유아 가구의 소비실태조사 및 양육비용 연구(Ⅱ)와 ‘KICCE 소비실태조사: 양육비용 및 육아서비스 수요 연구(Ⅲ)’에 기초하여 작성되었다. 논문의 수정을 위해 귀중한 조언을 해주신 학술대회 토론자 및 심사위원들께 감사드린다.

** 육아정책연구소 연구위원(hmchoi@kicce.re.kr).

I. 서론

인터넷의 보급 및 기술 발전에 힘입어 조사과정에서 오류를 최소화할 수 있는 과학적 도구의 활용이 쉬워진 데 반해, 개인정보의 보호 및 1인 가구의 증가 등으로 인해 대면조사가 어려워지고 있다. 이에 맞춰 조사방식 또한 다각화되고 있으며, 국내·외 많은 조사들이 하나의 조사 안에서 다양한 조사방식을 접목하여 활용하거나 조사차수가 달라짐에 따라 조사방식을 변경하는 등 조사자료의 품질개선을 위해 노력하고 있다(김유빈 외 2017; 박시내 외 2014; 백지선 외 2017; 장인성 외 2021 등).

혼합조사방식(mixed-mode survey)은 만나기 어려운 응답자들의 조사 참여를 용이하게 하고, 응답 편의 및 표본 편의를 감소시켜 데이터의 품질을 개선하는 긍정적 효과가 있을 수 있으나(Van der Laan, P. & W. van Nunspeet 2009), 반대로 조사방식에 따라 응답값이 차이를 보이는 모드효과(mode effects)를 발생시킬 우려가 높다(Dillman, A. 2009). 모드효과란 조사방식(면접조사, 전화조사, 우편조사, 인터넷조사 등)의 변화로 인해 발생하는 표본이탈, 무응답오차, 측정오차 등을 종합적으로 일컫는 용어로(최효미 외 2012: 27), 각기 다른 조사방식으로 수집된 데이터가 통계적으로 체계적인 차이를 갖게 되는 것을 말한다. 이같은 모드효과는 각각의 조사방식이 가지는 독특한 특성뿐 아니라 응답자가 각 조사방식에 대해 느끼는 인지적 부담의 차이에 따라서도 발생하는 것으로 알려져 있다(박시내 외 2014: 75). 다시 말해, 하나의 조사 안에서 다양한 조사방식이 혼재되어 있는 혼합조사방식의 경우에는 조사방식에 따라 응답값이 체계적인 차이를 보이는 문제가 발생할 우려가 있으며, 이는 조사결과의 신뢰성 등 데이터 품질에 직접적 영향을 미치게 된다.

한편, 전 세계적으로 코로나19가 확산됨에 따라 고강도의 사회적 거리두기가 시행되면서 대면면접 조사는 더욱 어려운 상황에 봉착하였으며, 조사기관들은 비대면 조사방식으로서의 전환 혹은 혼합조사방식의 도입 등을 적극적으로 검토해야 하는 상황에 놓이게 되었다. Carletto, G.(2020 재인용¹⁾)에 따르면, 코로나19로 인해 대면조사로만 조사를 완료한 비중은 2020년 5월 69%였으나 2020년 10월에는 31%까지 급감한 것으로 나타났다. 이때 조사기관 혹은 조사업체들이 코로나19로 인해 어려

1) 원문: UNSD-WB Survey of National Statistical Officer (May 2020).

워진 조사환경에 대응하여 대안적으로 사용한 연구방법 중 가장 큰 비중을 차지하는 것이 조사방식의 변화라 볼 수 있다(Carletto, G. 2020 재인용)²⁾. 또, 방대하고 복잡한 설문문항으로 인해 전화조사 등이 용이하지 않은 가구 패널조사들의 경우에도 코로나19 시기에는 표본응답률 제고를 위해 비대면 조사방식을 도입하여 혼합조사 방식으로 전환한 경우가 많았다(Burton, J. et al. 2020: 234; 장인성 외 2021: 118-120 등). 이때 감염병으로부터 취약한 집단, 예컨대 노약자와 영유아 등의 경우에는 일반적인 가구에 비해 더욱 대면조사가 어려웠을 것으로 예상된다.

이 연구는 ‘KICCE 소비실태조사’의 조사방식이 팬데믹 전후로 대면 면접조사에서 혼합조사방식(대면, 전화, 설문 유치 등)으로 전환되는 과정에서 데이터의 품질에 어떠한 변화가 발생하였는지를 고찰한 것이다. 조사방식에 따른 모드효과를 정확하게 측정하기 위해서는 사전에 엄밀하게 계획된 실험설계 등을 통한 측정이 필요하다(김유빈 외 2017; 백지선·민경아 2016; 백지선 외 2017; 최효미 외 2012). 그러나 코로나19라는 갑작스러운 사회적 충격으로 인해 다수의 조사들이 사전에 새로운 조사방식에 대한 면밀한 검토와 엄격한 실험설계 등을 하지 못한 채 부득이하게 조사방식을 변경해야만 하는 상황에 놓이게 되었다. 이전 조사자료와의 연속성이 중요한 종단면 조사(패널 조사) 자료들도 표본응답률을 제고하기 위해 서둘러 혼합조사방식을 도입하였다(장인성 외 2021; 최효미 외 2020).

실험설계 없이 혼합조사방식으로 조사된 자료의 경우 응답값의 차이가 조사방식의 차이에 따른 영향인지, 응답자 특성에 따른 영향인지를 구분하기가 쉽지 않다. 다만, 종단면 데이터를 활용하게 되면 개체의 독특한 특성(즉 미관측 이질성)을 통제함으로써, 횡단면 데이터에 비해 일치추정량을 도출할 수 있다는 장점이 있다(Wooldridge, J. 2002). 즉, 코로나19와 같은 강력한 외생적 충격이 발생한 상태에서 개체 특성으로 인한 영향을 상당 부분 통제하게 되면, 일종의 자연 실험과 유사한 상황에서의 분석이 가능하다. 또, 종단면 데이터를 활용하는 경우 각기 다른 조사 시점 사이의 변화를 관찰할 수 있다는 장점을 갖는다. 통상 실험설계 자료의 경우 횡단면 자료인 경우가 많아 조사방식에 따른 차이를 엄격하게 분리할 수 있다는 장점은 있지만, 조사방식의 변화에 따른 영향은 측정하기 어렵다는 한계가 있다.

이에 이 논문은 코로나19를 전후하여 조사방식에 급격한 변화(대면 면접조사에서 혼합조사방식으로 변화)가 발생했으나 패널조사의 성격을 가지고 있어서 종단면

2) 조사방식의 변경 이외에 코로나 관련 변수 추가, 설문 문항 축소, 표본 축소, 표본설계 변경 등의 순으로 응답되었다(Carletto, G. 2020).

분석이 가능한 육아정책연구소의 ‘KICCE 소비실태조사’를 분석자료로 활용하였다.³⁾ ‘KICCE 소비실태조사’는 종단 분석이 가능하다는 장점 외에도 응답자의 특성이 다른 가구 패널조사들에 비해 매우 동질적이라는 장점이 있다. ‘KICCE 소비실태조사’는 조사대상을 어린 자녀가 있는 육아 가구로 국한하고 있다. 즉, ‘KICCE 소비실태조사’의 조사대상은 감염병 확산 상황에서 대면 면접조사에 특히 민감하게 반응하는 집단이면서, 매년 반복조사에 참여함으로써 설문문항에 대한 기본적인 이해도가 유사하고, 가구 특성(예: 부모 연령 등)의 편차가 크지 않은 편이다.

따라서, 실험설계 상황은 아니지만 ‘KICCE 소비실태조사’ 자료를 종단면으로 구성하여, 동일한 사람이 서로 다른 조사방식으로 동일한 설문에 응답했을 때의 응답 차이를 살펴봄으로써, 실험설계와 유사한 상황에서의 모드효과를 측정해 보고자 하였다. 이때, 두 자료 간에는 1년이라는 시차가 존재하고, 코로나19 확산 이후 조사된 데이터의 응답값은 코로나19로 인한 실질적인 변동이 포함되어 있을 가능성이 크므로, 패널 분석 방법론을 적용하여 코로나19로 인한 변동(시차효과)을 분리함으로써, 조사방식에 따른 모드효과를 좀 더 엄격하게 추정해 보고자 하였다.

이 연구의 주요 연구내용은 다음 3가지로 요약된다.

연구내용 1: ‘KICCE 소비실태조사’의 조사방식 변화는 응답률에 어떠한 영향을 미쳤는가?

연구내용 2: 코로나19로 인해 조사방식이 다양화된 상황(혼합조사방식 도입)에서 육아 가구들의 조사방식 선택에 영향을 미친 요인은 무엇인가?

연구내용 3: 주요 변수(총 양육비용)의 응답값이 종단면 자료로서 안정적으로 구축되었는가? 코로나19로 인한 영향을 통제한 상태에서도 조사방식의 차이에 따른 영향(모드효과)이 발견되는가?

이를 통해 ‘KICCE 소비실태조사’ 데이터의 종단면 자료로서의 안정성과 신뢰성을 점검하고, 코로나19가 종식된 이후에도 육아 가구 대상의 조사에서 유용하게 적용될 수 있는 조사방식이 무엇인지를 모색해 보았다.

3) 분석자료인 ‘KICCE 소비실태조사’에 대한 구체적인 설명은 3장의 내용을 참조하기 바란다.

II. 선행 연구

선행연구는 크게 혼합조사방식으로 조사된 데이터의 모드효과에 관한 연구와 코로나19 전후로 혼합조사방식을 도입한 연구로 나누어 살펴보았다.

1. 혼합조사방식과 모드효과

혼합조사방식의 모드효과에 대해 살펴보기에 앞서, 다양한 조사방식과 조사방식별 장·단점을 간략히 정리해 보았다. 조사방식은 설문지 형태에 따라 종이조사와 전자조사로 구분할 수 있으며, 면접 방식에 따라 대면조사와 비대면조사로 구분할 수 있다. 이때 비대면조사는 다시 전화조사, 우편조사, 인터넷조사 등으로 구분된다. 한편, 응답값 기입 방식에 따라 면접원 기입식과 응답자의 자기기입식으로 구분되며, 조사문항을 청각으로 듣는 방식과 시각으로 보는 방식에 따라서도 구분된다(박시내 외 2014: 75-76 재인용⁴⁾).

각 조사방식별 장·단점은 <표 1>에 제시하였다. 조사방식별로 각각의 장점과 단점이 명확하기 때문에, 각 조사방식이 가진 장점을 결합한 혼합조사방식이 널리 사용되게 된 것이다. 그러나, 혼합조사방식을 적용할 경우 조사과정에서 조사방식에 따른 장점뿐 아니라 단점도 복합적으로 섞여서 발생하게 되며, 이로 인해 조사방식에 따른 모드효과 즉 비표본오차가 커질 위험이 높다. 따라서, 혼합조사방식을 적용한 조사들은 혼합조사방식이 조사과정에서 응답률 제고 등에 효과적으로 작용하였는지, 서로 다른 조사방식 사이에서 응답값의 체계적 차이가 발생하지는 않았는지 등을 점검하여, 데이터의 품질을 제고하기 위한 노력이 요구된다.

‘KICCE 소비실태조사’의 경우 코로나19 확산 이전에는 태블릿 PC를 활용한 대면 면접조사 방식으로 조사가 진행되었는데, 코로나19 확산 이후 대면 면접조사가 어려워짐에 따라 기존의 대면 면접조사와 함께 설문의 유치(자기기입식)와 전화 조사 등을 병행하여 조사하였다. ‘KICCE 소비실태조사’는 정부 지원금 수혜 및 각종 양육비용, 육아서비스 활용 실태 등을 조사하고 있어, 설문의 구조에 익숙하지 않을

4) 원문: Tourangeau, R., L.J. Rips, and K. Rasinski. 2000. *The Psychology of Survey Response*. Cambridge, England; Cambridge University Press.

경우 응답 오류가 발생할 위험이 높다고 판단되어, TAPI를 활용한 대면 면접조사가 조사가 수행되고 있었다. 그러나, 코로나19 이후 대면 면접조사가 어려워진 상황에서 ‘KICCE 소비실태조사’는 전년도 표본에 대한 추적조사를 실시하기 때문에, 표본 추적률이 매우 중요한 상황으로 전화조사 및 설문 유치 등 혼합조사방식을 도입하여 조사하였다. 이 과정에서 설문의 유치가 용이한 조사방식인 PAPI(종이설문)로 조사도구가 전면 변경되었다.

<표 1> 조사방식별 장·단점

조사방식	장점	단점
대면 면접조사	✓ 표본 대표성 확보 용이	✓ 대면 접촉이 어려운 집단(예: 청년, 직장인 등)에 대한 조사가 어려움
웹조사 (비대면)	✓ 대면 접촉이 어려운 집단에 조사 용이 ✓ 조사자료 처리가 신속	✓ 조사 참여 독려가 어렵고 응답시간이 길어질 수 있음 ✓ 인터넷 접근성에 따라 응답률 차이가 큼
자기 기입식 조사	✓ 응답자가 원하는 시간에 조사하는 등 응답 편의성 및 응답률 제고 가능	✓ 설문문항에 대한 이해도가 낮아 부정확한 응답 가능성이 높음
전화조사	✓ 비대면조사와 면접조사의 장점을 동시에 가짐 ✓ 응답률을 제고하면서도 응답값의 신뢰도가 높음	✓ 설문이 방대한 경우 장시간 청각에 의존한 조사로 인해 응답 피로도가 높음 ✓ 응답값의 정확도가 낮아질 우려가 큼
종이설문	✓ 컴퓨터를 이용한 전자조사에 비해 휴대성이 용이 ✓ 초기 비용이 저렴	✓ 복잡한 설문구조로 인한 응답 오류 위험이 커짐 ✓ 숙련된 면접원 필요, 조사자료 처리과정이 길
전자조사	✓ 기계적으로 응답 누락 등을 감소시키고, 응답값의 신뢰성 제고 가능 ✓ 자료 처리가 신속	✓ 초기 구축 비용이 큼 ✓ 응답자가 전자조사에 부정적 태도를 보일 수도 있음

자료: 김유빈 외(2017), 박시내 외(2014), 최효미 외(2012) 등을 참조하여 정리함.

혼합조사방식에 따른 모드효과에 관한 주요 선행연구의 결과는 <표 2>에 제시하였다. 지면 관계상 각 선행연구에 대한 구체적인 설명은 생략한다. 선행연구의 결과를 종합하여 요약하면, 조사방식에 따른 모드효과는 크게 조사과정 상에서의 차이(표본응답률, 항목 무응답, 응답시간 등)와 데이터 품질에 미치는 영향(주요 응답값 비교)을 검증하는 방식으로 구분된다. 조사방식에 따른 모드효과는 각 조사별로, 그

리고 연구방법에 따라 제각각 다른 결과를 보여주고 있다. 이를 통해 각각의 상황에 맞는 조사방식의 도입 필요성을 엿볼 수 있으며, 조사방식에 따른 모드효과가 발생한 경우 데이터 품질 제고를 위한 사후적 조치가 반드시 요구됨을 확인할 수 있었다. 이 연구 또한 코로나19 전후 조사방식의 변경이 데이터의 품질에 어떠한 영향을 미쳤는지 살펴보고, 향후 보다 원활한 조사를 위한 방안을 모색해 보고자 한다.

<표 2> 혼합조사방식과 모드효과

선행연구	조사명 (분석자료)	조사방식	주요결과
김유빈 외 (2017)	한국노동패널조사	대면 면접조사, 웹조사(자기기입식): 1차 실험설계	✓ CAWI는 CAPI에 비해 초기 진입은 쉽게 하나 조사를 끝까지 완료하는 비중이 매우 낮음. ✓ 항목무응답률, 분포의 차이, 항목 무응답률 등에서도 모드효과가 발견됨.
김유빈 외 (2018)	한국노동패널조사	대면 면접조사, 웹조사(자기기입식): 2차 실험설계	✓ CAPI와 CAWI 사이에 모드효과가 크게 감소, 조사도구에 대한 학습효과가 있음. ✓ 여전히 몇몇 민감한 내용을 묻는 문항에 대해서는 모드효과가 발견됨.
박시내 외 (2014)	경제활동인구조사	종이조사, CAPI, CATI, CASI 등	✓ CASI 응답률이 다른 조사와 상당한 차이를 보임. ✓ 단, 가중치 조정을 통해 모드효과를 상쇄시킬 수 있음.
탁현우 외 (2019)	공무원 인식조사	종이설문, TAPI	✓ 조사도구에 따른 응답결과의 차이는 발견되지 않음.
Denscombe, M. (2006)	Perception of Social Issue(PSI) 프로젝트에서 실시한 조사 자료	웹설문, 종이설문	✓ 웹설문이 종이설문에 비해 응답률과 항목 무응답률이 낮음. ✓ 응답값은 조사방식에 따라 큰 차이가 없음.
Mindell, S. et al. (2010)	‘Health Survey for England(HSE)’의 런던지역 부스터 표본	방문 면접조사, 설문 유치 (자기기입방식)	✓ 가구응답률은 부스터표본(유치설문)이 높음. ✓ 개인응답률은 방문 면접조사가 높음. ✓ 표본 특성은 거의 유사. ✓ 단, 교육수준에 따른 차이는 가중치로도 보정되지 않음.

2. 코로나19로 인한 조사방식 변화와 모드효과

한편, 코로나19의 확산에 따라 사회적 거리두기가 강화되면서 사회 조사들은 다양한 도전에 맞닥뜨리게 되었으며 조사방식에 있어서도 큰 변화가 발생하였다. 상대적으로 조사방식의 변경에 신중할 수밖에 없는 종단조사의 경우에도 예외는 아니어서, 국내·외 주요 가구 패널조사의 경우에도 서둘러 혼합조사방식을 도입한 사례를 쉽게 찾아볼 수 있었다. 그러나, 현재까지 진행된 코로나19 확산 전후의 모드효과에 관한 연구는 조사과정에서의 결과(표본응답율, 항목 무응답, 응답자 특성 등)에 국한된 경향을 보인다. 한편, 상대적으로 빠르게 조사방식 변경에 대한 데이터 품질 점검을 실시한 국내 두 개 패널조사(한국노동패널조사와 사업체패널조사)의 경우에는 패널 개체 특성을 통제한 종단분석을 실시하기는 했으나, 분석모형상에서 코로나19가 미친 영향을 통제하지 않아 조사방식에 따른 효과와 코로나19의 영향이 혼재되어 있는 한계를 보였다. 코로나19로 인한 조사방식 변화와 모드효과에 관한 선행연구의 주요 내용과 결과는 <표 3>에 정리하여 제시하였다. 이 연구는 선행연구의 경향을 따라가되, 진일보한 분석방법을 적용함으로써 조사방식에 따른 모드효과를 좀 더 정교하고 엄밀하게 추출하고자 노력하였다.

<표 3> 코로나19로 인한 조사방식 변화와 모드효과

선행연구	조사명 (분석자료)	조사방식 변경	주요결과
장인성 외 (2021)	한국노동 패널조사	대면 면접조사 위주 → 웹조사, 전화조사 확대	✓ 주관적 판단을 요하는 문항에서는 조사 방식에 따른 차이가 큼 ✓ 객관적 판단이 가능한 문항에서는 모드 효과가 크지 않음. ✓ 웹조사와 전화조사의 응답행태는 매우 유사하나, 민감한 문항에서는 웹조사의 응답값이 높게 나타남.
장인성 외 (2021)	사업체 패널조사	대면 면접조사 위주 → 웹조사 혼합	✓ CAWI가 CAPI에 비해 응답률이 낮고, 응답기간이 김 ✓ 무응답률 등 데이터 품질은 조사방법에 따라 큰 차이를 보이지 않음.
Burton, J. (2020)	‘Understanding Society’	대면 가정 방문조사 → 웹조사 혼합	✓ 응답률 측면에서 긍정적 효과
ESCAP (2021)	파푸아 뉴기니의 ‘인구통계 및 건강조사’	코로나19 이후 전화조사(휴대전화)	✓ 코로나19 전후 응답자 특성이 다름

III. 분석자료 및 분석방법

1. 분석자료의 특징

이 논문은 육아정책연구소의 ‘KICCE 소비실태조사’ 2019년(2차년도)조사와 2020년(3차년도)조사 원자료를 활용하여, 코로나19 전후 조사방식 변경과정에서 발견된 모드효과를 점검해 본 연구이다. ‘KICCE 소비실태조사’는 2018년부터 매해 육아 가구의 양육비용과 육아서비스 이용 실태 등을 조사하는 자료로, 전년도 조사표본에 대한 추적조사⁵⁾ 형태로 조사가 이뤄진다. 이때 인접한 두 개의 차수 간에는 상당수 표본이 추적조사에 성공하기 때문에 패널자료로 활용이 가능하다. 실제 이 논문에 활용된 분석자료의 경우, 2019년(2차년도) 조사 가구 1,902 가구 중 2020년(3차년도)에 추적조사에 성공한 가구는 1,634가구로 표본추적률은 85.9%에 달한다.

‘KICCE 소비실태조사’는 코로나19가 확산되기 이전인 2019년(2차년도)까지는 태블릿PC를 활용한 대면 면접조사 방식(TAPI)으로 조사되었으나, 2020년(3차년도) 조사부터는 코로나19로 인해 조사도구를 종이설문(PAPI)으로 변경하고 비대면 조사방식(전화, 설문유치 등)을 도입하여 혼합조사방식으로 조사를 실시하였다(최효미 외 2020: 68-69). 2020년(3차년도)의 조사방식은 응답자와의 첫 접촉시⁶⁾에는 가능한 기존과 동일한 대면 면접조사 방식으로 안내하되, 응답자가 대면조사에 대한 어려움을 토로하는 경우 응답자가 선호하는 조사방식을 선택할 수 있도록 하였다. 또, 2020년(3차년도) 조사는 종이설문으로 조사가 되었으나, 조사방식 변경에 따른 응답값 입력오류를 최소화하기 위하여 기존 태블릿PC의 입력시스템과 유사한 전산 시스템을 활용하여 조사자료를 더블 편칭하도록 하고, 항목 무응답 및 응답 오류가 발생한 경우 재조사를 실시하는 등의 과정을 거쳤다(최효미 외 2020: 68-69). 결과적으로 ‘KICCE 소비실태조사’의 경우 모든 설문문항에 대해 항목 무응답 결측이

-
- 5) ‘KICCE 소비실태조사’의 조사 대상은 막내자녀 연령을 기준으로 영유아기 및 초등저학년 자녀가 있는 육아 가구(매해 약간씩 상이)로, 매해 신생아 가구 표본을 새롭게 조사하고, 표본 대표성 확보를 위해 추적조사에 실패한 가구에 대한 대체표본 구축이 매해 이뤄지기 때문에, 1차년도 표본이 아닌 직전년도 표본에 대한 추적조사를 실시하는 형태로 되어 있다(최효미 외 2021: 85).
- 6) 3차년도 조사 이전에도 첫 접촉은 대부분 방문 일정 등을 상의하기 위해 휴대전화 혹은 이메일 등을 통해 사전 접촉을 하는 경우임.

거의 발생하지 않은 상태로, 조사방식에 따른 항목 무응답율의 차이는 발생하지 않는 구조이다.

이 논문은 앞서 서론에서 언급한 바와 같이 동일한 개인이 동일한 설문문항에 대해 1년의 시차를 두고 조사방식 선택이 가능해진 상황에서, 어떠한 응답자 특성이 조사방식 선택에 영향을 미치고 있는지를 확인하고, 데이터의 품질이 조사방식에 따라 체계적 차이를 갖지는 않는지 점검하는 것에 목적이 있다. 이 과정에서 패널 개체의 특성과 코로나19로 인한 시차효과를 엄격히 통제하여, 일치추정량을 도출하고자 하였다. 따라서, 분석대상으로 팬데믹 전후 시기의 2개년 모두 조사에 응답한 가구 즉 추적조사에 성공한 가구인 1,634가구를 분석대상으로 한다.⁷⁾

분석에 활용된 자료의 기초통계량은 <표 4>와 같다. 분석에 활용된 분석대상은 균형패널로 2차년도와 3차년도 조사 모두에 응답한 가구들이므로, 2개 차수 간 가구 특성의 변화가 크지 않고 안정적임을 확인할 수 있었다. 각 연도 1,634가구씩 총 3,268가구가 분석에 활용되었으며, 외별이가구가 맞별이가구에 비해 약 10%포인트 가량 많으며, 자녀가 2명인 가구가 가장 많았다. 한편 지역규모별로는 두 개 연도 사이에 거의 차이가 없으며, 막내자녀 연령별로도 특정 연령의 조사비중이 낮아지거나 하는 문제는 발견되지 않았다.

가구소득⁸⁾은 2019년(2차년도)에 월평균 491만9천 원이었으나, 2020년(3차년도)에는 466만9천 원에 그쳤다. 그러나, 소득구간별로 299만 원 이하 저소득 가구의 비중은 2019년(2차년도)가 2020년(3차년도)에 비해 높았다. 표로 제시하지는 않았으나, 2019년(2차년도)의 가구소득은 중위값 440만 원, 표준편차 267.96, 최소값 10만원, 최대값 4,490만 원이었으며, 2020년(3차년도)에는 중위값 430만 원 표준편차 194.72, 최소값 0원, 최대값 2,420만 원으로, 2019년(2차년도)가 2020년(3차년도)에

7) 조사에 응답한 가구(추적표본)와 거절한 가구(혹은 대체표본) 사이에 체계적인 표본이탈이 있는지, 이로 인해 비표본 오차가 발생했는지 등은 이 연구의 범위를 넘어서는 것으로 생략한다. 단, 2019년(2차년도) 응답 가구 중 2020년(3차년도) 조사 응답 여부는 로짓분석 결과 총 자녀 수에 따라서만 약간의 차이를 보일 뿐 맞별이가구 여부, 가구소득, 지역규모 등 다른 가구 특성에 따라서는 유의미한 차이가 발견되지 않았다. 한편, 대체표본과 신규표본은 해당 조사에 처음 참여하는 가구들로, 설문문항에 대한 이해도 등이 기존 가구와 차이가 날 것을 우려하여 조사 초기 면접가이드에 가능한 방문 면담조사로 진행하도록 되어 있었다. 결과적으로 기존 표본 가구에 비해 이들 가구의 방문 조사 비중이 높은 특징을 보인다(최효미 2020: 12).

8) 본인과 배우자, 기타 가구원의 근로소득 및 사업소득, 부동산소득, 사회보험소득, 이전소득을 합산한 값이며, 기타소득 및 지자체 긴급지원금은 한시적 지원에 해당하므로 제외하였다.

비해 가구소득이 약간 더 넓게 분포된 경향을 보이는 것으로 나타났다.

총 양육비용은 ‘KICCE 소비실태조사’ 가구용 설문문의 주요 변수로서, 개별 자녀별로 16개 비목(식비, 피복비, 교육/보육비 등, 세목 포함)에 응답한 값을 합산(=아동당 양육비용)한 다음, 가구 내 모든 자녀(초등학교 1학년 이상 자녀 포함)의 아동당 양육비용을 합산하여 산출한 값이다. 즉, 총 양육비용은 자녀가 1명인 가구의 경우 총 16개(세목포함) 응답값의 합산치이고, 자녀가 2명인 경우 32개 응답값의 합산치이다. 총 양육비용은 ‘KICCE 소비실태조사’를 통해 생산되는 가장 핵심적인 주요 변수일 뿐 아니라, 가구용 데이터의 절반 이상의 변수를 활용하여 산출되는 변수이다.

이 연구가 조사방식이 데이터의 품질에 미친 영향을 점검하기 위해 총 양육비용에 주목하는 이유는 앞서 설명한 바와 같이 총 양육비용이 ‘KICCE 소비실태조사’ 가구용 설문문의 절반 가량을 차지하는 변수를 활용하여 산출되는 변수라는 점 때문이다. 총 양육비용은 다수의 가계지출 현황을 자녀별 및 비목별로 세세하게 나누어 응답해야 하는 문항으로, 조사방식 변경에 따라 응답값의 측정오차가 커질 위험이 높은 변수이다. 다시 말해, 청각에 의존하여 조사가 실시되는 전화조사의 경우 응답자들이 세분된 양육비용의 구분점을 정확하게 이해하고 응답할 수 있을 것인지, 자기입력으로 응답되는 유치설문의 경우 응답자들이 비목의 분류를 면접원만큼 정확하게 할 수 있을 것인지 등이 우려되는 상황이었기 때문이다. 그런데 횡단 분석을 통해서는 이러한 변동이 조사방식 변경에 따른 모드효과인지, 코로나19로 인해 가구 상황이 급변하여 생긴 실질적인 양육비용 변화인지, 조사표본의 가구특성 변화에 기인한 변동인지를 구분하기 어렵다. 따라서, 패널분석 모형을 활용하여 코로나19로 인한 시차효과와 개체의 고유한 특성을 통제한 상태에서 조사방식에 따른 차이를 식별하고자 하였다.

2019년(2차년도)에 월평균 총 양육비용은 127만7천 원이었으며, 2020년(3차년도)에는 약간 하락하여 126만9천 원인 것으로 조사됐다. 2019년(2차년도) 총 양육비용의 중위값은 114만 원, 표준편차 72.87, 최소값 3만 원 최대값 654만 원이었으며, 2020년(3차년도)의 총 양육비용은 중위값 113만 원, 표준편차 70.08, 최소값 13만 원, 최대값 608만 원이었다.

<표 4> 기초통계량_가구특성

(단위 : 가구, %, 명, 만원, 세)

		2019년 (2차년도)		2020년 (3차년도)	
		빈도	비중	빈도	비중
전 체		1,634	100.0	1,634	100.0
맞벌이 여부	맞벌이 가구	721	44.1	741	45.4
	외벌이 가구	913	55.9	893	54.6
총 자녀 수	1명	541	33.1	490	30.0
	2명	865	52.9	896	54.8
	3명 이상	228	14.0	248	15.2
	평균(명)	1.8		1.9	
가구소득	299만 원 이하	235	14.4	113	6.9
	300-399만 원 이하	459	28.1	357	21.9
	400-499만 원 이하	424	26.0	462	28.3
	500-599만 원 이하	252	15.4	304	18.6
	600만 원 이상	264	16.2	398	24.4
	평균(만 원)	491.9		466.9	
지역규모	대도시	663	40.6	664	40.6
	중소도시	714	43.7	714	43.7
	읍·면지역	257	15.7	256	15.7
막내자녀 연령	0세	139	8.5	27	1.7
	1세	250	15.3	165	10.1
	2세	215	13.2	247	15.1
	3세	197	12.1	196	12.0
	4세	217	13.3	188	11.5
	5세	199	12.2	211	12.9
	6세	220	13.5	194	11.9
	7세	197	12.1	216	13.2
	8세	-	-	190	11.6
	평균(세)	4.6		4.4	
총 양육비용 평균(만 원)		127.7		126.9	

주: 1) 막내자녀 연령은 출생 당해년도를 0세, 출생 차년도를 1세로 표기함. 2020년(3차년도) 조사에서 막내자녀의 출생년도가 2019년인 경우 1세로 표기하였으며, 2013년생이 7세로 초등1학년을 의미함.

2) 2019년(2차년도) 조사 당시 막내자녀 연령이 8세(2011년생)인 가구는 조사 대상이 아님.

코로나19로 인해 대면 면접조사가 어려워지자 ‘KICCE 소비실태조사’는 2020년(3차년도)에 조사방식을 다양화하였으며, 결과적으로 추적조사에 성공한 1,634가구 중 667가구(40.8%)만이 전체 설문을 방문조사로 응답하였다. 다음으로 방문하여 일부 설문을 작성하고 이후 나머지 설문을 조사가구에 유치한 후(자기기입식) 수거하는 형태로 조사가 실시된 방문+유치 조사가 391가구(23.9%)였으며, 처음부터 끝까지 전화 면접을 통해 조사가 실시된 전화조사가 260가구(15.9%), 전체 설문을 가구에 유치한 후 자기기입하도록 한 후 수거한 유치조사가 237가구(14.5%)였다.

<표 5>조사방식(2020년, 3차년도)

(단위: 가구, %)

	전체	방문	전화	유치	방문+유치	기타
빈도	1,634	667	260	237	391	79
비중	100.0	40.8	15.9	14.5	23.9	4.8

주: 기타는 방문+전화, 전화+유치, 방문+전화+유치 등임.

2. 분석 방법⁹⁾

이 연구는 3가지 연구문제에 대한 검증으로 구성되는데, 첫째 조사방식이 변화된 상황에서 조사주간별 표본응답률이 어떻게 달라졌는가? 둘째, 영유아 가구들이 선호하는 조사방식은 어떤 것이며 어떠한 가구특성이 조사방식 선택에 영향을 미쳤는가? 셋째, 데이터 품질 점검 차원에서 실시된 분석으로 ‘KICCE 소비실태조사’의 주요 조사변수인 총 양육비용이 조사방식에 따라 체계적인 차이를 보이는가? 하는 것이다. 조사방식에 따른 모드효과를 검증하는 과정에서 주로 활용되는 항목 무응답률(문항별 결측)의 경우, ‘KICCE 소비실태조사’는 조사과정에서 항목 무응답이 발생할 경우 재조사를 통해 모든 문항의 응답값을 확인하도록 하고 있어, 실제로 결측치가 거의 존재하지 않으므로 별도의 분석은 생략하였다.

조사방식 선택에 영향을 미친 요인에 관한 분석은 차이 검정과 패널 확률효과 다항로지분석을 실시하였으며, 총 양육비용이 조사방식에 따라 체계적 차이를 보이는

9) 패널 분석 방법에 대한 논의는 Wooldridge, J.(2002), 민인식·최필선(2010), 민인식·최필선(2012)를 참고하였다.

지를 살펴보기 위해서는 차이 검정과 패널 이원고정효과 모형(two-way fixed effects)을 적용하였다. 패널 확률효과 다항로짓분석 모형의 경우 이 연구와 같이 종속변수가 비순서형 선택이면서 이항이 아닌 여러 개의 항인 경우에 적용 가능하다. 일반적인 다항로짓분석 모형에 패널(응답자)의 개체 특성을 추가적으로 통제한 모형이라고 볼 수 있다. 한편, 이원고정효과 모형의 경우 종속변수가 연속변수인 경우에 적용한다. 이원고정효과 모형은 일반적인 OLS에서 개체 특성뿐 아니라 시차효과까지를 고려한 모형이다.

IV. 분석결과

1. 코로나19 전후 응답률 비교

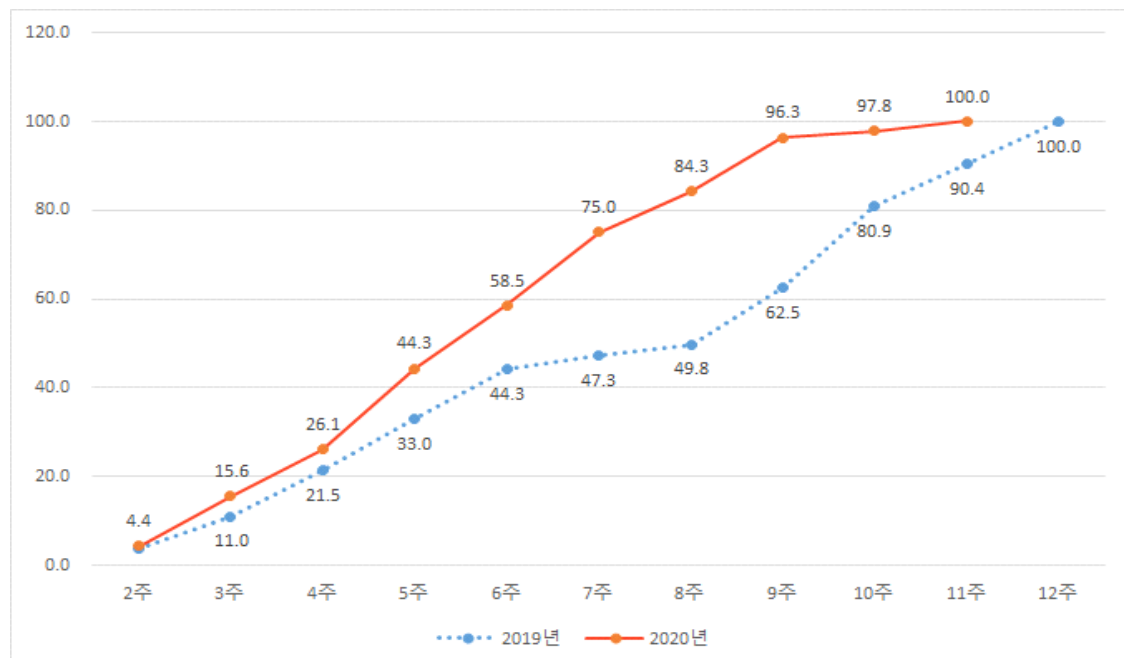
2019년(2차년도)과 2020년(3차년도)의 조사 주차별 조사 진행률을 비교해보면, 코로나19 확산으로 사회적 거리두기가 강화된 2020년(3차년도) 조사가 2019년(2차년도) 조사에 비해 빠른 속도로 진행되었음을 확인할 수 있었다. 심지어 2020년(3차년도) 조사는 조사대상 가구를 막내자녀가 초등학교 2학년인 가구까지로 설정함에 따라 조사대상 가구 수가 2019년(2차년도)에 비해 254가구나 많았음에도 불구하고, 전체 조사 완료까지 걸린 조사기간은 총 11주로 2019년(2차년도) 조사에 비해 1주일 가량 단축된 것으로 나타났다. 2019년(2차년도) 조사의 경우 TAPI를 활용한 대면 조사방식으로 조사가 진행되던데 반해, 2020년(3차년도)의 경우 응답자로 하여금 조사방식을 선택할 수 있도록 함에 따라 전반적인 조사 진행속도가 빨라진 것으로 보여진다.

한편 최효미 외(2019: 23)에 따르면 ‘KICCE 소비실태조사’ 2차년도 조사의 표본 추적률은 75.4%(추적대상 1,648 가구 중 1,243가구)로, 3차년도 조사의 표본추적률 85.9%(추적대상 1,902가구 중 1,634가구)에 비해 10%포인트 가량 낮았다. 이상의 결과를 통해 ‘KICCE 소비실태조사’의 경우 3차년도 조사에서 도입된 비대면 조사 방식(전화, 유치 등)이 응답율 측면에서는 긍정적 효과가 있었음을 유추해 볼 수 있었다.

<표 6> 조사 주차별 조사 진행률

(단위: 가구, %)

주차	2019년			2020년		
	날짜	완료 부수	진행율	날짜	완료 부수	진행율
2주	5/22	70	3.7	5/26	95	4.4
3주	5/30	209	11.0	6/01	336	15.6
4주	6/05	409	21.5	6/09	563	26.1
5주	6/13	627	33.0	6/16	956	44.3
6주	6/20	843	44.3	6/23	1,261	58.5
7주	6/27	900	47.3	6/30	1,616	75.0
8주	7/01	948	49.8	7/07	1,818	84.3
9주	7/11	1,188	62.5	7/14	2,077	96.3
10주	7/20	1,539	80.9	7/21	2,109	97.8
11주	7/25	1,720	90.4	7/31	2,156	100.0
12주	8/01	1,902	100.0	-	-	-



자료: 최효미(2020).

2. 조사방식 선택에 영향을 미친 요인

조사방식이 가구 특성에 따라 차이를 보이는지를 살펴본 결과, 총 자녀 수, 가구 소득, 지역규모 등에서 유의미한 차이가 발견되었다. 먼저 총 자녀 수가 1명인 경우에는 상대적으로 방문조사나 전화조사 비중이 높은 반면, 총 자녀 수가 3명 이상인 경우에는 방문+유치조사와 유치조사의 비중이 높게 나타났다. ‘KICCE 소비실태조사’는 총 자녀 수가 많을수록 응답해야 하는 항목 수가 많아지며 응답시간이 길어지는 특징이 있는데, 설문의 일부 혹은 전부를 유치하는 형태의 조사가 총 자녀 수가 많을수록 높아지는 것은 이에 따른 결과로 보인다.

한편 가구소득은 400~499만 원 이하의 중간소득 계층은 상대적으로 방문조사 비중이 높고, 299만 원 이하 저소득 가구와 600만 원 이상 고소득 가구는 방문+유치조사 비중이 높은 경향을 보였다. 지역규모에 따라서는 읍·면지역에 거주하는 경우 방문조사 비중이 높은 반면, 중소도시에 거주하는 가구들이 방문+유치조사 비중이 높은 편으로 조사됐다. 맞벌이가구 여부나 막내자녀의 연령에 따라서는 유의미한 차이는 보이지 않는 것으로 나타났다.

패널 확률효과 다항로지분석을 통해 조사방식 선택에 영향을 미친 요인을 살펴본 결과는 다음과 같다. 방문조사 대비 전화조사에 참여했을 확률은 가구소득이 높을수록 높아지는 경향을 보였다. 지역규모에 따라서는 중소도시에 거주하는 경우 대도시 거주자에 비해 방문조사보다는 전화조사로 응답할 확률이 유의미하게 높았다. 그러나, 읍·면지역에 거주하는 경우에는 대도시 거주자와 큰 차이를 보이지 않았다. 한편, 초등 이상 자녀가 있는 경우에 방문조사 대비 전화조사 참여확률이 높은 것으로 분석됐다. 반면, 총 양육비용이 높을수록 전화조사보다는 방문조사 참여 확률이 높은 것으로 나타났다.

전체 설문을 유치하여 자기기입하도록 한 유치조사의 경우에는 가구소득은 크게 영향을 미치지 않는 것으로 분석됐다. 반면, 방문조사 대비 유치조사에 참여할 확률은 자녀 수가 많을수록 높아졌으며, 초등 이상 자녀가 있는 경우 유치조사 참여 확률이 높은 것으로 나타났다. 방문조사 대비 유치조사의 경우에는 지역규모도 매우 유의미한 영향을 미친 요인인 것으로 분석됐는데, 중소도시 거주자 및 읍·면지역 거주자가 대도시 거주자에 비해 유치조사 확률이 낮았다. 즉, 대도시 거주자의 대면 면접조사가 더 어려웠다는 것으로, 이는 코로나19 확산 시기 대도시가 중소도시 및 읍·면지역에 비해 사회적 거리두기 단계가 높았던 것에 기인한 결과로 보인다.

<표 7> 가구 특성별 조사방식_2020년

(단위: 가구, %)

		전체	방문	전화	유치	방문+유치	기타	χ^2
전체		1,634 (100.0)	667 (40.8)	260 (15.9)	237 (14.5)	391 (23.9)	79 (4.8)	
맞벌이 여부	맞벌이 가구	741 (100.0)	305 (41.2)	118 (15.9)	105 (14.2)	181 (24.4)	32 (4.3)	1.030
	외벌이 가구	893 (100.0)	362 (40.5)	142 (15.9)	132 (14.8)	210 (23.5)	47 (5.3)	
총 자녀 수	1명	490 (100.0)	219 (44.7)	90 (18.4)	68 (13.9)	94 (19.2)	19 (3.9)	24.723***
	2명	896 (100.0)	360 (40.2)	143 (16.0)	129 (14.4)	224 (25.0)	40 (4.5)	
	3명 이상	248 (100.0)	88 (35.5)	27 (10.9)	40 (16.1)	73 (29.4)	20 (8.1)	
가구 소득	299만 원 이하	113 (100.0)	41 (36.3)	8 (7.1)	25 (22.1)	37 (32.7)	2 (1.8)	29.428**
	300-399만 원 이하	357 (100.0)	151 (42.3)	55 (15.4)	55 (15.4)	84 (23.5)	12 (3.4)	
	400-499만 원 이하	462 (100.0)	202 (43.7)	70 (15.2)	67 (14.5)	96 (20.8)	27 (5.8)	
	500-599만 원 이하	304 (100.0)	120 (39.5)	50 (16.5)	44 (14.5)	74 (24.3)	16 (5.3)	
	600만 원 이상	398 (100.0)	153 (38.4)	77 (19.4)	46 (11.6)	100 (25.1)	22 (5.5)	
지역 규모	대도시	664 (100.0)	291 (43.8)	77 (11.6)	138 (20.8)	130 (19.6)	28 (4.2)	96.439***
	중소도시	714 (100.0)	237 (33.2)	155 (21.7)	82 (11.5)	204 (28.6)	36 (5.0)	
	읍·면지역	256 (100.0)	139 (54.3)	28 (10.9)	17 (6.6)	57 (22.3)	15 (5.9)	
막내 자녀 연령	영아	635 (100.0)	263 (41.4)	105 (16.5)	88 (13.9)	147 (23.2)	32 (5.0)	6.731
	유아	593 (100.0)	22.7 (38.3)	102 (17.2)	88 (14.8)	144 (24.3)	32 (5.4)	
	초등저학년	406 (100.0)	177 (43.6)	53 (13.1)	61 (15.0)	100 (24.6)	15 (3.7)	

주: 1) * $p < .1$, ** $p < .05$, *** $p < .01$.

2) 영아는 0~3세, 유아 4~6세, 초등 7~8세를 의미함.

3) 기타는 방문+전화, 전화+유치, 방문+전화+유치 등임.

<표 8> 조사방식 선택에 영향을 미친 요인_패널 확률효과 다항로지

		전화	유치	방문+유치	기타
맞벌이 여부(맞벌이=1)		-0.231	0.070	-0.005	-0.445*
총 자녀 수	2명	0.116	0.287*	0.359**	0.258
	3명 이상	-0.070	0.517**	0.528***	0.937**
가구소득	300-399만 원 이하	0.759*	-0.442*	-0.452**	0.605
	400-499만 원 이하	1.071***	-0.289	-0.416*	1.407*
	500-599만 원 이하	1.187***	-0.328	-0.342	1.421*
	600만 원 이상	1.571***	-0.479	-0.280	1.656*
지역규모	중소도시	0.668***	-0.523***	0.465***	0.215
	읍·면지역	-0.105	-1.238***	0.046	0.265
막내자녀 연령	유아	0.211	0.222	0.158	0.235
	초등저학년	0.461**	0.719***	0.644***	0.374
총 양육비용		-0.004***	-0.001	0.002*	-0.002
상수		-3.255***	-1.974***	-2.336***	-4.818***

주 : 1) * $p < .1$, ** $p < .05$, *** $p < .01$.

2) 영아는 0~3세, 유아 4~6세, 초등 7~8세를 의미함.

3) 종속변수 base는 방문조사이며, 독립변수는 총 자녀 수는 1명, 가구소득 299만 원 이하, 지역규모는 대도시, 막내자녀 연령은 영아가 base임.

4) 기타는 방문+전화, 전화+유치, 방문+전화+유치 등임.

한편 방문조사 대비 방문+유치 조사가 이뤄졌을 확률은 자녀 수가 많을수록, 중소도시에 거주하는 경우, 초등자녀가 있는 경우, 총 양육비용이 높을수록 높은 것으로 분석됐다. 또, 가구소득이 300~399만 원 이하인 가구와 400~499만 원 이하인 가구의 경우는 가구소득이 299만 원 이하인 가구에 비해 방문조사 대비 방문+유치 조사 확률이 유의미하게 낮은 것으로 나타났다. 즉, 299만 원 이하 저소득 가구가 중위소득 가구에 비해 방문+유치조사 확률이 높다고 볼 수 있다.

이 외에 기타 조사방식(방문+전화, 전화+유치, 방문+전화+유치 등)으로 조사가 이뤄진 경우는 다른 조사방식과 달리 맞벌이 가구 여부가 유의미한 효과를 보였다. 외벌이 가구의 경우 맞벌이 가구에 비해 기타 조사방식으로 조사에 참여했을 확률

이 높은 것으로 분석됐다. 이는 외별이 가구가 맞별이 가구에 비해 여러 번 접촉하는 것에 좀 더 허용적임을 반증하는 결과라 볼 수 있다. 또, 자녀가 3명 이상인 경우, 가구소득이 높을수록 기타 조사방식으로 조사했을 확률이 높았다.

정리하면, 총 자녀 수가 많을수록 방문조사 대비 유치조사 혹은 방문+유치조사, 기타 조사방식으로 응답했을 가능성이 높았는데, 이는 자녀가 많을수록 응답문항이 많아지기 때문에 설문을 유치하여 조사하기를 희망하는 비중이 높기 때문으로 사료된다. 한편 자녀 중 초등자녀가 있는 경우에도 방문조사 확률이 낮아지는 경향을 보였으며, 가구소득은 높을수록 전화조사 혹은 기타 조사방식으로 응답할 확률이 높았다. 반대로 저소득 가구에 비해 중위소득 가구들은 유치조사 혹은 방문+유치조사 확률이 낮았다. 이는 저소득 가구들이 면접원이 자신의 거주지로 방문하여 조사하는 것을 더 꺼려하는 현상이 반영된 결과로 추측된다. 한편, 대도시 거주자에 비해 중소도시 거주자의 경우 방문조사 대비 전화조사와 방문+유치조사 참여 확률은 높은 반면, 설문 전체 유치조사 확률은 낮은 것으로 나타났다.

3. 조사방식이 총 양육비용에 미친 영향

이 절은 ‘KICCE 소비실태조사’를 통해 생산되는 주요 변수인 총 양육비용이 조사방식에 따라 체계적인 차이를 보이는지 점검한 결과를 제시하였다.

패널 분석에 앞서 먼저 총 양육비용이 조사방식과 가구특성에 따라 어떠한 차이를 보이는지 차이검정을 통해 살펴보았다. 2020년(3차년도) 총 양육비용은 평균 126만9천 원이며, 방문+유치조사를 한 경우에 총 양육비용이 142만3천 원으로 가장 높았다. 다음으로 기타 조사방식으로 응답된 경우에 월평균 135만7천 원, 유치조사 128만3천 원, 방문조사 119만9천 원, 전화조사 117만6천 원 순으로, 조사방식에 따라 총 양육비용이 상당한 차이를 보였다. 가구특성에 따라서는 맞별이 가구의 총 양육비용이 외별이 가구에 비해 높으며, 자녀 수가 많아질수록, 가구소득이 높을수록, 막내자녀의 연령이 많을수록 총 양육비용도 높아지는 경향을 띤다. 이러한 경향성은 다수의 양육비용 연구에서 나타나는 일반적인 현상이라 볼 수 있다.

한편 동일한 특성을 지닌 가구들 사이에서 조사방식에 따라 총 양육비용이 유의미한 차이를 보이는 경우는 다음과 같다. 먼저 맞별이 가구이든 외별이 가구이든 다른 조사방식에 비해 방문+유치조사를 한 경우에 총 양육비용이 높은 특징을 보였다. 또, 총 자녀 수가 2명인 경우에 조사방식에 따라 총 양육비용이 유의미한 차이

를 보이는 것으로 분석됐다. 동일한 소득구간 내에서는 방문+유치조사에 참여한 가구의 총 양육비용이 다른 조사방식으로 조사에 참여한 가구에 비해 높았으며, 중소도시나 읍·면지역에 거주하는 가구들의 경우에도 다른 조사방식에 비해 방문+유치조사를 한 경우 총 양육비용이 높은 것으로 분석됐다. 한편, 막내자녀가 영유아인 가구들은 다른 조사방식에 비해 방문조사와 전화조사를 한 경우에 총 양육비용이 유의미하게 낮은 것으로 나타났다.

<표 9> 가구특성 및 조사방식별 총 양육비용 차이_2020년

(단위: 만원)

		전체	방문	전화	유치	방문+유치	기타	F
전체		126.9	119.9	117.6	128.3	142.3	135.7	7.97***
맞벌이 여부	맞벌이 가구	137.7	130.3	125.3	133.1	157.4	157.3	5.52***
	외벌이 가구	117.9	111.1	111.2	124.5	129.2	121.0	3.37***
총 자녀 수	1명	84.2	80.7	85.0	84.3	92.6	77.6	1.32
	2명	136.4	133.3	125.4	136.6	148.2	137.0	3.28**
	3명 이상	176.7	162.5	184.9	176.5	187.8	188.3	1.03
가구 소득	299만 원 이하	91.1	80.7	104.6	96.2	95.3	110.0	0.71
	300-399만 원 이하	100.5	95.3	89.4	107.9	115.0	82.4	3.30**
	400-499만 원 이하	116.5	109.4	103.3	127.5	136.0	107.4	5.04***
	500-599만 원 이하	139.2	136.2	123.3	141.3	147.2	168.1	1.90
	600만 원 이상	163.3	155.8	148.5	159.1	184.9	178.2	2.83**
지역 규모	대도시	125.5	124.0	121.5	122.8	128.6	150.3	1.10
	중소도시	126.5	115.0	113.6	137.0	147.2	116.1	8.48***
	읍·면지역	131.6	119.7	129.1	131.1	155.7	155.5	3.05**
막내 자녀 연령	영아	114.9	106.2	107.5	116.6	132.7	125.9	4.89***
	유아	129.8	123.8	121.4	132.6	141.1	141.5	2.08*
	초등저학년	141.3	135.5	130.3	139.2	158.1	144.2	1.66

주 : 1) * $p < .1$, ** $p < .05$, *** $p < .01$.

2) 영아는 0~3세, 유아 4~6세, 초등 7~8세를 의미함.

3) 기타는 방문+전화, 전화+유치, 방문+전화+유치 등임.

조사방식의 변화가 총 양육비용에 미친 영향을 패널 이원고정효과 모형을 통해 추정해 본 결과, 방문조사 대비 방문+유치조사를 한 경우에만 양의 유의미한 영향이 포착되었다. 즉, 개인 특성 및 시차효과를 통제한 상태에서 방문+유치조사로 조사에 참여한 가구는 다른 조사방식으로 조사한 가구에 비해 총 양육비용을 좀 더 높게 응답한 것으로 분석됐다. 그러나 조사방식에 따른 모드효과는 방문+유치조사에 국한되며, 전화조사 혹은 유치조사의 경우에는 유의미한 차이를 보이지 않는 것으로 나타났다.

한편 코로나19로 인한 총 양육비용의 변동은 10% 유의수준에서 음의 효과가 발견되었는데, 이는 코로나19로 인해 육아가구의 총 양육비용이 소폭 감소하였음을 시사한다. 이는 코로나19로 인해 교육·보육기관의 이용이 용이치 않아, 총 양육비용 중 교육/보육비가 감소한 것에 기인한 것(최효미 외 2021: 115-116)으로 사료된다.

가구특성 중에서는 총 자녀 수가 총 양육비용에 가장 유의미한 차이를 가져온 변수로 분석되었는데, 미관측 이질성과 시차효과를 통제한 상태에서도 총 자녀 수가 많을수록 총 양육비용이 증가하는 양상이 뚜렷하게 관측되었다. 또 가구소득의 경우에는 600만 원 이상인 가구에서만 유의미한 양의 효과가 발견되었다.

앞서 패널 확률효과 다항로짓 분석에서 살펴본 바와 같이 총 자녀 수가 많을수록 방문조사 대비 방문+유치조사 확률이 높았던 것을 고려할 때, 총 자녀 수와 조사방식 사이에 높은 상관성이 존재함을 알 수 있다. 따라서, 조사방식에 따라 총 양육비용의 응답값이 체계적 차이를 보이는지 보다 면밀하게 추정해 보기 위해 자녀 수별로 집단을 구분하여 패널 이원고정효과 모형추정을 실시해 보았다.

분석의 결과, 자녀 수가 2명 혹은 3명 이상인 집단에서는 조사방식에 따라 총 양육비용이 체계적 차이를 보인다는 증거를 찾기 어려웠다. 오히려 자녀가 1명인 집단에서 유치조사를 한 경우 방문조사에 비해 총 양육비용이 유의미하게 낮아지는 현상이 포착되었다. 이를 통해 전체 표본 기준 방문조사 대비 방문+유치조사를 한 경우 총 양육비용이 높게 나타나는 현상은 측정오차라기보다는 총 자녀 수가 많은 응답 가구가 방문조사보다는 방문+유치조사를 선택할 확률이 높았던 선택효과에 기인함을 유추해 볼 수 있었다.

<표 10> 조사방식 변화가 총 양육비용에 미친 영향_이원고정효과

		전체	총 자녀 수별		
			1명	2명	3명 이상
맞벌이 여부(맞벌이=1)		6.487	15.331**	0.758	12.010
총자녀 수	2명	52.433***			
	3명 이상	90.197***			
가구 소득	300-399만 원 이하	0.524	0.781	-1.014	-3.712
	400-499만 원 이하	2.864	0.456	7.197	-8.822
	500-599만 원 이하	6.700	-10.291	17.971*	-7.337
	600만 원 이상	21.720***	-0.577	33.936**	3.701
지역 규모	중소도시	-20.256	-8.719	-29.769	-64.126
	읍·면지역	-20.030	-31.819	-24.444	-89.061
막내 자녀 연령	유아	4.746	16.455**	0.090	-9.396
	초등저학년	-1.038	6.466	2.095	-37.474
시차 효과(2020년=1)		-4.692*	-4.193	-2.726	-6.937
조사 방식	전화	2.642	7.623	-4.228	13.528
	유치	-7.885	-14.544**	-1.658	-22.332
	방문+유치	8.838**	2.388	6.342	22.449
	기타	-7.213	-1.970	-10.304	-1.843
상수		88.036***	83.027***	141.569***	234.367***
σ_u, σ_e		51.843, 46.239	41.627, 36.248	52.397, 48.523	80.141, 57.183
ρ		0.557	0.569	0.538	0.663
N		3,268	1,031	1,761	476

주 : 1) * $p < .1$, ** $p < .05$, *** $p < .01$.

2) 영아는 0~3세, 유아 4~6세, 초등 7~8세를 의미함.

3) 조사방식 base는 방문조사이며, 독립변수는 총 자녀 수는 1명, 가구소득 299만 원 이하, 지역규모는 대도시, 막내자녀 연령은 영아가 base임.

4) 이원고정효과 모형의 F test that all u_i 는 1% 수준에서 유의함.

V. 결론 및 논의

지금까지 ‘KICCE 소비실태조사’ 2차년도(2019년)과 3차년도(2020년) 자료를 활용하여, 혼합조사방식으로서의 변경에 따른 모드효과에 대해 살펴보았다. 이 연구는 코로나19의 확산에 따라 대면 면접조사가 극히 어려워진 상황에서 혼합조사방식으로서의 변경이 데이터의 품질에 어떠한 영향을 미쳤는지를 살펴보는 연구이다. 이를 위해 조사방식 변경 전후의 조사주차별 응답률을 비교해 보고, 조사방식 선택에 영향을 미친 요인이 무엇인지 살펴보았으며, 조사방식에 따라 주요 측정변수(총 양육비용)가 체계적인 차이를 보이는지 등을 고찰해 보았다.

조사주차별 조사 진행률은 혼합조사방식이 도입된 2020년(3차년도) 조사가 대면 면접조사 방식으로 진행됐던 2019년(2차년도) 조사에 비해 빨랐으며, 조사완료까지 걸린 기간도 2020년(3차년도) 조사가 짧았다. 또, 2차년도 표본추적률은 75.4%에 그친 반면, 3차년도 표본추적률은 85.9%에 달했다. 이를 통해 혼합조사방식의 도입이 응답률 측면에서는 긍정적 효과가 있음을 확인할 수 있었다.

한편, 조사방식 선택에는 가구소득과 지역규모, 총 자녀 수 등이 영향을 미친 것으로 분석됐는데, 특히 다른 조사방식에 비해 방문+유치조사의 경우 가구 특성에 따른 차이가 큰 것으로 나타났다. 패널 확률효과 다항로지분분석 결과를 통해서는 다음과 같은 현상을 관측할 수 있었다. 첫째, 총 자녀 수가 많을수록 방문조사 대비 유치조사 혹은 방문+유치조사 확률이 높은 것으로 분석됐다. 이는 자녀가 많을수록 응답해야 하는 문항이 많아져 응답시간이 길어지기 때문에, 장시간 면접을 꺼려하는 영유아 가구의 심리가 반영된 결과로 사료된다. 둘째, 저소득 가구(가구소득이 299만 원 이하)는 방문조사보다는 설문이 유치되는 조사방식(유치조사, 방문+유치조사)을 좀 더 선호하고, 고소득 가구일수록 전화조사를 선호하는 경향을 보였다. 셋째, 대도시에 거주하는 응답 가구에 비해 중소도시에 거주하는 가구들은 방문조사 대비 전화조사와 방문+유치조사 참여 확률이 높은 데 반해, 완전 유치조사 확률은 더 낮은 특징을 보였다. 읍·면지역 거주자의 경우에도 대도시 거주자에 비해 완전 유치조사 확률이 유의미하게 낮게 나타났는데, 이는 코로나19 방역조치로 인해 대도시 거주자가 중소도시 혹은 읍·면지역 거주자에 비해 대면 접촉이 더 어려웠

던 것에 기인한 결과로 사료된다.

마지막으로, 조사방식에 따라 주요 측정변수(총 양육비용)가 체계적 차이를 보이는지 검증해 본 결과에서는 방문조사 대비 방문+유치조사를 한 경우에만 유의미한 양의 영향이 포착되었다. 즉, 전화조사 혹은 유치조사로 응답된 응답값(총 양육비용)은 방문조사로 응답된 응답값과 통계적으로 큰 차이를 보이지 않는다고 볼 수 있다. 반면, 방문+유치조사로 조사된 응답값은 방문조사의 응답값에 비해서는 높게 응답되는 경향을 보인다고 생각할 수 있다.

그런데, 앞서 조사방식 선택 요인에서 살펴본 바와 같이 방문+유치조사의 경우 자녀 수가 많은 가구일수록 방문+유치조사 확률이 높았다. 총 양육비용 또한 자녀 수가 많을수록 높아지는 현상이 뚜렷하게 관측되기 때문에, 분석 대상 집단을 자녀 수별로 구분하여 동일한 분석을 실시해 보았다. 분석 결과, 자녀 수가 2명 혹은 3명 이상인 집단에서는 조사방식에 따라 총 양육비용에 유의미한 차이를 보이지 않는 것으로 분석됐다. 반면, 자녀 수가 1명인 집단에서는 방문조사 대비 유치조사를 한 경우 총 양육비용이 낮게 응답된 것으로 나타났다. 이를 통해 전체 표본 기준 방문+유치조사에서 총 양육비용이 높게 나타나는 현상은 측정오차라기보다는 다자녀 가구들의 장시간 대면 면접조사에 대한 부담감이 반영되어 나타난 선택효과일 가능성이 큼을 엿볼 수 있었다. 또 자녀가 1명인 경우 유치조사를 실시할 경우 총 양육비용을 낮게 응답하는 측정오류가 발생할 위험이 있어 조사 과정에서 주의가 요구되나, 이의 영향이 데이터 품질 자체를 위협할 정도로 큰 수준은 아님을 알 수 있었다.

코로나19라는 환경 변화로 조사방식의 변경이 있었으나, ‘KICCE 소비실태조사’의 경우 조사방식의 변경이 데이터 품질에는 큰 영향을 미치지 않은 것으로 판단된다. 그런데, 이때 총 자녀 수가 많은 육아 가구의 경우 방문조사보다는 방문+유치조사를 선호하는 경향을 보이는 것으로 나타났다. 한편, 혼합조사방식의 도입은 전반적으로 표본응답률 제고에는 긍정적인 효과를 보이는 것으로 분석됐다. 따라서, 향후 반복조사에 응하는 가구 조사에서 총 자녀 수가 많거나 대면 접촉이 어려운 가구의 경우 충분한 응답시간의 확보와 자유로운 응답시간 선택이 가능한 조사방식이 응답률 제고를 위한 효과적인 대안이 될 수 있을 것으로 사료된다.

다만 이 논문은 ‘KICCE 소비실태조사’ 2차년도와 3차년도 조사 모두에 응답한 가구만을 분석 대상으로 하고 있어, ‘KICCE 소비실태조사’를 처음 접하는 응답자들이 다른 조사방식에 대해 어떻게 반응했는지는 알 수 없다는 한계가 있다. 즉, 조

사방식 변경 이전에 ‘KICCE 소비실태조사’에 한 번 이상 참여함으로써 어느 정도 설문에 대한 이해를 가지고 있는 응답자들이기 때문에 응답 오류가 크게 발생하지 않았을 가능성이 있다. 김유빈 외(2018)는 응답자들이 조사방식과 설문문항 등에 대해 학습효과를 가지고 있어, 첫번째 조사에서는 응답오류가 발생했더라도 두번째 조사부터는 측정오류가 크게 감소한다고 보고한 바 있다. 따라서, 추후 자료의 안정적인 구축을 위해서는 기존 표본가구에 대해서는 응답자가 조사방식을 선택할 수 있도록 허용함으로써 응답 편의성을 제고해 주되, 새롭게 ‘KICCE 소비실태조사’에 참여하는 신규 가구에 대해서는 보다 신중한 접근이 요구된다 하겠다. 문제는 신규 가구의 상당수가 신생아 표본가구이기 때문에 대면 접촉을 더욱 꺼려하는 경향이 강하다는 점이다. 그러므로, 쌍방향 온라인 면접조사 혹은 Web+전화조사 등 비대면 면접 조사방식에 대해서도 적극적으로 검토해 볼 만할 것이다.

참고문헌

- 김유빈·박민규·변종석·최효미·권혁진·김우영·성재민·이지은·신선옥·이혜정·김기홍. 2017. 《패널자료 품질개선 연구(VII)》. 한국노동연구원.
- 김유빈·박인호·이지은·신선옥·이혜정·정현상·최세림·박경원. 2018. 《패널자료 품질개선 연구(VIII)》. 한국노동연구원.
- 민인식·최필선. 2010.《STATA 패널데이터 분석》. 한국STATA학회.
- 민인식·최필선. 2012.《STATA 고급 패널데이터 분석》. 한국STATA학회.
- 박시내·최유성·한승훈. 2014. “경제활동인구조사의 자료수집방법별 효과 분석: 응답자의 자료수집방법 선택효과를 중심으로.” 《조사연구》 15(1): 73-92.
- 백지선·민경아. 2016. “병행적 혼합조사의 모드효과 분석.” 《응용통계연구》 29(5): 787-806.
- 백지선·심규호·홍영택·장덕현. 2017. “혼합조사의 모드효과 추정: 선택효과와 측정효과 의 분리 방법.” 《통계연구》 22(3): 1-24.
- 장인성·이지은·김기민·김우영·이상혁·이지희·채민희·최효미. 2021. 《패널자료 품질개선 연구(XI)》. 한국노동연구원.
- 최효미. 2020. “코로나19 시대 TAPI 조사의 한계와 조사도구 변화에 따른 영향 : KICCE 소비실태조사를 중심으로.” 《2020 한국조사연구학회 추계학술대회》.
- 최효미·이상호·성재민·김기민·배기준. 2012. 《패널자료의 품질 개선을 위한 조사기법

- 연구 - 컴퓨터를 이용한 조사(CAPI) 방법». 한국노동연구원.
- 최효미·이정원·김자연·이재희·김태우. 2021. 《KICCE 소비실태조사: 양육비용 및 육아서비스 수요 연구(IV)». 육아정책연구소.
- 최효미·이정원·박은정·김태우. 2020. 《KICCE 소비실태조사 : 양육비용 및 육아서비스 수요 연구(III)». 육아정책연구소.
- 최효미·장혜원·김태우·고영우. 2019. 《영유아 가구의 소비실태조사 및 양육비용 연구 (Ⅱ)-부록». 육아정책연구소.
- 탁현우·고길곤·정다원. 2019. “설문조사 도구에 따른 비표본오차에 관한 연구: TAPI와 PAPI 조사방식의 비교를 중심으로.” 《행정논총》 57(1): 93-114.

- Burton, J., P. Lynn. and M. Benzeval. 2020. “How Understanding Society: The UK Household Logitudinal Study Adapted to the COVID0-19 Pandemic.” *Survey Research Methods* 14(2): 235-239.
- Carletto, G. 2020. “Survey Data Collection during and after COVID: Lessons and Recommendations.” *Joint ISWGHS-ECLAC Webinar on COVID-19: Assessment of the Mode Effect on Official Statistics*.
- Denscombe, M. 2006. “Web-Based Questionnaires and the Mode Effect.” *Social Science Computer Review* 24(2): 246-254.
- Dillman, A. 2009. “Some Consequences of Survey Mode Changes in Longitudinal Surveys.” *Methodology of Longitudinal Survey* WILEY, 127-140.
- Mindell, J., S. Tipping., K. Pickering., S. Hope., M. Roth., and B. Erens. 2010. “ The Effect of Survey Method on Survey Participation: Analysis of Data from the Health Survey for England 2006 and the Boost Survey for London.” *BMC Medical Research Methodology* Research Article, 1-8.
- Van der Laan, P. and W. van Nunspeet. 2009. “Modernising Household Surveys in the Netherlands: Design, Efficiency Gains and Perspectives.” Discussion Paper (09044), Statistics Netherlands. Hagur/Heerlen.
- Wooldridge, J. 2002. *Econometric Analysis of Cross Section and Panel Data*. Cambridge Massachusetts & London England, The MIT Press.

[홈페이지 자료]

- ESCAP¹⁰⁾ 홈페이지. 2021. https://www.unescap.org/sites/default/d8files/event-documents/WB_Mode_Effects_HFPS_in_PNG_Stats_Cafe_1Feb2021.pdf.(인출일: 2023.1.18).

10) Economic and Social Commision for Asia and the Pacific의 약어임.

[통계원 자료]

KICCE 소비실태조사 2019년(2차년도). 육아정책연구소.

KICCE 소비실태조사 2020년(3차년도). 육아정책연구소.

<접수 2023.01.25; 수정 2023.02.18; 게재확정 2023.03.10>

Changes in Data Collection Methods due to the COVID19 Pandemic and Mode Effects: The Case of the KICCE Spending Survey

Hyo-mi Choi

(Korea Institute of Child Care & Education)

The aim of this study is to ascertain how changes in data collection methods, which were necessitated by the near-impossibility of face-to-face interviews due to the COVID19 pandemic, might have affected data quality. Our analysis is based on the 2nd (2019) and 3rd (2020) waves of the KICCE Spending Survey.

Applying panel random effects multinomial logit estimation to gauge the effect of respondents' choices from among various data collection methods, we looked at whether changes in data collection methods affected data quality. In doing so, we employed a two-way fixed effects model that allowed us to control for both unobserved heterogeneity and time lag effects.

Analysis results indicate a positive effect with regards to response rates. Relative to face-to-face interviews, participation in 'face-to-face interview + leave a survey at home' type surveys was more likely among households with more children, low-income households with less than 3 million KRW in household income, and households residing in medium-sized areas. Finally, looking at whether data collection methods led to systematic changes in key measurement variables (e.g. total childcare costs), we found evidence of a significant positive effect only among 'face-to-face interview + leave a survey at home' cases, relative to face-to-face interviews. This mode effect appears attributable more to the effect of respondents' choice of data collection method, rather than to measurement error. Therefore, in the case of households with many children or households that are difficult to interview face-to-face, it appears that providing the option of freely choosing data collection methods will help improve convenience among respondents.

Key words: mode effect, choice of data collection method, pandemic-associated changes in data collection methods