

제3장

가구 면접조사에서 무응답률과 무응답 편향

- 지역별고용조사를 중심으로 -



김서영 · 안다영

제1절 서론

1. 연구배경

무응답은 사회조사에서 공통적으로 발생하는 현상이다. 과거 수십 년에 걸쳐 응답률이 크게 떨어졌다는 것은 많은 선행연구를 통해 알 수 있다(de Leeuw and de Heer, 2002). 동시에 조사기관의 노력이 증가하는 것에 비해 조사에 들어가는 예산은 크게 증가하고 있다(Curtin, Presser and Singer, 2005). 확률표본추출의 추출 패러다임은 조사 추정치의 불편성(unbiasedness)을 보장하기 위해 100%의 응답률을 요구하기 때문에 이러한 비용과 노력이 증가하는 것은 자연스러운 현상일 것이다.

현행 최선의 조사 실무에서 응답률을 최대화하고 무응답 오차의 위험을 최소화해야 한다고 연구자는 논의를 확장하고 있다. 2000년대 초반의 연구를 보면 무응답률의 변화가 반드시 조사 추정치를 바꾸는 것은 아니라는 의견을 제시하고 있다(Curtin et al., 2000; Keeter et al., 2000). 그러나 조사 연구자들의 가장 공통적인 처방은 무응답률을 최소화 하는 것이다. 예를 들면 Alreck 과 Settle(1995)은 무응답률은 가능한 최소로 줄이고 적절한 응답률을 장려하는 것이 중요하다고 강조하였다. Babbie(2007)는 사회과학 연구에 있어서 응답률의 수준을 언급하면서 70%의 응답률이 매우 좋다고 하였고, Singleton 과 Straits(2005)는 면접조사에서 최소한 85%의 응답률을 유지해야 한다는 의견을 제시하였다.

이처럼 무응답에 관한 논의는 과거는 물론이고 앞으로도 조사가 존재하는 한 끊임없이 논의될 수밖에 없는 주제임에 틀림없다. 그러나 다양한 연구를 통해서 볼 때 무응답에 관한 논점은 시대적 흐름과 무관하지 않은 것 같다. 조사협조가 잘 되었을 때는 조사 비용 측면을 고려하여 무응답률을 논의하였다. 반면에 개인정보 노출을 꺼려하거나 민

감한 조사내용에 대해서 조사를 거부하는 사회적 현상 또는 맞벌이 가구나 1인 가구가 증가하는 사회적 변화는 조사연구의 방향을 바꾸는데 기여했다고 볼 수 있다. 또한 특정 계층이나 가구에서 응답을 얻지 못하게 되는 상황이 점점 늘어나면서 이로 인한 무응답 오차가 커진다는 것에 연구의 초점이 맞춰지고 있다. 특히 최근에는 무응답률이 무응답 편향을 예측하지는 못한다는 주장도 나오고 있다(Groves, 2006; Olson, 2006). 또한 Heerweh 등 (2007)은 거부율과 비접촉률을 최소화하려는 노력이 조사품질에 어떤 영향을 미치는가에 관한 연구를 수행하기도 하였다. 이들의 연구에 따르면 자료를 많이 수집하는 것이 반드시 오차를 줄이는 것은 아니라는 결과와 함께 “양적으로 많은 자료와 질적으로 좋은 자료”를 혼동하지 않도록 주의해야 한다고 하였다.

이처럼 조사연구에 있어서 조사에 응답하기를 거부하는 응답대상자를 설득함으로써 응답률을 높히려는 노력이 중요해졌다. 실제로 이러한 노력은 응답전환율 향상과 더불어 응답률을 향상시키는 결과를 유도하였다(Retzer et al., 2004). 이처럼 응답전환율이 증가함에 따라 전환된 응답들이 조사 추정치와 자료품질이라는 두 측면에서 조사결과에 어떤 영향을 미치는지를 이해하는 것이 중요하게 되었다. Keeter 등 (2000)은 가구를 대상으로 하는 조사에서 추정치를 비교해 본 결과 한 번도 거부한 적이 없는 가구와 한두 번 이상 거부한 가구에서 조사된 항목 간에는 통계적으로 차이가 없다고 하였다. Curtin 등 (2000)은 자료에서 응답전환된 가구를 제외했을 때, 제외하기 전의 추정치와 큰 차이가 없다고 하였다. Miller와 Wedeking(2003)은 거부전환된 자료의 품질이 그렇지 않은 경우에 비해 더 좋지 않다는 결과를 보고하였다. 이는 거부경험이 있는 가구들이 어떤 이유에서든 부자연스럽게 응답으로 전환되었을 경우, 그들의 솔직한 응답여부에 대해서는 예측하기 어렵다는 것을 의미할 수도 있다. 이처럼 응답전환된 자료가 전체 자료에 미치는 영향은 클 수 있으며, 따라서 거부전환 노력은 보다 솔직한 응답을 얻을 수 있도록 신중하게 접근해야 할 것이다.

실제로 좋은 자료를 얻기 위해서는 응답률 못지않게 특정 그룹에서 편향이 없는 자료를 얻을 수 있어야 한다. 이를 위해서는 부차 그룹별로 무응답률을 구하고, 무응답 편향을 측정하여 적절하게 대처함으로써 자료의 품질 향상을 추구하는 것이 중요하다. 많은 통계 선진국들은 통계작성기관을 중심으로 학계의 학자들과 연계하여 무응답과 관련된 연구를 수행함으로써 조사통계의 발전을 꾀하고 있다. 특히 무응답 편향은 모든 조사에서 일관적으로 일어나는 현상이라기보다 각 조사에 따라 다르게 측정되는 것이 일반적이다. 그런 면에서 본다면 센서스 자료를 통해 얻을 수 있는 인구학적 특성변수가 아닌 조사의 핵심변수에 대해 무응답 편향을 평가하는 것이 중요하다.

우리나라의 경우도 무응답에 관한 관심이 높아지면서 조사통계 기관을 중심으로 다양한 국내의 활동을 통해 조사통계 발전에 기여하려는 노력들을 하고 있다. 이러한 연구는 자료 활용 또는 비용 측면에서 학계 또는 영리목적의 여론조사기관에서 수행하기는



잡지 않을 것으로 예상된다. 또한 우리나라도 다른 나라들과 마찬가지로 조사환경이 점점 어려워지고 있는 상황이다. 이런 시점에서 조사무응답에 관심을 갖고 적절한 대처방안을 세우기 위해서는, 우선적으로 무응답 유형별 특성을 파악하고, 무응답률과 무응답 오차의 관계를 분석함으로써 자료축적과 자료품질 간에 어떤 관계가 있는지를 분석할 필요가 있을 것이다. 이에 본 연구는 다양한 관련 문헌들을 리뷰하고, 실제 자료분석과 실험을 통해 다양한 방법으로 무응답 편향을 평가하며, 자료의 양과 질적 측면에서 무응답이 자료에 미치는 영향을 파악하는데 초점을 둔다.

2. 연구내용과 범위

무응답은 크게 표본이 조사에 응하지 않는 단위 무응답(unit nonresponse)과 조사에는 응하되 일부 항목에서 응답을 하지 않는 항목 무응답(item nonresponse)으로 나뉜다. 항목 무응답의 경우 무응답 대체(imputation)를 통해서 무응답 편향을 보정할 수 있다. 단위 무응답의 경우도 무응답 가중치를 통해서 무응답을 보정할 수 있다. 이때 무응답 가중치 사용은 응답자와 무응답자의 성향이 동일하다는 것을 가정한다. 이 가정이 위배될 경우, 무응답 가중치를 사용한 보정은 오히려 더 큰 편향을 유발할 수 있다. 본 연구에서는 무응답 대체 방법에 대해서는 논의하지 않을 것이다. 현실적으로 조사에서 응답자와 무응답자의 성향을 파악한다는 것이 쉬운 일은 아니다. 단위 무응답 발생은 응답률을 떨어뜨리는 결과를 낳기 때문에 실제 표본조사에서 단위 무응답에 더 큰 관심을 갖게 되는 것 같다. 본 연구의 관심도 단위 무응답에 있다.

조사가 완성되면 조사결과에 대해 서브 그룹에 걸쳐 응답률을 계산하고 비교해야 한다. 이러한 서브 그룹에는 성, 연령, 지역 등의 인구사회학적 특성변수들이 포함될 수 있다. 물론 이러한 방법이 무응답 편향의 정도를 결정하는데 도움이 되지는 않지만 이렇게 함으로써 무응답 편향이 있는지의 여부는 알 수 있다. 만약 응답률이 서브 그룹 간 매우 유사하다면 무응답 편향이 있더라도 전체 자료에는 제한적인 영향을 줄 가능성이 있다. 이러한 방법이 갖는 문제는 무응답이 발생하는 원인이 서브 그룹과 다른 곳에 있는 경우에는 파악하기 힘들다는 것이다. 즉, 응답성향이 서브 그룹 변수에만 영향을 받는 것이 아니라는 것이다(Canada, 2007). 응답률 비교는 좋은 출발점이긴 하나 무응답을 설명할 다른 방법을 대신하지는 못한다. 본 연구에서는 우리나라 전체 지역을 2개 권역(특·광역시와 대도시권 및 시군을 포함하는 소도시 및 농어촌 권역)으로 나누어 각 권역별로 응답률을 계산하고 비교한다.

통계적 조정 즉, 가중치(weighting) 작업은 무응답을 설명하기 위해 사용되는 가장 보편적인 방법이다(Groves, 2006). 자료에 대한 가중치 작업은 이용자가 자료의 대표성을 확신하도록 하는데 도움이 될 수 있다. 단, 무응답 편향이 전혀 없거나 있더라도 아주

작다고 가정할 수 있어야 한다. 가중치 작업은 무응답 편향을 설명하는데 있어서는 덜 효과적이다. 무응답 편향은 무응답률과 함께 무응답이 응답자와 무응답자 간의 차이를 초래할 때 발생한다. 무응답 편향을 조정하기 위해 연구자는 조사 참여자의 어떤 성향이 조사에 동의하게 하는지를 이해할 수 있어야 한다. 이러한 이해는 응답자와 무응답 간 체계적 차이는 관측 가능한 변수(즉, 성, 지역, 연령 등 가중치 작업에 의해 조정 가능한 변수)를 벗어난 다른 변수들로 확장될 수 있기 때문에 중요하다.

조사가 완성되면 응답률 비교와 함께 응답자와 무응답자 간의 특성을 비교해야 한다. 이러한 비교는 다른 알려진 변수나 인구사회학적 특성 변수를 이용하여 응답자와 무응답자의 특성이 어떻게 다른지를 알 수 있게 한다. 그러나 무응답자의 정보를 파악하는 것 자체가 그렇게 쉽지는 않다. 이런 경우 일반적으로 조사 외의 정보를 이용하는 방법이 있다. 행정자료와 같은 정보를 이용하여 부분적으로나마 무응답 가구의 정보를 파악할 수 있는 조사변수를 찾아야 한다. 또한 자료를 수집하는 과정에서 핵심 조사변수와 상관관계가 있는 변수를 찾아야 한다. 이때 조사변수와 상관관계가 있는 변수는 조사변수에 대해서 무응답 편향을 노출할 수 있어야 한다. 이를 위해 본 연구는 자료를 수집하는 과정에서 핵심변수와 관련이 될 만한 변수를 조사하였고, 이를 무응답 편향을 평가하기 위한 변수로 사용하였다.

이러한 선행연구들의 다양한 의견을 근간으로 하여 본 연구를 진행코자 하였다. 본 연구의 가장 큰 목적은 무응답률과 무응답 편향 간에 어떤 관계가 성립하는지를 확인하는 데 있다. 이 주제에 관한 많은 연구를 통해 각각 다른 주장들이 제기되어 왔고 지금도 진행 중이지만, 최근에는 전반적으로 무응답률과 무응답 편향 간에는 직접적인 관계가 성립하지 않는다는 것에 더 시선이 집중되고 있는 것 같다. 앞에서 언급하였듯이 무응답률과 무응답 편향 간에 특별한 관계가 성립되지 않는다는 것은 많은 의미를 내포할 수 있다. 대표적인 것이 자료의 양적·질적 측면에서의 관계, 확률표본추출과 응답률의 관계 및 무응답 편향을 고려한 조사설계 등에 대해 다양한 논의가 가능할 것이다. 실제로 이러한 의견들에 대한 일관성 있는 대답은 하기 어려운 것이 현실이고, 그렇기 때문에 많은 연구에서 이 둘 관계에 대한 명백한 결론을 내리지 못하는 것 같다. 이는 조사마다 또는 국가마다 처한 환경이 다르기 때문일 것이다.

3. 연구방법

본 연구는 무응답 오차와 관련된 여러 가지 일반적인 주제들 중 3가지 연구 질문에 대해서 분석하고자 한다. 본 연구에서 무응답 오차는 응답자와 무응답자 모두를 포함하는 전체 자료에 기초한 추정치와 응답자 자료에 기초한 특정 거주 유형의 추정치 비율 간에 발생한 불일치를 말한다.



위의 연구 질문에 답하기 위해 실제 자료분석과 자료를 이용한 모의실험을 수행한다. 구체적인 방법은 다음과 같다.

① 관심 추정치에 대해 발생하는 무응답 오차의 크기를 결정한다. 우리나라 국민들의 거주 유형을 크게 주택, 아파트, 기타로 나누었을 때, 관심 추정치는 주택에 거주하는 가구의 비율이다. 그리고 무응답 오차 성분별로도 그 크기를 측정한다. 본 연구에서 무응답 오차 성분은 거부, 비접촉, 기타 등 3가지 경우로 나눈다. 부차 그룹별 무응답 현황을 이해하기 위해 관심 추정치에 대한 무응답 오차 크기는 지역 특성을 고려하여 크게 특광역시를 포함하는 대도시권과 시군을 포함하는 소도시 권역으로 구분하여, 이로부터 권역별로 무응답 유형에 따른 무응답 오차의 구성 변화를 살핀다.

② 자료양이 축적됨에 따라 무응답 오차 구성이 바뀌는지 여부를 파악한다. 이를 위해 전체 현장 조사기간을 시간적 순서(날짜)에 따라 더 작게 쪼개서 시간 순에 따라 누적되는 자료의 품질을 평가한다. 이렇게 함으로써 현장 조사기간에 걸쳐 자료에 나타나는 오차의 경향을 감지할 수 있다. 본 연구에서는 현장 조사기간을 10개 슬라이스로 나눈다. 무응답 오차는 이 10개의 누적 슬라이스 구간을 통해 계산되고 평가된다.

③ 현장 조사기간 동안 접촉시도가 많으면 무응답 오차와 오차의 구성이 달라질 것인가에 대해 분석한다. 이는 연장된 조사원 노력이 무응답 편향에 미치는 영향을 파악하는 것과 관련된다(Lynn et al., 2002). 면접원의 연장된 노력이란, 면접원이 접촉하기 어려운 응답자를 접촉하거나 마지못해 응답하게 하는 응답전환 노력을 말한다. 이것은 ‘마지막 응답자 사냥’(Stoop, 2005)이라 한다. 조사에서 높은 응답률을 얻기 위해 접촉시도를 늘리는 것은 보편적인 절차이다. 이를 통해 질 높은 자료를 얻기를 바란다. 접촉시도가 많을수록 비접촉률은 낮아질 것이고, “비접촉률이 낮으면 비접촉 오차를 감소시킬 것이다”라는 가정에서, 만약 조사에서 발생하는 다른 모든 오차가 같다면 이것은 무응답 오차를 줄이게 될 것이다. 이 경우 무응답자는 마지못해 응답한 응답자와 유사하다고 가정한다. 그러나 다른 연구에서는 이 가정은 정당화 될 수 없다고 제안하고 있다(Lin and Schaeffer, 1995; Stoop, 2005). Keeter 등 (2000)과 Curtin 등 (2000)은 응답률을 높이거나 유지하기 위해 더 많은 노력을 할수록 대체로 다른 점추정치가 유도된다고 하였다. 더 많은 접촉시도에 의해 비접촉률이 감소하면서 무응답자와 응답자 간의 차이는 커진다고 할 수 있다. 왜냐하면 접촉하기 쉬운 응답자와 닳은 사람들은 응답자 그룹에 계속 쌓이게 되고 접촉하기 어려운 사람은 계속 접촉하지 못한 채로 남게 되기 때문이다. 그렇지만 응답자와 비접촉자 간의 차이가 커지는 것이 비접촉 오차를 충분히 커지게 할 수 있는 것인지의 여부는 알려져 있지 않다. 이것은 경험적인 방법에 의해 해결할 수 있는 문제인 것 같다.

4. 연구의 제한점과 기대효과

본 연구는 다음과 같은 측면에서 몇 가지 제한점을 갖는다. 본 연구를 통해 도출된 모든 결과는 이러한 제약을 고려하여 이해해야 할 것이다. 본 연구와 같이 무응답률과 무응답 편향의 관계 또는 방문회차별 무응답 편향의 변화를 평가하기 위해서는,

첫째, 조사변수와 관련된 변수를 이용할 수 있어야 한다. 즉, 무응답에 대해서 조사변수와 관련된 정보를 파악할 수 있어야 한다. 그러나 연구자의 판단에 의하면 이런 유형의 연구가 국내에서는 거의 시작단계라는 점에서 볼 때 연구목적 달성에 적합한 핵심변수를 얻기가 쉽지 않다. 본 연구도 우리나라 지역별고용조사의 가구관리종합표에 기록된 정보에 의존하기 때문에 연구변수 사용이 매우 제한적일 수밖에 없다.

둘째, 연구에 사용된 조사자료의 경우 조사설계가 이 연구를 목적으로 설계된 것이 아니고, 실제 조사 과정에서 발생하는 절차를 따르기 때문에 조사기간이나 방문회차를 유연하게 연장하거나 조정할 수 없었다는 한계가 있다. 이런 면에서 본 연구의 결과를 일반화하기에는 조심스러운 부분이 있을 수 있다.

셋째, 연구자가 조사 과정에 직접 참여하지 못했다는 점에서 조사 과정에서 발생할 수 있는 여러 가지 상황을 예측하기 어렵고, 자료분석 결과를 해석하는 과정에서 이러한 현상을 미루어 짐작할 수밖에 없다는 아쉬움이 있다.

본 연구는 다음과 같은 기대효과를 줄 수 있다.

직접적으로는 첫째, 본 연구의 핵심인 응답률과 무응답 편향과의 관계가 무응답 오차에 미치는 영향을 파악할 수 있다. 이러한 오차가 어떤 무응답 원인에 의해 발생하는 것인지를 파악함으로써 무응답 원인 그룹에 따른 무응답 대응 전략을 세우는 데 도움이 될 수 있을 것이다. 둘째, 방문회차에 따른 무응답 오차 구성의 변화를 통해 방문회차의 의미를 찾을 수 있다. 마지막으로 본 연구는 조사원의 노력에 의해 뒤늦게 조사에 응답한 응답전환자의 결과가 전체 조사품질에 어떤 영향을 미치는지를 파악하고, 이로부터 응답전환에 의해 응답률 향상을 꾀하고자 할 경우 지속적인 이들 자료의 분석과 관찰에 대한 필요성을 제시하는데 큰 의미가 있을 것이다.

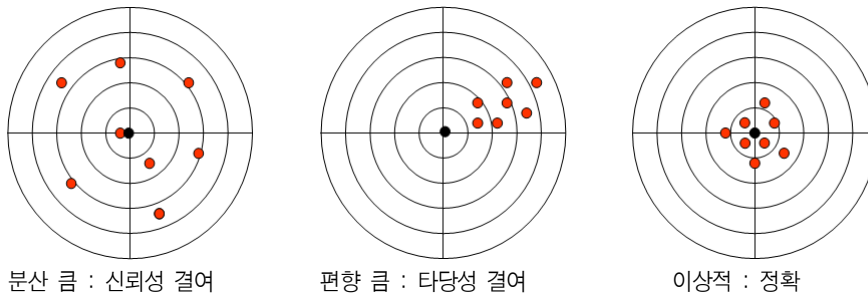
간접적으로는 파라미터의 중요성을 인지하는데 도움이 된다. 파라미터는 조사결과와 간접적인 관계를 통해 조사 전반을 이해하고 조사의 질을 높이는데 기여할 수 있는 자료이다. 따라서 본 연구가 현장 조사 기록 정보를 이용한 연구라는 점에서 그 의미가 있다 할 것이다.

제2절 무응답 오차

1. 무응답 오차와 오차 구성

다양한 방법을 통해 얻어지는 표본조사 결과치는 거의 대부분 오차를 포함하고 있다고 볼 수 있다. 무응답 오차는 무응답 오차의 분산과 편향이라는 두 가지 요소로 구성되어 있다. 무응답이 미치는 영향을 알아보기 위해 우선 확률표본론의 반복추출이론을 기초로 한 평균제곱오차(MSE: Mean Squared Error)에 대해 이해해 보기로 하자. MSE는 조사 추정량의 정확성 정도를 나타내기 위하여 가장 많이 사용되는 측도이다. 모집단(population)의 관심모수를 \bar{y}_p 라 하고, 표본(sample)에서 얻어지는 추정량을 \bar{y}_s 라 하면 MSE는 오차제곱의 기댓값으로 정의되며 식(3.1)과 같이 표현된다(김재광, 2008).

$$\begin{aligned} MSE(\bar{y}_s) &= E\{(\bar{y}_s - \bar{y}_p)^2\} = \sum_{s=1}^S (\bar{y}_s - \bar{y}_p)^2 & (3.1) \\ &= E\{(\bar{y}_s - E(\bar{y}_s) + E(\bar{y}_s) - \bar{y}_p)^2\} \\ &= var(\bar{y}_s) + (bias(\bar{y}_s))^2 \end{aligned}$$



[그림 3-1] 분산과 편향

위의 식(3.1)에서 \bar{y}_s 는 표본평균, \bar{y}_p 는 모집단 평균, $E(\bar{y}_s)$ 는 동일한 모집단에 대해 여러 번 표본추출을 했을 경우 \bar{y}_s 의 평균을 의미한다. 식(3.1)에서 보면 MSE는 표본평균의 분산과 편차로 구성되어 있다는 점에서 이 두 가지 요소 모두가 통계품질을 측정할 때 사용된다는 것이다. 일반적으로 표본조사에서 분산은 조사가 이론적으로 반복될 때 우리가 추정하고자 하는 조사 추정치가 변화하는 정도를 측정하는 것이다. 한편 편향은 표본조사의 반복에 상관없이 조사에서 일정하게 발생하는 구조적인 문제로 인해 야

기된다. 이런 측면에서 볼 때 오차는 분산과 편향이 더해진 개념으로 모집단에서 측정하고자 하는 값과 실제 조사를 통해 측정되는 값의 불일치를 나타내는 값으로 이해할 수 있다. 분산과 편향에 대한 개념적 해석은 분야마다 다르게 적용될 수 있다. 사회과학 분야의 경우, 분산은 측정에 대한 신뢰성(reliability) 또는 추정치의 정도(precision)를 나타내고, 편향은 측정하고자 하는 내용에 대한 타당성(validity)과 관계가 있다. 어쨌건 오차는 정확성과 연관시켜 생각할 수 있고, 분산과 편향 모두가 작을 때 조사에 대한 정확성이 성립될 수 있다(이승희, 2010).

앞에서 언급한 바와 같이 무응답 오차는 분산과 편향으로 구성되고, 조사에서 발생하는 다른 오차들과 마찬가지로 무응답 편향에 더 비중을 두는 것이 일반적이다(Groves, 2006; 이승희, 2010). 또한 현장에서 대부분의 관심은 무응답 편향에 있기 때문에 본 연구에서는 무응답 편향을 중심으로 무응답 오차를 설명해 나갈 것이다. 또한 경우에 따라서는 무응답 편향을 무응답 오차와 같은 개념으로 사용할 수 있다.

응답자 평균에 대해서 무응답 편향은 식(3.2)와 같이 표현될 수 있다.

$$bias(\bar{y}_r) = \frac{m}{n}(\bar{y}_r - \bar{y}_m) \quad (3.2)$$

여기서 \bar{y}_r 는 표본에서의 응답자 평균, \bar{y}_m 는 표본에서의 무응답자 평균, n 은 표본수, m 은 무응답수, $\frac{m}{n}$ 은 무응답률을 각각 나타낸다. 결과적으로 식(3.2)가 의미하는 무응답 편향은 무응답률과 응답자 평균과 무응답자 평균 간의 차에 대한 함수로써 응답자 평균이 모집단 평균과 다른 정도를 측정할 수 있다. 식(3.2)와 같은 표현식은 측정오차와 같이 조사에서 편향을 일으킬 수 있는 다른 근원은 없다고 가정하게 된다(Groves, 2006).

그런데 이 표현식은 절대적인 무응답 그룹의 개념과 양립될 수 있다. 응답률이 동일한 모집단이더라도 조사설계가 바뀔 때 따라 크게 달라질 수 있다는 근거에서 많은 연구자들은 조사대상자 개개는 환경에 의해 잠재적으로 응답자가 될 수도 있고, 무응답자가 될 수도 있다는 사실에 관심을 갖게 되었다(Groves, 2006). 즉, 조사대상자 개개는 응답자 또는 무응답자가 될 수 있다는 관측할 수 없는 성향(propensity)을 가지고 있고, 그 성향은 ρ_i 로 표현될 수 있다. 이러한 관점에서 또 다른 형태의 무응답 편향은 다음과 같이 표현될 수 있다(Bethlehem, 2002).

$$bias(\bar{y}_r) \approx \frac{s_{yp}}{\rho}$$

여기서 s_{yp} 는 표본에서의 조사변수 y 와 응답성향 ρ 간의 공분산(covariance), $\bar{\rho}$ 는 표본 전체의 응답성향에 대한 평균을 나타낸다. 이러한 표현식은 응답 가능성이 조사대상



자 모집 규칙의 구체성에 따라 변하는 랜덤변수라는 것을 내포한다. 즉, 무응답 편향은 응답성향이 연구자가 측정하려는 특성에 얼마나 상관되었느냐의 함수로써 발생한다. 동일 조사 내에서 다른 표본 추정치는 다른 무응답 편향에 관계할 수 있다. 성향이 응답과 연관되지 않은 경우는 무응답의 편향 효과에 영향을 받지 않을 수 있다. 그렇지 않은 경우는 동일한 조사 내에서도 큰 편향에 관계될 수 있다(Grove, 2006).

지금까지의 무응답률과 무응답 편향과의 관계 및 Groves(2006)의 내용을 토대로 다음과 같이 요약할 수 있다.

첫째, 무응답 편향은 같은 조사 내의 다른 통계 추정치에 대해 달라질 수 있다. 따라서 응답률이 낮은 조사가 반드시 나쁜 것이 아니지만, 응답률이 낮으면 무응답 편향이 큰 통계를 만들 수도 있다.

둘째, 무응답률 감소가 항상 무응답 편향을 작게 하는 것은 아니다. 무응답을 줄이는 방법 중 어떤 것은 조사변수와 응답성향 간의 공분산 s_{yp} 를 증가시킬 수 있다.

셋째, 무응답 성분은 조사변수 y 가 주어졌을 때 공분산 s_{yp} 에 서로 다르게 영향을 줄 수 있다. 비접촉 성향은 건강상의 문제로 나타나는 거부 성향보다 다른 y 에 상관되어 있을 수 있다.

넷째, 무응답률이 언제 무응답 편향을 예고하는지를 알기 위해서 우리는 반드시 조사에 참여한 개체와 참여하지 않은 개체의 영향이 조사에서 측정하고자 하는 변수와 얼마나 관련되어 있는지를 이해해야 한다. 이것은 응답률을 높이기 위해 사용하는 방법, 자료수집 방법, 조사모집단 등에 따라 변할 수 있다.

다섯째, 조사 추정치가 반드시 편향에 영향을 받는다는 전제하에서 최소 응답률을 규정하기는 어렵다.

2. 무응답 유형별 무응답률과 무응답 편향 계산

가. 무응답 유형

단위 무응답이 발생하는 이유는 다양하고, 각각의 이유가 전체 자료에 미치는 영향도 다르게 나타날 수 있다. 따라서 무응답을 유형별로 나누어 무응답을 이해할 필요가 있다. 대체로 무응답 유형을 크게 네 가지로 나누면, 응답을 시작했으나 부분적으로 응답하거나 중도에 포기하는 경우, 표본가구를 접촉할 수 없는 경우(비접촉, noncontact), 표본가구가 조사에 참여하기를 거부하는 경우(거부, refusal), 건강상 또는 여러 가지 이유로 조사에 응답할 수 없는 기타 경우와 같다. 이러한 개념 정의는 유형별 비율을 계산하는데 매우 중요하며 특히 첫 번째의 부분 응답이나 중도 포기 경우를 응답률 계산에 반영할 것인가는 사전에 명확히 정의하고 명시해야 한다. 이해를 돕기 위해 AAPOR(the

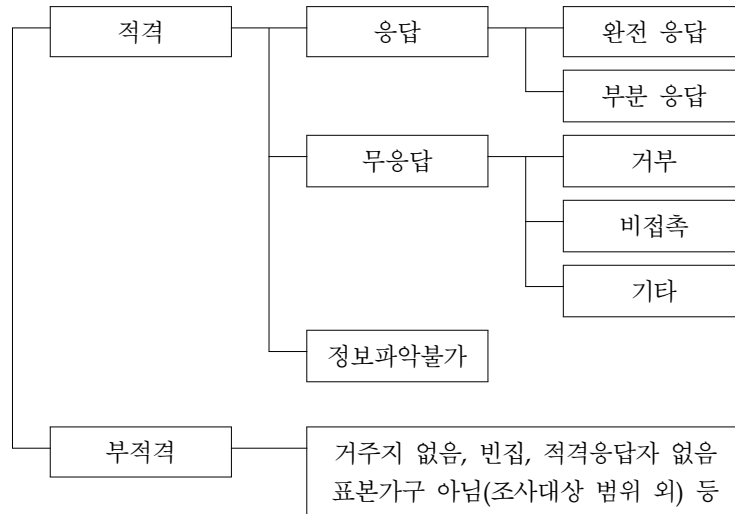
American Association for Public Opinion Research)의 표준 정의를 설명하기로 한다. 일반적으로 다음에서 열거하는 네 가지 방법은 응답자를 완전 응답(complete interview), 부분 응답(partial interview), 중도포기(break-off)를 조사에 포함시킬 것인지를 결정하는 표준이라 할 수 있다(AAPOR, 2009).

- ① 전체 질문 중 응답한 항목이 전체 질문의 50% 미만이면 중도 포기(비접촉, 거부 이외), 50%-80%이면 부분 응답, 80% 이상이면 완전 응답과 같다.
- ② 전체 질문 중 응답한 항목이 전체 질문의 50% 미만이면 중도 포기, 50%-80%이면 부분 응답, 80% 이상이면 완전 응답과 같다.
- ③ 핵심적 질문 중 응답한 항목이 전체 핵심 질문의 50% 미만이면 중도 포기, 50%-99%이면 부분 응답, 80% 이상이면 완전 응답과 같다.
- ④ 위 세 가지 경우를 조합하여 사용될 수 있다. 즉, 핵심 질문 모두에 응답하고 나머지 질문 중 80%에 응답했으면 완전 응답으로 간주한다.

위의 정의가 완전 응답, 부분 응답 또는 중도 포기를 명확히 정의한 것은 아니지만, AAPOR에서도 조사에서 반드시 이들 상황에 대한 명백한 정의를 제공해야 한다고 말하고 있다(AAPOR, 2009).

표본조사에서 적격표본(eligible sample)에 대한 응답 유형은 크게 응답, 무응답, 이용불가로 나눌 수 있다. 응답은 완전이든 부분이든 최종적으로 질문지를 완성한 경우를 말한다. 이때 위에서 설명한 바와 같이 부분 응답의 경우를 최종 응답으로 간주할 것인지에 대한 정의를 명확히 해야 한다. 무응답은 여러 가지 형태로 나누고 있지만 일반적으로 크게 비접촉, 거부, 기타로 구분하고 있다. 비접촉은 표본에 대해 접촉을 시도하였으나 어떤 이유에서든 최종적으로 표본을 접촉하지 못해서 응답을 얻지 못한 경우를 말한다. 거부는 표본을 접촉하였으나 표본이 조사질문에 대해 응답하기를 거부하여 최종적으로 응답을 얻지 못한 경우를 말한다. 기타는 조사대상자가 표본으로 선택되어 표본임을 알고 있으나, 조사기간 동안 외출, 건강, 언어, 사망 등의 문제로 응답을 못한 경우를 말한다. 불능은 표본으로서 적합하지 않은 경우로, 가구의 경우 모집단 명부로부터 표본가구로 추출되었지만, 빈집이거나 조사목적에 부합하지 않은 경우를 말한다. 응답 유형에 대한 정의는 여기에서 언급한 것 외에도 세부적으로 구분하여 사용하고 있다. 자세한 내용은 AAPOR(2009)을 참고하기 바란다.

〈표 3-1〉 유형별 응답형태(AAPOR, 2009)



〈표 3-1〉은 지금까지 설명한 AAPOR(2009)의 유형별 응답형태를 정리한 것이다. 표본 정보의 이용가능 여부에 따라 크게 표본을 적격과 부적격(not eligible)로 나눈다. 적격 표본은 다시 응답, 무응답으로 나누고, 무응답은 거부, 비접촉, 기타로 나눌 수 있다. 〈표 3-1〉에서 표본의 정보를 알 수 없는 경우(unknown eligibility)는 표본 가구인 것은 확실하지만 그 표본에 대한 정보를 파악할 수 없어서 조사를 할 수 없는 경우를 말한다. 즉, 가구가 존재하나 적격가구 여부가 분명하지 않거나, 또는 적격가구가 있는데 그 가구에 적격 대상자가 있는 지 여부를 알 수 없는 경우 등이 해당한다. 이런 가구는 조사를 하지 못하기 때문에 조사에 대한 응답여부도 파악할 수 없게 된다. 하지만 표본 자체는 적격 표본으로서 모집단의 명백한 한 단위이기 때문에 이에 대한 응답 유형 자체도 추정하여 사용할 수 있다. 부적격표본은 표본 자체를 이용할 수 없는 경우로, 대상표본이 아니거나(out of sample), 주거지역(hot-a-housing-unit)이 아닌 회사, 정부기관, 요양소, 군대 등 시설이거나, 장기적 혹은 정기적으로 비어있는 집(트레일러, 별장 등), 또는 적격대상자가 없는(housing unit with no eligible respondents) 가구 등이 해당된다.

나. 응답률, 무응답률, 응답전환율

응답률, 무응답률, 응답전환율은 조사결과에 대한 정보를 보여줄 수 있는 비율 지표로 각종 문헌에서 공통으로 사용되고 있다. 경우에 따라서는 근본적으로 다른 비율을 설명하는데 같은 이름이 사용된다거나 같은 비율을 설명하는데 다른 이름들이 사용되기도 한다. Groves와 Lyberg(1988)는 “응답률을 계산하는 방법이 너무 많고 따라서 조사 간에

응답률을 비교하는 것은 오해를 일으키기에 충분하다”고 말했다. 가장 공통적으로 사용되는 용어로는 응답(response), 협조(cooperation), 접촉(contact), 거부(refusal)가 있다.

본 연구에서 사용한 다양한 비율 계산은 APPOR(2009)에서 정의한 기준들 중 하나를 따른다. 본 연구에서 정의하고 사용할 방법을 설명하기에 앞서 몇 가지 공통적으로 사용되는 용어를 설명하면 다음과 같다.

- I = 완전 응답
- P = 부분 응답
- R = 거부 또는 중도 포기
- NC = 비접촉
- O = 기타

다음에서 사용한 모든 비율 계산식에서 부분 응답(P)은 응답을 정의하는 방법에 따라 응답률 계산에 사용될 수도 있고, 제외될 수도 있다.

① 응답률

응답률 계산 방법은 APPOR(2009)의 경우 6가지로 정의하고 있으면, 우리는 다음의 응답률을 사용한다.

$$\text{응답률} = \frac{I+P}{(I+P)+(R+NC+O)}$$

이와 같은 응답률은 파악되지 않은 조사대상자 중 이용 가능한 경우는 하나도 없다는 가정이 유효할 때 적절한 방법이다. 또한 이 응답률은 최대 응답률을 나타내기도 한다. 정의에 따라 부분응답(P)은 응답률 계산에서 제외될 수 있다.

② 거부율

거부율은 이용 가능한 표본 중 응답자가 면접에 응하기를 거부한 표본 또는 중도에 면접을 포기한 표본의 비율을 말한다. AAPOR(2009)은 3가지 거부율 계산 방법을 정의하고 있고, 세 번째 정의는 다음과 같다.

$$\text{거부율} = \frac{R}{(I+P)+(R+NC+O)}$$

이 거부율은 위에서 설명한 응답률 계산과 유사하다. 즉, 파악되지 않은 조사대상

자는 정확히 조사에서 제외될 수 있어야 한다. 비접촉률과 기타 무응답률로 거부율과 비슷한 방식으로 계산될 수 있다. 무응답률은 거부, 비접촉, 기타를 더한 것과 같아지게 된다.

③ 접촉률

접촉률은 이용 가능한 표본 중 가구 내의 응답 가능한 누군가를 접촉한 가구의 비율을 말한다. AAPOR(2009)은 3가지 접촉률 계산방법을 정의하고 있고, 세 번째 정의는 다음과 같다.

$$\text{접촉률} = \frac{(I+P)+R+O}{(I+P)+R+O+NC}$$

응답률, 거부율과 마찬가지로 이 접촉률도 조사대상에 대해 오직 알려진 이용 가능한 경우만을 포함한다.

④ 협조율

협조율은 접촉한 표본 중 조사에 응답한 모든 표본들의 비율을 말한다. 이 경우 가구 수준과 응답자 수준의 협조율이 있다. 여기서 협조율은 가구 수준의 협조율을 말한다. APPOR(2009)은 4가지 협조율 계산 방법을 정의하고 있다. 두 번째 방법은 다음과 같다.

$$\text{협조율} = \frac{(I+P)}{(I+P)+R+O}$$

본 연구에서는 협조율은 구하지 않고, 위의 정의는 참고를 위해 제시한다. 위의 협조율 공식에서 분모의 기타(O) 또는 분자의 P를 제외하기도 한다.



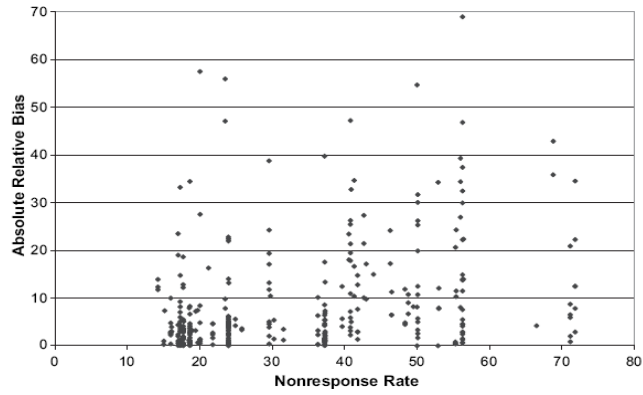
다. 무응답률과 무응답 편향의 관계

본 절은 Groves (2006)에서 제시한 사례연구의 결과를 적절하게 인용하여 설명하였다. Groves 와 Peytcheva(2007)는 30편의 연구문헌에서 발표한 결과들에 대해 메타 분석(meta analysis)을 함으로써, 조사에서 형성되는 무응답률과 무응답 편향과의 관계를 설명하고 그 결과들을 중심으로 의견을 제시하였다. 본 절에서도 상당 부분 이 의견을 토대로 정리하고자 하였다.

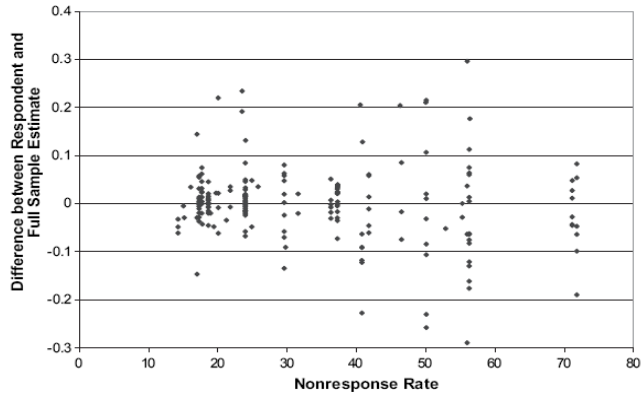
이와 관련하여 Keeter(2000) 등, Curtin 등 (2000), Merkle 와 Murray(2002)는 무응답률이 생각했던 것만큼 조사 추정치에 위협적이지는 않다는 의견을 제시하였다. 그러나 Groves(2006)에 따르면 이들 문헌들의 그 시기에서의 방법론 연구라는 점을 주목할 필요가 있고, 아주 드문 일이긴 하지만, 극단적으로는 무응답 편향을 걱정할 이유가 없다는 의미를 내포함으로써 잘못 해석될 여지가 있다고 하였다. 무응답률과 무응답 편향에 관한 연구는 많고, 그 결과도 일관적이지 않을 뿐더러 다양한 결론을 도출하고 있는 실정이다.

Groves(2006)는 무응답률과 무응답 편향의 관계를 설명하기 위해 30편의 유사한 형태의 연구결과를 분석하여 다음의 4가지 결과를 유도하였다. 또한 4가지 결과들은 Groves(2006, pp659-661)의 [그림]을 인용하여 유도된 결과임을 밝힌다([그림 3-2]). 위의 그림을 통해서 알 수 있는 사실은 조사에서 무응답 편향은 반드시 발생한다는 것이다. 그리고 무응답 편향은 분명히 발생하는데, 무응답률은 편향의 크기를 예측하기에 그다지 충분한 요인은 아니라는 것이다. 즉, 그림에서 보면 무응답률별 편향의 변동 정도가 매우 낮다. 무응답률은 무응답 편향을 민감하게 하는 추정치의 특성을 반영하지 않기 때문에 상관관계가 낮게 된다.

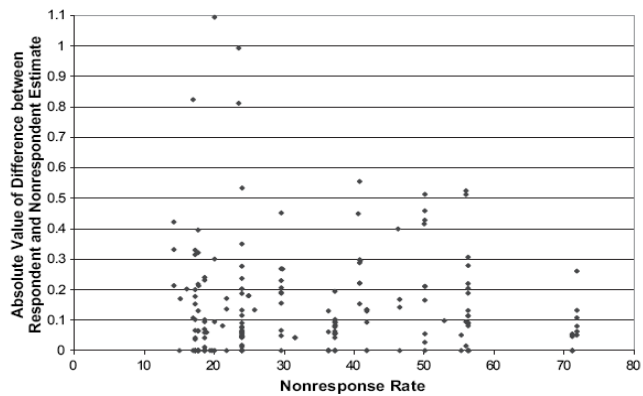
정리하면, ① 높은 응답률이 조사 추정치에 대해 반드시 무응답 편향을 줄이는 것은 아니라는 것이다. ② 무응답률과 무응답 편향과의 관계에 대해서 어떤 통계추정치인 경우, 무응답 편향은 크지만 경험적으로 볼 때 응답률과 무응답 편향 간의 강한 관계는 성립되지 않는다. ③ 어떤 조사의 무응답률은 정해졌더라도 같은 조사 내에서 무응답 편향은 추정치에 따라 변할 수 있다. 편향은 응답성향과 측정 조사변수 간의 공분산에 매우 영향을 받게 되기 때문이다. 간단히 말해서 무응답 편향은 단순한 무응답률보다 훨씬 더 복잡한 현상이다(Groves, 2006).



(a) 상대편향



(b) 응답자와 전체 추정치 간의 차이



(c) 응답자와 무응답자 추정치 간의 차이

출처: Groves(2006) pp. 659-661의 Figure 2.3.4

[그림 3-2] 무응답률과 무응답 편향의 관계

무응답 유형에 따른 무응답 편향은 식(3.2)이용하여 다음과 같이 표현될 수 있다.

$$\begin{aligned} bias(\bar{y}_r) &= \bar{y}_r - \bar{y}_n = \frac{m}{n}(\bar{y}_r - \bar{y}_m) \\ &= \frac{m_{ref}}{n}(\bar{y}_r - \bar{y}_{m_{ref}}) + \frac{m_{nc}}{n}(\bar{y}_r - \bar{y}_{m_{nc}}) + \frac{m_{oth}}{n}(\bar{y}_r - \bar{y}_{m_{oth}}) \end{aligned} \quad (3.3)$$

여기서 r 은 표본 중 응답 개수, n 은 전체 표본 개수, m_{ref} 는 거부에 의한 무응답 개수, m_{nc} 는 비접촉에 의한 무응답 개수, m_{oth} 는 기타에 의한 무응답 개수를 각각 나타낸다. 식(3.3)은 무응답 유형을 거부, 비접촉, 기타의 3가지 형태로 사용하고 있고, 총체적 편향은 각 무응답 유형에 대해 영향을 받게 된다. 각 유형별 무응답 편향은 이미 식(3.2)를 통해 설명한 바와 같이 해당 유형의 무응답률과 응답자와 무응답자 간 차에 대한 곱의 형태이다. 즉, 유형별 무응답 편향은 각 무응답 유형별 무응답 비율과 응답자와 해당 유형의 무응답자 간 차의 정도에 따라 다르게 작용할 수 있다.

제3절 자료분석 및 결과

1. 분석 자료

본 절에서는 실제 자료분석과 모의실험을 통해 무응답 오차가 갖는 특성을 파악하고자 한다. 본 연구에 사용된 자료는 지역별고용조사의 가구관리종합표에 기록된 정보이다. 지역별고용조사는 우리나라 시군 고용현황을 파악하기 위한 것으로, 이 조사자료 중 약 14만 가구의 정보를 분석에 사용하였다. 가구관리종합표에 포함된 자료는 가구단위의 정보를 포함하고 있으며, 세부적으로는 행정코드, 거주 유형, 가구방문회차별 응답현황, 가구원수, 조사대상자수, 취업자수, 실업자수 등의 정보가 기록되어 있다. 거주 유형은 주택, 아파트, 기타로 구분하였으며, 한 가구의 방문회차는 최대 4회까지로 제한되었다. 이는 조사목적 달성을 위해 조사기간에 조사를 완료해야 하기 때문에 어쩔 수 없는 제약이라고 볼 수 있다.

분석과 관련하여 본 조사의 목적이 거주 유형과 관련된 사항을 파악하고자 하는 것이 아니라는 점에 주목할 필요가 있다. 이미 서론에서 연구의 한계로서 언급하였지만, 이런 유형의 연구는 무응답가구에 대한 정보가 필요하기 때문에 행정자료나 모집단 또는 조사에서 얻은 정보를 이용하는 것이 일반적이다. 그렇지만 많은 경우, 조사목적에 부합되는 무응답가구의 정보를 찾기는 쉽지 않은 상황이다.



2. 기초분석 결과

조사연구에서 무응답 편향 측정은 매우 어려운 일이다. 실제로 무응답에 대한 정보를 파악하기 힘들기 때문이다. 특히 조사목적에 적합한 추정 변수는 파악이 거의 불가능하기 때문에 추정 변수와 연관이 높은 변수를 사용하는 것이 일반적이다. 물론 무응답 편향 측정을 위한 실험설계에 의한 것이 아니라면 완전하진 않지만 차선의 방법을 이용하여 연구를 진행하는 것은 좋은 연구방법의 대안이라 할 수 있다.

1) 주택유형과 실업률의 관계

본 연구에서는 무응답 편향 측정을 위한 핵심변수(key variable)로 주택유형(주택/아파트/기타)을 고려하였다. 우선 주택유형과 조사목적과 관련한 실업률(고용률) 간의 관계를 분석하였다. <표 3-2>는 주택유형과 고용특성과의 연관성을 파악하기 위한 것으로, 주택유형에 따라 실업률(고용률)에 차이가 있는지를 분석하였다. 이때 실업률은 연구내용 확인 수단으로 가구당 실업률을 계산하여 분석에 이용하였다. <표 3-2>의 분산분석 결과에 따르면, 유의수준 0.05에서 실업률과 고용률은 모두 주택유형에 따라 통계적으로 유의한 차이가 있는 것으로 나타났다($p=0.0001$).

<표 3-2> 주택유형별 실업률·고용률 간 차이검정

	주택유형	평균(%)	표준편차	F값	p값
실업률	주택	1.59	0.1105	55.18	0.0001
	아파트	2.25	0.1289		
	기타	2.92	0.1497		
고용률	주택	59.82	0.4101	96.19	0.0001
	아파트	56.35	0.3385		
	기타	57.78	0.3809		

2) 응답 유형과 주택유형 분포

① <표 3-3>은 주택유형별 응답분포를 나타낸다. 조사대상가구인 140,180가구 중 56.51%(79,212가구)는 주택, 34.24%(47,992가구)는 아파트, 9.26%(12,976가구)는 기타유형에 거주하는 것으로 나타났다. 전체 140,180가구 중 응답가구 비율은 78.39%, 비접촉은 5.35%, 거부는 4.33%, 불응은 10.19%로 나타났다. 불응가구 중 약 78.3%는 주택에 거주하는 가구인 것으로 나타났다.

〈표 3-3〉 전체 자료의 응답 및 무응답 원인별 주택유형 분포

응답 유형	주택		아파트		기타		전체	
	n	%	n	%	n	%	n	%
응답	61,307	77.40	38,553	80.33	10,027	77.27	109,887	78.39
비접촉	3,155	3.98	3,473	7.24	869	6.70	7,497	5.35
불능	11,185	14.12	2,031	4.23	1,067	8.22	14,283	10.19
거부	1,670	2.11	3,714	7.74	685	5.28	6,069	4.33
기타	1,895	2.39	221	0.46	328	2.53	2,444	1.74
	79,212	(56.51)	47,992	(34.24)	12,976	(9.26)	140,180	(100)

② 실제로 응답률 계산에 조사 불능가구는 포함되지 않는다. 따라서 불능가구를 제외하고 총 125,897가구에 대해 응답 유형별 응답률을 계산하면 <표 3-4>와 같다. 그 결과 응답률은 87.3%, 무응답 중 비접촉과 거부율은 각각 약 6.0%와 4.8%로 비율 측면에서는 비접촉률이 거부율보다 높게 나타났다.

③ 실제로 응답률 계산에 조사 불능가구는 포함되지 않는다. 따라서 불능가구를 제외하고 총 125,897가구에 대해 응답 유형별 비율을 계산하면 응답률은 87.3%, 무응답 중 비접촉률과 거부율은 각각 약 6.0%와 4.8%로 비접촉률이 거부율보다 높게 나타났다(<표 3-4>).

〈표 3-4〉 불능가구를 제외한 응답 및 무응답 원인별 주택유형 분포

응답 유형	주택		아파트		기타		전체	
	n	%	n	%	n	%	n	%
응답	61,307	90.12	38,553	83.88	10,027	84.20	109,887	87.28
비접촉	3,155	4.64	3,473	7.56	869	7.30	7,497	5.95
거부	1,670	2.45	3,714	8.08	685	5.75	6,069	4.82
기타	1,895	2.79	221	0.48	328	2.75	2,444	1.94
	68,027	(54.03)	45,961	(36.51)	11,909	(9.46)	125,897	(100)

④ <표 3-5>는 실제 본 연구의 분석 또는 모의실험에 사용된 자료현황을 나타낸다. 모의실험을 위해 자료는 조사기간인 2009년 10월 19일부터 10월 30일에 걸쳐 조사가 완성된 일자를 기준으로 자료를 정렬하여 사용하였다. 따라서 조사기간에 대한 입력오류가 있는 경우는 분석에서 제외하였다. 아래 <표 3-5>는 응답 유형이 불능인 경우를 제외하고, 질문항목의 응답범위가 아닌 경우를 삭제한 후의 최종 자료 분포표이다. 이후 분석은 모두 이 자료를 최종 자료로 활용하였다. 최종 116,224가구 가운데 주택 거주율은 53.93%, 아파트 거주율은 36.74%, 기타는 9.34%인 것으로 나타났다. 분석에 사용된 자료

② 무응답 유형별로 무응답 편향을 계산해 보자(식(3.3) 이용).

$$0.0184 = \frac{3,573}{116,224} \left(\frac{59,827}{107,276} - \frac{910}{3,573} \right) + \frac{5,271}{116,224} \left(\frac{59,827}{107,276} - \frac{1,909}{5,271} \right) + \frac{104}{116,224} \left(\frac{59,827}{107,276} - \frac{29}{104} \right)$$

(a)
(b)
(c)

여기서 ①은 거부에 의한 무응답 편향으로 0.0093(0.93%), ②는 비접촉에 의한 무응답 편향으로 0.0089(0.89%), ③은 기타 응답에 의한 무응답 편향으로 0.0002(0.02%)이다. 이와 같이 추정치에 대한 무응답 편향은 무응답 유형별 무응답 편향의 합으로 표현될 수 있다. 이를 토대로 전체 무응답 편향의 구성 비중을 보면, 주택 거주율이 추정 변수일 경우 거부에 의한 무응답 편향이 50.54%, 비접촉에 의한 무응답 편향이 48.37%, 기타에 의한 무응답 편향이 1.35%로 무응답 편향의 대부분은 거부 및 비접촉에 의해 발생하는 편향임을 알 수 있다(<표 3-6>). 만약 아파트 거주율이 추정 변수라면 비접촉 무응답 편향이 차지하는 비중은 45.29%, 거부에 의한 무응답 편향이 차지하는 비중은 54.69%로 주택 거주율 추정의 경우와 마찬가지로 거부에 의한 편향이 50% 이상을 차지하는 것을 알 수 있다(<표 3-6>). 즉 분석에 사용된 자료의 경우 비율 측면에서는 비접촉률이 약 6%로 거부율 4.9%보다 높지만, 편향 측면에서는 거부 편향이 더 높게 나타나고 있음을 알 수 있다. 따라서 우선순위를 둔다면 비접촉가구보다는 거부가구 그룹에 대한 성향 분석을 통해 대응방안을 찾는 것이 전체 자료의 무응답 편향을 줄이는 데 기여도가 더 클 것이다.

<표 3-6> 무응답 원인별 오차 크기

	주택		아파트		기타	
	오차크기	비율(%)	오차크기	비율(%)	오차크기	비율(%)
비접촉	0.0089	48.11	-0.0072	45.29	-0.0017	65.90
거부	0.0093	50.54	-0.0087	54.69	-0.0006	24.36
기타	0.0002	1.35	-3.86E-06	0.02	-0.0002	9.74
무응답 편향	0.0184		-0.0159		-0.0025	

* 오차의 크기가 음수인 경우는 응답자의 해당가구 유형 비율이 전체 표본이 차지하는 해당가구의 비율보다 높기 때문에 나타나는 현상으로, 그림에서는 절대값을 취해서 사용하였다.

나. 자료축적에 따른 무응답 오차 구성의 변화

이 분석은 조사기간에 걸쳐 자료의 양이 축적됨에 따라 무응답 오차 구성이 어떻게 변하는지를 살펴보기 위한 것이다. 분석 방법은 다음과 같다. 분석에 이용된 자료가 지역별고용조사라는 점을 고려하여 우리나라 전체를 크게 특광역시/시군의 2개 권역으로



임의로 나누고 각 권역 자료를 100%라 한다. 그리고 최종 응답이 이루어진 날짜를 기준으로 자료를 정렬하여 순서대로 10%씩 쪼개서 10% 슬라이스 자료를 만든다. 이 10% 슬라이스 자료를 누적하여 얻은 10개 지점에서 무응답 구성 성분별 오차를 계산한다. 본 연구에서 활용되는 자료의 권역별 비중은 특광역시 19,178가구(16.5%), 시군 97,046가구(83.5%)이다. 이 중 이미 언급한 바와 같이 최종 응답 유형 중 불응을 제외한 후의 주택 유형별 응답 및 무응답분포는 <표 3-7>과 같다.

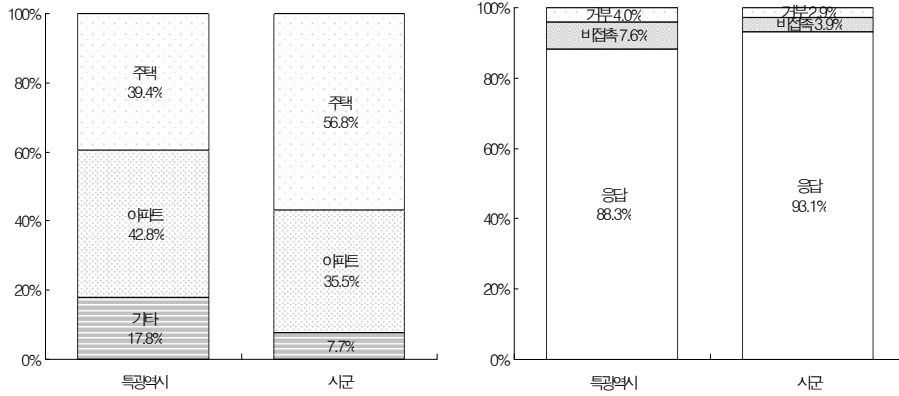
<표 3-7>에서 보면 전체적으로 시군지역의 응답률이 높고, 각 권역 내에서는 비접촉률이 거부율보다 높다. 특광역시와 같은 대도시의 경우는 주택 응답률이 시군보다 낮고, 아파트 응답률은 특광역시와 시군 모두 유사한 것으로 나타났다.

<표 3-7> 특광역시, 시군 권역의 응답 및 무응답 원인별 주택유형 분포

응답 유형	특광역시							
	주택		아파트		기타		전체	
	n	%	n	%	n	%	n	%
응답	6,777	89.69	7,138	87.04	3,028	88.51	16,943	88.35
비접촉	552	7.36	644	7.85	261	7.63	1,457	7.60
거부	227	3.00	418	5.60	126	3.68	771	4.02
기타	0	0	1	0.01	6	0.18	7	0.04
	7,556	(39.40)	8,201	(42.76)	3,421	(17.84)	19,178	(100)

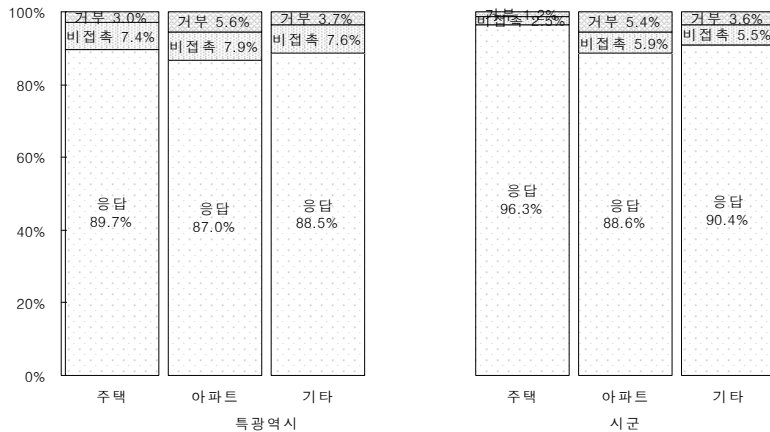
응답 유형	시군							
	주택		아파트		기타		전체	
	n	%	n	%	n	%	n	%
응답	53,050	96.25	30,565	88.60	6,718	90.41	90,333	93.08
비접촉	1,357	2.46	2,046	5.93	411	5.53	3,814	3.93
거부	683	1.24	1,849	5.36	270	3.63	2,802	2.89
기타	29	0.05	36	0.10	32	0.43	97	0.01
	55,119	(56.80)	34,496	(35.55)	7,431	(7.66)	97,046	(100)

[그림 3-3]은 위 표로부터 권역 간 응답 유형별 분포 또는 주택유형별 분포를 동시에 이해하는데 도움이 된다.



[그림 3-3] 권역별 주택유형 분포

[그림 3-3]은 <표 3-7>의 결과를 간략하게 도식화한 것이다. 이미 설명한 바와 같이 특광역시는 아파트가 주택의 비중보다 높고, 특광역시의 아파트는 시군의 아파트 비중보다 약 1.2배 크며, 시군의 주택은 특광역시의 주택 비중보다 약 1.4배 크다. 각 권역의 응답 유형별 비중은 특광역시보다 시군의 응답률이 커짐을 알 수 있다. 이는 주거형태가 주택인 경우에서 응답률이 높았던 것으로 볼 때(<표 3-5>), 주택의 비중이 높은 시군 권역에서 응답률이 높은 것으로 해석할 수 있다. 무응답 유형이 비접촉인 경우는 특광역시가 시군보다 높다. 이는 대도시일수록 조사 방문 시간에 집에 있는 사람이 소도시에 비해 적은 것으로 이해할 수 있다. 무응답 유형이 거부인 경우도 특광역시가 시군보다 높게 나타났다.



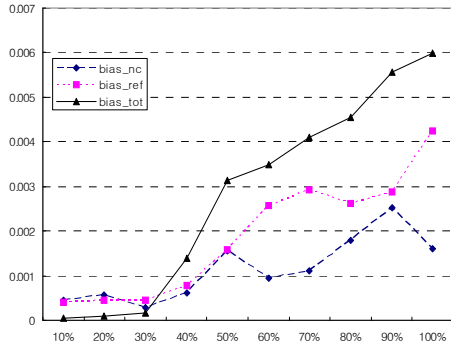
[그림 3-4] 권역 주택유형별 응답 및 무응답 유형 분포



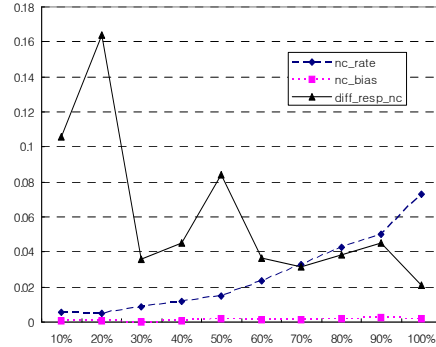
[그림 3-4]는 각 권역에서의 응답률을 주택유형별로 나타낸 것이다. 특광역시와 시군 모두 응답률은 모두 주택이 높고, 거부율은 두 권역 모두 아파트가 주택보다 더 높게 나타났다.

[그림 3-5]부터 [그림 3-10]은 특광역시와 시군에 따른 무응답률과 무응답 편향을 자료의 누적정도에 따라 나타낸 것이다. [그림 3-5]부터 [그림 3-7]은 특광역시의 경우로 우선 [그림 3-5]를 보면 자료양의 증가에 따른 특광역시의 전체 편향은 초기 30%까지는 약간 감소 혹은 안정화된 상태이다. 그 이후부터 90% 지점까지는 단계적으로 증가하는 모습이다. 90%에서는 비접촉 편향은 감소하고 총 편향과 거부 편향은 더 커지는 추세를 보인다. 분석 자료를 시간 순으로 정렬하였음을 감안할 때 특광역시 권역은 조사 종료일에 가까울수록 무응답 편향이 커진다고 볼 수 있다. 이것을 무응답 유형별로 살펴보자. [그림 3-6]을 보면 특광역시는 자료양이 많아질수록 비접촉률은 증가하는 반면, 비접촉자와 응답자 간 추정치의 차이는 증가와 감소를 반복하다가 마지막 지점에서 감소하는 경향이 있다. 이로부터 비접촉자의 성향과 응답자의 성향이 크게 차이가 없을 것으로 짐작할 수 있다. 또한 자료양의 증가에 따라 특광역시에서의 거부율은 지속적으로 증가하는 반면, 거부자와 응답자 간 추정치 차이는 증가와 감소를 반복하다가 60%지점부터 90%지점까지 감소하고 이후 약간 증가하는 경향이 있다. 60%~90% 구간에 축적된 자료의 거부자 성향은 응답자 성향과 닮아있다고 볼 수 있으나 90%지점 이후 축적된 자료의 거부자는 응답자 성향과 다르다고 볼 수 있다([그림 3-7]).

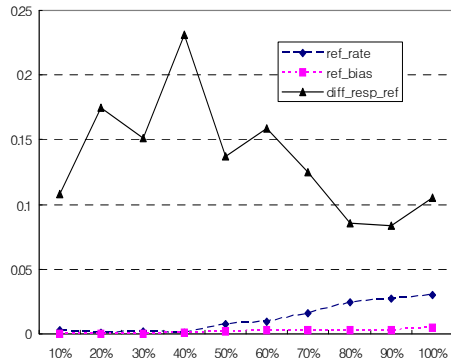
[그림 3-8]부터 [그림 3-10]은 시군의 모습을 보여준다. 우선 [그림 3-8]을 보면 전체적으로 자료가 축적됨에 따라 총 편향을 비롯한 비접촉 편향, 거부 편향이 증가하고, 자료의 양이 40%가 넘어가면서 급격하게 증가하며 자료의 모든 지점에서 거부 편향은 비접촉 편향보다 커지는 경향이 있다. [그림 3-9]를 보면 시군의 비접촉률은 증가하고, 비접촉자와 응답자 간 추정치 차이는 60% 지점까지 지속적으로 증가하다가 그 이후부터는 안정된 양상을 보인다. 이로부터 자료가 축적됨에 따라 비접촉에 의한 무응답 편향은 증가하지만, 추정치의 차이가 감소한다는 점에서 비접촉자의 성향이 응답자 성향과 닮을 가능성이 높다고 볼 수 있다. 또한 시군에서 거부율은 증가하고 거부자와 응답자 간 추정치의 차이는 누적자료의 50% 지점에서 감소하다가 다시 80% 지점에서 약간 증가하는 경향이 있다. 거부 편향은 증가추세이다. 자료양의 증가에 따라 거부율과 거부자와 응답자 간 추정치 차이 모두가 증가하는 것으로 볼 때, 거부자 성향이 응답자 성향과는 약간 다를 것으로 예상된다([그림 3-10]). 이들 그림을 통해 현장 조사기간에 걸쳐 비접촉률과 거부율은 꾸준히 증가하고, 이와 관련된 편향도 증가하는 것을 알 수 있다. 종합적으로 본 분석 결과에 따르면 자료양이 증가할수록 무응답 오차 측면에서의 품질 향상은 기대하기 어려울 것으로 보인다.



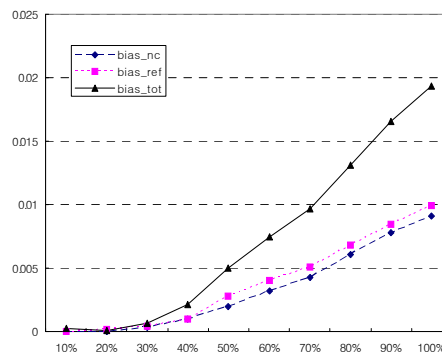
[그림 3-5] 무응답 구성성분별오차(특광역시)



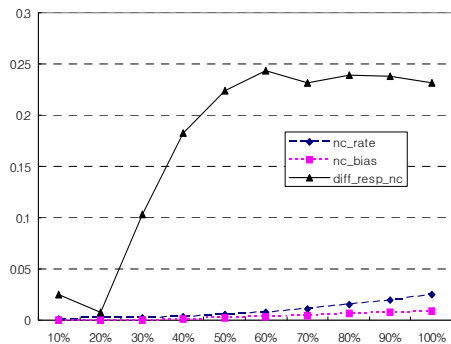
[그림 3-6] 비접촉 무응답 편향(특광역시)



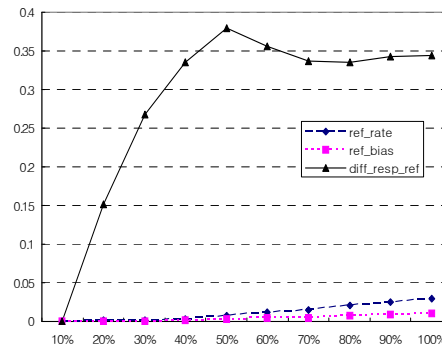
[그림 3-7] 거부 무응답 편향(특광역시)



[그림 3-8] 무응답 구성성분별오차(시군)



[그림 3-9] 비접촉 무응답 편향(시군)



[그림 3-10] 거부 무응답 편향(시군)

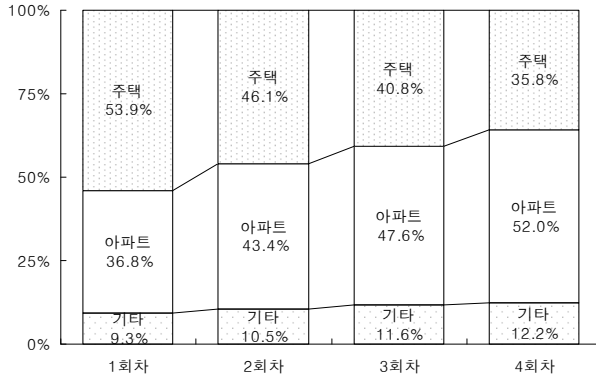


그러나 본 연구에서 사용된 자료는 조사기간이 짧고 방문회차도 최대 4번으로 제한되었기 때문에 조사 초기에 응답이 대폭적으로 이뤄졌다고 볼 수 있다. 실제로 전체 응답 중 약 50% 가량이 1회 방문에서 응답이 완성된 것으로 나타났다(3절 참고). 따라서 조사에 늦게 협조한 가구는 조사 초기에 응답한 가구와 성향이 다를 수 있다. 이처럼 응답 성향이 초기 응답자와 다른 가구가 누적 지점 50% 이상부터 자료에 대폭 쌓이면서 무응답 편향이 점점 더 커지는 상황이 발생한다고 볼 수 있다. 이는 이 조사의 특성으로서 늦게 조사에 협조한 가구들의 성향 분석을 통해 이들 가구가 전체 자료에 미치는 영향과 함께 이해하는 것이 바람직할 것이다(3절 참고).

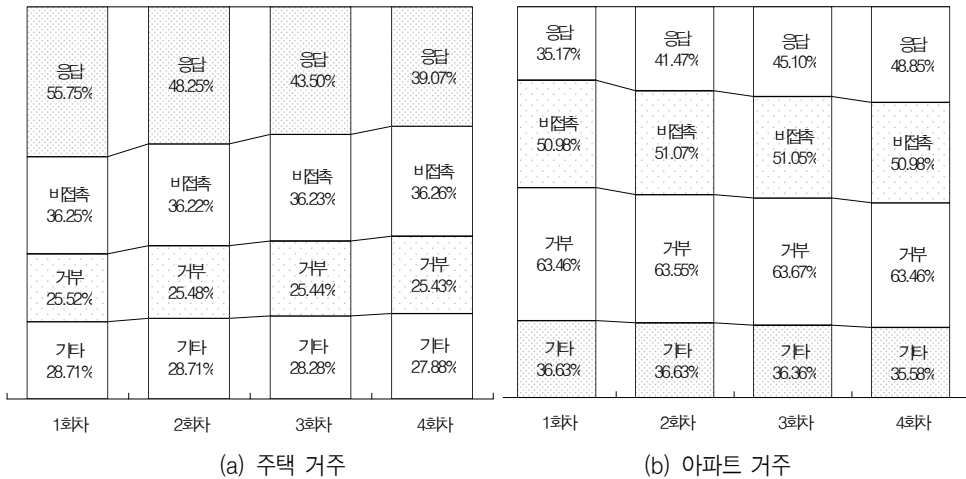
4. 방문회차에 따른 무응답 편향과 오차 구성

가구 대상의 면접조사에서 응답률을 높이기 위한 전략 중 하나가 방문횟수를 늘리는 것이다. 즉 조사원의 응답률 향상을 위한 노력의 성과는 응답전환율로서 측정될 수 있다. 여기서 응답전환이란 용어는 조사 초기에 응답하기를 거부했거나 접촉하지 못했던 가구를 두 번 또는 세 번 이상 방문함으로써 초기 조사에 비협조적인 가구가 재접촉하거나 설득에 의해 응답으로 전환되는 경우를 의미한다. 이렇게 하여 전체 조사의 응답률을 올리는데 기여할 수 있게 된다.

방문회차에 따른 무응답 편향 분석에 앞서 회차별, 주택유형별, 응답 유형별 응답률 현황을 살펴보면 다음 [그림 3-11], [그림 3-12]과 같다. [그림 3-11]을 통해 초기 방문 시 주택의 방문 비율이 더 높는데 비해 회차가 거듭될수록 격차가 줄어들면서 마지막 4회차에서는 아파트의 방문 비율이 더 높은 것을 알 수 있다. 분석에 사용된 자료는 주택 거주자가 아파트 거주자보다 많다. 이런 점으로 볼 때 아파트 거주자가 주택 거주자보다 무응답이 많이 발생하고, 그렇기 때문에 회차가 증가할수록 아파트 거주 비율이 높아진다고 볼 수 있다. 또한 거주 유형별로 방문회차에 따른 응답 유형 분포를 보면([그림 3-12]), 주택 거주자의 경우 회차가 증가할수록 응답을 얻기 어려운 반면, 아파트의 경우는 비접촉과 거부가 줄면서 회차별 응답률이 증가하는 것을 알 수 있다. 이러한 응답률 분포의 변화는 사후적으로 무응답 편향에 영향을 줄 수 있다.

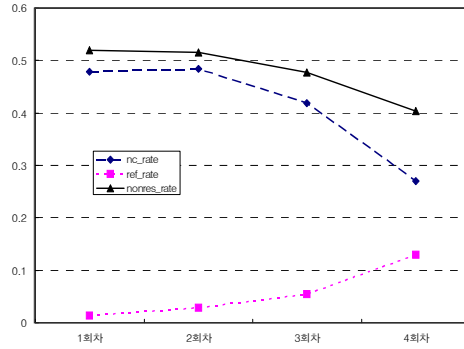


[그림 3-11] 방문회차별 주택유형 분포

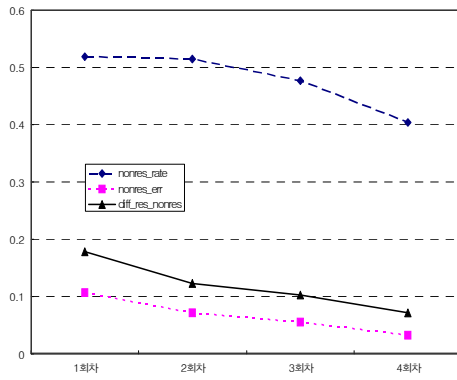


[그림 3-12] 주택과 아파트 거주자의 회차별 응답 비중

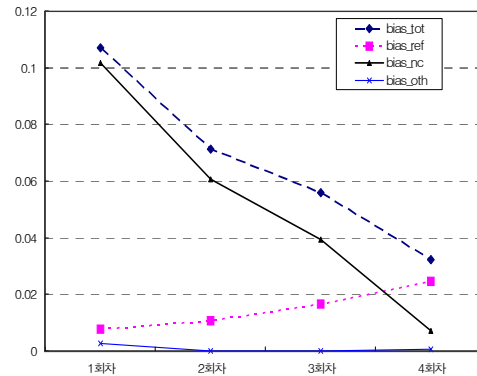
[그림 3-13]은 방문횟수가 많아질수록 전체 무응답률은 감소한다(51.9%에서 40.3%). 무응답 유형별로 보면 방문횟수가 많아지면 비접촉률은 감소하고 거부율은 증가하는 것을 알 수 있다. 이 결과는 가구를 더 많이 접촉하면 할수록 더 많은 가구로부터 응답을 유도할 수 있지만(무응답률 감소), 접촉이 이루어진 많은 가구들에서 조사에 협조하기를 거부하는 가구도 증가한다(거부율 증가)는 사실을 보여준다. 즉, 순수하게 맞벌이나 일시적으로 조사기간 중 부재로 인해 접촉할 수 없었던 가구들에 대해서는 조사원들의 꾸준한 재방문 노력은 추가적인 응답을 얻어내는데 매우 효과적인 방법이 될 수 있다. 그렇지만 그러한 가구들 중에는 여전히 조사에 부정적이거나 비협조적인 가구가 포함되어 있어서 접촉에는 성공하더라도 조사를 거부하게 될 가능성도 무시할 수 없다.



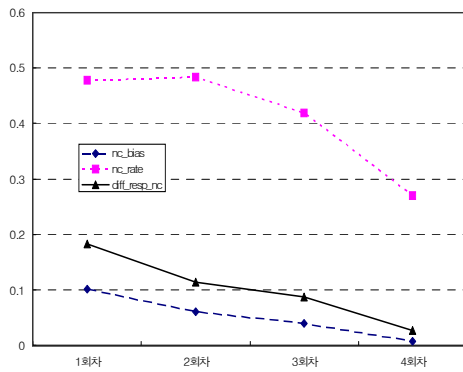
[그림 3-13] 방문회차에 따른 전체 무응답률, 비접촉률, 거부율, 기타 무응답률



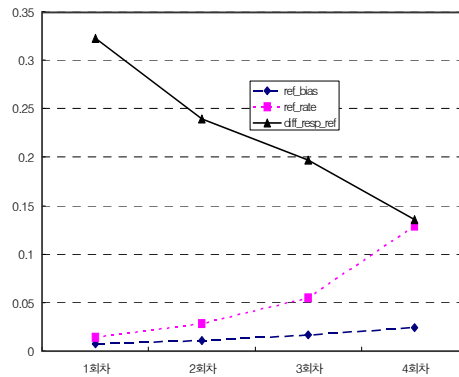
(a) 무응답 편향



(b) 무응답 편향 분해



(c) 비접촉 편향



(d) 거부 편향 분해

[그림 3-14] 방문회차별 무응답 편향

[그림 3-14]는 방문횟수에 따른 무응답 편향을 나타낸다. 그림에서 (a)는 방문횟수별로 응답자와 무응답자 간의 차이(diff_res_nonres)를 표시함으로써 무응답률과 무응답 편향과의 관계를 보여준다. 2번 이상 재방문 시도에 의해 무응답 편향은 0.07에서 0.03으로 감소한다. 이는 주택에 거주하는 가구 비율 추정치가 만약 2번만 방문한다면 7% 포인트 과대 추정되고, 4번 방문을 하더라도 여전히 3% 포인트 과대 추정된다는 것을 의미한다. 이는 재접촉시도를 통해 약 57.1% 정도 과대 추정을 줄일 수 있다는 것을 나타낸다.

[그림 3-14]의 (b)는 방문횟수에 따라 발생하는 무응답 편향의 구성 변화를 나타낸다. 이 그림은 방문횟수가 비접촉에 의한 무응답 편향에 큰 영향을 준다는 것을 보여준다. 거부 편향은 회차가 증가함에 따라 비접촉 편향에 비해 상대적으로 적은 영향을 받는 것을 알 수 있다. [그림 3-14]의 (c)와 (d)의 비접촉과 거부 편향에 대한 그림도 (a) 또는 (b)와 비슷한 방법으로 해석할 수 있다. 즉, 방문횟수가 증가하면 비접촉률은 감소하고 응답자와 비접촉자 간 차이도 감소하고, 따라서 비접촉 편향은 점점 감소하는 경향이 있다. 거부 편향은 4회 방문까지 지속적으로 증가하는 경향이 있다. 방문횟수가 늘어날수록 응답자와 거부자 간 추정치 차이가 현저히 감소하는데 불구하고 거부 편향이 증가하는 것은 그만큼 거부율이 거부 편향에 강하게 영향을 주기 때문으로 이해할 수 있다.

결과적으로 방문횟수의 증가는 전체적으로는 무응답 편향에 영향을 주고 있으며, 세부적으로는 비접촉 무응답의 편향은 줄어드는 반면, 거부 편향은 증가하는 경향이 있다. 현재 4회차 방문횟수가 그렇게 많은 것이 아니라는 점을 고려하더라도, 이 결과만으로 회차의 증감의 필요성을 논하기는 어려울 것이다. 하지만 최소한 방문횟수의 증가를 통해 무응답 오차를 줄일 수 있을 것으로 예상된다. 이와 동시에 거부자의 경우 거부율이 거부 편향에 강한 영향을 준다는 점에서 거부자 설득을 위한 체계적인 전략이 필요할 것으로 보인다. 한편으로 이 거부가구의 성향을 분석하고, 이들 거부 성향이 강한 가구가 전체 자료에는 어떻게 반영되고 있는지를 살펴봄으로써 실현가능한 대응 전략을 모색할 수 있을 것으로 판단된다.

5. 응답전환 분석

지금까지는 무응답 편향과 무응답률과의 관계를 분석하는데 초점을 맞추었다. 그 결과 전체적인 무응답 편향은 자료양의 증가에 따라 커지고, 방문횟수의 증가에 따라서는 감소하였다. 늦게 응답하는 가구와 초기 응답하는 가구는 그 성향이 다를 수 있고, 이는 응답전환 가구들의 특성 분석을 통해 알 수 있다. 우선 초기 응답자(1회차 응답가구)와 마지막 응답자(4회차 응답가구)를 비교함으로써 응답전환가구의 성향을 파악하고자 하였다. 분석에 사용된 자료는 최종 응답결과가 불능인 가구를 제외하기 전인 118,039가구



의 자료에서 각 회차별 응답 결과가 불능이거나 기타인 가구를 제외하였다(114,763가구). 이 분석에 사용된 자료와 슬라이스별 분석을 수행했던 자료(116,224가구)와의 차이점은 슬라이스 자료는 최종 응답이 불능인 가구를 제외한 것이고, 응답전환 분석을 위한 이 자료는 각 방문회차에서 응답 결과가 불능 및 기타인 가구를 제외한 것이다.

가. 응답전환 현황

<표 3-8>에 따르면 1회차에 응답으로 조사가 완성된 가구는 전체(114,763) 중 48,603(42.35%)이고, 1회차에 무응답이었던 가구가 2회차에 응답으로 조사가 완성된 가구는 26,126(25.38%), 동일한 방법으로 1, 2회차에서 무응답이고 3회차에 응답으로 조사가 완성된 표본은 17,954(15.64%)이다. 마지막으로 4회차에서 응답으로 조사가 완성된 경우는 10,453(9.11%)가구이다. 이러한 무응답 상황을 무응답 층별로 구분하여 최종 응답현황 비율을 살펴보자.

<표 3-8> 방문회차별 응답률과 응답전환율

응답률 (총수)	응답전환율			
	1회차 응답률	2회	3회	4회
92.48 (103,136)	42.35 (48,603)	25.38 (26,126)	15.64 (17,954)	9.11 (10,453)

<표 3-9> 무응답 층의 최종 응답전환율

1회차	2회차	3회차	4회차	표본수	%1	전체 %2) **
비접촉	비접촉	비접촉	응답	10,014	95.80	8.73
		거부		118	1.13	0.10
	거부	비접촉		73	0.70	0.06
		거부		59	0.56	0.05
거부	비접촉	비접촉		98	0.94	0.09
		거부		16	0.15	0.01
	거부	비접촉		22	0.21	0.02
		거부		53	0.51	0.05
총계				10,453	100	-
전체 총계				114,763	-	9.11

- 1) 4회차에 응답전환된 10,453가구에 대한 각 무응답그룹이 차지하는 비율 - 4회차에 응답전환된 가구 10,453가구 중 3번 연속 비접촉이다가 4회차에서 응답으로 전환된 가구 비율은 95.80%에 해당된다.
- 2) 총 조사대상자 114,763가구에 대한 각 무응답 그룹이 차지하는 비율 - 총 조사대상가구 중 4회차에 응답전환된 가구는 10,453가구로 전체의 8.73%에 해당되고, 그 중 3번 연속 접촉하지 못하다가 4회차가 접촉하여 응답을 얻어낸 가구는 전체의 약 8.73%(10,014가구)에 해당된다.

<표 3-9>는 1회차에 무응답(비접촉 또는 거부)하고, 마지막 4회차에서 응답을 한 경우에 대해 4회차에 걸쳐 나올 수 있는 모든 응답 유형의 조합을 정리한 것이다. 다시 말하면 방문 4회차에 와서 응답전환된 표본이다. 응답전환은 4회차에 이르지 않고도 발생할 수 있으나, 위 표에서는 4회차에 와서 응답전환된 경우로 한정하였다. 이는 지속적으로 동일 패턴의 응답 가구 유형을 찾고, 그것이 차지하는 비중을 살펴보기 위함이다.

위와 같이 8개 응답 조합에서 4회차에 응답전환된 비율은 전체의 9.11%이다. 표에서 보면 1, 2, 3회차 모두 비접촉한 가구를 4회차 방문에서 응답으로 전환된 가구는 10,014 가구로 4회차 응답가구의 95.8%를 차지한다. 그리고 4회 방문 동안 한 번도 접촉을 하지 못한 가구는 5,048가구로 전체 114,763가구의 4.4%를 차지한다(<표 3-10>). 한편 3회까지 모두 조사에 협조하지 않다가 마지막 4회차에서 응답한 가구는 53가구로 응답전환이 매우 어려운 조합이라 할 수 있다. 또한 4번 모두 접촉은 하였으나 모두 거부한 가구도 764가구에 해당된다(<표 3-10>). 전체적으로 보면 조사 초기의 비접촉가구는 재차 방문을 통해 응답을 얻어낼 가능성이 있는 반면, 처음부터 거부하거나 접촉 중에 1번이라도 조사를 거부한 가구는 응답전환이 어려울 것으로 보인다. 이런 가구들이 조사를 왜 거부하는지에 대한 이유를 조사하고, 이러한 분석을 통해 무응답층별로 차별화된 응답전환 전략을 세우고 적용하는 것이 응답률 향상에 도움이 될 것이다.

<표 3-10>는 방문 3회차까지 비접촉 혹은 거부로 일관한 가구들의 최종 응답현황을 나타낸 분포이다. 3회차까지 지속적으로 접촉할 수 없었던 가구(15,062) 중 33.51%는 4차 방문에도 비접촉이었으며, 66.5% 가구는 최종 응답을 하였다. 한편 3회차까지 지속적으로 거부한 가구(817) 중 6.5%만이 최종 응답으로 전환되었고 나머지 가구는 여전히 거부한 채로 남겨졌다.

<표 3-10> 지속적 무응답가구들의 마지막 회차에서의 응답 분포

방문 1-3회차		4회차	빈도	그룹 내 %	전체 %	
비접촉그룹	비접촉-비접촉-비접촉	응답	15,062	10,014	66.49	63.06
		비접촉		5,048	33.51	31.79
거부그룹	거부 - 거부 - 거부	응답	817	53	6.49	0.33
		거부		764	93.51	4.81
합계			15,879		100	

<표 3-11>은 방문 3회차까지 무응답(비접촉/거부)으로 일관한 가구들의 거주 유형 분포를 나타낸다. 전체적으로 방문 3회차까지 무응답인 가구는 주택보다 아파트 비율이 더 높다(38.16%<49.69%). 4회차 응답상황에 따라 3회차까지 한 번도 접촉하지 못했던 가



가구가 4회차에 응답(비접촉-비접촉-비접촉-응답)한 비율을 거주 유형별로 보면 주택보다 아파트가 더 높다(39.62%<48.35%). 마찬가지로 3회차까지 계속 거부했던 가구가 최종 응답으로 전환(거부-거부-거부-응답)되는 비율은 역시 주택보다 아파트가 더 높다(41.51%<50.94%). 통상적으로 주택보다 아파트에 젊은 사람들이 많이 거주하며 이들은 주로 맞벌이가구가 많아 접촉하기 어렵거나, 혹은 어르신들에 비해 상대적으로 사생활 보호 등의 이유로 조사에 대한 거부심이 높을 수도 있음을 생각해볼 수 있겠다. 방문 4회차 모두 비접촉인 경우는 주택보다 아파트가 더 높으며(36.63%<50.59%), 4회차 모두 거부인 경우 역시 주택보다 아파트가 더 높은 것으로 나타났다(28.93%<61.13%).

<표 3-11> 회차별 응답의 주택유형 분포

가구수(%)

유형	비접촉-비접촉-비접촉-응답	비접촉-비접촉-비접촉-비접촉	거부-거부-거부-응답	거부-거부-거부-거부	총합
주택	3,968(39.62)	1,849(36.63)	22(41.51)	221(28.93)	6,060(38.16)
아파트	4,842(48.35)	2,554(50.59)	27(50.94)	467(61.13)	7,890(49.69)
기타	1,204(12.02)	645(12.78)	4(7.55)	76(9.95)	1,929(12.15)
총합	10,014(63.06)	5,048(31.79)	53(0.33)	764(4.81)	15,879(100)

<표 3-10>과 <표 3-11>의 연속 4차까지의 응답현황으로부터 최종 비접촉자는 계속 비접촉이었다가 최종 응답한 가구들과 유사하고, 연속 4번 거부한 가구들은 계속 거부하다가 마지막에 응답한 가구들과 유사한 성향을 가졌을 것으로 짐작할 수 있다. <표 3-11>에서 알 수 있듯이 연속 3회 접촉하지 못한 비접촉 그룹의 거주 유형은 주택과 아파트 거주율이 거의 유사한 것을 알 수 있다. 한편 연속 3회 거부 후 최종 응답전환 한 가구의 주택 거주율은 최종 거부한 가구보다 다소 높다(41.51%>28.93%). 한편 연속 3회 거부한 가구들도 최종 응답그룹의 주택 거주율이 최종까지 응답거부한 가구보다 다소 높다(41.51%>28.93%). 전체적으로 보면 아파트 거주비율이 두 그룹 모두에서 50% 이상을 차지하고 상대적으로 주택 거주율이 낮다는 점에서 이들 그룹의 거주 성향이 비슷하다고 볼 수 있다. 가구 내 가구원 누군가를 만나기 위해 계속적으로 접촉을 시도하였으나 접촉할 수 없었던 가구들의 특성은 도시와 비도시지역에 따라 차이가 있을 것으로 보인다. 즉 지속적 대도시지역의 비접촉 가구는 일상이 바쁜 젊은 세대이거나 맞벌이가구일 가능성이 높고, 농어촌 지역의 비접촉 가구는 노인가구로서 집을 자주 비우거나 농사일에 바쁜 가구일 가능성이 높다고 볼 수 있다. 물론 이 자료만으로 이에 대한 명쾌한 해석은 어렵다 하더라도 최소한 이러한 가능성이 있다는 사실을 충분히 이해할 수 있는 부분이라 할 수 있다.

나. 응답전환기구의 성향 분석

<표 3-12>는 본 자료에서 파악 가능한 무응답가구의 특성을 응답 회차에 대해서 정리한 결과이다.

<표 3-12> 응답전환가구 분포

	응답을 얻은 차수				전체응답전환	전체응답
	방문 1회차	방문 2회차	방문 3회차	방문 4회차		
응답자수	48,603 45.31%	29,579 27.57%	18,342 17.10%	10,753 10.02%	58,673 54.69%	107,276 100%
주택	31,516 64.84%	15,657 52.93%	8,453 46.09%	4,202 39.08%	28,311 48.25%	59,827 56%
아파트	13,374 27.52%	11,211 37.90%	7,866 42.89%	5,252 48.84%	24,329 41.47%	37,703 35%
기타	3,713 7.64%	2,711 9.17%	2,023 11.03%	1,299 12.08%	6,033 10.28%	9,746 9%
실업률	1.88%	1.92%	2.05%	2.22%	2.02%	1.96%
1인 가구수	12,211 25.12%	6,919 23.39%	4,600 25.08%	3,004 27.94%	14,523 24.75%	26,819 25%
15세 이상 가구수	48,495 99.78%	29,521 99.80%	18,299 99.77%	10,700 99.51%	58,519 99.74%	107,545 100%
특별시	1,987 4.09%	1,711 5.78%	1,262 6.88%	934 8.69%	3,907 6.66%	5,894 5.49%
광역시	3,568 7.34%	3,360 11.36%	2,404 13.11%	1,717 15.97%	7,481 12.75%	11,049 10.30%
시	19,109 39.32%	13,762 46.53%	8,892 48.48%	5,542 51.54%	28,195 48.05%	47,304 44.10%
군	23,939 49.25%	10,746 36.33%	5,784 31.53%	2,560 23.81%	19,090 32.54%	43,029 40.11%

<표 3-12>에 보면 조사기간에 걸쳐 1번도 무응답하지 않고 조사 초기에 조사를 완성한 응답률은 전체 자료(116,224)의 41.82%로 1번 이상 접촉을 못하거나 거부한 후 응답한 가구의 전환율 50.48%에 비해 낮다. 이는 조사대상 가구의 50% 이상이 조사기간 중 최소한 1번 이상은 조사를 거부하거나 만날 수 없는 가구들로 조사가 그만큼 어렵게 진행된다는 것을 의미한다. 주택유형별로 보면 늦게 응답하는 가구의 주택유형별 거주 분포는 초기에 응답하는 가구의 주택유형과 차이가 있음을 알 수 있다. 즉 1차 방문에서 응답한 가구의 주택 거주율은 64.84%, 아파트 거주율은 27.52%로 주택이 2배 이상 높다. 시간이 지남에 따라 이 거주율의 차이는 점점 줄어드는 경향이 있으며 마지막 4차에서 응답전환된 가구 비율은 오히려 아파트 거주율이 48.84%로 주택 거주율 39.08%보다 높았다. 실업률의 경우 조사 초기에 협조한 가구들은 1.88%인 반면, 4차에 조사에 협조한



가구는 2.22%로 늦게 응답한 가구일수록 실업률이 높은 것을 알 수 있다. 1인 가구 비율은 큰 차이는 없지만 4회차에 응답한 1인 가구의 비율이 27.94%로 전체 응답한 가구 중 1인 가구 비율인 25%보다 약간 높은 경향이 있는 것으로 나타났다. 가구 내 15세 이상 가구원의 비율은 회차와는 무관한 것으로 나타났다. 종합하면 아파트가구 또는 1인 가구일수록 늦게 응답하는 경향이 있고, 늦게 응답한 가구는 초기에 응답한 가구에 비해 실업률이 높은 성향이 있는 가구일 확률이 높다고 볼 수 있다. 이 표에서는 설명하지 않았지만 1인 가구는 군지역의 주택에 거주하는 경향이 있고, 이러한 가구라면 많은 경우 농어촌 노인 가구일 가능성이 높다고 짐작할 수 있다.

<표 3-13>은 주택유형과 특별시, 광역시, 시, 군지역별로 나누어 거부횟수와 비접촉 횟수에 따라 가구의 특성을 살펴본 것이다.

<표 3-13> 거부 및 비접촉횟수에 따른 분포(주택유형/4개 권역별)

	거부				비접촉				
	1번	2번	3번	4번	1번	2번	3번	4번	
주택유형별	주택	716 29.08%	285 26.71%	160 25.04%	221 28.93%	15,578 51.92%	8,495 44.73%	4,346 37.71%	1,849 36.63%
	아파트	1,445 58.69%	658 61.67%	426 66.67%	467 61.13%	11,670 38.90%	8,382 44.13%	5,772 50.09%	2,554 50.59%
	기타	301 12.23%	124 11.62%	53 8.29%	76 9.95%	2,753 9.18%	2,116 11.14%	1,406 12.20%	645 12.78%
	합계	2,462	1,067	639	764	30,001	18,993	11,524	5,048
권역별	특별시	329 13.36%	94 8.81%	51 7.98%	51 6.68%	1,758 5.86%	1,374 7.23%	1,111 9.64%	579 11.47%
	광역시	319 12.96%	155 14.53%	60 9.39%	68 8.90%	3,426 11.42%	2,462 12.96%	1,692 14.68%	810 16.05%
	시	1,442 58.57%	665 62.32%	433 67.76%	495 64.79%	14,233 47.44%	9,413 49.56%	6,063 52.61%	2,672 52.93%
	군	372 15.11%	153 14.34%	95 14.87%	150 19.63%	10,584 35.28%	5,744 30.24%	2,658 23.06%	987 19.55%
합계	2,462	1,067	639	764	30,001	18,993	11,524	5,048	

<표 3-13>에서 알 수 있듯이 거부횟수가 많아지면 주택 거주율은 약간 감소하는 반면, 아파트 거주율은 58.69%에서 61.13%로 약간 증가하는 것으로 나타났다. 지역별로 보면 시지역은 1번 거부한 가구 비율은 58.57%, 4번 거부한 가구 비율은 64.79%로 거부횟수가 많은 가구의 비율이 점점 증가하는 것을 알 수 있다. 군지역도 시와 유사하게 1번 거부에서 4번 거부한 가구의 비율은 15.11%에서 19.63%로 약 4.52% 차이가 있는 것으로 나타났다. 특광역시는 거부를 많이 하는 가구의 비율이 시간이 지남에 따라 감소하는 경향이 있는 것으로 나타났다. 비접촉의 경우는 비접촉횟수가 1번에서 4번으로 증가하면

아파트 거주자의 비접촉 비율은 38.90%에서 50.09%로 크게 증가하는 경향이 있다. 지역별로는 군지역은 비접촉횟수가 1번에서 4번으로 증가하면 비접촉률이 35.28%에서 19.55%로 대폭 감소하는 반면, 다른 지역들에서는 지속적으로 증가하는 것으로 나타났다. 종합하면 아파트 거주자는 주택 거주자에 비해 거부나 비접촉횟수가 많고, 특히 거부횟수에 비해 비접촉횟수가 훨씬 많다. 또한 시지역은 특광역시에 비해 거부횟수별 거부가구의 비율에 대한 증가폭이 크고, 군지역은 비접촉횟수별 비접촉가구 비율이 다른 지역들에 비해 감소하였다. 즉 군지역은 농어촌 지역으로 주택표본이 시지역 이상에 비해 많고, 이들 농어촌 주택 거주자는 노인가구일 가능성이 높다. 따라서 비접촉이나 거부 가능성이 상대적으로 낮을 것으로 예상된다. 한편 시 이상의 도시지역들은 아파트가구가 많고, 아파트가구는 농어촌 지역에 비해 상대적으로 젊은 맞벌이 가구들이 많을 가능성이 높은 관계로, 집에 사람이 없거나 있더라도 분명한 사유에 의해 거부할 가능성이 높을 것이다.

이미 앞에서 분석한 바에 의하면, 무응답 편향에 대해서는 비접촉에 의해 무응답가구는 응답가구의 성향과 유사하다고 볼 수 있다는 결과를 제시하였다. 최종 가구는 거부한 경험이 있는 가구와 그 성향이 유사하다는 전제하에 이 거부가구들이 전체 실업률 추정치에 어떤 영향을 미치는 살펴보기로 한다. <표 3-14>는 1번 이상 거부한 후에 응답으로 전환된 가구와 1번 이상 접촉이 안 된 이후에 응답을 얻게 된 가구들의 가구 평균 실업률을 나타낸 것이다. 전체 응답한 가구들의 가구 평균 실업률은 1.96%이다.

<표 3-14> 거부/비접촉 후 응답전환된 가구의 실업률 및 고용률

	()번 거부 후 응답전환된 가구			()번 비접촉 후 응답전환된 가구		
	1번	2번	3번	1번	2번	3번
실업률(%)	2.43	1.54	5.95	1.91	2.09	2.17
고용률(%)	54.54	54.69	45.19	58.90	59.95	61.20

<표 3-14>에서 보면, 1번 거부한 경험이 가구들의 평균 실업률은 2.43%, 2번 거부한 가구들은 1.54%, 3번 연속 거부한 가구들의 평균 실업률은 5.95%로 거부횟수가 많은 가구에서 실업률은 전체 평균보다도 2배 이상 높게 나타나는 것을 알 수 있다. 비접촉의 경우는 4번 방문 중 1번만 접촉이 안된 가구들의 평균 실업률은 1.91%, 2번은 2.09%, 3번 연속 접촉이 안된 가구는 실업률이 2.17%로 전체 평균과 크게 차이가 없는 것을 알 수 있다. 고용률에 있어서는 3번 연속 거부한 가구들의 평균 고용률은 45.19%로 1, 2번 거부한 가구들에 비해 약간 낮고, 비접촉가구들의 고용률은 비접촉횟수가 많을수록 약간 증가하는 경향이 있지만 큰 차이는 아니라고 볼 수 있다(58.90%에서 61.20%). 따라서



전체적으로 비접촉가구에 비해 거부 경험이 있는 가구들의 실업률이 높고, 고용률이 낮게 추정되는 경향이 있는 것으로 나타났다. 또한 여러 번 방문을 통해서도 잘 접촉할 수 있는 가구의 고용률이 증가하는 것으로 보아, 이들 가구들은 맞벌이 가구이거나 취업상태로 인해 가구를 접촉하지 못했을 가능성이 높을 것으로 예상된다.

그렇다면 이 거부 경험이 있는 가구들의 응답이 전체 자료에는 어떤 영향을 미칠 수 있을 것인가에 대해 분석해 보기로 하자. 거부가구들이 자료에 미치는 효과를 살펴보기 위하여 우선 거부가구들은 최소한 거부 경험이 있는 가구들의 성향과 유사하다고 가정한다. 실제로 전체 자료에서 최종 거부에 의해 무응답이 발생한 거부율은 3.1%(약 3,602 가구)로 전체 자료에서 차지하는 비중은 그렇지 높지 않다고 볼 수 있다. 그러나 이 가구들을 설득하기 위한 끊임없는 노력의 양은 측정하기는 어렵지만 상당히 소비적인 작업이 될 수 있을 것이다.

<표 3-15> 2차 이상 방문에서 무응답가구들을 제외했을 때의 실업률

	응답전환자 제외		
	2차 이상 방문	3차 이상 방문	4차 방문
실업률(%)	1.88	1.90	1.93
총수(명)	37,257	61,105	76,106

* 전체 실업률 = 1.92%

<표 3-15>는 2차 이상 방문에서 응답전환된 가구들을 전체 자료에서 제외했을 때의 실업률을 나타낸다. 이 표에서 응답전환된 가구들은 해당 회차 이전에 거부 또는 비접촉에 의해 무응답이었던 가구들을 의미한다. 2차 이상 방문에서 응답전환된 가구들을 제외한 가구들의 평균 실업률은 1.88%로 전체 실업률 1.92%보다 0.04% 포인트 낮다. 최종적으로 마지막 4차 방문에서 응답전환된 가구들을 제외한 나머지 76,106 가구들에 대한 평균 실업률은 1.93%로 전체 응답률과 거의 유사한 것을 알 수 있다. 이처럼 늦게 응답전환된 가구들이 전체 자료에서 제외되면 실업률은 초기 응답자들에 비해 약간씩 증가하는 경향이 있다. 이는 비접촉가구들의 성향이 응답가구들의 성향과 유사하다고 볼 때, 양적으로 많은 비접촉가구들이 3차까지에서 많이 탈락되면서 전체적인 응답성향에 영향을 준다고 볼 수 있다. 실제로 마지막 4차에서 약 10,000가구 이상(<표 3-9>)이 응답으로 전환되었고, 이 가구들의 실업률은 전체 실업률보다 약간 높다. 이런 가구들이 전체 자료에서 대거 탈락되면서 전체 실업률에 영향을 미칠 수 있을 것으로 예상된다.

그러나 거부가구들은 그 성향이 응답자 성향과 다르다고 볼 수 있고, 전체 자료에서 차지하는 비중이 약 3% 정도로 그렇게 많지 않은 숫자이다. 실제로 이들 거부가구들을

제외하고 실업률을 추정하면, 전체 실업률과 거의 변화가 없음을 알 수 있다(<표 3-16>). 다음의 <표 3-16>은 1번 이상 거부한 경험이 있는 가구들을 제외하였을 때의 각 실업률을 나타낸다.

<표 3-16> 1번 이상 거부 경험이 있는 가구를 제외하였을 때의 실업률

	1번 이상 거부 제외	2번 이상 거부 제외	3번 이상 거부 제외
실업률(%)	1.96	1.96	1.96

<표 3-16>은 전국 단위의 가구 평균 실업률로서, 전국적으로는 초기 응답가구보다 실업률이 높은 거부가구를 제외하여도 전체 실업률에는 전혀 영향을 미치는 것은 보여주지 않는다. 그러나 분석에 사용된 조사는 지역통계생산을 위한 조사로 지역별 추정치에 거부가구가 미치는 영향력을 살피는 것이 더 중요할 수 있다. <표 3-17>과 <표 3-18>은 각 방문화차에서 거부가구를 제외하였을 때의 시도 수준의 가구 평균 실업률과 시군 수준의 가구 평균 실업률을 나타낸 것이다. <표 3-18>은 전국 158개 시군지역 중 일부 지역에 대해서 제시한 것으로 모든 시군구에 대한 결과는 <부록>에 제시하였다.

<표 3-17> 16개 시도에서 거부 경험이 있는 가구를 제외했을 때의 실업률

지역	1번 이상 거부 제외	2번 이상 거부 제외	3번 이상 거부 제외	전체 실업률
A	1.00	1.01	1.01	1.01
B	2.47	2.48	2.47	2.47
C	1.62	1.62	1.62	1.62
D	1.55	1.56	1.55	1.55
E	4.06	3.95	3.92	3.92
F	3.08	3.25	3.24	3.23
G	2.73	2.72	2.72	2.72
H	3.76	3.70	3.71	3.71
I	3.14	3.12	3.12	3.13
J	3.19	3.14	3.14	3.14
K	3.95	3.97	3.98	3.98
L	1.01	1.01	1.01	1.01
M	1.52	1.51	1.51	1.53
N	1.41	1.44	1.43	1.43
O	1.83	1.82	1.82	1.82
P	1.71	1.73	1.72	1.72

〈표 3-18〉 158개 시군에서 거부 경험이 있는 가구를 제외했을 때의 실업률(일부)

지역	1번 이상 거부 제외	2번 이상 거부 제외	3번 이상 거부 제외	전체 실업률
A	1.75	1.72	1.72	1.71
B	1.67	1.66	1.66	1.66
C	2.30	2.29	2.27	2.27
D	2.56	2.53	2.53	2.52
E	3.16	3.14	3.14	3.14
F	3.51	3.41	3.38	3.38
G	2.64	2.54	2.53	2.52
H	3.08	3.13	3.13	3.13
I	2.92	2.93	2.91	2.89
J	1.10	1.09	1.09	1.09
K	2.54	2.50	2.49	2.49
L	2.50	2.49	2.49	2.49
M	1.75	1.75	1.75	1.75
N	2.73	2.69	2.68	2.68

먼저 <표 3-17>에서 보면 M지역을 제외한 15개 시도 모두에서 3번 이상 거부한 가구들을 제외해도 실업률은 전체 실업률과 거의 동일한 것을 알 수 있다. 2번 이상 거부한 가구들을 제외하더라도 거의 대부분 지역에서 각 지역의 전체 실업률과 동일하거나 0.01% 포인트의 아주 작은 차이를 보였다. 그렇지만 1번 이상 거부한 가구들을 제외하였을 때는 거의 대부분 지역에서 각 지역의 전체 실업률과 약간의 차이를 보이는 것을 알 수 있다. M지역은 다른 15개 시도 지역과는 약간 다른 결과를 보인다. M은 거부가구들을 제외하면 전체 실업률보다 오히려 커지는 경향이 있다. 이와 같은 현상은 시군지역의 결과에서도 비슷하게 나타난다(<표 3-18>). 시군지역은 시도처럼 명백하게 횡수에 의해 구분되는 것은 아니지만, 대부분 지역에서 3번 이상 거부한 가구를 제외하더라도 전체 실업률에는 영향을 미치지 않고, 2번 이상 거부가구들의 경우는 경우에 따라서는 약간씩 불안정하게 나타나는 지역들도 있다. 명백한 것은 최소한 1번 이상 거부한 가구들을 제외하게 되면, 각 시군에서의 실업률은 전체 실업률에 대해 큰 차이를 보일 수 있다는 사실이다.

결과적으로 비접촉가구는 응답가구들과 성향이 유사하고, 이 비접촉가구들이 늦게 응답전환이 되더라도 이 가구들의 전체적인 응답성향은 초기 응답가구들의 성향과 큰 차이를 보이지 않는다. 또한 이 가구들의 전반적인 실업률이 전체 응답가구들의 실업률 보다는 약간 높고, 이러한 가구들이 수적으로 많기 때문에 이들 가구들은 전체적인 자료에 어느 정도 영향을 미칠 수 있다고 보인다. 한편 거부가구들은 응답가구들과 응답 성



향이 다르고, 거부 경험이 있는 가구들의 실업률은 전체 응답가구들의 실업률보다 훨씬 높고, 거부횟수가 많은 가구일수록 실업률이 매우 높았다. 하지만 이 거부가구들은 수적으로 전체가구에서 차지하는 비중이 낮고, 거부가구들이 전체 자료의 추정치에 미치는 영향은 그다지 크지 않다.

제4절 결론 및 논의

지금까지 조사통계 작성에 있어 무응답률과 무응답 편향을 중심으로 무응답가구의 성향을 파악하고, 이로부터 무응답가구에 대한 대처 방안을 찾고자 하였다. 본 연구의 목표는 가구를 대상으로 하는 면접조사에서 크게 두 가지 효과를 측정하는 데 있다. 첫째, 자료의 양과 무응답 편향은 어떤 관계가 있는지, 둘째, 재접촉 또는 거부전환 노력이 자료의 질적·양적 측면에 어떤 영향을 미치는지를 파악하는 것이다. 본 연구에 사용된 조사는 최대 4번의 재접촉시도를 허용하고 있고, 무응답 편향을 측정하기 위한 핵심변수로는 주택 거주율을 사용하였다. 응답률은 조사자료의 품질을 측정할 수 있는 주요 지표로서 사용되고 있다. 일반적으로 확률표본을 사용한 조사에서 높은 응답률을 유지하는 것은 통계의 신뢰성 측면에서 매우 중요하다고 할 수 있다. 따라서 조사통계를 작성하는 모든 기관에서는 응답률을 높이기 위한 다양한 노력을 하고 있다. 대표적인 사례로서 여러 번의 방문, 응답자 인센티브, 조사방법 혼합 및 변경, 조사기간의 연장 등 여러 가지 방법을 사용하고 있다. 이 노력은 응답률을 높이는 데 효과적일 수 있지만, 이로 인해 오히려 자료의 오차가 증가하는 부작용이 발생할 수 있다는 우려의 목소리도 점점 높아지고 있다.

1. 연구내용 요약

지금까지 연구결과를 요약하면 다음과 같다. 본 연구에 사용된 자료는 지역별고용조사의 가구관리종합표에 기록된 정보를 활용하였으며, 원래 약 14만 가구 중 유효표본가구에 해당하는 약 11만 가구를 실제 분석 자료로 사용하였다. 이러한 분석을 통해 다음과 같은 몇 가지 사실을 확인할 수 있다.

① 거부율이 비접촉률보다 낮지만, 거부에 의한 무응답 편향이 더 크다. 아파트가구가 주택가구보다 조사에 덜 협조적이다.

분석 결과에 의하면, 조사자료는 전체가구 중 주택가구가 53.9%로 아파트가구 36.7%에 비해 높은 비중을 차지하였다. 또한 전체 무응답률은 7.3%로 이 중 비접촉에 의한 무



응답률이 4.5%로 거부에 의한 무응답률 3.1%보다 높았다. 주택유형별 무응답에서는 아파트보다 주택가구가 더 협조적이고, 아파트가구의 비접촉률과 거부율은 각각 6.3%와 5.3%로 주택가구에 대해 약 2배와 약 4배 높은 수치를 보여주었다. 그리고 비접촉가구에 대해 상대적으로 적은 거부가구들은 무응답 편향에는 더 많은 기여를 하였다. 이것은 전체 자료의 오차는 비접촉보다는 거부의 의해 더 많은 영향을 받는다는 것을 나타낸다.

② 자료의 양적 증가가 무응답 편향을 줄이는 것은 아니다. 조사기간이 짧은 경우 조사가 늦게 완성된 가구는 오히려 편향 발생의 원인이 된다.

이 조사는 표본이 아주 큰 조사이기 때문에 전체 조사기간을 더 짧은 기간으로 쪼개어 분석할 수 있다. 이렇게 조사기간에 걸쳐 나눈 구간에 대해 각 구간에 해당하는 표본가구의 크기는 약 10,000여 가구로 여전히 큰 규모에 해당된다. 이로부터 시간에 걸쳐 발생하는 오차의 크기를 파악한다. 분석 자료는 지역통계 생산을 위한 자료라는 점을 고려하여 자료를 다시 특광역시와 시군의 2개 권역으로 나누어 분석하였다. 이 결과는 전체적으로 조사가 진행될수록 거부와 비접촉에 의한 무응답 편향이 증가하는 것을 보였다. 이러한 증가는 무응답률의 증가도 있지만, 해당 구간에서 응답자와 무응답자 간 추정치의 차에 더 많이 기인하였다. 이런 현상은 그렇게 이상한 것은 아니다. 왜냐하면 짧은 조사기간의 자료를 10개 구간으로 나누어 자료의 축적 정도에 따른 무응답 편향을 측정하였기 때문이다. 즉, 짧은 조사기간 중 조사 초기에는 조사에 협조적인 가구들이 많고, 이때는 무응답률도 낮을뿐더러 무응답가구와 응답가구 간 주택 거주율 차이도 그렇게 크지 않다. 하지만 시간이 지남에 따라 조사에 비협조적인 가구들이 점점 조사에 참여하게 되고, 이 가구들의 성향은 초기 응답가구들이 성향과 점점 다르게 나타난다. 조사가 진행됨에 따라 아파트가구일수록, 1인 가구일수록 조사에 늦게 협조하는 경향이 있고, 이러한 가구들의 특성은 자료가 축적됨에 따라 자료의 편향을 형성하는 데 더 많은 기여를 하게 된다. 대도시(특광역시)의 누적자료의 80% 지점에서 거부와 비접촉에 의한 무응답 편향 모두 크게 증가하고, 소도시(시군)는 30% 누적 기점을 중심으로 무응답 편향이 꾸준히 증가하는 경향을 보였다. 이 분석 결과는 처음 기대했던 바와 달리 짧은 조사기간의 자료라는 점에서 자료의 양적 증가가 무응답 편향을 줄이는 데 기여하지 못한다는 사실을 증명하기에는 다소 부족하였다. 오히려 우리의 분석에서는 늦게 응답 전환된 가구들이 조사자료에 포함되면서 이 가구들의 성향이 누적 자료에서 편향을 증가하게 하는 원인이라는 것이 확인되었다.

③ 가구 방문횟수의 증가에 따라 비접촉 편향은 감소하고, 거부 편향은 증가한다.

무응답 편향에 대하여 재방문의 효과 평가에서, 더 많은 방문은 전체 무응답 편향을 줄이는데 효과가 있다는 것을 보여준다. 이에 대한 주요 이유에 대해서 비접촉 패턴을

보면 알 수 있다. 더 많은 접촉시도는 비접촉률을 줄이고, 비접촉에 의한 무응답가구와 응답가구와의 차이를 다소 줄어든다. 따라서 이러한 영향에 의해 무응답 편향이 줄어드는 경향이 있다. 거부가구의 경우는 더 많은 접촉시도가 거부율을 증가시키는 반면, 거부가구와 응답가구 간 주택 거주율의 차이는 감소하긴 하더라도 전체적인 거부 편향은 약간 증가하는 경향이 있다. 접촉시도가 많아짐에 따라 거부가구와 응답가구의 추정치 간 차이가 줄어드는 이유는, 짧은 조사기간 중 나타나는 조사 특성에 의한 자연스런 현상이라고 볼 수 있다. 즉, 조사 초기에는 주택가구에서 응답을 많이 한 반면 후반에는 응답전환 노력에 의해 아파트가구에서의 응답이 많아지면서 자연스레 거부가구와 응답가구 간 추정치 차이가 줄어들었다고 볼 수 있다.

④ 거부가구의 자료가 전체 자료에 미치는 영향은 크지 않다.

조사 초기에 응답한 가구와 응답전환된 가구들을 비교해보면 주택유형과 지역분포에서 차이를 보인다. 특히 최종 거부가구들은 이전에 거부한 경험이 있는 가구가 응답으로 전환된 가구들의 특성과 유사하다고 가정할 때, 실업률이 상당히 높은 가구일 가능성이 높다. 하지만 실제로 거부가구의 비율은 전체의 3.1%에 불과하여 이들 가구들이 전체 자료에서 제외된다고 하더라도 전체 실업률 추정에는 크게 영향을 미치지 않는 것으로 나타났다. 최소한 2번 이상 연속 거부한 가구들에 대해서는 이후 방문에 대해서도 응답을 얻기는 매우 어렵다는 것을 쉽게 예상할 수 있다. 설령 어렵게 응답을 유도한다고 해도 들어간 노력에 비해 자료품질에는 크게 영향을 주지 않을 것이라는 것도 생각할 수 있다. 이러한 결과는 조사에 따라 또는 조사 내 항목에 따라서는 달라질 수 있는 내용이기 때문에 일관성 있게 답하기에는 매우 곤란한 문제이기도 하다. 그렇다 하더라도 최소한 지속적인 거부가구들에 대해 단순히 응답률을 높이기 위한 수단으로 갖은 노력을 한다는 것은 여러 측면에서 단순 소모적인 일에 불과하다는 것도 염두에 둘 필요가 있을 것이다.

2. 무응답가구 대응 방법

이미 서론에서도 설명한 바와 같이 높은 응답률을 유지하고 질 좋은 통계를 만들어야 한다는 점에서는 모두가 공감할 것이다. 그러나 조사 현실은 이러한 바람을 뒷받침하는데 많은 어려움을 동반한다. 굳이 언급하지 않아도 응답자들은 바쁘다, 사생활 보호 등 여러 가지 사유로 조사에 협조하지 않으려는 성향이 강해지고 있다. 즉, 특별한 성향을 가진 무응답가구가 점점 증가한다는 것이다. 이런 현상은 우리나라뿐 아니라 이미 국제사회에서 나타나고 있는 현상이고, 유럽, 호주, 미국 등 통계선진국들에 비하면 다행스럽게도 우리의 조사 현실은 아직은 좋은 편이라 할 수 있다. 이들 국가는 무응답가



구에 대한 관심이 대단하며, 단순 설득이나 응답률 향상 차원이 아닌 자료품질과 조사비용 차원에서 무응답가구에 대한 대응 노력을 찾고 있다. 이러한 시점에서 우리는 단순한 응답률 향상이 아닌 비용대비 자료품질 향상 차원에서, 우리의 상황에 적합한 무응답 처리 방법을 찾을 필요가 있을 것이다. 본 분석 결과를 토대로 몇 가지 현실적인 대안들을 다음과 같이 생각해볼 수 있겠다.

① 현장에서는 비접촉가구에 대한 접촉시도를 꾸준히 하고, 거부가구를 설득할 수 있는 전문 조사원 교육을 끊임없이 해야 한다.

이것은 너무 당연한 방법이고, 지금도 현장에서는 끊임없이 고민하고 있는 문제이다. 하지만 현장에서는 이것 이상 할 수 있는 방법은 없을 것 같다. 그만큼 가장 기본적이고 최선의 방법이라는 의미일 수 있다. 현장에서는 인센티브 인상이나 여러 가지 조사방법을 병행하여 조사 협조율을 높이고자 할 것이다. 그만큼 응답자의 성향이 다양해졌기 때문이다. 그러나 이러한 모든 결정은 조사원 개인에 의해서가 아니라, 체계적 시스템하에 공식적으로 이루어지는 것이 바람직하다. 즉, 슈퍼바이저라는 현장 조사 전반의 관리자에 의해 단순한 응답률 향상이 아닌 자료품질 향상을 목적으로 이루어지는 것이 중요하다. 왜냐하면 이러한 무응답을 줄이기 위한 노력이 오히려 측정오차라든가 다른 차원의 조사오차를 커지게 하는 요인이 될 수 있기 때문이다. 그렇기 때문에 조사원 개개인은 체계적인 현장 조사 매뉴얼에 의해 움직여야 하고, 한두 번 이상 거부한 가구를 지속적으로 접촉시도할 것인가에 대한 결정도 조사원 개인이 아닌 슈퍼바이저의 몫이 되어야 한다. 그러나 현재 우리의 현장 조사는 상당부분 조사원의 결정에 의존하는 것이 현실이 아닌가 싶다. 조사에 협조할 의사가 전혀 없는 가구를 계속해서 방문하는 것은 오히려 조사에 대한 불신과 불만만 높이는 원인이 되고, 이는 조사오차 전반에 영향을 주게 된다는 점도 생각해볼 만하다.

물론 복잡한 현장의 상황을 수학공식처럼 정해놓고 움직일 수는 없다. 그렇지만 최대한 유연함 속에서 일관성 있게 행동하는 것은 매우 중요하며, 이는 통계조사에 대한 신뢰성을 높이고 조사오차를 줄이는 데 큰 도움이 될 수 있다. 이처럼 유연함 속에서 일관성 있는 조사행동규칙을 만들기 위해서는 조사원들에 대한 전문성 향상을 위한 다양한 훈련과 교육이 필요하다. 이미 선진통계국가들은 전문조사원 양성을 위한 다양한 교육프로그램을 만들고, 이 교육프로그램의 효과를 측정하는 다양한 연구가 진행되고 있다. 조사원 훈련프로그램은 현장 조사원들의 경험에만 의존해서도 안되며 객관적인 실험을 통해 확인된 결과를 중심으로 만들어야 한다. 이제는 조사원들이 조사 자체를 완성하는 것에만 의미를 둘 것이 아니라, 현장 조사 과정에서 발생하는 일련의 상황들과 조사오차와의 관계를 인지하고 조사에 임하는 것이 매우 중요할 것이다. 즉, 조사 기획자

의 의도와 현장에서 움직임은 매우 유기적이어야 하며, 그렇게 함으로써 통계품질에 대한 공동의 책임을 져야 할 것이다. 따라서 조사원들에 조사방법, 거부가구 설득방법과 같은 교육과 함께 현장 조사 과정에서 발생하는 조사원들의 행동하나하나가 자료품질에 어떤 형태로 반영되는지에 관한 꾸준한 교육이 필요할 것으로 보인다.

② 무응답 유형별 맞춤형 대응을 위한 조사 실천 매뉴얼이 작성되고, 현장 조사 과정이 관리되어야 한다.

무응답가구의 발생 원인은 매우 다양하지만, 그 중에서도 전혀 접촉이 안되는 경우와 접촉을 하더라도 조사를 거부하는 경우가 가장 큰 원인이 되고 있다. 특히 거부가구에 대해서 그 이유가 정확히 파악되어야 하며, 이와 함께 무응답 원인별 가구들의 특성을 정밀하게 분석해야 한다. 또한 그 분석 결과를 현장에 피드백함으로써 다양한 무응답 이유와 유형별 맞춤형 대응 논리를 체계적으로 정리하고 실천하도록 해야 한다. 분명한 것은 거부자와 비접촉자의 성향이 다르고, 거부 이유 또한 다양하다는 것이다. 이에 대해 이미 현장 조사원들은 자신만의 경험과 노하우에 의해 무응답에 적절하게 대응하고 있으나 이것이 단순한 응답률 향상을 위한 것인지 조사의 품질을 높이기 위한 것인지에 대한 명백한 구분이 없다고 본다. 또한 이러한 각자의 대응 방법은 조사원에 따라서 매우 주관적일 수 있고, 가구마다 처리기준이 다를 수 있기 때문에 조사전체를 관리하는 슈퍼바이저에 의해 매 단계 단계의 처리기준이 결정되는 것이 바람직할 것이다.

통계품질 향상을 염두하고 무응답에 대한 대응논리를 찾는다면, 모든 현장 조사 과정의 관리는 현장의 몫이기 이전에 조사 기획자의 책임이 크다 할 수 있다. 정확한 지침에 의해 현장을 관리해야 하며, 이 모든 것은 기록으로 남겨져야 한다. 이러한 꾸준한 현장 조사 과정에 대한 정보를 축적하고 분석함으로써 조사 방문횟수에 대한 결정, 거부가구의 지속적 설득 여부, 비접촉가구들의 접촉 방법에 대한 구체적인 근거를 찾을 수 있을 것이다. 지금까지는 현장 조사 과정을 이해할 수 있는 객관적인 자료들이 많지 않았기 때문에 현장 조사 개선을 위한 연구가 상당히 어려운 실정이다. 따라서 현장 조사 과정 기록, 내검이전 자료 확보 등 파라미터 구축과 같은 작업은 어려운 조사 현장을 개선하는 좋은 방법이 될 것으로 기대한다.

③ 단순한 응답률 증가보다 자료의 질적 향상을 목적으로 무응답에 대처해야 할 것이다.

많은 학술연구에서 그랬듯이 무응답이 반드시 자료의 품질을 떨어뜨리는 것은 아니며, 응답률이 높다고 해서 반드시 자료의 질이 좋다고 단언하기도 어렵다. 응답률과 오차와의 관계는 조사마다 또는 조사 내 항목마다의 상황이 다르기 때문에 꾸준한 현장 조사 기록을 분석함으로써 적절한 대응논리를 찾을 수 있어야 한다. 그러기 위해서는 파



라데이터와 같은 현장 조사 기록을 통해 무응답가구에 관한 정보를 구축하고, 이로부터 응답가구들의 특성을 꾸준히 비교분석하는 것이 중요하다. 인센티브 인상, 최소한의 방문횟수 결정, 조사방법 변경에 대한 문제도 이로부터 상당부분 해결할 수 있을 것으로 본다. 무응답 대응 방법은 조사마다 다르기 때문에 모든 조사에 일관되게 적용할 수 없으며, 적용하여서도 안되는 문제이다. 본 연구에서 분석한 바와 같이 어떤 경우는 무응답이 발생하더라도 전체 추정치에 크게 영향을 미치지 않을 수 있고, 어떤 경우는 오히려 응답전환 노력이 추정치의 질을 떨어뜨리는 결과를 초래할 수 있다는 사실을 인지할 수 있어야 한다. 높은 응답률과 좋은 통계가 항상 긍정적인 상관관계가 있는 것은 아니기 때문에 양적 측면과 함께 질적 측면에서 무응답 편향을 줄이는 노력을 할 필요가 있다고 본다. 이렇게 하기 위해서는 응답률이 조사기관을 평가하는 절대적인 척도가 되기 보다는 오히려 현장 관리 기록을 통한 분석 결과에 의해서 자료의 질적 향상을 위해 얼마만큼 노력하였는가를 평가 기준으로 삼는 것도 좋은 방법이 될 것으로 생각한다.

④ 재조사와 통계적 무응답 가중치조정법을 통해 무응답 편향을 줄일 수 있다.

단위 무응답에 대해서 가장 쉽게 생각할 수 있는 방법은 재조사일 것이다. 무응답 유형에 대해 무응답가구들의 일부를 다시 조사하여 조사결과를 얻어낸다면 무응답의 편향을 줄이는 데 효과적일 것이다. 전체 무응답가구 중에서 다시 표본추출 방법에 의해 추출된 일부 무응답가구에 대해 재조사를 실시하는 방법은 많은 경우 1차 조사보다 재조사가 비용이 적게 든다(김재광, 2008). 만약 1차 조사에서 표본가구를 직접 대면하지 못하여 무응답한 가구는 전화, 인터넷 또는 이메일 조사를 통해 응답을 유도할 수 있을 것이다. 거부에 의한 무응답인 경우라면 그 이유를 충분히 파악해야 한다. 예를 들어, 무응답 이유가 사생활 침해나 일과시간이 바빠서 등의 이유라면 면접이 아닌 우편이나 이메일 조사방법을 취할 수 있을 것이다. 또 정부나 조사를 신뢰하기 어렵고 조사결과를 어디에 사용하는지 믿을 수가 없다면 보다 노련한 조사원으로 바뀌주고, 작성된 통계에 대한 정보를 사후적으로 안내해 줌으로써 조사 참여를 유도할 수 있을 것이다. 이처럼 재조사의 경우 1차 조사와 다른 조사방법을 통해 응답을 얻어 냄으로써 무응답 편향을 줄일 수 있을 것이다. 그러나 국가통계는 공표기간이 정해져 있기 때문에 재조사 일정은 매우 촉박하게 진행되어야 하며, 추가적인 비용부담도 고려해야 될 문제로 남게 된다.

다른 방법으로 무응답 가중치조정법을 생각할 수 있다. 무응답 가중치조정법은 재조사를 하지 않고 응답자만의 가중치를 조정하여 추정량의 편향을 줄이고자 하는 방법이다. 간단하게 무응답 가중치조정법은 무응답층 간의 무응답률이 같다고 가정하는 동질적인 응답모형을 사용하게 된다. 그러나 실제로 무응답층 간에 무응답률이 동일하지 않은 경우가 많다. 본 연구에서도 주택유형별 또는 지역별 무응답률이 다르게 나타났다.

따라서 이러한 응답 성향을 반영할 수 있는 무응답 가중치조정법을 사용하면 무응답 편향을 줄이는 데 더 많은 기여를 할 수 있을 것이다. 예로 응답자 성향 모형을 이용할 수 있다(Cobben, 2009). 조사대상가구가 응답을 하고 안하고는 응답자의 선택사항이 될 수 있으므로 이러한 내용을 반영할 수 있는 통계적 모형을 이용하여 무응답을 조정할 수 있다. Cobben은 응답과 무응답의 상황이 응답대상자의 선택 성향이라고 간주하고, 응답 선택모형(response selection model)과 응답성향방법(response propensity methods)을 제시하였다. 이 방법은 무응답 유형을 거부와 비접촉으로 세분화하여 모형을 구축하고, 무응답 가중치를 산출하는 방법이다. 이러한 방법은 기존의 단순히 무응답률에 의해 조정되는 방법에 비해 다양한 응답 유형을 고려하여 응답자 성향을 반영함으로써 더 정밀한 가중치 작업을 할 수 있다는 점에서 무응답 편향을 줄이는 데 매우 유용한 방법이 될 수 있을 것이다. 이 외에도 모형에 의한 무응답가중치 조정을 위한 커널 스무딩(kernel smoothing)과 같은 방법들도 사용할 수 있다(Silva and Opsomer, 2006).

3. 향후 연구방향

많은 연구자들은 무응답가구에 대한 정보를 얻을 수 없기 때문에 무응답 연구가 매우 어렵다고 말하고 있다. 본 연구에서도 분석에 사용된 자료가 정확히 연구를 위해 설계된 자료가 아니라는 점에서 연구결론을 일반화하기에 제약이 있을 것으로 보인다. 그러나 이러한 무응답에 관한 지속적인 연구는 통계품질 향상에 매우 필요한 과정이라 할 수 있다. 그런 측면에서 본 연구는 단위 무응답 연구의 좋은 출발점이 될 것으로 기대한다. 본 연구와 관련하여 무응답 편향을 줄일 수 있는 많은 연구가 후속적으로 진행됨으로써 증가하는 조사무응답의 고민을 줄일 수 있을 것이다.

우선 본 연구의 후속으로 가구방문회차별 거부 이유를 파악하여 무응답 층별 거부 이유를 분석하고 해결 방법을 모색할 필요가 있다. 이를 위해 본 연구에 사용된 조사에서 재차 가구관리종합표 작성을 의무화하고 관리하여 본 연구결과에 대한 타당성 검증을 수행해야 할 것이다. 그리고 무응답률과 무응답 편향과의 관계는 조사마다 다르고, 조사 내 항목마다 다를 수 있기 때문에 가구 대상의 다른 조사로 확장하여 동일한 연구를 수행하는 것도 좋을 것이다. 또한 무응답 편향 감소의 방법으로 무응답 가중치조정법을 활용하여 앞으로 증가하는 무응답 대응 방안을 찾는 것도 매우 중요할 것이다. 또한 경우에 따라서는 무응답가구들이 새로운 표본으로 대체되는 경우도 발생한다. 대체 표본 사용은 이미 학술계에서 사용하지 않는 것을 권장하고 있지만, 리서치 회사등 통계작성기관은 조사의 현실성 측면에서 대체 표본을 사용하기도 하는 것 같다. 이에 대해 대체 표본이 전체 자료에 미치는 효과를 측정하는 것이 중요하고, 이를 통해 대체 표본

의 사용여부 및 응답률 계산방법 등에 대한 논의가 있어야 할 것이다. 따라서 대체 표본 효과 측정에 관한 연구도 향후 진행할 예정이다.

무응답 관련 연구는 매우 어려운 분야인 만큼, 관련자들의 지속적인 관심이 필요하며 무응답 연구를 위한 부단한 노력 없이 무응답 대응 방안을 찾는 것은 쉽지 않을 것이다. 무응답 관리는, 사후적으로 발생하기 때문에 관리하고 대응하는 것보다, 조사 전 단계에서부터 무응답 가능성을 예측하고 무응답 편향을 줄이기 위한 대응 논리로 만들어 가는 것이 더 중요할지도 모른다.



참고문헌

- 김재광(2008), 표본조사론, 자유아카데미.
- 이승희(2010), 총조사 오차의 패러다임으로 이해하는 표본조사, *통계연구* 15-1, 44-74.
- AAPOR(2009), Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys, The American Association for Public Opinion Research.
- Alreck, P. L. and Settle, R. B.(1995), *The Survey research handbook: Guidelines and Strategies for Conducting a Survey*, Chicago: Irwin Professional Publishing.
- Schouten, B., Cobben, F. and Bethlehem, J.(2009), Indicators for the Representativeness of survey response, *Survey Methodology*, Vol.35, No.1, pp.101-113
- Bethlehem, J.(2002), Weighting Nonresponse Adjustments Based on Auxiliary Information. In *Survey Nonresponse*, ed. Robert M. Groves, Don A. Dillman, John L. Eltinge and Roderick J. A. Little, 275-288, New York: Wiley.
- Canada(2007), *Best Practice in Public Opinion Research*.
- Curtin, R., Presser, S. and Singer, E.(2000), The Effects of Response Rate Changes on the Index of Consumer Sentiment, *Public Opinion Quarterly*, 64, 413-428.
- de Leeuw, E. and de Heer, W.(2002), Trends in Household Survey Nonresponse: A Longitudinal and International Comparison, *In Survey Nonresponse*, ed. Robert M. Groves, Don A. Dillman, John L. Eltinge and Roderick J. A. Little, pp.41 - 54, New York: Wiley.
- Babbie, E. R.(2007), *The Practice of Social Research*, Wadsworth Publishing Company.
- Groves, R.M. and Lyberg, L.E.(1988), An Overview of Nonresponse Issues in Telephone Surveys, *In Telephone Survey Methodology*, edited by Robert M. Groves, et al. New York: John Wiley & Sons.
- Groves, R.M.(2006), Nonresponse Rates and Nonresponse Bias in Household Surveys, *Public Opinion Quarterly*, 70, 646-675.
- Groves, R.M. and Peytcheva, E.(2006), The Impact of Nonresponse Rates on Nonresponse Bias: A Meta-Analysis, Draft for review.
- Keeter, S., Miller, C., Kohut, A., Groves, R.M. and Presser, S.(2000), Consequences of Reducing Nonresponse in a National Telephone Survey, *Public Opinion Quarterly*, 64, 125-148.
- Olson, K.(2006) Survey Participation, Nonresponse Bias, Measurement error Bias and Total Bias, *Public Opinion Quarterly*, Vol.70, No.5, Special Issue 2006, pp.737 - 758.
- Lin I-F. and Schaeffer, N.C.(1995), Using Participants to Estimate the Impact of Nonparticipation, *Public Opinion Quarterly*, 59, 236-258.
- Lynn, P., P. Clarke, J. Martin and P. Sturgis(2002), The Effect of Extended Interview Efforts on Nonresponse Bias. In R. Groves, D. Dillman, J. Eltinge, & R. Little(Eds.), *Survey Nonresponse*, 135-147 New York: Wiley.

- Merkle, D. and Murray, E.(2002), Nonresponse in Exit Polls: A Comprehensive Analysis. In Survey Nonresponse, ed. Robert M. Groves, Don A. Dillman, John L. Eltinge and Roderick J.A. Little, PP. 243-258. New York: Wiley.
- Miller, J.M. and Wedeking, J.(2003), The Harder We Try, The Worse It Gets? Examining the Impact of Refusal Conversions and High Callback Attempts on the Quality of Survey Data, Unpublished Paper Presented at AAPOR 2003.
- Retzer, K.F., Schipani, D. and Cho, Y.I.(2004), Refusal Conversion: Monitoring the Trends, American Association for Public Opinion Research Section on Survey Research Methods, 4984-4990.
- Silva, D.N. and Opsomer, J.D.(2006), A kernel smoothing method of adjusting for unit non-response in sample surveys, *The Canadian Journal of Statistics*, 34, 563-579.
- Singleton, R. A. and Straits, B. C.(2005), Approaches to Social Research 4th ed. New York: Oxford University Press.
- Stoop, I.(2005), Nonresponse in Sample Surveys, The Hunt for the Last Respondent, The Hague: Social and Cultural Planning Office.



<부 록>

<부표> 158개 시군에서 거부 경험이 있는 가구를 제외했을 때의 실업률(%)

연번	1번 이상 거부 제외	2번 이상 거부 제외	3번 이상 거부 제외	전체 실업률
1	4.17	4.12	4.18	4.18
2	1.75	1.72	1.72	1.71
3	1.67	1.66	1.66	1.66
4	2.30	2.29	2.27	2.27
5	2.56	2.53	2.53	2.52
6	3.16	3.14	3.14	3.14
7	3.51	3.41	3.38	3.38
8	2.64	2.54	2.53	2.52
9	3.08	3.13	3.13	3.13
10	2.92	2.93	2.91	2.89
11	1.10	1.09	1.09	1.09
12	2.54	2.50	2.49	2.49
13	2.50	2.49	2.49	2.49
14	1.75	1.75	1.75	1.75
15	2.73	2.69	2.68	2.68
16	3.94	4.09	4.08	4.08
17	2.34	2.41	2.40	2.40
18	4.72	4.63	4.60	4.60
19	2.96	2.91	2.88	2.87
20	1.71	1.68	1.66	1.66
21	3.07	3.07	3.06	3.06
22	1.86	2.30	2.30	2.29
23	2.21	2.20	2.20	2.20
24	1.36	1.35	1.35	1.35
25	2.46	2.43	2.41	2.41
26	3.03	2.99	2.99	2.99
27	0.66	0.66	0.66	0.66
28	1.08	1.07	1.07	1.07
29	1.48	1.47	1.47	1.47
30	2.39	2.61	2.61	2.61
31	2.03	2.02	2.01	2.01
32	2.23	2.19	2.19	2.19
33	1.89	2.11	2.11	2.11
34	2.46	2.40	2.39	2.39
35	0.59	0.69	0.69	0.69
36	0.91	0.91	0.91	0.91
37	1.48	1.48	1.48	1.48
38	0.46	0.46	0.45	0.45

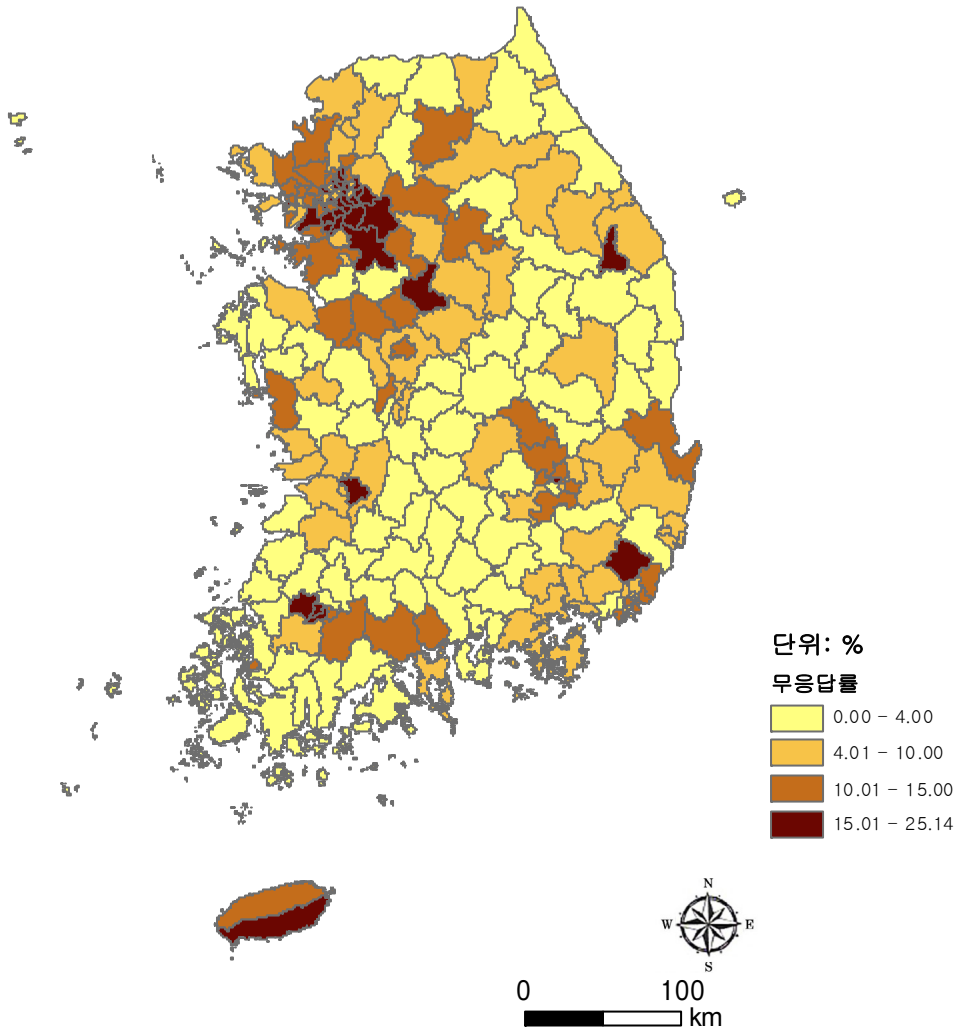
연번	1번 이상 거부 제외	2번 이상 거부 제외	3번 이상 거부 제외	전체 실업률
39	0.35	0.35	0.35	0.35
40	0.40	0.40	0.40	0.40
41	0.25	0.25	0.25	0.25
42	1.52	1.52	1.52	1.52
43	2.70	2.68	2.68	2.68
44	0.45	0.45	0.45	0.45
45	0.48	0.48	0.48	0.48
46	0.30	0.30	0.30	0.30
47	0.35	0.35	0.35	0.35
48	0.69	0.69	0.69	0.69
49	1.11	1.10	1.10	1.10
50	2.42	2.40	2.40	2.40
51	1.93	1.93	1.93	1.93
52	2.24	2.24	2.24	2.24
53	1.97	1.96	1.95	1.95
54	1.04	1.04	1.04	1.04
55	2.00	1.98	1.98	1.98
56	0.47	0.46	0.46	0.46
57	1.72	1.70	1.69	1.69
58	0.84	0.84	0.84	0.84
59	2.10	2.33	2.33	2.33
60	0.34	0.34	0.34	0.34
61	2.88	2.87	2.87	2.87
62	4.08	4.00	3.98	3.98
63	1.39	1.39	1.39	1.39
64	1.24	1.22	1.22	1.22
65	2.65	2.60	2.60	2.60
66	1.77	1.75	1.75	1.75
67	2.41	2.40	2.40	2.40
68	3.22	3.22	3.22	3.22
69	0.41	0.40	0.40	0.40
70	2.66	2.65	2.65	2.65
71	0.85	0.85	0.85	0.85
72	1.23	1.23	1.23	1.23
73	1.10	1.09	1.09	1.09
74	1.97	1.96	1.96	1.96
75	1.15	1.15	1.15	1.15
76	2.50	2.50	2.50	2.50
77	1.48	1.48	1.48	1.48
78	2.41	2.37	2.37	2.37
79	1.46	1.46	1.45	1.45
80	3.96	3.92	3.88	4.15



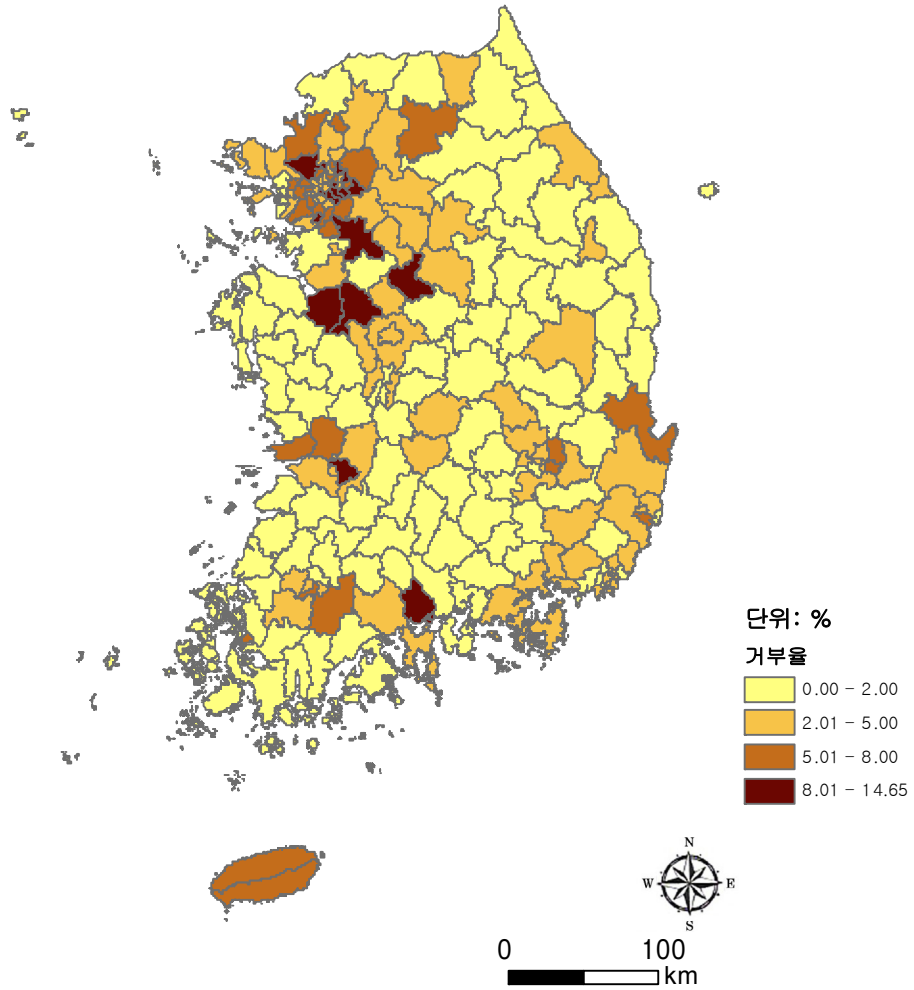
연번	1번 이상 거부 제외	2번 이상 거부 제외	3번 이상 거부 제외	전체 실업률
81	1.68	1.67	1.67	1.67
82	2.72	2.66	2.65	2.64
83	1.75	1.72	1.72	1.72
84	1.35	1.34	1.34	1.34
85	1.27	1.27	1.27	1.27
86	1.73	1.73	1.73	1.73
87	1.00	1.00	1.00	1.00
88	0.62	0.62	0.62	0.62
89	0.14	0.14	0.14	0.14
90	0.64	0.64	0.64	0.64
91	0.69	0.69	0.69	0.69
92	3.53	3.61	3.61	3.61
93	2.37	2.40	2.36	2.35
94	1.79	1.77	1.77	1.77
95	2.08	2.06	2.06	2.06
96	1.36	1.34	1.33	1.33
97	0.68	0.68	0.68	0.68
98	0.31	0.31	0.31	0.31
99	0.20	0.20	0.20	0.20
100	0.47	0.47	0.47	0.47
101	3.08	3.07	3.07	3.07
102	0.44	0.44	0.44	0.44
103	0.81	0.81	0.81	0.81
104	0.26	0.26	0.26	0.26
105	0.67	0.67	0.67	0.67
106	0.92	0.91	0.91	0.90
107	0.23	0.23	0.23	0.23
108	0.52	0.52	0.52	0.52
109	1.25	1.24	1.24	1.24
110	0.90	0.90	0.90	0.90
111	0.22	0.22	0.22	0.22
112	0.52	0.51	0.51	0.51
113	2.68	2.67	2.67	2.67
114	1.38	1.37	1.37	1.37
115	2.41	2.37	2.37	2.37
116	1.07	1.06	1.06	1.05
117	3.57	3.53	3.53	3.53
118	1.67	1.66	1.66	1.66
119	1.87	1.86	1.86	1.86
120	1.65	1.78	1.77	1.77
121	1.28	1.37	1.37	1.37
122	2.74	2.93	2.92	2.92

연번	1번 이상 거부 제외	2번 이상 거부 제외	3번 이상 거부 제외	전체 실업률
123	0.46	0.45	0.45	0.45
124	0.61	0.61	0.61	0.61
125	1.74	1.74	1.74	1.74
126	1.73	1.73	1.73	1.73
127	0.55	0.54	0.54	0.54
128	0.89	0.89	0.89	0.89
129	1.37	1.35	1.35	1.35
130	1.24	1.24	1.24	1.24
131	2.21	2.20	2.19	2.19
132	0.76	0.74	0.73	0.73
133	1.03	1.03	1.03	1.03
134	1.98	1.97	1.97	1.97
135	0.42	0.42	0.42	0.42
136	4.21	4.15	4.14	4.14
137	2.41	2.38	2.38	2.37
138	2.06	2.05	2.05	2.05
139	3.97	3.92	3.91	3.90
140	0.66	0.66	0.66	0.66
141	1.60	1.58	1.58	1.75
142	1.54	1.54	1.54	1.54
143	1.43	1.43	1.43	1.43
144	1.18	1.16	1.16	1.16
145	1.54	1.54	1.54	1.54
146	1.11	1.11	1.11	1.11
147	2.56	2.56	2.56	2.56
148	1.31	1.31	1.31	1.31
149	0.69	0.69	0.69	0.69
150	0.74	0.73	0.73	0.73
151	0.82	1.08	1.08	1.08
152	0.76	0.76	0.76	0.76
153	0.23	0.23	0.23	0.23
154	0.61	0.60	0.60	0.60
155	0.82	0.81	0.81	0.81
156	2.91	3.05	3.00	2.98
157	1.50	1.48	1.48	1.48
158	0.21	0.21	0.21	0.21

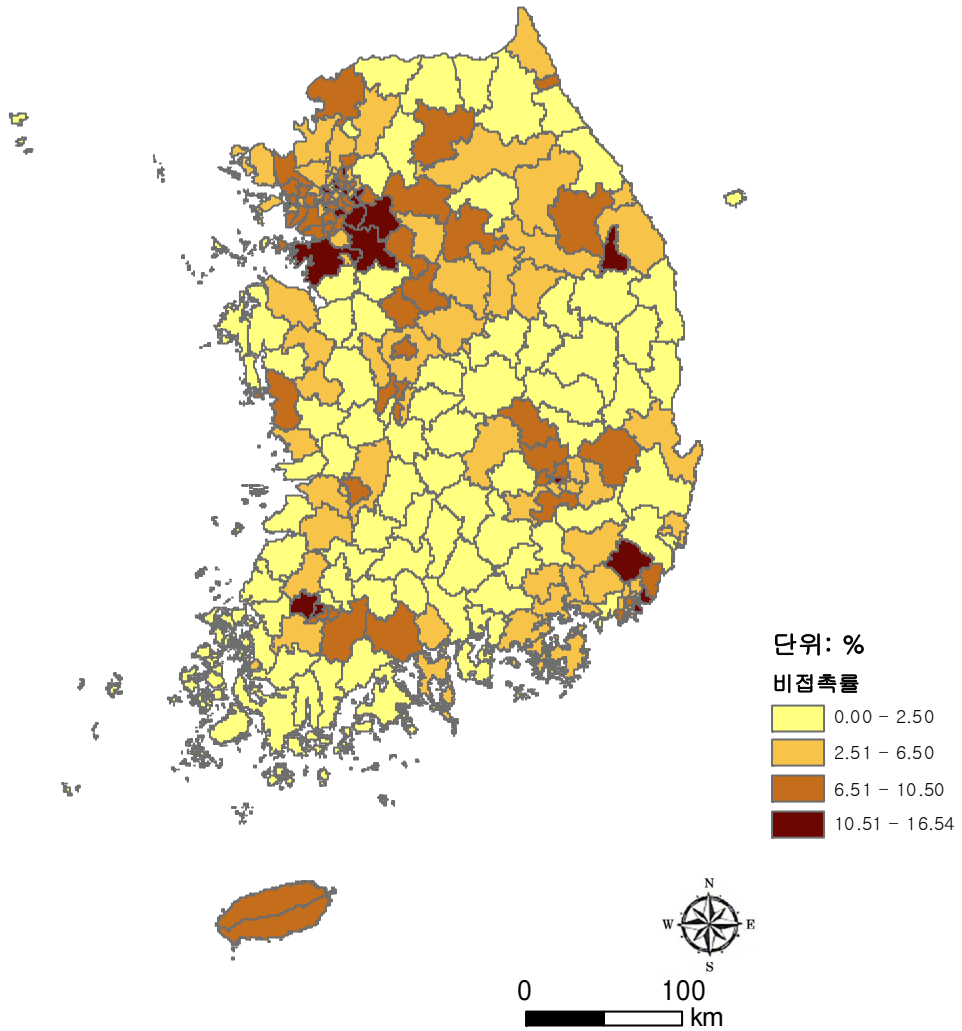




[부그림 1] 시군구 무응답률



[부그림 2] 시군구 거부율



[부그림 3] 시군구 비접촉률