발간등록번호
11-1240000-001442-01

『북한이탈주민실태조사』 품질개선 컨설팅 최종결과보고서

Final Report on Quality Improvement Consulting for
『Survey of North Korean Refugees in South Korea』

한국조사연구학회

2021.6.

통계청
Statistics Korea

제 출 문

통계청장 귀하

본 보고서를 「북한이탈주민실태조사」 품질개선 컨설팅의 최종결과 보고서로 제출합니다.

2021년 6월

한국조사연구학회

연구진

책 임 연 구 원	이기성, 우석대학교 아동복지학과 교수
연 구 원	손창균, 동국대학교 빅데이터응용통계학과 부교수
연 구 보 조 원	우영제, 동국대학교 통계학과 박사과정 수료

최종결과보고서 요약문

연구과제명	북한이탈주민실태조사 품질개선 컨설팅
주 제 어	북한이탈주민실태조사, 표본설계, 조사항목
연 구 기 간	2021.4.1 - 2021.6.30.
연 구 기 관	한국조사연구학회
연구진구성	이기성, 우석대학교 아동복지학과 교수 손창균, 동국대학교 빅데이터응용통계학과 부교수

2021년은 북한이탈주민실태조사 10년 차를 맞이하는 해로 앞으로 새로운 10년을 대비하여 기존 표본설계 검토와 조사항목 재정비 등 조사재설계 및 통계 품질 개선에 대한 요구가 지속적으로 제기되어 왔다. 또한 2020년 통계청 정기통계품질 진단에서 시간 흐름에 따른 정착실태 및 이탈주민 지원정책 효과 등을 파악하기에 효과적인 패널조사로 조사방식을 변경하는 것을 검토하고, 중요도 및 활용도에 따라 조사 문항을 조정할 필요가 있는 것으로 나타났다.

이에 본 연구에서는 2020년 북한이탈주민실태조사의 표본설계, 조사결과 및 2020년 정기통계 품질진단 결과보고서를 토대로 패널조사로의 전환 가능성을 검토하고, 새로운 표본설계 방안(표본배분, 가중치 작성, 모수추정 등) 마련 및 조사항목 개선을 통해 북한이탈주민실태조사의 품질을 제고하고자 한다.

본 연구의 주요 결과를 정리하면 다음과 같다.

첫째, 중복 조사되는 표본의 결과를 적절히 활용할 수 있도록 조사방식을 고정패널이나 분할패널설계로 전환하는 문제를 검토하였다.

둘째, 통일부에서 작성 및 관리하고 있는 최신 북한이탈주민명부를 표본추출틀로 사용하여 새로운 표본설계 방안을 마련하였다. 지역(2개 층), 성별(2개 층), 연령(5개 층), 남한거주기간(2개 층)을 층화기준으로 사용하여 조사모집단에 대한 충분한 분석을 통해 모집단의 특성을 잘 반영한 효과적인 표본배분방법을 제시하고, 응답부담 경감을 위해 허용오차 내에서 표본크기 축소가 가능한 지 검토하였다. 그리고 정확한 통계산출을 위해 표본설계에 따른 설계가중치, 조사과정에서 발생하는 무응답 조정을 위한 무응답조정가중치 등을 적용한 가중치를 부여하는 과정을 구체적으로 기술하였고, 모수추정식과 상대표준오차식을 제시하였다. 또한 패널 이탈 방지를 위한 패널관리방안을 마련하고, 통계작성목적을 고려하여 현행 조사방법을 유지하면서 심층면접 방법을 병행하는 방안도 추가적으로 검토하였다.

셋째, 조사항목 수가 너무 많아 모든 항목을 매해 조사하기보다 조사영역을 나누어 정착실태와 같은 주요 분야는 매년 조사하고 사회통합과 관련된 조사항목은 사회통합조사 I (보건 및 건강, 교육, 가족)과 II(노동, 사회활동, 복지, 사회적 관계망, 가구 경제상태)로 영역을 구분하여 짝수년에는 사회통합조사 I 을, 홀수년에는 사회통합조사II를 조사하는 방안을 제안하였다.

차 례

요약 ····· 1

제 1 장 서론 ····· 5

제 1 절 연구배경 및 목적 ····· 5

제 2 절 연구내용 및 방법 ····· 7

제 3 절 연구의 기본 목표와 절차 ····· 10

제 2 장 북한이탈주민실태조사 및 작성현황 검토 ····· 11

제 1 절 북한이탈주민실태조사 개요 및 표본설계 ····· 11

제 2 절 북한이탈주민실태조사 현황 및 주요 조사항목 분석 ····· 19

제 3 절 북한이탈주민실태조사 검토 결과 ····· 24

제 3 장 패널조사로의 전환 가능성 검토 ····· 26

제 1 절 북한이탈주민실태조사의 중복 표본 현황 ····· 26

제 2 절 패널설계방법 검토 및 패널조사 사례 ····· 29

제 3 절 패널조사로의 전환 가능성 검토 ····· 41

제 4 장 표본설계 개선 방안 ····· 44

제 1 절 표본설계 개선안의 특징 ····· 44

제 2 절 층화 ····· 44

제 3 절 표본크기 결정 및 표본배분 ····· 46

제 4 절 표본추출 및 대체 ····· 68

제 5 절 가중치 작성 및 모수 추정 ····· 69

제 5 장 조사표 개선 ····· 76
제 1 절 조사방식 및 조사항목 검토 ····· 76
제 2 절 조사항목 개선안 ····· 82

제 6 장 결론 및 제언 ····· 89

참고문헌 ····· 93
부록: 조사표 ····· 97

표 차 례

<표 1-1> 2021년 기준 북한이탈주민실태조사 조사항목(115개) ····· 6
<표 2-1> 거주 지역별 북한이탈주민 모집단 분포 ····· 13
<표 2-2> 성별 북한이탈주민 모집단 분포 ····· 14
<표 2-3> 연령 그룹별 북한이탈주민 모집단 분포 ····· 14
<표 2-4> 남한 거주기간별 북한이탈주민 모집단 분포 ····· 14
<표 2-5> 각 층별 모집단 크기 ····· 15
<표 2-6> 각 층별 표본크기 ····· 16
<표 2-7> 특성별 북한이탈주민실태조사 응답 현황 ····· 19
<표 2-8> 특성별 남한생활 만족도 비율 및 상대표준오차 ····· 21
<표 2-9> 특성별 경제활동지표 및 상대표준오차 ····· 22
<표 3-1> 중복표본 현황 ····· 27
<표 3-2> 원표본과 대체표본 현황 ····· 28
<표 3-3> BHPS 1차 웨이브 응답자의 응답률 ····· 36
<표 3-4> SHF의 과대표집 비율 ····· 40
<표 4-1> 층별 비례배분에 따른 허용오차(표본크기=3,000) ····· 49
<표 4-2> 층별 제곱근비례배분에 따른 허용오차(표본크기=3,000) ····· 50
<표 4-3> 층별 네이만배분에 따른 허용오차(표본크기=3,000) ····· 51
<표 4-4> 층별 멱등배분에 따른 허용오차(표본크기=3,000) ····· 52
<표 4-5> 공표수준별 표본크기(3,000) 및 허용오차(제곱근비례배분) ····· 54
<표 4-6> 공표수준별 표본크기(3,000) 및 허용오차(멱등배분) ····· 54
<표 4-7> 층별 비례배분에 따른 허용오차(표본크기=2,500) ····· 56
<표 4-8> 층별 제곱근비례배분에 따른 허용오차(표본크기=2,500) ····· 57
<표 4-9> 층별 네이만배분에 따른 허용오차(표본크기=2,500) ····· 58
<표 4-10> 층별 멱등배분에 따른 허용오차(표본크기=2,500) ····· 59

<표 4-11> 공표수준별 표본크기(2,500) 및 허용오차(제곱근비례배분) … 60
<표 4-12> 공표수준별 표본크기(2,500) 및 허용오차(멱등배분) …… 61
<표 4-13> 층별 비례배분에 따른 허용오차(표본크기=2,000) …… 62
<표 4-14> 층별 제곱근비례배분에 따른 허용오차(표본크기=2,000) …… 63
<표 4-15> 층별 네이만배분에 따른 허용오차(표본크기=2,000) …… 64
<표 4-16> 층별 멱등배분에 따른 허용오차(표본크기=2,000) …… 65
<표 4-17> 공표수준별 표본크기(2,000) 및 허용오차(제곱근비례배분) … 66
<표 4-18> 공표수준별 표본크기(2,000) 및 허용오차(멱등배분) …… 67
<표 5-1> 북한이탈주민 정착실태조사 조사항목 …… 77
<표 5-2> 북한이탈주민 사회통합실태조사 조사항목 …… 79
<표 5-3> 북한이탈주민 정착실태조사 조사항목 개선안 …… 82
<표 5-4> 북한이탈주민 사회통합조사 Ⅰ 조사항목 개선안 …… 85
<표 5-5> 북한이탈주민 사회통합조사 Ⅱ 조사항목 개선안 …… 86

요약

북한이탈주민실태조사는 북한이탈주민의 남한정착정도를 파악하고, 북한이탈주민에 대한 지원정책 수립 등에 필요한 기초자료를 제공하기 위해 실시되고 있다.

2020년 북한이탈주민실태조사의 표본설계와 관련하여 검토한 결과 '패널조사로의 전환 검토', '표본설계 개선방안 마련', '조사표 검토' 등이 필요한 것으로 파악되었다.

따라서 본 연구에서는 2020년 북한이탈주민실태조사 결과와 2020년 정기통계품질진단 결과를 토대로 새로운 표본설계 방안(표본배분, 가중치 작성 및 모수추정, 패널관리방안 등)과 조사항목 등의 개선안을 마련하여 북한이탈주민실태조사의 품질을 개선하였다.

연구결과 현행 표본설계와 표본설계 개선안을 비교하여 요약하면 다음과 같다.

신구 표본설계 요약 비교

구분	현행 표본설계	표본설계 개선안
모집단과 표본추출틀	- 목표모집단 : 북한이탈주민의 보호 및 정착지원에 관한 법 제2조 제1항에서의 정의된 바와 같이, "군사분계선 이북지역에 주소, 직계가족, 배우자, 직장 등을 두고 있는 사람으로서 북한을 벗어난 후 외국 국적을 취득하지 아니한 자" 들 중 2021년 5월 기준 남한에 정착한 만 15세 이상 북한이탈주민 전체 - 조사모집단 : 2021년 5월 기준 통일부에 등록되어 있는 만 15세 이상 북한이탈주민 중 남한에 실제 거주하는 자 - 표본추출틀 : 2021년 5월 기준 통일부에 등록되어 있는 만 15세 이상 북한이탈주민 명부	좌동

구분	현행 표본설계	표본설계 개선안
층화기준	- 층화 : 지역(2), 성(2), 연령그룹(5), 남한거주기간(2) 변수를 이용한 총 40개의 층을 구성	좌동
표본크기 및 표본배분	- 표본크기 : 3,000개 - 표본크기 산출식 미제공 - 표본배분 : 제곱근비례배분	**- 표본크기 : 2,500개** **- 표본크기 산출식 제공** - 표본배분 : 제곱근비례배분
표본추출	- 층화계통추출법	좌동
가중치	- 층 h 및 사후 층 l에 속한 이탈주민 i에 부여된 표본가중치 : $\alpha_{hli} = \frac{N_h}{n_h}$ N_h: 층 h의 북한이탈주민 모집단 수 n_h: 층 h의 북한이탈주민 표본 수 - 사후층화가중치 : $w_{hli} = \frac{N_l}{\hat{N}_l}\alpha_{hli}$ N_l : 사후 층 l의 북한이탈주민 모집단 수, $\hat{N}_l = \sum_h \sum_i \alpha_{hli}$	**- 기본가중치 :** $w_h^B = \frac{N_h}{n_h}$ N_h **: h층의 북한이탈주민 모집단 수** n_h **: h층의 북한이탈주민 표본 수** **- 무응답 조정 가중치 :** $w_h^R = \frac{n_h}{r_h}$ r_h **: h층에서 응답한 응답자의 수** **- 사후층화에 의한 최종가중치 :** $w_{hi} = w_h^B \times w_h^R \times \frac{X_h}{\hat{X}_h}$ X_h **: h층의 북한이탈주민 모집단 총수** $\hat{X}_h$ **: h층의 응답자 가중합**
모수추정	- 총합추정량 : $\hat{t}_y = \sum_h \sum_l \sum_i w_{hli} y_{hli}$ - 총합의 분산추정량 : $\hat{V} = \sum_h \left(1 - \frac{n_h}{N_h}\right) \frac{n_h}{n_{h-1}} \sum_l \sum_i (z_{hli} - \bar{z}_{h..})^2$, $z_{hli} = w_{hli} y_{hli}$	**(총계 추정)** **- 총계 추정량** $\hat{Y}$ **:** $\hat{Y} = \sum_{h=1}^{L} \sum_{i=1}^{n_h} w_{hi} y_{hi}$ **- 총계 추정량** $\hat{Y}$**의 분산추정량 :** $\hat{V}(\hat{Y}) = \sum_{h=1}^{L} \frac{n_h(1-f_h)}{n_h - 1} \sum_{i=1}^{n_h} (\tau_{hi} - \bar{\tau}_{h.})^2$ $\tau_{hi} = w_{hi} y_{hi}$, $\tau_{h.} = \frac{1}{n_h} \sum_{i=1}^{n_h} \tau_{hi}$ **- 총계 추정량** $\hat{Y}$**의 상대표준오차 :** $\widehat{RSE}(\hat{Y}) = \frac{\sqrt{\hat{V}(\hat{Y})}}{\hat{Y}} \times 100(\%)$

구분	현행 표본설계	표본설계 개선안
	- 모평균의 경우, 총합추정량과 총합의 분산추정량을 총 북한이탈주민의 수인 $\widehat{N}_h$로 나누어 계산됨	**(평균 추정)** **- 평균 추정량 :** $$\bar{y} = \frac{\sum_{h=1}^{L}\sum_{i=1}^{n_h} w_{hi}\, y_{hi}}{\sum_{h=1}^{L}\sum_{i=1}^{n_h} w_{hi}}$$ $h = 1,2,\cdots, L$: **층의 수** $i = 1, 2, \cdots, n_h$: **h층 내의 표본 수** w_{hi} : **h층의 i번째 응답자에 대한 최종 가중치** y_{hi} : **h층의 i번째 응답자로부터 얻은 변수값** f_h : **추출률** **- $\bar{y}$의 분산추정량:** $$\widehat{V}(\bar{y}) = \sum_{h=1}^{L} \frac{n_h(1-f_h)}{n_h-1} \sum_{i=1}^{n_h}(e_{hi} - \bar{e}_{h.})^2$$ $e_{hi} = w_{hi}(y_{hi} - \bar{y})/w_{..}$, $\bar{e}_{h.} = \left(\sum_{i=1}^{n_h} e_{hi}\right)/n_h$ $w_{..} = \sum_{h=1}^{L}\sum_{i=1}^{n_h} w_{hi}$ **- 평균 추정량 $\bar{y}$의 상대표준오차 :** $$\widehat{RSE}(\bar{y}) = \frac{\sqrt{\widehat{V}(\bar{y})}}{\bar{y}} \times 100\ (\%)$$

구분	현행 표본설계	표본설계 개선안
모수추정	- 모비율의 추정은 y의 값이 1과 0을 갖는 범주형 변수로 정의하여 모평균의 추정을 위한 공식을 사용함	**(비율 추정)** **- 비율 추정량 $\hat{p}$:** $$\hat{p} = \frac{\sum_{h=1}^{L}\sum_{i=1}^{n_h} w_{hi}\, y_{hi}}{\sum_{h=1}^{L}\sum_{i=1}^{n_h} w_{hi}}$$ **y_{hi}는 특정한 속성을 가지면 1, 그렇지 않으면 0** **- 비율 추정량 $\hat{p}$의 분산추정량 :** $$\hat{V}(\hat{p}) = \sum_{h=1}^{L} \frac{n_h(1-f_h)}{n_h-1} \sum_{i=1}^{n_h}(e_{hi}-\bar{e}_{h.})^2$$ $$e_{hi} = w_{hi}(y_{hi}-\hat{p})/w_{..},\ \bar{e}_{h.} = \left(\sum_{i=1}^{n_h} e_{hi}\right)/n_h$$ $$w_{..} = \sum_{h=1}^{L}\sum_{i=1}^{n_h} w_{hi}$$ **- 비율 추정량 $\hat{p}$의 상대표준오차 :** $$\widehat{RSE}(\hat{p}) = \frac{\sqrt{\hat{V}(\hat{p})}}{\hat{p}} \times 100\ (\%)$$

결론적으로 본 연구에서는 통일부에서 작성 및 관리하고 있는 최신의 표본추출틀(북한이탈주민명부)을 이용하여 새로운 표본설계 방안을 마련하였다. 그리고 패널조사로 전환시 필요한 패널 이탈 방지를 위한 패널관리방안을 마련하였다.

또한 응답자들의 응답부담을 줄이기 위하여 활용도가 높은 정착실태 관련 항목은 매년 생산될 수 있도록 하되, 사회통합과 관련된 조사항목은 사회통합조사 Ⅰ(보건 및 건강, 교육, 가족)과 Ⅱ(노동, 사회활동, 복지, 사회적 관계망, 가구 경제상태)로 영역을 구분하여 짝수년에는 사회통합조사 Ⅰ을, 홀수년에는 사회통합조사 Ⅱ를 조사하는 방안을 제안하였다.

제 1 장 서 론

제 1 절 연구배경 및 목적

‘북한이탈주민실태조사’는 북한이탈주민의 남한정착정도를 파악하고, 북한이탈주민에 대한 지원정책 수립 등에 필요한 기초자료를 제공하기 위해 실시되고 있는 조사통계로 북한이탈주민에 대한 유일한 국가통계이자 북한이탈주민을 대상으로 하는 대표성 높은 통계자료이다. 경제활동 지표(고용률, 경제활동참가율, 임금근로자 근속기간, 평균임금 등), 필요로 하는 지원, 가구경제상태 등 실태조사결과는 정책 사각지대를 파악하고 맞춤형 지원정책을 개발하는데 활용되고 있다.

본 조사는 2011년 성인(19세 이상) 및 청소년(8세~18세) 대상으로 전수조사를 실시한 것을 시작으로 2012년 북한이탈주민실태조사로 명칭이 변경된 후, 2013년 표본조사를 실시하면서 경제활동 중심의 경제활동 실태조사와 사회활동 중심의 사회조사로 분리 조사되었다.

2016년 조사항목이 추가되면서 북한이탈주민 실태조사 명칭이 정착실태조사와 사회통합조사로 변경되었고, 조사대상도 통계청 경제활동 인구조사에 맞춰 15세 이상 북한이탈주민으로 통일되었다.

이후 북한이탈주민정착실태조사(승인번호 437001, 표본조사)라는 명칭으로 국가통계 승인을 받았으며('17.5.), 2019년 「북한이탈주민실태조사」로 통계명칭이 변경되었다.

북한이탈주민실태조사의 조사주기는 1년으로 재단 소속 조사원(북한이탈주민 전문상담사 등)이 조사대상자를 만나 대인면접조사를 실시하고 있으며, 일반국민과의 비교가능성 제고를 위해 조사항목, 문항 규칙을 관련통계(주민등록인구, 통계청 경제활동인구조사 및 사회조사)와 최대한 동일하게 구성하였다. 또한 북한이탈주민 특성을 잘 반영할 수 있는 문항 추가 및 특정 이슈에

대해서는 추가 질의나 보기 항목을 개발하여 조사하되, 시계열 비교가 가능하도록 핵심 조사항목을 매년 유지하고 있다.

<표 1-1> 2021년 기준 북한이탈주민실태조사 조사항목(115개)

기본사항(3문항)	정착실태(55문항)	사회통합(57문항)
- 성별 - 생년월 - 남한에서 최종학력	- 입국 전 사항(4문항) - 생활안정(10문항) - 직업교육훈련(4문항) - 일에 관한 사항(26문항) - 구직에 관한 사항(4문항) - 기타 활동에 관한 사항(5문항) - 이전 직장(일)에 관한 사항(2문항)	- 노동(7문항) - 사회활동(8문항) - 복지(6문항) - 보건 및 건강(11문항) - 교육(6문항) - 가족(11문항) - 사회적 관계망(4문항) - 가구 경제상태(4문항)

북한이탈주민실태조사에서 공표되는 주요 분류 수준은 성별(남성, 여성), 연령대(15~29세, 20대, 30대, 40대, 50대, 60대 이상), 거주지역(수도권(서울/경기/인천), 비수도권), 남한거주기간(3년 미만, 3~5년 미만, 5~10년 미만, 10년 이상) 등으로 되어 있다.

결과보고서는 남북하나재단 홈페이지(www.koreahana.or.kr)에서 보고서 다운로드 서비스를 제공하고 있으며, KOSIS에 결과표를 업로드하여 서비스하고 있다.

2020년 북한이탈주민실태조사의 정기통계품질진단 결과와 표본설계를 검토한 결과 '패널조사로의 전환 검토', '표본설계 개선방안 마련', '조사표 검토' 등이 필요한 것으로 파악되었다.

따라서 본 연구에서는 고정패널이나 분할패널로의 전환 검토, 새로운 표본설계 방안(표본배분, 가중치 작성 및 모수추정, 패널관리방안 등)과 조사항목 등의 개선안을 마련하여 북한이탈주민실태조사의 품질을 개선하고자 한다.

제 2 절 연구내용 및 방법

본 연구에서는 북한이탈주민실태조사 통계의 신뢰성 제고를 위한 개선방안을 제시한다. 먼저 최신의 모집단 분석을 통해 표본이 모집단을 잘 대표하는지, 모집단의 특성이 반영되도록 표본설계가 되어 있는지, 중복 조사되는 표본의 조사결과를 적절히 활용할 수 있는 방법이 있는지 등에 대해 검토하고, 2020년 북한이탈주민실태조사 자료 분석을 통해 표본배분을 위한 최적의 변수를 찾아 표본설계(안)에 반영한다. 이를 위해서 다음과 같은 단계에 걸쳐서 세부적인 내용을 다룬다.

1. 현행 표본설계 및 패널조사 전환 가능성 검토

북한이탈주민실태조사의 현행 표본설계와 표본관리의 문제점에 대하여 구체적으로 검토한다. 그리고 2020년 정기통계품질진단 결과, 현재 표본추출틀 대비 표본크기가 큰 편이어서 매해 중복 응답자가 상당히 존재할 것으로 여겨지므로 정착기간에 따른 실태파악 및 지원정책 효과 파악에 효과적인 패널조사로 조사방법을 변경하는 것이 타당한지 검토한다.

2. 표본설계 개선 방안

2020년에 조사된 북한이탈주민실태조사 원시자료(raw-data)를 이용해서 주요 문항별 특성을 분석한다. 그리고 2021년의 표본설계 전 과정을 상세히 검토해서 현행 조사의 문제점, 개선내용 등을 종합적으로 파악하여 표본설계안에 반영한다.

표본설계안에서는 기존의 표본설계에서 미흡했던 점이나 한계점들을 찾아서 보완토록 한다. 2020년 북한이탈주민실태조사 자료와 2021년 모집단 자료를

분석해서 이를 표본배분에 반영한다. 최종적으로 연구진, 외부전문가 그리고 작성기관 실무진들과의 협의를 통해 표본배분 기준을 정한다. 표본추출은 특성별로 층화하여, 표본배분 원칙에 따라 표본추출프로그램(SAS)을 이용하여 배분한다. 모수 추정을 위한 가중치와 추정치의 오차계산을 위해 분산추정식을 제시하고 아울러 가중치 계산과정을 다룬다.

또한 패널조사로 전환시 필요한 패널 이탈 방지를 위한 패널관리방안을 마련한다.

3. 조사항목 개선

2021년 기준 북한이탈주민실태조사의 조사항목은 크게 기본사항(3문항), 정착실태(55문항), 사회통합(57문항)으로 총 115개 문항으로 구성되어 있다. 문항 수가 상당하여 응답자들에게 응답부담이 높은 것으로 파악되고 있다. 따라서 일부 중복되거나 활용도 낮은 문항을 통합하거나 삭제하는 것에 대한 검토가 필요하며, 조사부문별로 조사주기를 달리하는 방안에 대한 검토가 필요하다.

상기의 단계적인 세부 내용을 수행하기 위해서 다음과 같은 방법을 사용한다.

1) 조사연구

최근 북한이탈주민실태조사 보고서와 현행 표본설계방안 등을 분석한다. 그리고 모집단 분석, 표본오차의 추정을 위하여 연관된 조사 자료를 수집한다. 또한 추정량, 추정량의 오차를 계산하기 위한 가중치 계산, 표본추출방법, 표본배분 공식, 분산추정식, 패널설계 등에 대한 통계이론들을 살펴보고 문헌검토 및 자료 수집을 통해서 여러 사례들을 검토한다.

2) 관련전문가의 의견 수렴 및 검토회의 개최

본 조사를 총괄적으로 담당한 작성기관 관련자와 북한이탈주민실태조사 관련 전문가, 통계청 관계자, 표본설계 전문가[1]와의 자문회의 등을 통해 수시로 문제점 등을 협의한다.

1) 홍기학(동신대학교 교수), 남기성(한국고용정보원 선임연구위원)

제 3 절 연구의 기본 목표와 절차

본 연구에서는 다음 사항들을 기본목표로 한다.

(1) 최신의 모집단의 특성을 분석을 한다.
(2) 북한이탈주민실태조사의 패널화 적정 여부를 검토한 후, 최적 패널 조사 및 패널 이탈 방지를 위한 패널관리방안을 제시한다.
(3) 2021년 표본설계의 장·단점을 분석하고 문제점을 보완해서 조사결과의 정확성과 신뢰성을 높일 수 있도록 표본설계를 한다.
(4) 조사부문별 조사주기를 달리하는 방안 및 조사항목에 대한 개선안을 마련한다.

한편, 효율적인 연구를 위해서 다음과 같은 단계적 절차를 밟는다.

(1) 최신의 북한이탈주민실태조사에 필요한 표본추출틀을 정비한다.
(2) 북한이탈주민실태조사의 패널 전환 의미, 패널 유지 가능성(응답자와 라포형성), 전년도 표본과 중복률 확인 등 패널화 가능성을 검토한다.
(3) 2020년 북한이탈주민실태조사 자료로부터 특성별 분석을 실시하고, 표본배분을 위한 변수 등을 찾아 주어진 조사여건 하에서 층화방법, 표본배분방법, 표본관리방안 등을 종합적으로 검토한 다음 새로운 표본설계에 반영한다. 또한 확률추출을 근간으로 해서 모수추정이 가능하도록 가중치와 추정량, 그리고 추정량에 대한 오차공식 등을 제시한다.
(4) 조사부문(정착실태조사, 사회통합조사)별 조사주기 검토, 활용성이 낮은 문항과 서술형 문항 등 조사항목에 대한 개선안을 마련한다.

제 2 장 북한이탈주민실태조사 작성현황 검토

제 1 절 북한이탈주민실태조사 개요 및 표본설계

1. 2021년 북한이탈주민실태조사 개요

1) 조사목적

- 북한이탈주민의 남한생활 정착에 관한 전반적인 실태조사를 통해 이들의 정착 정도를 파악하고, 성공적 남한 정착을 이끌어내기 위한 정책적 지향점 도출에 활용하기 위한 기초자료 구축

2) 모집단 정의

- 목표모집단 : 북한이탈주민의 보호 및 정착지원에 관한 법 제2조 제1항에서 정의된 바와 같이 “군사분계선 이북지역에 주소, 직계가족, 배우자, 직장 등을 두고 있는 사람으로서 북한을 벗어난 후 외국 국적을 취득하지 아니한 자”들 중 2021년 5월 기준 남한에 정착한 만 15세 이상 북한이탈주민 전체
- 조사모집단 : 2021년 5월 기준 통일부에 등록되어 있는 만 15세 이상 북한이탈주민 중 남한에 실제 거주하는 자
- 표본추출틀 : 2021년 5월 기준 통일부에 등록되어 있는 만 15세 이상 북한이탈주민 명부

2. 2021년 북한이탈주민실태조사 모집단 분포 및 표본설계 현황

1) 표본설계 기본방향

- 모집단인 『조사 시점 대한민국에 거주하는 만 15세 이상의 북한이탈주민』을 대표할 수 있도록 거주지역, 성, 연령, 남한거주기간을 층화변수(stratification variable)로 사용
- 조사의 효율성을 높이고 표본대체로 인한 표본의 왜곡을 막기 위한 표본대체 과정 제시

2) 모집단 및 표본 분포

- 통일부에서 제공하는 북한이탈주민 자료를 사용하여 층화 변수인 거주지역별, 성별, 연령별 그리고 남한거주기간별 모집단과 표본의 분포를 살펴봄(<표 2-1>~<표 2-6> 참고)

- 조사 여건을 고려하여 표본크기는 전국 약 3,000여 명 내외의 북한이탈주민을 목표로 하고 각 층화 변수 별로 정의되는 모비율에 대한 허용오차가 약 5% 이내가 되도록 산정

<표 2-1> 거주 지역별 북한이탈주민 모집단 분포

지역		빈도(명)	백분율(%)	누적	
				빈도(명)	백분율(%)
수도권	서울특별시	6,441	21.6	6,441	21.6
	인천광역시	2,823	9.4	9,264	31.0
	경기도	10,093	33.8	19,357	64.8
비수도권	부산광역시	939	3.1	20,296	67.9
	대구광역시	652	2.2	20,948	70.1
	광주광역시	563	1.9	21,511	72.0
	대전광역시	574	1.9	22,085	73.9
	울산광역시	475	1.6	22,560	75.5
	세종특별자치시	92	0.3	22,652	75.8
	강원도	800	2.7	23,452	78.5
	충청북도	1,302	4.4	24,754	82.8
	충청남도	1,569	5.3	26,323	88.1
	전라북도	512	1.7	26,835	89.8
	전라남도	636	2.1	27,471	91.9
	경상북도	1,063	3.6	28,534	95.5
	경상남도	1,039	3.5	29,573	99.0
	제주특별자치도	307	1.0	29,880	100.0

<표 2-2> 성별 북한이탈주민 모집단 분포

성별	빈도(명)	백분율(%)	누적	
			빈도(명)	백분율(%)
남	7,157	24.0	7,157	24.0
여	22,723	76.0	29,880	100.0

<표 2-3> 연령 그룹별 북한이탈주민 모집단 분포

연령	빈도(명)	백분율(%)	누적	
			빈도(명)	백분율(%)
15 ~ 19세	706	2.4	706	2.4
20대	3,672	12.3	4,378	14.7
30대	7,314	24.5	11,692	39.1
40대	8,515	28.5	20,207	67.6
50대 이상	9,673	32.4	29,880	100.0

<표 2-4> 남한 거주기간별 북한이탈주민 모집단 분포

남한 거주 기간	빈도(명)	백분율(%)	누적	
			빈도(명)	백분율(%)
1~3년 미만	2,321	7.8	2,321	7.8
3~5년 미만	2,436	8.2	4,757	15.9
5~10년 미만	7,823	26.2	12,580	42.1
10년 이상	17,300	57.9	29,880	100.0

<표 2-5> 각 층별 모집단 크기

층 번호	층화변수				모집단(명)
	지역	성별	연령 그룹	남한거주기간	
1	수도권	남	10대	5년 미만	52
2				5년 이상	200
3			20대	5년 미만	242
4				5년 이상	739
5			30대	5년 미만	165
6				5년 이상	1,151
7			40대	5년 미만	110
8				5년 이상	913
9			50대 이상	5년 미만	137
10				5년 이상	1,242
11		여	10대	5년 미만	73
12				5년 이상	182
13			20대	5년 미만	722
14				5년 이상	1,077
15			30대	5년 미만	944
16				5년 이상	2,697
17			40대	5년 미만	674
18				5년 이상	3,572
19			50대 이상	5년 미만	573
20				5년 이상	3,892
21	비수도권	남	10대	5년 미만	16
22				5년 이상	77
23			20대	5년 미만	54
24				5년 이상	284
25			30대	5년 미만	49
26				5년 이상	434
27			40대	5년 미만	41
28				5년 이상	415
29			50대 이상	5년 미만	55
30				5년 이상	781
31		여	10대	5년 미만	25
32				5년 이상	81
33			20대	5년 미만	176
34				5년 이상	378
35			30대	5년 미만	399
36				5년 이상	1,475
37			40대	5년 미만	369
38				5년 이상	2,421
39			50대 이상	5년 미만	318
40				5년 이상	2,675
총합계					29,880

<표 2-6> 각 층별 표본크기

층 번호	층화변수				표본크기(명)
	지역	성별	연령 그룹	남한거주기간	
1	수도권	남	10대	5년 미만	23
2				5년 이상	48
3			20대	5년 미만	50
4				5년 이상	92
5			30대	5년 미만	41
6				5년 이상	115
7			40대	5년 미만	34
8				5년 이상	101
9			50대 이상	5년 미만	37
10				5년 이상	118
11		여	10대	5년 미만	28
12				5년 이상	45
13			20대	5년 미만	86
14				5년 이상	113
15			30대	5년 미만	98
16				5년 이상	176
17			40대	5년 미만	84
18				5년 이상	201
19			50대 이상	5년 미만	77
20				5년 이상	209
21	비수도권	남	10대	5년 미만	12
22				5년 이상	30
23			20대	5년 미만	24
24				5년 이상	57
25			30대	5년 미만	21
26				5년 이상	70
27			40대	5년 미만	19
28				5년 이상	69
29			50대 이상	5년 미만	23
30				5년 이상	94
31		여	10대	5년 미만	16
32				5년 이상	30
33			20대	5년 미만	41
34				5년 이상	67
35			30대	5년 미만	63
36				5년 이상	130
37			40대	5년 미만	61
38				5년 이상	166
39			50대 이상	5년 미만	57
40				5년 이상	174
총합계					3,000

3) 표본추출방법

- 층화 : 거주지역(2), 성별(2), 연령그룹(5), 남한거주기간(2)을 기준으로 40개 층을 구성
- 층화계통추출 : 모집단 정보를 활용하여 각 층 내에서 지역(17개 시도)으로 정렬한 다음 계통추출

4) 표본의 대체

- 조사과정에서 무응답이 발생할 경우 무응답 북한이탈주민의 특성(거주지역, 성별, 연령, 남한거주기간 등)과 동일한 속성을 가진 북한이탈주민으로 대체하는 것을 원칙으로 함

5) 가중치 및 추정

- 표본설계 시 사용된 층 h 및 추정을 위한 사후 층 l에 속한 이탈주민 i에 부여된 표본가중치

$$\alpha_{hli} = \frac{N_h}{n_h}$$

N_h : 층 h의 북한이탈주민 모집단 수, n_h: 층 h의 북한이탈주민 표본 수

- 조사 완료 후 적용되는 사후층화가중치

$$w_{hli} = \frac{N_l}{\hat{N}_l}\,\alpha_{hli}$$

N_l : 사후 층 l의 북한이탈주민 모집단 수, $\hat{N}_l = \sum_h \sum_i \alpha_{hli}$

- 총합추정량

$$\hat{t}_y = \sum_h \sum_l \sum_i w_{hli} y_{hli}$$

- 총합의 분산추정량

$$\hat{V} = \sum_h \left(1 - \frac{n_h}{N_h}\right) \frac{n_h}{n_h - 1} \sum_l \sum_i (z_{hli} - \bar{z}_{h..})^2, \quad z_{hli} = w_{hli} y_{hli}$$

- 모평균의 경우, 총합추정량과 총합의 분산추정량을 총 북한이탈주민의 수인 $\hat{N}_h$로 나누어 계산
- 모비율의 추정은 y의 값이 1과 0을 갖는 범주형 변수로 정의하여 모평균의 추정을 위한 공식을 사용함

제 2 절 북한이탈주민실태조사 현황 및 주요 조사항목 분석

1. 특성별 북한이탈주민실태조사 응답 현황

2020년 기준 특성별(지역, 성별, 연령, 남한거주기간) 북한이탈주민실태조사 응답 현황을 살펴보면 다음과 같다.

<표 2-7> 특성별 북한이탈주민실태조사 응답 현황

층 번호	층화변수				표본크기(명)	응답자수(명)	응답률(%)
	지역	성별	연령 그룹	남한거주기간			
1	수도권	남	10대	5년 미만	25	23	92.0
2				5년 이상	47	42	89.4
3			20대	5년 미만	52	37	71.2
4				5년 이상	93	65	69.9
5			30대	5년 미만	46	39	84.8
6				5년 이상	108	55	**50.9**
7			40대	5년 미만	34	32	94.1
8				5년 이상	98	70	71.4
9			50대 이상	5년 미만	38	37	**97.4**
10				5년 이상	110	95	86.4
11		여	10대	5년 미만	30	23	76.7
12				5년 이상	46	42	91.3
13			20대	5년 미만	96	78	81.3
14				5년 이상	111	80	72.1
15			30대	5년 미만	104	88	84.6
16				5년 이상	170	126	74.1
17			40대	5년 미만	89	81	91.0
18				5년 이상	195	149	76.4
19			50대 이상	5년 미만	78	71	91.0
20				5년 이상	193	160	82.9
21	비수도권	남	10대	5년 미만	15	12	80.0
22				5년 이상	31	28	90.3
23			20대	5년 미만	25	21	84.0
24				5년 이상	56	42	75.0
25			30대	5년 미만	24	21	87.5
26				5년 이상	68	54	79.4
27			40대	5년 미만	22	19	86.4
28				5년 이상	68	58	85.3
29			50대 이상	5년 미만	27	21	77.8
30				5년 이상	88	71	80.7
31		여	10대	5년 미만	18	15	83.3
32				5년 이상	33	29	87.9
33			20대	5년 미만	51	46	90.2
34				5년 이상	69	55	79.7
35			30대	5년 미만	71	69	97.2
36				5년 이상	124	99	79.8
37			40대	5년 미만	67	62	92.5
38				5년 이상	161	143	88.8
39			50대 이상	5년 미만	61	59	96.7
40				5년 이상	158	145	91.8
총합계					3,000	2,462	82.1

3,000명의 표본 중 2,462명(원표본 1,684명, 대체표본 778명)이 응답 완료되어 응답률(응답완료수/표본크기*100)은 82.1%로 나타났다. 특성별 응답률 분포는 50.9~97.4%로 나타났다.

2. 주요 조사항목 분석

1) 특성별 남한생활만족도 비율 및 상대표준오차

2020년 기준 특성별 남한생활만족도 비율 및 상대표준오차 현황을 살펴보면 다음과 같다. 남한생활만족도 조사에서 만족한다라고 응답한 비율은 76.4%로 나타났고, 상대표준오차는 0.3%로 나타났다. 특성별 남한생활만족도 비율 분포는 52.4~92.9%로 나타났으며 상대표준오차는 4.0~25.5%로 나타났다.

<표 2-8> 특성별 남한생활 만족도 비율 및 상대표준오차

층 번호	층화변수				만족도 비율 (%)	상대표준오차 (%)
	지역	성별	연령 그룹	남한거주기간		
1	수도권	남	10대	5년 미만	73.9	12.7
2				5년 이상	92.9	4.3
3			20대	5년 미만	70.3	10.8
4				5년 이상	78.5	6.5
5			30대	5년 미만	76.9	8.9
6				5년 이상	89.1	4.8
7			40대	5년 미만	81.3	8.6
8				5년 이상	68.6	8.2
9			50대 이상	5년 미만	75.7	9.4
10				5년 이상	75.8	5.8
11		여	10대	5년 미만	65.2	15.6
12				5년 이상	76.2	8.7
13			20대	5년 미만	78.2	6.0
14				5년 이상	83.8	5.0
15			30대	5년 미만	75.0	6.2
16				5년 이상	76.2	5.0
17			40대	5년 미만	77.8	6.0
18				5년 이상	75.2	4.7
19			50대 이상	5년 미만	84.5	5.1
20				5년 이상	79.4	**4.0**
21	비수도권	남	10대	5년 미만	58.3	**25.5**
22				5년 이상	92.9	5.3
23			20대	5년 미만	76.2	12.5
24				5년 이상	69.0	10.5
25			30대	5년 미만	52.4	21.3
26				5년 이상	74.1	8.1
27			40대	5년 미만	78.9	12.2
28				5년 이상	77.6	7.1
29			50대 이상	5년 미만	85.7	9.1
30				5년 이상	71.8	7.5
31		여	10대	5년 미만	86.7	10.5
32				5년 이상	69.0	12.7
33			20대	5년 미만	69.6	9.9
34				5년 이상	80.0	6.8
35			30대	5년 미만	72.5	7.5
36				5년 이상	73.7	6.0
37			40대	5년 미만	79.0	6.6
38				5년 이상	70.6	5.4
39			50대 이상	5년 미만	76.3	7.3
40				5년 이상	75.2	4.8
총합계					76.4	0.3

2) 특성별 경제활동지표 및 상대표준오차

2020년 기준 특성별 경제활동지표와 그 상대표준오차를 살펴보면 다음과 같다.

<표 2-9> 특성별 경제활동지표 및 상대표준오차

층번호	층화변수				취업자		실업자		비경제활동	
	지역	성별	연령그룹	남한거주기간	비율 (%)	상대표준오차 (%)	비율 (%)	상대표준오차 (%)	비율 (%)	상대표준오차 (%)
1	수도권	남	10대	5년 미만	8.7	69.1	0.0	.	91.3	6.6
2				5년 이상	7.1	56.3	2.4	100.0	90.5	5.1
3			20대	5년 미만	62.2	13.0	0.0	.	37.8	21.4
4				5년 이상	58.5	10.5	4.6	56.8	36.9	16.3
5			30대	5년 미만	87.2	6.2	2.6	100.0	10.3	48.0
6				5년 이상	80.0	6.8	3.6	70.1	14.5	33.0
7			40대	5년 미만	78.1	9.5	3.1	100.0	18.8	37.4
8				5년 이상	82.9	5.5	2.9	70.2	12.9	31.3
9			50대 이상	5년 미만	64.9	12.3	5.4	69.7	29.7	25.6
10				5년 이상	61.1	8.2	3.2	57.1	34.7	14.1
11		여	10대	5년 미만	0.0	.	4.3	100.0	95.7	4.5
12				5년 이상	16.7	34.9	0.0	.	83.3	7.0
13			20대	5년 미만	34.6	15.7	9.0	36.3	55.1	10.3
14				5년 이상	47.5	11.8	8.8	36.3	40.0	13.8
15			30대	5년 미만	43.2	12.3	9.1	33.9	46.6	11.5
16				5년 이상	51.6	8.7	4.8	40.0	42.1	10.5
17			40대	5년 미만	56.8	9.8	8.6	36.4	33.3	15.8
18				5년 이상	55.0	7.4	6.7	30.6	37.6	10.6
19			50대 이상	5년 미만	38.0	15.3	5.6	48.9	56.3	10.5
20				5년 이상	40.6	9.6	4.4	37.1	52.5	7.5
21	비수도권	남	10대	5년 미만	33.3	42.6	8.3	100.0	58.3	25.5
22				5년 이상	3.6	100.0	0.0	.	92.9	5.3
23			20대	5년 미만	81.0	10.8	19.0	46.1	0.0	.
24				5년 이상	61.9	12.3	9.5	48.1	28.6	24.7
25			30대	5년 미만	95.2	5.0	0.0	.	4.8	100.0
26				5년 이상	88.9	4.9	3.7	70.0	7.4	48.6
27			40대	5년 미만	84.2	10.2	5.3	100.0	10.5	68.7
28				5년 이상	87.9	4.9	0.0	.	12.1	35.8
29			50대 이상	5년 미만	71.4	14.1	9.5	68.9	19.0	46.1
30				5년 이상	74.6	7.0	2.8	70.2	22.5	22.2
31		여	10대	5년 미만	6.7	100.0	6.7	100.0	86.7	10.5
32				5년 이상	17.2	41.4	3.4	100.0	79.3	9.7
33			20대	5년 미만	45.7	16.3	2.2	100.0	52.2	14.3
34				5년 이상	45.5	14.9	3.6	70.1	45.5	14.9
35			30대	5년 미만	39.1	15.1	5.8	48.9	55.1	11.0
36				5년 이상	53.5	9.4	7.1	36.6	39.4	12.5
37			40대	5년 미만	62.9	9.8	8.1	43.2	27.4	20.8
38				5년 이상	57.3	7.2	9.1	26.5	32.2	12.2
39			50대 이상	5년 미만	44.1	14.8	6.8	48.7	45.8	14.3
40				5년 이상	50.3	8.3	4.1	40.1	44.1	9.4
총합계					53.7	0.5	5.6	2.4	39.4	0.7

경제활동지표 중 취업자 비율은 53.7%로 나타났고, 상대표준오차는 0.5%로 나타났다. 특성별 취업자 비율 분포는 0.0~95.2%로 나타났으며, 상대표준오차는 4.9~100.0%로 나타났다. 특히 비수도권, 남자, 10대이면서 거주기간이 5년 이상 응답자의 취업률은 3.6%, 비수도권, 여자, 10대이면서 거주기간이 5년 미만 응답자의 취업률은 6.7%로 낮게 파악되어 상대표준오차가 100.0%로 크게 나타나는 현상을 보였다. 이는 연령층의 특성상 취업자 수가 적은 반면에 층내 변동은 커서 나타나는 것으로 분석할 수 있다.

그리고 경제활동지표 중 실업자 비율은 5.6%로 나타났고, 상대표준오차는 2.4%로 나타났다. 특성별 실업자 비율 분포는 0.0~19.0%로 나타났으며, 상대표준오차는 26.5~100.0%로 나타났다.

또한 경제활동지표 중 비경제활동 비율은 39.4%로 나타났고, 상대표준오차는 0.7%로 나타났다. 특성별 비경제활동 비율 분포는 0.0~95.7%로 나타났으며, 상대표준오차는 4.5~100.0%로 나타났다.

제 3 절 북한이탈주민실태조사 검토 결과

2020년 북한이탈주민실태조사 정기통계품질진단을 통해 도출된 문제점 및 개선사항과 본 연구에서 검토하여 개선이 필요한 사항을 정리하면 다음과 같다.

1. 패널조사로의 전환 가능성 검토

2020년 기준 북한이탈주민실태조사의 표본 3,000명 중 실제 조사완료된 표본 2,462명의 19.2%(472명)가 2019년에도 응답하였으며, 사례관리를 통해 전문상담사와 과거 응답자간 라포가 형성되어 있어 패널관리도 용이할 것으로 보이므로 패널화에 대한 논의가 의미가 있을 것으로 판단된다. 단, 현재의 횡단조사 방식도 '북한이탈주민의 남한정착 실태 파악'이라는 통계작성 목적을 충족하고 있으므로 패널조사로 전환했을 때의 장단점 및 조사목적과의 정합성을 먼저 충분히 검토한 다음 최적의 조사 방안을 선택하여야 한다.

2. 표본설계 개선 필요

조사부담 경감 및 조사 효율성·신뢰성 제고를 위해 최신의 모집단 정보를 활용하여 새로운 표본설계안이 마련될 필요가 있다. 표본설계 시 표본규모 및 층별 표본크기에 대한 적정성의 검토와 함께 적절한 표본배분이 필요하며 가중치를 적용한 모수추정식과 분산추정식 및 상대표준오차가 제시되고 산출되어야 한다. 그리고 패널조사로의 전환을 고려해 패널 이탈 방지를 위한 패널 관리방안을 마련하여야 한다.

3. 조사주기 및 조사항목 개선 필요

북한이탈주민실태조사는 조사내용 및 분량 측면에서 응답자 조사 부담이 상당한 조사로 평가되고 있음에도 불구하고, 기존 정착해 있는 이탈주민의 비중이 점차 높아짐에 따라 지원정책과 정착 측면의 내용을 보다 깊게 파악할 수 있는 새로운 문항 추가에 대한 수요도 지속적으로 발생하고 있다.

또한 자영업자만을 대상으로 하는 문항과 같이 활용도 낮은 문항은 매년 조사항목에 포함시키는 것이 아니라 특정 시기나 필요한 시기에 부가조사를 하는 형태로 변경할 필요가 있다.

더불어 조사항목을 늘리거나 줄이는 문제는 그리 간단한 사항은 아닌 것으로 판단되므로 조사영역을 정책실태와 사회통합으로 나누어 격년 조사하도록 조사주기를 변경하는 방안을 고려해 볼 필요가 있다. 물론 주요 항목에 대해서는 매년 조사하도록 하는 방안도 함께 검토되어야 할 것이다.

제 3 장 패널조사로의 전환 가능성 검토

제 1 절 북한이탈주민실태조사의 중복 표본 현황

북한이탈주민실태조사의 기본 표본추출틀은 통일부에서 작성 및 관리하고 있는 북한이탈주민명부이다. 2020년 5월 기준 북한이탈주민실태조사의 표본수(3,000명)는 모집단 크기(29,400명)에 비해 표본크기가 큰 편이라 할 수 있다. 현재 북한이탈주민실태조사는 신규입국자 수가 점점 줄어 조사대상자는 한정적인데 비해 매해 표본으로 추출되는 응답자가 중복 선정되고 있다.

실제 2020년 북한이탈주민실태조사를 통해 얻은 2,462명의 완료 표본 중 472명이 2019년에도 조사되어 중복 비율이 19.2%로 파악되었으며, 특성별 표본 중복비율 분포는 6.7~61.9%로 나타나 특정 층의 경우 매우 높은 중복 비율을 나타내고 있었다. 구체적인 중복 비율 현황은 <표 3-1>과 같다.

원표본과 대체표본에 대한 관리현황은 <표 3-2>와 같이 완료표본 2,462명 중 원표본이 1,684명, 대체표본이 778명으로 나타나 조사완료된 원표본 비율은 68.4%인 것으로 파악되었으며, 특성별 원표본 비율 분포는 43.8%~100.0%로 나타나 특정층의 경우 조사완료된 원표본 비율이 낮게 나타나고 있음을 알 수 있었다.

이에 중복 조사되는 표본을 보다 효율적으로 관리하고 이탈주민의 남한 정착과정을 종단적으로 분석할 수 있는 패널조사로 조사방식을 변경하는 안에 대한 검토 필요성이 대두되었다.

<표 3-1> 중복표본 현황

층 번호	층화변수				모집단 (명)	표본 크기 (명)	완료 표본 수(명)	완료 표본중 중복수 (명)	완료 표본 중복 비율(%)
	지역	성별	연령 그룹	남한거주기간					
1	수도권	남	10대	5년 미만	57	25	23	5	21.7
2				5년 이상	202	47	42	11	26.2
3			20대	5년 미만	249	52	37	8	21.6
4				5년 이상	798	93	65	20	30.8
5			30대	5년 미만	198	46	39	9	23.1
6				5년 이상	1,077	108	55	14	25.5
7			40대	5년 미만	109	34	32	12	37.5
8				5년 이상	880	98	70	13	18.6
9			50대 이상	5년 미만	132	38	37	8	21.6
10				5년 이상	1,149	110	95	17	17.9
11		여	10대	5년 미만	85	30	23	5	21.7
12				5년 이상	192	46	42	15	35.7
13			20대	5년 미만	858	96	78	6	7.7
14				5년 이상	1,130	111	80	15	18.8
15			30대	5년 미만	1,009	104	88	7	8.0
16				5년 이상	2,676	170	126	15	11.9
17			40대	5년 미만	732	89	81	10	12.3
18				5년 이상	3,502	195	149	10	6.7
19			50대 이상	5년 미만	559	78	71	10	14.1
20				5년 이상	3,462	193	160	14	8.8
21	비수도권	남	10대	5년 미만	20	15	12	6	50.0
22				5년 이상	89	31	28	13	46.4
23			20대	5년 미만	58	25	21	8	38.1
24				5년 이상	286	56	42	15	35.7
25			30대	5년 미만	54	24	21	5	23.8
26				5년 이상	428	68	54	15	27.8
27			40대	5년 미만	46	22	19	7	36.8
28				5년 이상	422	68	58	20	34.5
29			50대 이상	5년 미만	67	27	21	13	61.9
30				5년 이상	714	88	71	15	21.1
31		여	10대	5년 미만	29	18	15	6	40.0
32				5년 이상	102	33	29	10	34.5
33			20대	5년 미만	237	51	46	12	26.1
34				5년 이상	436	69	55	20	36.4
35			30대	5년 미만	470	71	69	10	14.5
36				5년 이상	1,431	124	99	17	17.2
37			40대	5년 미만	410	67	62	11	17.7
38				5년 이상	2,390	161	143	19	13.3
39			50대 이상	5년 미만	341	61	59	7	11.9
40				5년 이상	2,314	158	145	19	13.1
총합계					29,400	3,000	2,462	472	19.2

<표 3-2> 원표본과 대체표본 현황

층 번호	층화변수				모집단 (명)	표본 크기 (명)	완료 표본 수(명)	원표본 (명)	대체 표본 (명)	원표본 비율 (%)
	지역	성별	연령 그룹	남한거주기간						
1	수도권	남	10대	5년 미만	57	25	23	20	3	87.0
2				5년 이상	202	47	42	31	11	73.8
3			20대	5년 미만	249	52	37	24	13	64.9
4				5년 이상	798	93	65	38	27	58.5
5			30대	5년 미만	198	46	39	31	8	79.5
6				5년 이상	1,077	108	55	26	29	47.3
7			40대	5년 미만	109	34	32	23	9	71.9
8				5년 이상	880	98	70	36	34	51.4
9			50대 이상	5년 미만	132	38	37	36	1	97.3
10				5년 이상	1,149	110	95	65	30	68.4
11		여	10대	5년 미만	85	30	23	16	7	69.6
12				5년 이상	192	46	42	28	14	66.7
13			20대	5년 미만	858	96	78	53	25	67.9
14				5년 이상	1,130	111	80	35	45	43.8
15			30대	5년 미만	1,009	104	88	63	25	71.6
16				5년 이상	2,676	170	126	66	60	52.4
17			40대	5년 미만	732	89	81	61	20	75.3
18				5년 이상	3,502	195	149	87	62	58.4
19			50대 이상	5년 미만	559	78	71	57	14	80.3
20				5년 이상	3,462	193	160	104	56	65.0
21	비수도권	남	10대	5년 미만	20	15	12	11	1	91.7
22				5년 이상	89	31	28	24	4	85.7
23			20대	5년 미만	58	25	21	16	5	76.2
24				5년 이상	286	56	42	28	14	66.7
25			30대	5년 미만	54	24	21	18	3	85.7
26				5년 이상	428	68	54	36	18	66.7
27			40대	5년 미만	46	22	19	15	4	78.9
28				5년 이상	422	68	58	39	19	67.2
29			50대 이상	5년 미만	67	27	21	17	4	81.0
30				5년 이상	714	88	71	50	21	70.4
31		여	10대	5년 미만	29	18	15	15	0	100.0
32				5년 이상	102	33	29	18	11	62.1
33			20대	5년 미만	237	51	46	36	10	78.3
34				5년 이상	436	69	55	31	24	56.4
35			30대	5년 미만	470	71	69	55	14	79.7
36				5년 이상	1,431	124	99	63	36	63.6
37			40대	5년 미만	410	67	62	55	7	88.7
38				5년 이상	2,390	161	143	96	47	67.1
39			50대 이상	5년 미만	341	61	59	52	7	88.1
40				5년 이상	2,314	158	145	109	36	75.2
총합계					29,400	3,000	2,462	1,684	778	68.4

제 2 절 패널설계방법 검토 및 패널조사 사례

1. 패널설계방법 검토

북한이탈주민실태조사의 패널화 검토를 위해 우선 다양한 패널설계방법에 대하여 살펴보고자 한다.

패널설계는 시간의 흐름에 따라 조사대상 개체의 변화를 측정하고, 이를 바탕으로 변화요인이 개체 내부의 시간변동에 따른 것인지 아니면 외부의 정책적 효과에 기인한 것인지를 분석할 수 있는 매우 효과적인 설계방법이다. 국내의 통계작성기관은 이와 같은 목적을 달성하기 위해 복지, 보건, 노동, 연금, 재정, 청소년, 여성 등의 사회적 이슈에 따라 다양한 패널조사를 수행하고 있다.

패널설계 방법론에 대해 Duncan(1987)은 시간의 경과에 따른 측정을 위한 조사설계방법으로 반복조사(Repeated Survey), 고정패널조사(Fixed Panel Survey), 순환패널조사(Rotating Panel Survey), 분할패널조사(Split Panel Survey)와 같이 크게 4가지를 제시하고 있다.

먼저 반복조사는 동일한 모집단으로부터 선택된 표본에서 유사한 내용을 서로 다른 시점에 조사하는 방법이다. 따라서 반복조사에서는 어떠한 개체가 연속적으로 한 번 이상 추출되는 것을 보장하지는 않는다.

반면에 고정패널조사는 동일한 표본에서 유사한 내용을 서로 다른 시점에 조사하는 방법이다. 고정패널 조사의 경우 시간이 경과함에 따라 패널 탈락과 노후화의 문제가 발생한다.

순환패널조사는 조사기간 동안 표본개체들을 패널로 유지하는 고정패널 조사와는 달리 패널에 포함된 개체들의 패널포함 기간을 제한하는 방법이다. 순환패널조사에서는 패널 지속기간이 지나면 일부 패널을 탈락시키고, 새로운 개체를 추가하는 방식이다. 이 방법은 일정기간마다 패널을 교체함으로써 모집단 대표성을 확보할 수 있기 때문에 횡단면 모수추정에 주로 사용된다.

마지막으로 분할패널조사는 고정패널조사에 반복조사 또는 순환패널조사를 결합시킨 방법으로 각 설계형태에 따른 장·단점을 절충하여 설계할 수 있다.

1) 반복패널 설계

반복패널 설계는 매년 또는 매월 동일한 모집단으로부터 서로 다른 표본을 추출하여 조사하는 주기성 조사로서 매년 새롭게 표본을 추출하는 방법으로 횡단적 분석에 적합한 표본설계이다. 반복조사 설계는 표본의 중첩을 허용하는 경우와 그렇지 않은 설계로 구분할 수 있으며, 만일 기존의 표본이 지속적으로 패널에 남아있을 경우 중첩된 반복패널로 볼 수 있다.

예를 들어 1차년도 원가구 및 원가구원이 2차년도에도 지속적으로 조사되는 경우 중첩을 허용한 조사방법이며, 패널의 일부는 기존의 패널을 유지하고, 나머지는 새롭게 추출된 가구 및 가구원에 대한 조사가 수행되는 설계로 고려할 수 있다.

반복조사의 대표적인 사례로서는 미국의 ACS(American Community Survey) 조사가 있으며 ACS의 경우 연간통계 및 3년주기 통계와 5년주기 통계를 산출하는데, 그에 필요한 가중치는 별도로 산출하여 제시하고 있다.

즉, ACS 표본설계는 10년주기 센서스에 따른 인구변동을 반영하기 위한 조사 성격이 강하며, ACS 조사목적 중의 하나로서 10년주기 센서스에 따른 모집단의 변화를 반영하기 위한 것이며, 또한 ACS의 표본설계는 횡단적 인구변동을 반영하기 위한 것으로서 종단적 연구를 위한 패널설계와는 목적성이 다르다고 할 수 있다.

2) 고정패널 설계

고정패널설계는 동일한 표본에서 유사한 내용을 서로 다른 시점에 조사하는 방법으로서 장기간의 관찰을 통한 개체의 변동을 추정하는 경우에 적용되는 설계방법이며 현재 국내의 대다수 패널설계가 고정패널설계를 따르고 있다.

고정패널설계의 특징으로는 첫째, 조사기간동안 표본개체를 지속적으로 패널로 유지하는 패널조사로서 정해진 패널 유지기간은 별도로 없으며, 패널마다 특징은 있지만, 대체로 원표본과 더불어 연계된 단위를 신규패널로 흡수하는 개방형 패널과 신규표본의 진입을 허용하지 않는 폐쇄형 패널로 구분한다. 이때, 개방형 고정패널은 패널이 유지되는 동안 일정 정도의 패널규모를 유지할 수 있어 원표본의 탈락에도 불구하고, 장기간 패널을 유지할 수 있는 장점이 있는 반면, 시간이 흐름에 따라 원표본의 특성을 상실하는 단점이 있다. 폐쇄형 고정패널은 신규 단위의 진입을 허용하지 않기 때문에 시간이 지남에 따라 원표본의 규모가 지속적으로 감소하며, 전체 패널 규모도 같이 감소하기 때문에 장기간 패널을 구축할 경우 횡단적 대표성에 문제가 발생할 수 있다.

둘째, 고정패널의 자료수집방법은 추적조사로 인한 비용이 많이 소요되나, 1차 패널이 구축된 이후 별도의 패널 설계가 불필요하기 때문에 패널설계에 대한 부담이 적다.

셋째, 고정패널은 장기간 패널을 유지해야 하기 때문에 다른 패널 설계에 비해 응답부담이 높은 특징을 갖는다.

넷째, 고정패널은 1차년도 진입한 패널의 시간의 흐름에 따라 장기간 추적 관찰하기 때문에 횡단적 분석보다는 종단적 분석에 유리하며, 장기간 추세를 관찰할 수 있는 장점이 있으나, 장기간 패널에 포함되기 때문에 응답자의 패널조건화*가 다른 패널에 비해 크게 나타나는 단점이 있다.

대표적인 조사로서는 캐나다의 Canadian Population National Health Survey, 미국의 PSID, SIPP, U.S.MEP 등이 있으며, 국내에서는 여성가족패널, 복지패널, 한국의료패널, 노동패널 등 다수가 있다.

* 패널조건화(panel conditioning)
- 이전 조사에 참여했던 경험으로 인해 현재 조사의 응답에 영향을 주는 것을 의미
- 패널 피로감에 의해 불성실한 응답을 할 가능성이 있으며, 이전의 응답내용을 기억하고 의식적으로 일관성을 유지하려 하거나 반대로 의식적으로 변화를 보이는 방향으로 응답할 가능성이 있어 연구 결과의 정확성이 낮아질 수 있음

3) 순환패널 설계

순환패널설계는 표본개체들의 패널수명을 제한하는 방법으로 일정 기간 동안 패널에 포함된 표본은 탈락시키고 새로운 패널로 교체하는 방법으로 월별·분기별 변동을 추정하는 경우에 많이 활용되며 다음과 같은 특징을 갖는다.

첫째, 조사기간 동안 표본개체를 지속적으로 패널로 유지하는 고정패널과는 달리 표본개체들의 수명을 일정기간으로 제한하여 패널 지속기간이 지나면 표본을 탈락시키고 새로운 표본개체를 추가하여 주기적으로 패널표본을 교체하는 방법으로서 특정시점의 모수 또는 일정 기간 동안의 평균적 변화, 일정 기간 동안의 순변동을 측정할 목적으로 설계된다.

둘째, 순환패널의 자료수집방법은 고정패널에 비해 추적조사로 인한 비용을 줄일 수 있으나, 매 주기마다 표본설계 작업이 요구된다.

셋째, 순환패널은 패널 유지기간이 정해져 있기 때문에 응답자 부담이 고정패널에 비해 훨씬 낮다.

넷째, 순환패널은 고정패널에 비해 패널탈락률이 낮기 때문에 모집단 대표성이 상대적으로 높고, 패널교체를 통해 모집단의 새로운 개체반영이 가능하다.

마지막으로 고정패널과 비교할 때, 응답자의 패널조건화를 감소시킬 수 있다.

이를 위해 순환패널에서는 다음과 같은 사항을 고려해야 한다.

① 표본의 효율을 높이기 위해 표본의 대체율을 어떻게 정할 것인가?
② 현시점의 모집단 평균을 추정하기 위해 과거의 자료를 이용하는 것이 정도를 높일 수 있는가?
③ 대체되는 시점의 표본설계를 얼마나 효과적으로 할 수 있는가?
④ 비용대비 추정의 효과는 고정패널설계와 비교하여 월등한가?
⑤ 순환율에 따른 조사의 속도는 어느 정도로 정할 것인가?

순환패널의 대표적인 조사로서는 캐나다의 National Crime Victimization Survey Canadian Labour Force Survey, SLID 등이 있으며, 국내의 경우 순환패널 방식은 아니지만, 유사한 방법으로 설계된 조사로 통계청의 가계동향조사가 있다.

4) 분할패널 설계

분할(또는 분리)패널은 고정패널조사에 반복조사 또는 순환패널조사를 결합한 조사방법으로 설계자체가 복잡하고 패널관리의 문제점이 있으나, 설계목적에 따라 이와 같은 방법을 적용할 수 있다. 분할패널은 횡단적 비교연구와 종단적 분석이 모두 필요할 때 사용 가능하며, 표본의 일부는 유지(고정패널)하여 횡단적 비교를, 나머지 일부는 패널을 변경(표본추출 또는 순환패널 구축)하여 동적 변화를 살펴볼 수 있다. 그리고 횡단적으로 유지되는 단위와 패널로 유지되는 단위 간 비교도 가능하다. 대표적인 분할패널 설계로 이탈리아 은행의 SHIW(Survey on Household Income and Wealth)와 스페인의 SHF가 있으며, 두 패널의 특성은 유사하지만, 패널유지과정에서 차이가 있다.

2. 패널조사 사례

1) 고정패널 : 영국 BHPS(British Household Panel Study)

(1) 조사개요

영국의 에섹스대학교(University of Essex) 내 연구소(ISER)에서는 영국의 개인과 가구 수준에서의 사회 경제적 변화를 이해하고자 BHPS를 조사하고 있다. 1991년 조사를 시작한 이래로 약 5,000가구에 대해 매년 조사를 실시한다. 조사방법은 PAPI(Paper And Pencil Interviewing)로 시작하였으며 9차 웨이브부터는 CAPI(Computer Assisted Personal Interviewing)로 변경하였다. 전국을 대표하는 16세 이상 개인에 대해 가구구성, 노동력, 소득, 재산, 주택, 건강, 사회-경제적 변수들을 조사한다.

(2) 표본설계

BHPS에서는 표집틀(sampling frame)인 우편번호 주소 파일(PAF)로부터 2단계 층화집락계통추출법에 의해 표본을 선정하였다. 1단계에서는 1차추출단위(PSU: Primary Sampling Unit)인 우편번호 섹터를 250개 선정하였다. 우편번호 섹터란 최소 500가구가 포함되도록 구성된 일종의 우편번호 집락이다. 이와 같이 구성된 PSU에 대해 1981년 인구센서스 결과로부터 지역과 가구주의 사회경제적 그룹(직업이나 경영위치) 비율과 같은 사회인구학 변수를 사용하여 1차(major) 층화를 수행하였다. 다음으로 1차 층 내에서 연금수혜연령(여자는 60세, 남자는 65세) 비율에 의해 2차(minor) 층화를 하였다. 각 층이 구성된 후 비도시지역은 농업종사자 비율, 도시지역은 연금수혜연령 미만인 1인가구의 비율을 분류기준으로 설정하고, 이 분류기준에 의해 정렬 후 계통추출하였다.

2단계에서는 2차추출단위(SSU: Second Sampling Unit)인 주소를 8,167개 선정하였는데, PSU내에서 평균적으로 33개의 주소가 선정될 수 있도록 계통추출한다. 선택된 주소 내에 3가구 이하의 가구가 거주하는 경우에는 모든 가구를 표본으로 선정하고, 주소 내 3가구를 초과하는 경우에는 3가구만을 랜덤추출하였다.

패널형태는 고정패널로 원표본(original sample) 가구원으로부터 출생하거나 입양된 가구원, 원표본가구로 전입해온 가구원, 원표본가구가 전출하여 속하게 된 가구의 가구원을 모두 원표본으로 간주하였다. 또한 임시적으로 원표본가구에 거주하게 된 사람의 일부도 원표본으로 간주하고 추적조사를 함으로써 장기 고정패널에 의한 표본마모를 일부 보충할 수 있도록 하였다. 1차 웨이브의 응답자에 대해 표본마모 현황은 <표 3-3>과 같다. 1차 웨이브 응답자 중 2차 웨이브 부적격자는 1.4%였으며, 적격자 중 응답을 한 경우가 2차에서 87.7%, 10차에서는 70.4%였다. 또한 항목무응답의 경우에는 핫덱(Hot-Deck)과 회귀대체방법 이용하여 무응답 대체를 수행하였다.

<표 3-3> BHPS 1차 웨이브 응답자의 응답률

	% ineligible	% of eligible responding	% of eligible responding at all waves	Wave on wave response rate(interviewed all waves up to previous)
Wave 2	1.4	87.7	87.7	87.7
Wave 3	2.9	81.5	79.1	90.3
Wave 4	4.3	79.9	74.8	94.9
Wave 5	5.6	76.8	70.6	94.8
Wave 6	6.9	77.3	68.7	97.6
Wave 7	8.4	76.0	66.7	97.6
Wave 8	9.5	74.1	64.7	97.4
Wave 9	10.5	72.1	62.4	97.0
Wave 10	12.0	70.4	60.0	96.7
Wave 11	12.8	68.4	59.3	96.1
Wave 12	13.7	66.6	57.1	96.5
Wave 13	14.8	64.9	55.1	96.8

2) 순환(연동)패널 : 미국 SIPP(Survey of Income and Program Participation)

(1) 조사개요

미국 센서스국(Census Bureau)에서는 개인과 가구의 소득, 프로그램 참여, 그리고 이에 대한 주요 결정요인에 대한 정확하고 포괄적인 정보를 제공하기 위해 SIPP조사를 수행하고 있다. 1984년 처음으로 조사가 시작한 이래로 1996년에는 종단면 추정값의 품질을 향상시키기 위해 설계를 변경하였다. 1996년 설계 변경의 주요 내용으로는 초기표본의 수를 약 37,000가구로 증가시키고, 32개월(8차 웨이브)의 중복(overlapping) 패널 대신 4년(12차 웨이브)의 단일 패널로 변경하였으며, CAI(Computer-Assisted Interviewing)를 도입하고, 빈곤층 밀집지역에서 가구를 과대표집하였다. 표본의 크기는 약 4만 가구이며, 조사 주기는 4개월이다.

이때 조사원의 업무 부담을 매월 균등하게 배분하기 위해서 전체 표본을 1/4씩 랜덤하게 선정해 4개의 연동그룹을 구성하여 4개월 동안 매월 한 개의 연동그룹을 조사할 수 있도록 하였다. 설문지는 매조사마다 질문하는 핵심(core)파트와 조사의 빈도와 주기는 일정하지 않지만 특정주제에 대한 깊이 있는 질문이 가능한 주제별 모듈(topical module)로 구성되어 있다. 핵심파트는 인구통계학적 특성, 노동력 참여, 프로그램 참여, 이전소득과 다양한 프로그램에서 받은 비화폐성 혜택을 포함한 근로소득과 불로소득, 자산 소유권, 사적의료보험 등의 항목으로 구성되어 있다. 주제별 모듈에서는 자산과 부채, 학교 입학, 결혼, 출산, 이주, 장애 및 직업 기록 등을 조사한다.

(2) 표본설계

1996년 표본설계의 내용을 위주로 살펴보면 SIPP의 표본설계는 2단계 집락 추출방법으로 1단계에서는 PSU인 지역을 선정한다. 총 322개의 PSU 중에 112개는 자기대표성이 있는(SR : Self-Representing) PSU이고, 자기대표성이 없는(NSR : Non-SR) PSU가 210개이다. SR PSU는 모두 선정되며, NSR PSU는 4개의 센서스 지역(S/NE/MW/W)과 15개의 층화 변수에 의해 105개 층으로 구분되고, 각 층에서 1990년 센서스의 주택수에 비례하는 확률로 2개씩 추출된다. 이때 선형 알고리즘을 적용함으로써 1980년 센서스와 1990년 센서스로부터 선정된 PSU가 최대한 중복될 수 있도록 하였다. 1996년 표본설계에서는 이전 설계에 비해서 PSU 수가 증가(1985-1991 : 230개, 1992-1993 : 284개)하였는데, 이로 인하여 PSU간 분산은 줄어들고 추정값의 분산추정에 대한 자유도는 늘어나는 효과가 나타났다.

2단계에서는 주소로 집락을 구성하고 SSU인 주소 집락을 선정하였다. 만약 집락 내 예상된 거처보다 많은 거처가 있을 때에는 집락 내에서 부표집(sub-sampling)을 실시하였다. 먼저 주소로 집락을 구성하기 위해 5개의

개별적인 추출틀을 이용하였다. 첫 번째는 Unit 추출틀(the address enumeration districts frame)로 건축허가가 발급되고 최소한 96%의 주소가 완전한 지역에 있는 센서스 조사구 내에서 거처 주소가 있는 추출틀이며, 하나의 거처가 집락이 된다. 두 번째는 Area 추출틀(area EDs frame)인데, Unit 추출틀의 나머지 센서스조사구들로부터 건축허가가 발급되지 않았거나 4%이상의 주소가 누락된 지역의 주소 추출틀이다. 대부분 시골지역이 해당되며, 4개 거처가 예상되도록 집락을 구성하였다. 세 번째로 Group Quarters 추출틀(special places frame)은 기숙사나 호텔, 보호시설과 같이 센서스에서 주택으로 카운트되지 않은 거처의 주소 추출틀로서 하나의 거처에 상당하는 집락으로 구성되었다. 네 번째는 New Construction 추출틀로 지난 센서스 이래로 건축허가가 발급된 새로운 건축물에 대한 추출틀이며, 4개 거처 또는 8개 거처가 한 집락이 되도록 구성되어 있다. 다섯 번째는 Coverage Improvement 추출틀로 센서스에서는 누락되었지만 사후조사에서 발견된 거처의 주소(0.1%) 추출틀로서 집락은 구성되지 않는 추출틀이다.

SIPP는 저소득 가구의 표본크기를 증가시켜 빈곤 추정값의 분산을 줄이기 위해서 저소득 가구에 대한 과대표집을 실시하였다. 센서스 정보를 이용하여 각 PSU내의 거처를 빈곤층과 아닌 경우로 구분하고, 각 PSU에서 고소득층에 비해 저소득층을 1.6배 높게 선정될 수 있도록 선정하였다.

패널의 형태는 연동패널로 하나의 패널을 4년 동안 유지하다가 4년이 지나면 일시에 표본을 교체하는 단일 패널로 설계되어 있다. 이 단일 패널에서 표본을 1/4씩 랜덤하게 추출하여 동일한 크기가 되도록 4개의 연동그룹을 구성하였다. 각 연동그룹은 4개월마다 인터뷰에 참여하게 되고, 매월 하나의 연동그룹을 조사하는 방식으로 운영하고 있다.

3) 분할패널 : 스페인 SHF(Survey of Household Finances)

(1) 조사개요

스페인 은행에서는 스페인 경제의 금융 계정의 집계 데이터를 보완하고, 스페인 가구의 투자 및 자금 조달 결정과 이들의 금융 포지션을 만드는 연구를 위해 SHF를 실시하고 있다. SHF로 인하여 수행되는 연구들은 스페인 경제의 더 나은 이해와 공공 정책을 적절히 디자인하는 역할을 하고 있다. 이 조사는 이탈리아 은행의 SHIW를 기초로 하여 2002년에 처음으로 시작하였다. 조사 주기는 3년으로 2005년에 2차, 2008년에 3차 웨이브가 진행되었다. 표본크기는 약 7천 가구로 CAPI방법에 의해 조사를 수행하고 있다.

인터뷰에 소요되는 시간은 중앙값이 55분이고 전체 응답자의 약 95%가 2시간 미만에 조사를 완료한다. 조사되는 내용으로는 인구통계, 실물자산 및 그와 관련된 채무, 기타부채, 금융자산, 연금 및 보험, 고용시장 상황과 근로소득, 직전연도의 비근로소득, 결제수단, 소비와 저축 등이 있다.

(2) 표본설계

SHF의 표집틀 작성을 위해 통계청에서는 먼저 주소에 의해 정의된 가구단위의 도시등록부(Municipal Register)를 국세청으로 이송한다. 그 다음 국세청에서 재산과 소득세금환급액에 기초한 정보인 다음의 세 변수를 각 주소마다 추가하여 표집틀을 구축한다. 첫 번째 변수는 가구 총 과세재산을 기준으로 구성된 재산층(wealth strata)이다. 두 번째 변수는 소득세는 있으나 재산세는 없는 가구에 대하여 가구가 속해있는 전국 소득분포의 사분위수이며, 세 번째 변수는 1인당 가구소득이다. 이와 같은 소득관련 변수로 정렬을 한 다음 계통추출(systematic sampling)을 하는 방식으로 표본을 선정한다. 각 웨이브마다 실제로 설계된 내용은 뒤에서 좀 더 설명하겠다.

SHF 표본설계의 특징은 부유층에 대한 과대표집이라 할 수 있다. 일반적으로 부유층은 인구에 있어서는 낮은 비율을 차지하고 있지만, 전체 가구 재산에 대해서는 큰 비중을 보인다. 인구에 대한 대표성 확보뿐만 아니라 재산의 총계를 내기 위하여 과대표집을 실시하였다. 이를 위해 스페인 은행은 국세청과 통계청의 협조를 받아 개인의 재산세 정보를 기초로 과대표집을 실시하였다. <표 3-4>에는 2002년과 2005년 과대표집 비율이 나와 있는데, 2005년 상위 1%의 부유층의 경우 랜덤추출에 비해 약 9배정도 과대표집 한 것을 알 수 있다.

<표 3-4> SHF의 과대표집 비율

Net worth decile group	EFF 2002		EFF 2005	
	Number of observation	Oversampling rate	Number of observation	Oversampling rate
Bottom 50%	1,878	0.73	2,234	0.75
50% to 90%	1,944	0.94	2,036	0.85
90% to 95%	429	1.67	481	1.61
95% to 99%	524	2.55	675	2.83
Top 1%	368	7.16	536	8.99

제 3 절 패널조사로의 전환 가능성 검토

1. 패널조사 전환 필요성 및 최적 패널설계 방법

북한이탈주민실태조사를 패널조사로 변경하여 조사내용을 추적조사할 수 있다면, 정착기간에 따른 실태파악 및 다양한 지원 정책의 효과 등을 파악하고 북한이탈주민을 위해 보다 더 필요하고 효과적인 지원정책을 마련하는 데에 도움이 될 수 있을 것으로 보인다.

특히 본 조사는 탈북민 생활·심리상담을 담당하는 재단 소속의 전문상담사가 조사를 담당하고 있어 조사대상과 일정수준의 유대관계를 구축하고 있으며, 조사원 외 보호경찰관, 복지기관 및 주변 북한이탈주민 등을 통해 조사대상을 최대한 추적하는 등 표본관리 노하우가 축적되어 있어 패널 유지 활동을 통한 효율적인 패널관리가 가능할 것으로 보인다.

따라서 여러 가지 패널설계방법 중 북한이탈주민실태조사에 적합한 방법을 선택하는 것이 중요한데, 패널설계방법에 따른 특성의 차이와 북한이탈주민 실태조사의 상황을 고려할 때 개방형 고정패널설계나 분할패널설계가 적절할 것으로 판단된다.

개방형 고정패널설계의 경우는 매년 표본설계를 하지 않고 패널의 노후화가 시작되는 때 표본을 보완해도 된다는 장점이 있으나, 시간이 흐름에 따라 원표본의 특성이 상실되어 남한정착실태 파악에는 한계가 있을 수 있다. 분할패널설계를 할 경우는 북한이탈주민의 횡단적 비교연구와 종단적 분석이 모두 가능하다는 장점이 있으나 매년 또는 일정한 주기를 가지고 표본설계를 새로 하여야 하고 두 종류의 패널(고정/변경)로 구성되어 패널관리에 더 많은 주의를 기울여야 하며, 설계 자체가 복잡하다는 단점이 있다.

이 중 어떤 패널설계방법을 선택하더라도 패널 관리 및 유지를 위해 조사 답례품 수준을 현재보다 높일 필요가 있으며, 표본설계 및 패널분석이 가능한 인력이 필요하므로 이에 따른 추가 예산 확보가 선행되어야 한다.

2. 패널관리방안

패널조사는 매 주기에 조사를 실시하고, 비교적 조사의 양이 많기 때문에 응답 거부 상황이 많이 발생하나, 동일한 응답자가 계속해서 조사에 응답하여야만 자료의 가치와 유용성이 있으므로 응답자가 계속 조사에 참여하도록 하는 패널 유지 활동이 매우 중요하다.

패널 유지 활동은 지속적인 패널 관리를 통해 조사협조를 유지하는 방법과 매 조사에서 응답 거부 시 조사에 응하도록 설득하는 방안이 함께 고려되어야 하는데 최근에는 모든 응답자를 대상으로 패널 유지활동을 하려면 비용이 많이 들기 때문에 특정한 속성을 가진 응답자에 대해 집중적으로 유지 활동을 실시하기도 한다.

먼저 조사협조를 위해 일반적으로 시행되고 있는 패널 유지 방법은 다음과 같다.

첫째, 관련 정책수립을 위한 기초자료를 제공하고 있다는 자긍심, 사회활동의 하나로 가치있는 일이라는 위안 등의 정서적 보상을 한다.

둘째, 패널 응답에 따른 적절한 보상을 통해 지속적인 조사참여를 유인하고, 조사 후 경품행사를 통해 조사에 대한 관심을 유지할 수 있도록 한다.

- 생일 및 기념일에 카드 보내기
- 주기적인 메일 보내기
- 이사 시에 연락처를 보낸 응답자에 대해 선물하기
- 장기적인 관점에서 응답자와 면접원 간 신뢰관계 유지로서 동일 면접원이 조사하기 등

다음으로 패널 응답 거절에 대한 설득 방식은 아래와 같다.

첫째, 조사의 공공성과 신뢰성에 대한 의심 패널의 경우, 조사안내 팸플릿, 협조 공문 등을 제시하고 조사주체와 조사목적, 활용 방안 등을 자세히 설명하도록 한다.

둘째, 조사 참여로 인한 불이익에 대해 우려를 표현하는 패널의 경우, 개인정보보호와 관련된 통계법 제33조를 설명하고 조사 진행 후 비밀보장 각서를 제공하도록 한다.

셋째, 조사 참여에 대한 실질적인 혜택을 요구하는 패널에게는 물질적 보상과 동시에 패널조사 결과자료를 공유하는 등 정보 제공과 정서적 보상도 함께 제공하도록 한다.

넷째, 조사 진행 과정에 대한 불만 패널에 대해서는 면접원 교체, 시기 조정 등 적극적으로 맞춤식 조사를 진행하도록 한다.

3. 현행 조사방법을 유지하면서 심층면접 방법을 병행하는 방안 검토

북한이탈주민실태조사는 북한이탈주민실태 파악을 주목적으로 하고 있는데, 현재의 횡단조사만으로도 통계작성목적을 충족하고 있다고 판단된다. 추가적으로 남한정착형태 변화 양상이나 이탈주민 정책 효과성 확인 등 심층분석이 필요한 경우 현재 조사방법처럼 매해 표본을 독립적으로 추출하여 횡단적 분석을 실시하되 일정 수의 주민들을 대상으로 심층면접을 추가 실시하는 질적방법을 병행하는 방안도 고려해 볼 수 있다.

제 4 장 표본설계 개선 방안

제 1 절 표본설계 개선안의 특징

본 연구에서 제시하는 북한이탈주민실태조사 표본설계 개선방향은 다음과 같다.

(1) 조사의 목적을 충실히 달성하기 위해서 합리적인 표본배분 원칙과 실사의 효율성을 높이도록 한다.

(2) 최신의 모집단 분석 결과와 2020년 조사데이터 분석결과를 토대로 표본배정을 위한 층화변수 등을 찾고, 표본배분을 하는데 비례배분, 제곱근비례배분, 멱등배분 등을 검토하고, 조사여건을 고려하여 최종 결정한다.

(3) 가중치 부여 과정을 포함해서 모수추정식과 추정식의 오차계산 공식을 제시한다.

제 2 절 층화

본 연구에서는 이전 조사(2015년~2021년)와의 일관성을 유지하기 위하여 2020년 조사와 동일한 층화기준을 사용하고자 한다.

북한이탈주민실태조사는 주요관심사인 북한이탈주민 경제활동상태 양상이 구분되어지는 지역(2개 층), 성별(2개 층), 연령(5개 층), 남한거주기간(2개 층)의 4개 변수를 층화변수로 구성하여 총 40개의 층으로 구분되어 있다. 거주지역은 수도권(서울, 경기, 인천)과 비수도권 지역으로 층을 구성하였으며, 성별은 남성과 여성의 층으로 구성하였다. 그리고 연령은 15~19세, 20대, 30대, 40대, 50대 이상으로 층을 구성하였으며, 남한거주기간은 5년 미만과 5년 이상으로 층을 구성하였다.

층화기준에 따른 응답 특성 분석 결과 연령이 15~19세인 경우 모집단 크기가 크지 않고 다른 연령대에 비해 응답 협조도가 낮은 특성을 나타내고 있어 20대 이하 연령층으로 합치는 안에 대한 검토가 필요한 것으로 나타났다. 이때 북한이탈주민 관련 전문가와 통계이용자들의 의견을 충분히 수렴하여 반영하여야 한다.

한편 현재 층화기준과 공표수준(KOSIS)의 범주가 달라 층별 가중치 적용 시 세분화 된 공표수준별 추정이 불가능한 상황이다. 층화기준에서는 남한거주기간을 5년미만/5년이상으로, 공표수준에서는 3년미만/3~5년미만/5~10년미만/10년이상으로 범주화하고 있으며, 거주지역도 층화기준에서는 수도권/비수도권으로, 공표수준에서는 수도권(서울/경기/인천)/비수도권으로 범주화하고 있다. 따라서 공표수준을 층화기준에 맞게 일치시킬 필요가 있다.

제 3 절 표본크기 결정 및 표본배분

1. 표본크기 결정

2020년 5월 기준 북한이탈주민실태조사는 모집단 크기(29,400명)에 비해 표본크기(3,000명)가 큰 것으로 평가받고 있는데 특정 층의 경우 모집단에 비해 표본이 과도하게 할당되어 과거 조사부터 연속적으로 표본으로 선정되거나 대체표본으로 충당하기 어려운 경우가 발생하고 있어 표본크기를 줄이는 방안 마련이 필요해 보인다.

일반적으로 조사를 계획할 때에는 조사에서 얻고자 하는 추정값의 오차의 한계를 미리 정해주는데, 이것을 목표오차(목표정도, target precision)라고 한다. 목표오차는 절대적인 값으로 정해줄 수도 있고, 상대적인 값으로 정해줄 수도 있는데 이 기준을 충족하는 최소 표본수를 구한 다음 가용 조사인력, 예산, 조사 소요 시간 등 조사에 필요한 제반 여건을 종합적으로 검토하여 표본크기를 결정하게 된다.

본 연구에서는 지역별, 성별, 연령별, 남한거주기간에 따라 추정값에 차이가 크게 나타날 수 있다는 점을 고려하여 상대표준오차를 기준으로 표본크기를 결정하고자 한다. 이 조사의 주요 관심 조사항목은 지역별, 성별, 연령별, 남한거주기간별 응답자 수이므로 목표 상대표준오차를 달성하기 위해 필요한 표본의 크기는 다음의 식을 통해 얻을 수 있다.

$$n = \frac{N\sum_{h=1}^{L} N_h P_h Q_h}{N^2 D + \sum_{h=1}^{L} N_h P_h Q_h}$$

여기서 $D = (B/z_{\alpha/2})^2$이고, B는 전체 북한이탈주민의 통계를 생산하는 데 사용할 허용오차이다. h층의 모비율 P_h, 그리고 N과 N_h는 조사모집단 자료를 사용한다.

북한이탈주민실태조사에서 표본크기가 2,000명일 때, 목표 상대표준오차는 2.1%이며, 2,500명일 때 1.9%, 3,000명일 때 1.7%이다.

2. 표본배분방법 검토

1) 비례배분

각 층별로 모집단의 구성비만큼 표본을 배분하는 방법으로서 표본 층의 구성비를 모집단층의 구성비에 맞추어 주는 방법으로 모집단에 대한 이용 가능한 정보가 거의 없을 경우 활용되는 배분방법이다. 비례배분의 장점으로는 추정량의 식이 자체가중추정량식으로 변환되어 추정식이 간단해지며, 단순임의 추출에 비해 분산이 작다.

$$n_h = n\frac{N_h}{N}$$

2) 제곱근($\sqrt{N}$) 비례배분

비례배분은 층의 크기가 상이한 경우, 층의 크기가 큰 층에 지나치게 많은 표본을 배분하여 층의 크기가 작은 층의 표본 대표성에 문제가 발생할 여지가 많은 것이 단점이다. 제곱근비례배분 방법은 이와 같은 비례배분의 단점을 보완한 것으로 다음과 같은 제곱근비례배분을 사용한다. 이 방법은 조사에서 전체 추정치에 대한 정도보다는 각 층별로 추정치에 대한 정도를 개별적으로 관리하고자 할 때 유용하게 쓰인다.

$$n_h = n \times \frac{\sqrt{N_h}}{\sum_h \sqrt{N_h}}$$

3) 네이만배분

네이만배분은 층의 크기와 층내 변동을 함께 고려한 배분방법으로 층의 크기가 크고 층내 변동이 큰 층에 표본을 많이 배분하도록 하는 방법이다. 네이만 표본배분 식은 다음과 같다.

$$n_h = n \times \frac{N_h \sqrt{P_h Q_h}}{\sum_h N_h \sqrt{P_h Q_h}}$$

여기서 모비율 P_h를 모르면 표본비율 $\hat{p}_h$을 이용한다.

4) 멱등배분

멱등배분은 네이만배분식에 멱수를 주어 최적배분의 형식을 취하면서, 층의 크기와 표준편차의 곱에 비례하도록 표본을 배분함으로서 특정 층에 지나치게 많은 표본이 배분되는 것을 방지할 수 있는 방법이다. 멱등배분은 네이만배분의 장점과 비례배분의 장점을 동시에 가지고 있어 최근에 널리 이용되고 있는 배분방법이다. p 값에 따라 다양한 배분이 가능하기에 최적의 p 값을 찾는데 어려움을 가질 수 있으나, 일반적으로 강제배분을 최소화하는 p 값을 최적값으로 선택한다. 층별 멱등배분 식은 다음과 같다.

$$n_h = n \times \frac{(N_h \sqrt{P_h Q_h})^p}{\sum_h (N_h \sqrt{P_h Q_h})^p}, \quad 0 < p \leq 1$$

3. 표본배분

1) 표본크기가 3,000명 일 때

(1) 비례배분

3,000명의 표본을 총 40개 층에 대해 비례배분을 한 결과, 각 층별 모집단 크기 및 표본크기는 다음과 같다. 각 층별 허용오차 분포는 4.8~69.2%로 나타났다.

<표 4-1> 층별 비례배분에 따른 허용오차(표본크기=3,000)

층 번호	층화변수				모집단(명)	비례배분 (명)	허용오차 (%)
	지역	성별	연령 그룹	남한거주기간			
1	수도권	남	10대	5년 미만	52	5	44.0
2				5년 이상	200	20	21.2
3			20대	5년 미만	242	24	19.4
4				5년 이상	739	74	10.9
5			30대	5년 미만	165	17	22.8
6				5년 이상	1,151	116	8.7
7			40대	5년 미만	110	11	29.0
8				5년 이상	913	92	9.8
9			50대 이상	5년 미만	137	14	25.0
10				5년 이상	1,242	125	8.4
11		여	10대	5년 미만	73	7	34.3
12				5년 이상	182	18	22.4
13			20대	5년 미만	722	72	11.1
14				5년 이상	1,077	108	9.1
15			30대	5년 미만	944	95	9.6
16				5년 이상	2,697	271	5.7
17			40대	5년 미만	674	68	11.4
18				5년 이상	3,572	359	5.0
19			50대 이상	5년 미만	573	58	12.3
20				5년 이상	3,892	391	4.8
21	비수도권	남	10대	5년 미만	16	2	58.0
22				5년 이상	77	8	32.1
23			20대	5년 미만	54	5	40.1
24				5년 이상	284	28	17.6
25			30대	5년 미만	49	5	39.8
26				5년 이상	434	43	14.3
27			40대	5년 미만	41	4	44.1
28				5년 이상	415	42	14.4
29			50대 이상	5년 미만	55	5	40.1
30				5년 이상	781	78	10.6
31		여	10대	5년 미만	25	2	69.2
32				5년 이상	81	8	34.4
33			20대	5년 미만	176	18	22.4
34				5년 이상	378	38	15.3
35			30대	5년 미만	399	40	14.9
36				5년 이상	1,475	148	7.7
37			40대	5년 미만	369	37	15.5
38				5년 이상	2,421	243	6.0
39			50대 이상	5년 미만	318	32	16.7
40				5년 이상	2,675	269	5.7
총합계					29,880	3,000	1.7

(2) 제곱근비례배분

3,000명의 표본을 총 40개 층에 대해 제곱근비례배분을 한 결과는 다음과 같고, 각 층별 허용오차 분포는 6.9~15.4%로 나타났다.

<표 4-2> 층별 제곱근비례배분에 따른 허용오차(표본크기=3,000)

층 번호	층화변수				모집단(명)	제곱근비례 배분(명)	허용오차 (%)
	지역	성별	연령 그룹	남한거주기간			
1	수도권	남	10대	5년 미만	52	24	15.1
2				5년 이상	200	47	12.7
3			20대	5년 미만	242	52	12.2
4				5년 이상	739	90	9.8
5			30대	5년 미만	165	43	13.0
6				5년 이상	1,151	112	8.82
7			40대	5년 미만	110	35	13.9
8				5년 이상	913	100	9.4
9			50대 이상	5년 미만	137	39	13.5
10				5년 이상	1,242	117	8.7
11		여	10대	5년 미만	73	28	14.5
12				5년 이상	182	45	12.8
13			20대	5년 미만	722	89	9.8
14				5년 이상	1,077	109	9.0
15			30대	5년 미만	944	102	9.2
16				5년 이상	2,697	172	7.3
17			40대	5년 미만	674	86	10.0
18				5년 이상	3,572	198	6.9
19			50대 이상	5년 미만	573	79	10.4
20				5년 이상	3,892	207	6.9
21	비수도권	남	10대	5년 미만	16	13	12.7
22				5년 이상	77	29	14.4
23			20대	5년 미만	54	24	14.8
24				5년 이상	284	56	11.8
25			30대	5년 미만	49	23	14.8
26				5년 이상	434	69	10.9
27			40대	5년 미만	41	21	15.4
28				5년 이상	415	68	11.0
29			50대 이상	5년 미만	55	25	14.4
30				5년 이상	781	93	9.7
31		여	10대	5년 미만	25	17	14.2
32				5년 이상	81	30	14.5
33			20대	5년 미만	176	44	13.0
34				5년 이상	378	64	11.3
35			30대	5년 미만	399	66	11.1
36				5년 이상	1,475	127	8.4
37			40대	5년 미만	369	64	11.3
38				5년 이상	2,421	163	7.5
39			50대 이상	5년 미만	318	59	11.6
40				5년 이상	2,675	171	7.3
총합계					29,880	3,000	1.7

(3) 네이만배분

북한이탈주민실태조사에서 중요도가 높은 층별 남한생활만족도 표본비율을 이용하여 네이만배분을 한 결과는 다음과 같고, 각 층별 허용오차 분포는 4.9~ 59.4%로 나타났다.

<표 4-3> 층별 네이만배분에 따른 허용오차(표본크기=3,000)

층 번호	층화변수				모집단(명)	네이만배분 (명)	허용오차 (%)
	지역	성별	연령 그룹	남한거주기간			
1	수도권	남	10대	5년 미만	52	5	44.0
2				5년 이상	200	12	27.1
3			20대	5년 미만	242	26	18.2
4				5년 이상	739	73	11.0
5			30대	5년 미만	165	17	22.4
6				5년 이상	1,151	86	10.2
7			40대	5년 미만	110	10	29.1
8				5년 이상	913	101	9.3
9			50대 이상	5년 미만	137	14	24.6
10				5년 이상	1,242	127	8.3
11		여	10대	5년 미만	73	8	31.9
12				5년 이상	182	19	21.2
13			20대	5년 미만	722	71	11.1
14				5년 이상	1,077	95	9.7
15			30대	5년 미만	944	98	9.4
16				5년 이상	2,697	275	5.7
17			40대	5년 미만	674	67	11.4
18				5년 이상	3,572	369	4.9
19			50대 이상	5년 미만	573	50	13.3
20				5년 이상	3,892	376	4.9
21	비수도권	남	10대	5년 미만	16	2	57.5
22				5년 이상	77	5	40.7
23			20대	5년 미만	54	6	36.4
24				5년 이상	284	31	16.6
25			30대	5년 미만	49	6	36.2
26				5년 이상	434	45	13.9
27			40대	5년 미만	41	4	44.0
28				5년 이상	415	41	14.6
29			50대 이상	5년 미만	55	5	40.1
30				5년 이상	781	84	10.2
31		여	10대	5년 미만	25	2	59.4
32				5년 이상	81	9	30.2
33			20대	5년 미만	176	19	21.1
34				5년 이상	378	36	15.6
35			30대	5년 미만	399	43	14.2
36				5년 이상	1,475	155	7.5
37			40대	5년 미만	369	36	15.6
38				5년 이상	2,421	264	5.8
39			50대 이상	5년 미만	318	32	16.5
40				5년 이상	2,675	276	5.6
총합계					29,880	3,000	1.7

(4) 멱등배분

층별 남한생활만족도 표본비율을 이용하여 멱등배분을 한 결과, 각 층별 모집단 크기 및 표본크기는 다음과 같다.

<표 4-4> 층별 멱등배분에 따른 허용오차(표본크기=3,000)

층번호	층화변수				모집단(명)	p=0.3		p=0.4		p=0.5	
	지역	성별	연령그룹	남한거주기간		멱등배분(명)	허용오차(%)	멱등배분(명)	허용오차(%)	멱등배분(명)	허용오차(%)
1	수도권	남	10대	5년 미만	52	41	7.2	32	10.8	25	14.1
2				5년 이상	200	52	11.7	44	13.1	37	14.6
3			20대	5년 미만	242	65	10.5	60	11.0	54	11.8
4				5년 이상	739	89	9.8	90	9.7	89	9.8
5			30대	5년 미만	165	57	10.5	50	11.6	43	12.9
6				5년 이상	1,151	93	9.8	96	9.7	97	9.6
7			40대	5년 미만	110	49	10.4	41	12.1	34	14.0
8				5년 이상	913	98	9.4	102	9.2	106	9.0
9			50대 이상	5년 미만	137	54	10.4	46	11.8	39	13.3
10				5년 이상	1,242	105	9.2	112	8.9	118	8.7
11		여	10대	5년 미만	73	46	8.8	38	11.0	30	13.7
12				5년 이상	182	59	10.5	52	11.5	45	12.7
13			20대	5년 미만	722	88	9.9	89	9.8	89	9.8
14				5년 이상	1,077	96	9.6	100	9.4	102	9.3
15			30대	5년 미만	944	97	9.5	101	9.3	104	9.1
16				5년 이상	2,697	132	8.4	152	7.8	174	7.3
17			40대	5년 미만	674	87	9.9	87	9.9	86	9.9
18				5년 이상	3,572	144	8.1	172	7.4	202	6.8
19			50대 이상	5년 미만	573	79	10.3	77	10.5	74	10.7
20				5년 이상	3,892	145	8.1	173	7.4	204	6.7
21	비수도권	남	10대	5년 미만	16	16	.	16	.	14	8.0
22				5년 이상	77	39	11.0	30	13.9	23	17.0
23			20대	5년 미만	54	41	7.4	32	11.0	25	14.3
24				5년 이상	284	69	10.3	64	10.8	59	11.4
25			30대	5년 미만	49	42	5.5	33	9.6	25	13.6
26				5년 이상	434	77	10.2	74	10.4	71	10.7
27			40대	5년 미만	41	37	4.7	28	10.2	21	14.7
28				5년 이상	415	75	10.3	72	10.6	68	10.9
29			50대 이상	5년 미만	55	39	8.4	30	12.0	22	16.1
30				5년 이상	781	93	9.6	95	9.5	96	9.4
31		여	10대	5년 미만	25	25	.	22	6.6	15	15.5
32				5년 이상	81	47	9.2	39	11.3	31	13.8
33			20대	5년 미만	176	60	10.3	53	11.3	46	12.5
34				5년 이상	378	72	10.5	68	10.8	63	11.3
35			30대	5년 미만	399	76	10.2	73	10.4	69	10.8
36				5년 이상	1,475	111	9.0	121	8.6	131	8.2
37			40대	5년 미만	369	72	10.4	68	10.8	63	11.3
38				5년 이상	2,421	131	8.4	150	7.8	171	7.3
39			50대 이상	5년 미만	318	70	10.4	65	10.9	60	11.5
40				5년 이상	2,675	132	8.4	153	7.8	175	7.2
총합계					29,880	3,000	1.7	3,000	1.7	3,000	1.7

p=0.3일 때, 각 층별 허용오차 분포는 4.7~11.7%로 나타났고, p=0.4일 때, 각 층별 허용오차 분포는 6.6~13.9%로 나타났으며, p=0.5일 때, 각 층별 허용오차 분포는 6.7~17.0%로 나타났다.

(5) 배분방법 비교 결과

비례배분의 경우 표본크기가 너무 작은 층들이 있어 허용오차가 매우 큰 층이 존재하는 것으로 파악되어, 층별 통계를 생산하는데 한계가 있는 것으로 나타났다.

제곱근비례배분은 층별 허용오차의 차이가 매우 컸던 비례배분의 문제점을 어느 정도 해결하고 있는 것으로 나타났다.

북한이탈주민실태조사의 층별 남한생활만족도 표본비율을 이용한 네이만배분은 표본크기가 너무 작은 층들이 나타났다. 멱등배분은 p=0.3인 경우 특정 층에 표본이 적게 배분되는 이러한 네이만배분의 문제점을 보완하는 것으로 나타났다.

결론적으로 네 가지 표본배분 방법들을 비교 검토한 결과 제곱근비례배분과 p=0.3인 경우 멱등배분이 적절한 것으로 파악되었다.

제곱근비례배분에 의한 공표수준별 표본크기 및 허용오차를 구해보면 다음과 같다.

<표 4-5> 공표수준별 표본크기(3,000) 및 허용오차(제곱근비례배분)

공표수준		모집단(명)	표본크기(명)	허용오차(%)
거주지역	수도권	19,357	1,774	2.2
	비수도권	10,523	1,226	2.7
성별	남성	7,157	1,080	2.8
	여성	22,723	1,920	2.2
연령그룹	10대	706	233	5.3
	20대	3,672	528	4.0
	30대	7,314	714	3.5
	40대	8,515	735	3.5
	50대 이상	9,673	790	3.4
남한거주기간	5년 미만	5,194	933	3.0
	5년 이상	24,686	2,067	2.1

공표수준별 허용오차 분포는 2.1~5.3%로 나타났다.

다음으로 p=0.3인 경우 멱등배분에 의한 공표수준별 표본크기 및 허용오차를 구해보면 다음과 같다.

<표 4-6> 공표수준별 표본크기(3,000) 및 허용오차(멱등배분)

공표수준		모집단(명)	표본크기(명)	허용오차(%)
거주지역	수도권	19,357	1,676	2.3
	비수도권	10,523	1,324	2.5
성별	남성	7,157	1,231	2.6
	여성	22,723	1,769	2.3
연령그룹	10대	706	325	4.0
	20대	3,672	580	3.8
	30대	7,314	685	3.6
	40대	8,515	693	3.6
	50대 이상	9,673	717	3.6
남한거주기간	5년 미만	5,194	1,141	2.6
	5년 이상	24,686	1,859	2.2

공표수준별 허용오차 분포는 2.2~4.0%로 나타났다.

제곱근비례배분과 p=0.3인 경우 멱등배분을 비교해 본 결과, 10대 연령층의 706명의 모집단 크기에 대하여 제곱근비례배분에서는 233명, p=0.3인 경우 멱등배분에서는 325명이 배분되었다. 따라서 멱등배분이 제곱근비례배분에 비해 너무 많은 표본이 배분되어 다른 연령대에 비해 응답협조도가 낮은 특성을 고려했을 때 조사에 어려움이 나타날 것으로 예상되어 제곱근비례배분이 더 적절한 방법이라고 생각된다. 특히 모집단이 크기가 작은 경우에는 응답률에 더욱 많은 신경을 쓸 필요가 있다.

2) 표본크기가 2,500명 일 때

(1) 비례배분

2,500명의 표본을 총 40개 층에 대해 비례배분을 한 결과 각 층별 모집단 크기 및 표본크기는 다음과 같다. 이때 허용오차는 모집단과 표본크기를 고려한 단순임의표본추출 가정 하에 산출되었다. 각 층별 허용오차 분포는 5.3~77.0%로 나타났다.

<표 4-7> 층별 비례배분에 따른 허용오차(표본크기=2,500)

층 번호	층화변수				모집단(명)	비례배분 (명)	허용오차 (%)
	지역	성별	연령 그룹	남한거주기간			
1	수도권	남	10대	5년 미만	52	4	51.0
2				5년 이상	200	17	23.0
3			20대	5년 미만	242	20	21.4
4				5년 이상	739	62	12.0
5			30대	5년 미만	165	14	25.8
6				5년 이상	1,151	96	9.7
7			40대	5년 미만	110	9	32.6
8				5년 이상	913	76	10.9
9			50대 이상	5년 미만	137	12	27.0
10				5년 이상	1,242	104	9.3
11		여	10대	5년 미만	73	6	38.0
12				5년 이상	182	15	24.1
13			20대	5년 미만	722	60	12.2
14				5년 이상	1,077	90	10.0
15			30대	5년 미만	944	79	10.6
16				5년 이상	2,697	226	6.3
17			40대	5년 미만	674	56	12.7
18				5년 이상	3,572	299	5.5
19			50대 이상	5년 미만	573	48	13.6
20				5년 이상	3,892	326	5.3
21	비수도권	남	10대	5년 미만	16	1	77.0
22				5년 이상	77	6	37.2
23			20대	5년 미만	54	5	40.1
24				5년 이상	284	24	19.2
25			30대	5년 미만	49	4	44.5
26				5년 이상	434	36	15.7
27			40대	5년 미만	41	3	60.6
28				5년 이상	415	35	15.9
29			50대 이상	5년 미만	55	5	40.1
30				5년 이상	781	65	11.8
31		여	10대	5년 미만	25	2	69.2
32				5년 이상	81	7	37.2
33			20대	5년 미만	176	15	24.9
34				5년 이상	378	32	16.8
35			30대	5년 미만	399	33	16.6
36				5년 이상	1,475	123	8.5
37			40대	5년 미만	369	31	17.1
38				5년 이상	2,421	203	6.7
39			50대 이상	5년 미만	318	27	18.4
40				5년 이상	2,675	224	6.3
총합계					29,880	2,500	1.9

(2) 제곱근비례배분

2,500명의 표본을 총 40개 층에 대해 제곱근비례배분 한 결과는 다음과 같고, 각 층별 허용오차 분포는 7.4~17.7%로 나타났다.

<표 4-8> 층별 제곱근비례배분에 따른 허용오차(표본크기=2,500)

층 번호	층화변수				모집단(명)	제곱근비례 배분(명)	허용오차 (%)
	지역	성별	연령 그룹	남한거주기간			
1	수도권	남	10대	5년 미만	52	20	17.7
2				5년 이상	200	39	14.2
3			20대	5년 미만	242	43	13.7
4				5년 이상	739	75	10.8
5			30대	5년 미만	165	35	14.8
6				5년 이상	1,151	94	9.8
7			40대	5년 미만	110	29	15.9
8				5년 이상	913	84	10.3
9			50대 이상	5년 미만	137	32	15.3
10				5년 이상	1,242	97	9.6
11		여	10대	5년 미만	73	24	16.4
12				5년 이상	182	37	14.5
13			20대	5년 미만	722	74	10.9
14				5년 이상	1,077	91	9.9
15			30대	5년 미만	944	85	10.1
16				5년 이상	2,697	143	8.1
17			40대	5년 미만	674	72	11.0
18				5년 이상	3,572	165	7.6
19			50대 이상	5년 미만	573	66	11.5
20				5년 이상	3,892	172	7.4
21	비수도권	남	10대	5년 미만	16	11	15.7
22				5년 이상	77	24	16.6
23			20대	5년 미만	54	20	17.3
24				5년 이상	284	47	13.2
25			30대	5년 미만	49	19	17.4
26				5년 이상	434	58	12.1
27			40대	5년 미만	41	18	17.5
28				5년 이상	415	56	12.3
29			50대 이상	5년 미만	55	21	16.7
30				5년 이상	781	77	10.7
31		여	10대	5년 미만	25	14	17.2
32				5년 이상	81	25	16.7
33			20대	5년 미만	176	37	14.5
34				5년 이상	378	54	12.5
35			30대	5년 미만	399	55	12.4
36				5년 이상	1,475	106	9.2
37			40대	5년 미만	369	53	12.6
38				5년 이상	2,421	136	8.2
39			50대 이상	5년 미만	318	49	13.0
40				5년 이상	2,675	143	8.1
총합계					29,880	2,500	1.9

(3) 네이만배분

층별 남한생활만족도 표본비율을 이용하여 네이만배분을 한 결과는 다음과 같고, 각 층별 허용오차 분포는 5.4~59.4%로 나타났다.

<표 4-9> 층별 네이만배분에 따른 허용오차(표본크기=2,500)

층 번호	층화변수				모집단(명)	네이만배분 (명)	허용오차 (%)
	지역	성별	연령 그룹	남한거주기간			
1	수도권	남	10대	5년 미만	52	5	44.0
2				5년 이상	200	10	29.7
3			20대	5년 미만	242	22	19.9
4				5년 이상	739	61	12.1
5			30대	5년 미만	165	14	24.8
6				5년 이상	1,151	71	11.3
7			40대	5년 미만	110	9	30.7
8				5년 이상	913	84	10.3
9			50대 이상	5년 미만	137	12	26.7
10				5년 이상	1,242	106	9.2
11		여	10대	5년 미만	73	7	34.3
12				5년 이상	182	15	24.1
13			20대	5년 미만	722	59	12.3
14				5년 이상	1,077	79	10.7
15			30대	5년 미만	944	81	10.5
16				5년 이상	2,697	229	6.3
17			40대	5년 미만	674	56	12.6
18				5년 이상	3,572	307	5.4
19			50대 이상	5년 미만	573	41	14.8
20				5년 이상	3,892	314	5.4
21	비수도권	남	10대	5년 미만	16	2	57.5
22				5년 이상	77	4	45.3
23			20대	5년 미만	54	5	40.0
24				5년 이상	284	26	18.3
25			30대	5년 미만	49	5	39.8
26				5년 이상	434	38	15.2
27			40대	5년 미만	41	3	50.6
28				5년 이상	415	34	16.1
29			50대 이상	5년 미만	55	4	44.7
30				5년 이상	781	70	11.3
31		여	10대	5년 미만	25	2	59.4
32				5년 이상	81	7	34.4
33			20대	5년 미만	176	16	23.2
34				5년 이상	378	30	17.2
35			30대	5년 미만	399	36	15.6
36				5년 이상	1,475	129	8.3
37			40대	5년 미만	369	30	17.2
38				5년 이상	2,421	220	6.4
39			50대 이상	5년 미만	318	27	18.1
40				5년 이상	2,675	230	6.2
총합계					29,880	2,500	1.9

(4) 멱등배분

층별 남한생활만족도 표본비율을 이용하여 멱등배분을 한 결과는 다음과 같고, 각 층별 모집단 크기 및 표본크기는 다음과 같다.

<표 4-10> 층별 멱등배분에 따른 허용오차(표본크기=2,500)

층번호	층화변수				모집단(명)	p=0.3		p=0.4		p=0.5	
	지역	성별	연령그룹	남한거주기간		멱등배분(명)	허용오차(%)	멱등배분(명)	허용오차(%)	멱등배분(명)	허용오차(%)
1	수도권	남	10대	5년 미만	52	34	10.4	27	13.1	20	17.1
2				5년 이상	200	43	13.3	37	14.6	31	16.2
3			20대	5년 미만	242	54	11.8	50	12.4	45	13.2
4				5년 이상	739	73	11.0	75	10.8	75	10.8
5			30대	5년 미만	165	47	12.1	41	13.3	36	14.5
6				5년 이상	1,151	77	10.9	80	10.7	81	10.6
7			40대	5년 미만	110	41	12.1	34	14.0	28	16.0
8				5년 이상	913	81	10.5	85	10.2	88	10.0
9			50대 이상	5년 미만	137	45	12.0	39	13.3	33	14.9
10				5년 이상	1,242	87	10.2	94	9.8	99	9.5
11		여	10대	5년 미만	73	38	11.0	31	13.3	25	15.8
12				5년 이상	182	49	12.0	43	13.1	38	14.2
13			20대	5년 미만	722	73	11.0	74	10.9	74	10.9
14				5년 이상	1,077	80	10.6	83	10.4	85	10.3
15			30대	5년 미만	944	80	10.6	84	10.3	87	10.1
16				5년 이상	2,697	109	9.2	127	8.6	145	8.0
17			40대	5년 미만	674	72	11.0	72	11.0	72	11.0
18				5년 이상	3,572	120	8.9	143	8.1	168	7.5
19			50대 이상	5년 미만	573	65	11.5	64	11.6	62	11.8
20				5년 이상	3,892	120	8.9	144	8.1	170	7.4
21	비수도권	남	10대	5년 미만	16	16	.	16	.	12	1.3
22				5년 이상	77	33	12.9	25	16.0	19	19.4
23			20대	5년 미만	54	35	9.7	27	13.2	20	17.2
24				5년 이상	284	58	11.5	54	12.1	49	12.8
25			30대	5년 미만	49	36	8.3	27	12.5	21	16.0
26				5년 이상	434	65	11.3	62	11.6	59	11.9
27			40대	5년 미만	41	32	7.9	23	13.4	18	17.1
28				5년 이상	415	63	11.4	60	11.8	56	12.3
29			50대 이상	5년 미만	55	33	10.7	25	14.4	19	18.0
30				5년 이상	781	77	10.7	79	10.5	80	10.5
31		여	10대	5년 미만	25	25	.	18	11.7	12	19.8
32				5년 이상	81	40	11.0	32	13.4	26	15.8
33			20대	5년 미만	176	49	11.9	44	12.8	38	14.1
34				5년 이상	378	60	11.7	57	12.0	53	12.6
35			30대	5년 미만	399	63	11.4	60	11.7	57	12.1
36				5년 이상	1,475	92	10.0	101	9.5	109	9.1
37			40대	5년 미만	369	59	11.8	56	12.1	52	12.7
38				5년 이상	2,421	108	9.3	125	8.6	142	8.1
39			50대 이상	5년 미만	318	58	11.7	54	12.2	50	12.8
40				5년 이상	2,675	110	9.2	128	8.5	146	8.0
총합계					29,880	2,500	1.9	2,500	1.9	2,500	1.9

p=0.3일 때, 각 층별 허용오차 분포는 7.9~13.3%로 나타났고, p=0.4일 때, 각 층별 허용오차 분포는 8.1~16.0%로 나타났으며, p=0.5일 때, 각 층별 허용오차 분포는 1.3~19.8%로 나타났다.

(5) 배분방법 비교 결과

네 가지 표본배분방법을 비교 검토한 결과 제곱근비례배분과 p=0.3인 경우 멱등배분이 적절한 것으로 파악되었다.

제곱근비례배분에 의한 공표수준별 표본크기 및 허용오차를 구해보면 다음과 같다.

<표 4-11> 공표수준별 표본크기(2,500) 및 허용오차(제곱근비례배분)

공표수준		모집단(명)	표본크기(명)	허용오차(%)
거주지역	수도권	19,357	1,477	2.5
	비수도권	10,523	1,023	2.9
성별	남성	7,157	899	3.1
	여성	22,723	1601	2.4
연령그룹	10대	706	194	6.1
	20대	3,672	441	4.4
	30대	7,314	595	3.9
	40대	8,515	613	3.9
	50대 이상	9,673	657	3.7
남한거주기간	5년 미만	5,194	777	3.3
	5년 이상	24,686	1,723	2.3

공표수준별 허용오차 분포는 2.3~6.1%로 나타났다.

다음으로 p=0.3인 경우 멱등배분에 의한 공표수준별 표본크기 및 허용오차를 구해보면 다음과 같다.

<표 4-12> 공표수준별 표본크기(2,500) 및 허용오차(멱등배분)

공표수준		모집단(명)	표본크기(명)	허용오차(%)
거주지역	수도권	19,357	1,388	2.6
	비수도권	10,523	1,112	2.8
성별	남성	7,157	1,030	2.9
	여성	22,723	1,470	2.5
연령그룹	10대	706	278	4.6
	20대	3,672	482	4.2
	30대	7,314	569	4.0
	40대	8,515	576	4.0
	50대 이상	9,673	595	3.9
남한거주기간	5년 미만	5,194	955	2.9
	5년 이상	24,686	1,545	2.4

공표수준별 허용오차 분포는 2.4~4.6%로 나타났다.

제곱근비례배분과 p=0.3인 경우 멱등배분을 비교해 본 결과, 10대 연령층의 706명의 모집단 크기에 대하여 제곱근비례배분에서는 194명, p=0.3인 경우 멱등배분에서는 278명이 배분되었다. 따라서 멱등배분이 제곱근비례배분에 비해 너무 많은 표본이 배분되어 다른 연령대에 비해 응답협조도가 낮은 특성을 고려했을 때 조사에 어려움이 나타날 것으로 예상되어 제곱근비례배분이 더 적절한 방법이라고 생각된다.

3) 표본크기가 2,000명 일 때 표본배분

(1) 비례배분

2,000명의 표본을 총 40개 층에 대해 비례배분을 한 결과 각 층별 모집단 크기 및 표본크기는 다음과 같다. 이때 허용오차는 모집단과 표본크기를 고려한 단순임의표본추출 가정 하에 산출되었다. 각 층별 허용오차 분포는 5.9~77.0%로 나타났다.

<표 4-13> 층별 비례배분에 따른 허용오차(표본크기=2,000)

층 번호	층화변수				모집단(명)	비례배분 (명)	허용오차 (%)
	지역	성별	연령 그룹	남한거주기간			
1	수도권	남	10대	5년 미만	52	3	61.0
2				5년 이상	200	14	25.6
3			20대	5년 미만	242	16	24.3
4				5년 이상	739	49	13.7
5			30대	5년 미만	165	11	29.0
6				5년 이상	1,151	77	10.9
7			40대	5년 미만	110	7	35.0
8				5년 이상	913	61	12.2
9			50대 이상	5년 미만	137	9	32.0
10				5년 이상	1,242	83	10.5
11		여	10대	5년 미만	73	5	45.1
12				5년 이상	182	12	28.0
13			20대	5년 미만	722	48	13.8
14				5년 이상	1,077	72	11.3
15			30대	5년 미만	944	63	12.1
16				5년 이상	2,697	181	7.1
17			40대	5년 미만	674	45	14.2
18				5년 이상	3,572	239	6.2
19			50대 이상	5년 미만	573	39	15.4
20				5년 이상	3,892	261	5.9
21	비수도권	남	10대	5년 미만	16	1	77.0
22				5년 이상	77	5	40.7
23			20대	5년 미만	54	4	44.7
24				5년 이상	284	19	21.7
25			30대	5년 미만	49	3	51.0
26				5년 이상	434	29	17.6
27			40대	5년 미만	41	3	60.6
28				5년 이상	415	28	17.9
29			50대 이상	5년 미만	55	4	44.8
30				5년 이상	781	52	13.3
31		여	10대	5년 미만	25	2	69.2
32				5년 이상	81	5	45.3
33			20대	5년 미만	176	12	28.2
34				5년 이상	378	25	19.3
35			30대	5년 미만	399	27	18.5
36				5년 이상	1,475	99	9.6
37			40대	5년 미만	369	25	19.3
38				5년 이상	2,421	162	7.5
39			50대 이상	5년 미만	318	21	21.1
40				5년 이상	2,675	179	7.1
총합계					29,880	2,000	2.1

(2) 제곱근비례배분

2,000명의 표본을 총 40개 층에 대해 제곱근비례배분을 한 결과는 다음과 같고, 각 층별 허용오차 분포는 8.3~23.2%로 나타났다.

<표 4-14> 층별 제곱근비례배분에 따른 허용오차(표본크기=2,000)

층 번호	층화변수				모집단(명)	제곱근비례 배분(명)	허용오차 (%)
	지역	성별	연령 그룹	남한거주기간			
1	수도권	남	10대	5년 미만	52	16	21.0
2				5년 이상	200	31	16.2
3			20대	5년 미만	242	34	15.7
4				5년 이상	739	60	12.3
5			30대	5년 미만	165	28	17.2
6				5년 이상	1,151	75	11.1
7			40대	5년 미만	110	23	18.5
8				5년 이상	913	67	11.6
9			50대 이상	5년 미만	137	26	17.3
10				5년 이상	1,242	78	10.8
11		여	10대	5년 미만	73	19	19.2
12				5년 이상	182	30	16.5
13			20대	5년 미만	722	59	12.4
14				5년 이상	1,077	73	11.2
15			30대	5년 미만	944	68	11.6
16				5년 이상	2,697	115	9.1
17			40대	5년 미만	674	57	12.6
18				5년 이상	3,572	132	8.6
19			50대 이상	5년 미만	573	53	13.0
20				5년 이상	3,892	138	8.3
21	비수도권	남	10대	5년 미만	16	9	19.1
22				5년 이상	77	19	19.4
23			20대	5년 미만	54	16	20.4
24				5년 이상	284	37	15.1
25			30대	5년 미만	49	16	19.9
26				5년 이상	434	46	13.8
27			40대	5년 미만	41	14	21.2
28				5년 이상	415	45	13.9
29			50대 이상	5년 미만	55	16	21.3
30				5년 이상	781	62	12.1
31		여	10대	5년 미만	25	11	23.2
32				5년 이상	81	20	19.5
33			20대	5년 미만	176	29	16.9
34				5년 이상	378	43	14.3
35			30대	5년 미만	399	44	14.1
36				5년 이상	1,475	85	10.4
37			40대	5년 미만	369	43	14.2
38				5년 이상	2,421	109	9.3
39			50대 이상	5년 미만	318	40	14.7
40				5년 이상	2,675	114	9.0
총합계					29,880	2,000	2.1

(3) 네이만배분

층별 남한생활만족도 표본비율을 이용하여 네이만배분 한 결과는 다음과 같고, 각 층별 허용오차 분포는 6.0~78.7%로 나타났다.

<표 4-15> 층별 네이만배분에 따른 허용오차(표본크기=2,000)

층 번호	층화변수				모집단(명)	네이만배분 (명)	허용오차 (%)
	지역	성별	연령 그룹	남한거주기간			
1	수도권	남	10대	5년 미만	52	4	51.0
2				5년 이상	200	8	33.2
3			20대	5년 미만	242	18	22.1
4				5년 이상	739	48	13.8
5			30대	5년 미만	165	11	28.2
6				5년 이상	1,151	57	12.7
7			40대	5년 미만	110	7	34.9
8				5년 이상	913	67	11.6
9			50대 이상	5년 미만	137	9	31.0
10				5년 이상	1,242	85	10.3
11		여	10대	5년 미만	73	6	37.1
12				5년 이상	182	12	27.0
13			20대	5년 미만	722	48	13.7
14				5년 이상	1,077	63	12.1
15			30대	5년 미만	944	65	11.8
16				5년 이상	2,697	183	7.1
17			40대	5년 미만	674	45	14.2
18				5년 이상	3,572	246	6.1
19			50대 이상	5년 미만	573	33	16.6
20				5년 이상	3,892	251	6.0
21	비수도권	남	10대	5년 미만	16	1	77.0
22				5년 이상	77	3	51.7
23			20대	5년 미만	54	4	44.7
24				5년 이상	284	21	20.5
25			30대	5년 미만	49	4	44.5
26				5년 이상	434	30	17.3
27			40대	5년 미만	41	3	50.6
28				5년 이상	415	28	17.9
29			50대 이상	5년 미만	55	3	51.2
30				5년 이상	781	56	12.7
31		여	10대	5년 미만	25	1	78.7
32				5년 이상	81	6	37.2
33			20대	5년 미만	176	13	25.9
34				5년 이상	378	24	19.4
35			30대	5년 미만	399	28	17.9
36				5년 이상	1,475	103	9.4
37			40대	5년 미만	369	24	19.3
38				5년 이상	2,421	176	7.2
39			50대 이상	5년 미만	318	22	20.1
40				5년 이상	2,675	184	7.4
총합계					29,880	2,000	2.1

(4) 멱등배분

층별 남한생활만족도 표본비율을 이용하여 멱등배분을 한 결과, 각 층별 모집단 크기 및 표본크기는 다음과 같다.

<표 4-16> 층별 멱등배분에 따른 허용오차(표본크기=2,000)

층번호	층화변수				모집단(명)	p=0.3		p=0.4		p=0.5	
	지역	성별	연령그룹	남한거주기간		멱등배분(명)	허용오차(%)	멱등배분(명)	허용오차(%)	멱등배분(명)	허용오차(%)
1	수도권	남	10대	5년 미만	52	27	13.1	21	16.5	16	20.2
2				5년 이상	200	35	15.1	29	16.8	25	18.3
3			20대	5년 미만	242	43	13.6	40	14.2	36	15.1
4				5년 이상	739	59	12.3	60	12.2	60	12.2
5			30대	5년 미만	165	38	14.0	33	15.3	29	16.5
6				5년 이상	1,151	62	12.2	64	12.0	65	11.9
7			40대	5년 미만	110	33	14.3	27	16.8	23	18.1
8				5년 이상	913	65	11.8	68	11.5	71	11.2
9			50대	5년 미만	137	36	14.0	31	15.5	26	17.3
10			이상	5년 이상	1,242	70	11.5	75	11.1	79	10.8
11		여	10대	5년 미만	73	31	13.3	25	15.8	20	18.6
12				5년 이상	182	39	13.9	35	14.9	30	16.4
13			20대	5년 미만	722	58	12.4	59	12.3	59	12.3
14				5년 이상	1,077	64	12.0	67	11.7	68	11.6
15			30대	5년 미만	944	64	11.9	67	11.6	69	11.4
16				5년 이상	2,697	88	10.4	102	9.6	116	9.0
17			40대	5년 미만	674	57	12.5	58	12.4	57	12.5
18				5년 이상	3,572	96	10.0	114	9.1	135	8.4
19			50대	5년 미만	573	52	13.0	51	13.2	49	13.5
20			이상	5년 이상	3,892	96	10.0	115	9.1	136	8.3
21	비수도권	남	10대	5년 미만	16	16	.	14	8.0	10	17.9
22				5년 이상	77	26	15.6	20	18.7	15	22.5
23			20대	5년 미만	54	27	13.2	21	16.6	16	20.3
24				5년 이상	284	46	13.3	43	13.8	39	14.6
25			30대	5년 미만	49	28	12.0	22	15.4	17	19.0
26				5년 이상	434	51	13.0	50	13.1	47	13.6
27			40대	5년 미만	41	25	12.1	19	16.0	14	20.9
28				5년 이상	415	50	13.1	48	13.4	45	13.9
29			50대	5년 미만	55	26	13.9	20	17.3	15	21.3
30			이상	5년 이상	781	61	12.1	63	11.9	64	11.8
31		여	10대	5년 미만	25	20	9.3	14	16.9	10	23.3
32				5년 이상	81	31	13.8	26	15.8	21	18.3
33			20대	5년 미만	176	40	11.0	35	12.5	31	13.8
34				5년 이상	378	48	12.1	45	12.6	42	13.2
35			30대	5년 미만	399	50	13.0	48	13.3	46	13.6
36				5년 이상	1,475	74	10.4	81	9.8	87	9.4
37			40대	5년 미만	369	48	13.3	45	13.8	42	14.3
38				5년 이상	2,421	86	10.5	100	9.7	114	9.0
39			50대	5년 미만	318	46	13.4	43	14.0	40	14.5
40			이상	5년 이상	2,675	88	10.4	102	9.6	116	9.0
총합계					29,880	2,000	2.1	2,000	2.1	2,000	2.1

p=0.3일 때, 각 층별 허용오차 분포는 9.3~15.6%로 나타났고, p=0.4일 때, 각 층별 허용오차 분포는 8.0~18.7%로 나타났으며, p=0.5일 때, 각 층별 허용오차 분포는 8.3~23.3%로 나타났다.

(5) 배분방법 비교 결과

네 가지 표본배분방법을 비교 검토한 결과 제곱근비례배분과 p=0.3인 경우 멱등배분이 적절한 것으로 파악되었다.

제곱근비례배분에 의한 공표수준별 표본크기 및 허용오차를 구해보면 다음과 같다.

<표 4-17> 공표수준별 표본크기(2,000) 및 허용오차(제곱근비례배분)

공표수준		모집단(명)	표본크기(명)	허용오차(%)
거주지역	수도권	19,357	1,182	2.8
	비수도권	10,523	818	3.3
성별	남성	7,157	718	3.5
	여성	22,723	1,282	2.7
연령그룹	10대	706	155	7.0
	20대	3,672	351	5.0
	30대	7,314	477	4.4
	40대	8,515	490	4.3
	50대 이상	9,673	527	4.2
남한거주기간	5년 미만	5,194	621	3.7
	5년 이상	24,686	1,379	2.6

공표수준별 허용오차 분포는 2.6~7.0%로 나타났다.

다음으로 p=0.3인 경우 멱등배분에 의한 공표수준별 표본크기 및 허용오차를 구해보면 다음과 같다.

<표 4-18> 공표수준별 표본크기(2,000) 및 허용오차(멱등배분)

공표수준		모집단(명)	표본크기(명)	허용오차(%)
거주지역	수도권	19,357	1,113	2.9
	비수도권	10,523	887	3.2
성별	남성	7,157	824	3.3
	여성	22,723	1,176	2.8
연령그룹	10대	706	225	5.4
	20대	3,672	385	4.8
	30대	7,314	455	4.5
	40대	8,515	460	4.5
	50대 이상	9,673	475	4.4
남한거주기간	5년 미만	5,194	765	3.3
	5년 이상	24,686	1,235	2.8

공표수준별 허용오차 분포는 2.8~5.4%로 나타났다.

제곱근비례배분과 p=0.3인 경우 멱등배분을 비교해 본 결과, 10대 연령층의 706명의 모집단 크기에 대하여 제곱근비례배분에서는 155명, p=0.3인 경우 멱등배분에서는 225명이 배분되었다. 따라서 멱등배분이 제곱근비례배분에 비해 너무 많은 표본이 배분되어 다른 연령대에 비해 응답협조도가 낮은 특성을 고려했을 때 조사에 어려움이 나타날 것으로 예상되어 제곱근비례배분이 더 적절한 방법이라고 생각된다.

4. 최적 표본크기 및 배분방법

북한이탈주민실태조사의 현행 표본크기 3,000명을 유지할 경우 허용오차는 1.7%이고 2,500명으로 줄일 경우 1.9%로 나타나, 허용오차 측면과 조사여건 등을 고려했을 때 표본크기를 2,500명으로 줄이는 것을 제안하고자 한다. 그리고 표본배분방법으로는 공표수준별 허용오차가 2.3~6.1%이면서 10대 연령층의 특성을 고려한 제곱근비례배분이 적절할 것으로 판단된다.

제 4 절 표본추출 및 대체

표본추출은 특성별로 층화를 한 후, 각 층에 표본을 할당하고, 주어진 40개의 층으로부터 지역(17개 시도)으로 정렬한 후 계통추출법으로 표본을 추출한다. 실제 조사를 수행했을 때 표본으로 추출된 대상의 조사가 불가능한 경우 동일한 속성(거주지역, 성별, 연령, 남한거주기간 등)을 가진 북한이탈주민으로 대체하는 것을 원칙으로 한다. 표본추출과정에서 사전에 원표본과 대체표본을 추출한 후, 원표본 → 1차 대체표본 → 2차 대체표본 → 3차 대체표본 등의 순으로 정해진 대체 순서에 따라 대체한다. 조사원의 임의적인 표본대체를 방지하기 위하여 대체지침에 따라 표본을 대체하고 이전 차수에서 조사 불가사항이 발생하면 구체적인 대응방안과 사유를 기록하도록 한다.

제 5 절 가중치 작성 및 모수 추정

기존 표본설계에서는 표본가중치와 사후층화가중치에 대한 설명을 하고 있으며, 모수추정에서는 총합 추정량과 분산추정량을 제시하고 평균과 비율 추정에 대하여 간략하게 언급하고 있다. 따라서 3단계로 이루어지는 가중치 조정방법에 대하여 구체적으로 다룰 필요가 있으며, 평균, 비율, 총합 추정량과 그 분산추정량 그리고 상대표준오차도 자세하게 제시할 필요가 있다.

1. 가중치 조정방법

일반적으로 조사단위에 대해 가중의 절차는 (1) 추출가중치 또는 설계가중치 (2) 무응답조정가중치 (3) 사후층화가중치의 단계로 이루어진다.

1) 설계가중치

설계가중치는 표본추출설계로부터 직접적으로 얻어지는 값이다. 임의의 모집단으로부터 적절한 크기의 표본을 추출할 때 단위가 표본에 포함될 확률의 역수로 설계가중치를 계산할 수 있다. 이때 포함확률은 기지(known)의 값이다.

2) 무응답조정가중치

무응답조정가중치의 중요한 역할은 조사로부터 발생한 무응답으로 인한 무응답 편향을 제거하기 위한 것이다. 무응답 편향은 무응답자들이 응답자들과 차이가 있을 때 발생하며, 모집단에 대해 매우 높은 비율로 추정치에 영향을 줄 때 발생한다.

무응답조정가중치를 계산하는 방법으로는 이용 가능한 정보의 근원에 따라 '표본에 기초한 무응답조정 방법'과 '외부정보를 이용한 무응답 조정방법'이 있다. 표본에 기초한 무응답 조정방법은 이용 가능한 정보가 표본으로 한정되며, 전체 모집단에 대한 정보는 알 수 없고, 무응답 단위들의 설계가중치를 표본응답자들에게 배정하여 응답단위들에 대해 조정된 가중치의 합이 전체 표본단위들에 대한 설계가중치의 합이 된다. 외부정보를 이용한 무응답 조정방법은 표본에 기초한 무응답조정가중치를 계산한 후 외부자료를 이용하여 사후층화, raking, 또는 calibration 등의 방법으로 이 가중치를 조정한다.

3) 사후층화 가중치

사후층화조정은 표본추출틀의 불완전으로 인한 포괄성의 차이, 표본의 불균형 또는 비대표성, 무응답에 의한 차이 등을 조정하기 위해 광범위하게 이용되는 방법이다. 즉, 표본응답자들의 가중치를 조정함으로서 가중된 표본분포가 기지의 모집단분포와 같아지도록 하는 방법이다.

이러한 사후층화조정을 실시하는 주된 이유는 첫째, 추정치의 정도를 개선할 수 있으며, 둘째, 추정분산을 줄이기 위해 층화와 비추정을 사용할 수 있고, 셋째, 부차모집단간의 포괄성과 무응답에 따른 추정치의 편향을 감소시키며, 넷째, 모집단의 다양한 그룹에 대해 추정치의 일치성을 보장한다.

2. 북한이탈주민실태조사의 가중치 산정

일반적으로 추정단계에서 가중치를 이용하면 모집단에 대한 특성치인 모수에 대한 비편향추정량(unbiased estimator)을 얻을 수 있다. 만약 통계분석 과정에서 가중치를 무시하고 계산한 추정치는 심각한 편향이 발생할 수 있다.

표본의 크기가 큰 대규모 조사에서 문제가 되는 것은 추정량의 편향이기 때문에 추정과정에서 반드시 가중치를 이용해야 한다. 일반적으로 복합표본조사의 가중치는 기본 가중치, 무응답에 대한 조정, 사후층화에 대한 조정 등의 세 가지 요인을 통합하여 산정된다.

이번 표본설계에서는 특성별 평균 및 비율에 대한 추정치 계산이 가능하도록 하였고, 추정량과 추정량의 표준오차공식을 제시하였다.

북한이탈주민실태조사에서 각종 통계치는 사후층화 조정을 통해서 얻어진 가중치를 이용하여 계산된 것으로 추정량의 분산 계산이 복잡하다. 최근에는 복합표본조사 데이터를 전문적으로 분석할 수 있는 다양한 통계소프트웨어가 개발되어 널리 활용되고 있다. 대표적인 복합표본조사 분석용 통계소프트웨어는 SAS 9.4판 이상, SUDAAN, WestVarPC, Stata 등이다. 북한이탈주민실태조사 자료에 대한 고급분석을 위해서는 복합표본조사 분석용 통계소프트웨어를 활용하는 것이 필요하다.

북한이탈주민실태조사에서 각 응답자에 부여되는 가중치 산출은 총 3단계에 걸쳐 이루어진다. 제 1단계에서는 다음과 같은 방법으로 설계가중치를 계산한다. 지역(2개 층), 성별(2개 층), 연령(5개 층), 남한거주기간(2개 층)을 층화변수로 사용하여 총 40개의 층을 고려하고 있으므로, 각 층에 부여되는 설계가중치는 각 층에 속한 모집단 응답자 수와 표본 응답자의 비율로 표현할 수 있다.

h층에 속한 표본응답자의 기본가중치는 다음과 같이 표현할 수 있다.

$$\text{기본가중치} : w_h^B = \frac{N_h}{n_h}$$

여기서 N_h는 h층의 북한이탈주민 모집단 수이고, n_h는 h층의 북한이탈주민 표본 수이다.

2단계에서는 다음과 같이 무응답 조정 가중치를 계산한다.

$$\text{무응답 조정 가중치 : } w_h^R = \frac{n_h}{r_h}$$

여기서 r_h는 h층에서 응답한 응답자 수이다.

3단계에서는 사후층화에 의한 최종가중치를 다음과 같이 산출한다.

$$\text{사후층화에 의한 최종가중치 : } w_{hi} = w_h^B \times w_h^R \times \frac{X_h}{\hat{X}_h}$$

여기서 X_h는 h층의 북한이탈주민 최신 모집단 총수이며, $\hat{X}_h$는 h층의 가중합이다.

3. 북한이탈주민실태조사의 모수 추정

1) 평균 추정

복합표본조사 데이터를 분석할 때 가중치를 무시하고 분석하면 모수 추정에 심각한 편향(bias)이 발생할 수 있고, 추정량의 분산이 과소평가되어 문제가 된다. 따라서 북한이탈주민실태조사에서 모집단의 특성치에 대한 추정 시 가중치를 이용해야 하고 만약 단순총계를 사용하면 추정치에 편향이 발생할 수 있다.

북한이탈주민실태조사에서 모집단 특성치의 추정에 사용될 기호들은 다음과 같다.

$h = 1,2,\cdots,L$: 층의 수

$i = 1, 2, \cdots, n_h$: h층 내의 표본 수

w_{hi} : h층의 i번째 응답자에 대한 최종가중치

y_{hi} : h층의 i번째 응답자로부터 얻은 변수값

f_h : 추출률

따라서 북한이탈주민실태조사의 주요변수에 대한 평균 추정량 $\bar{y}$는 다음과 같다.

$$\bar{y} = \frac{\sum_{h=1}^{L}\sum_{i=1}^{n_h} w_{hi}\, y_{hi}}{\sum_{h=1}^{L}\sum_{i=1}^{n_h} w_{hi}}$$

그리고 북한이탈주민실태조사의 주요변수에 대한 평균 추정량 $\bar{y}$의 분산 추정량은 다음과 같다.

$$\hat{V}(\bar{y}) = \sum_{h=1}^{L} \frac{n_h(1-f_h)}{n_h-1} \sum_{i=1}^{n_h}(e_{hi}-\bar{e}_{h.})^2$$

여기서 $e_{hi} = w_{hi}(y_{hi}-\bar{y})/w_{..}$, $\bar{e}_{h.} = \left(\sum_{i=1}^{n_h} e_{hi}\right)/n_h$, $w_{..} = \sum_{h=1}^{L}\sum_{i=1}^{n_h} w_{hi}$이다.

또한 북한이탈주민실태조사의 주요변수에 대한 평균 추정량 $\bar{y}$의 상대표준오차는 다음과 같다.

$$\widehat{RSE}(\bar{y}) = \frac{\sqrt{\hat{V}(\bar{y})}}{\bar{y}} \times 100\,(\%)$$

2) 총계 추정

북한이탈주민실태조사의 주요변수 대한 총계 추정량 $\hat{Y}$는 다음과 같다.

$$\hat{Y} = \sum_{h=1}^{L}\sum_{i=1}^{n_h} w_{hi} y_{hi}$$

그리고 북한이탈주민실태조사의 주요변수 대한 총계 추정량 $\hat{Y}$의 분산 추정량은 다음과 같다.

$$\hat{V}(\hat{Y}) = \sum_{h=1}^{L} \hat{V}(\hat{Y}_h) = \sum_{h=1}^{L} \frac{n_h(1-f_h)}{n_h-1} \sum_{i=1}^{n_h} (\tau_{hi} - \bar{\tau}_{h.})^2$$

여기서 $\tau_{hi} = w_{hi} y_{hi}$, $\tau_{h.} = \frac{1}{n_h} \sum_{i=1}^{n_h} \tau_{hi}$이다.

또한 북한이탈주민실태조사의 주요변수 대한 총계 추정량 $\hat{Y}$의 상대표준오차는 다음과 같다.

$$\widehat{RSE}(\hat{Y}) = \frac{\sqrt{\hat{V}(\hat{Y})}}{\hat{Y}} \times 100(\%)$$

3) 비율 추정

북한이탈주민실태조사의 주요변수 대한 비율 추정량 $\hat{p}$는 다음과 같다.

$$\hat{p} = \frac{\sum_{h=1}^{L} \sum_{i=1}^{n_h} w_{hi}\, y_{hi}}{\sum_{h=1}^{L} \sum_{i=1}^{n_h} w_{hi}}$$

여기서 y_{hi}는 특정한 속성을 가지면 1, 그렇지 않으면 0이다.

그리고 북한이탈주민실태조사의 주요변수 대한 비율 추정량 $\hat{p}$의 분산 추정량은 다음과 같다.

$$\hat{V}(\hat{p}) = \sum_{h=1}^{L} \frac{n_h(1-f_h)}{n_h-1} \sum_{i=1}^{n_h} (e_{hi} - \bar{e}_{h.})^2$$

여기서 $e_{hi} = w_{hi}(y_{hi} - \hat{p})/w_{..}$, $\bar{e}_{h.} = \left(\sum_{i=1}^{n_h} e_{hi}\right)/n_h$, $w_{..} = \sum_{h=1}^{L}\sum_{i=1}^{n_h} w_{hi}$이다.

또한 북한이탈주민실태조사의 주요변수 대한 비율 추정량 $\hat{p}$의 상대표준오차는 다음과 같다.

$$\widehat{RSE}(\hat{p}) = \frac{\sqrt{\hat{V}(\hat{p})}}{\hat{p}} \times 100\,(\%)$$

제 5 장 조사표 개선

제 1 절 조사방식 및 조사항목 검토

1. 조사방식

북한이탈주민실태조사에 대한 조사표 설계는 대내외 전문가 자문회의 실시를 통해 개발 의견을 수렴하고 있다.

2021년 기준 북한이탈주민실태조사의 조사항목은 크게 기본사항(3문항), 정착실태(55문항), 사회통합(57문항)으로 총 115개 문항으로 구성되어 있다. 이처럼 북한이탈주민실태조사는 조사내용 및 분량 측면에서 응답자의 응답 부담이 매우 큰 조사로 평가되고 있다.

모든 항목을 표본으로 선정된 응답자들에 대해 매해 조사하기보다 조사 영역을 조사시기에 따라 나누어 로테이션하며 조사하는 등 현재의 조사방식을 변경하는 것이 바람직하다고 생각된다. 정착실태와 같이 주요한 통계는 매년 생산될 수 있도록 하고 사회통합과 관련된 조사항목은 영역별로 격년으로 조사하는 방안 등을 고려해 볼 수 있겠다.

2. 조사항목

1) 북한이탈주민 정착실태조사 조사항목

2021년 기준 북한이탈주민실태조사에서 정착실태 관련 문항은 55개 문항으로 입국 전 사항(4문항), 생활안정(10문항), 직업교육훈련(4문항), 일에 관한 사항(26문항), 구직에 관한 사항(4문항), 기타 활동에 관한 사항(5문항), 이전 직장(일)에 관한 사항(2문항)으로 구성되어 있다.

<표 5-1> 북한이탈주민 정착실태조사 조사항목

조사부문	조사항목
입국 전 사항 (4개)	1. 탈북동기 2. 탈북시기 3. 북한에서의 직업 4. 제3국에서의 직업
생활안정 (10개)	1. 현재 거주 주택 2. 남한 생활 만족도 3. 남한 생활 만족 이유 4. 남한 생활 불만족 이유 5. 북한에서의 사회경제적 지위 6. 남한에서의 사회경제적 지위 7. 사회·경제적 성취 만족도 8. 사회·경제적 지위 향상 가능성 9. 자식세대 사회·경제적 지위 향상 가능성 10. 더 나은 남한생활을 위해 필요한 지원
직업교육훈련 (4개)	1. 직업교육훈련 경험 여부 2. 직업교육훈련 경험 분야 3. 직업교육훈련 수료 여부 4. (직업교육훈련 수료한 경우)동일분야 근무 여부
일에 관한 사항 (26개)	1. 지난주 수입을 목적으로 1시간 이상 경제활동을 하였는지 여부 2. 지난주 함께 사는 가족의 사업체 또는 농장에서 보수를 받지 않고 일을 하였는지 여부 3. 지난주 일을 하지는 않았지만 직장(일)이 있었는지 여부 4. 지난주 주된 일 이외에 다른 일을 하였는지 여부 5. 평소 1주일간 총 근무 시간 6. 평소 1주일간 36시간 미만 근무 이유 7. 지난주 실제 총 근무 시간 8. 지난주 36시간 미만 근무 이유 9. 지난주 더 많은 시간 근무하기 원하였는지 여부 10. 지난주 더 많은 시간 근무할 수 있는 기회가 있었다면 할 수 있었는지 여부 11. 지난 4주간 더 많은 시간 일하기 위해 현재 하는 일이나 다른 일을 찾아 본 적이 있었는지 여부 12. 지난주 주된 일을 한 직장(사업체)의 사업체 유형 13. 지난주 주된 일을 한 직장(사업체)의 직원 수 14. 지난주 주된 일을 한 직장(사업체)에서의 본인 업무 유형 15. 지난주 주된 일을 한 직장(사업체) 근무 시작 년월 16. 지난주 주된 일을 한 직장(사업체)에서의 종사상 지위

	17. 지난주 주된 일을 한 직장(사업체)와 고용계약기간 정하였는지 여부 및 고용계약기간 18. 지난 3개월간 세금공제 전 월평균 임금 19. 지난주 사회보험 가입여부 20. 지난주 일을 한 직장(사업체) 구직 경로 21. 지난주 일을 한 직장(사업체) 구직 기간 22. 현재 사업의 사업자금 조달방법 23. 현재 사업 처음 시작시 사업자금 규모 24. 현재 사업 처음 시작하기까지 준비기간 25. 현재 사업 시작 동기 26. 현재 사업 시작시 가장 어려웠던 점
구직에 관한 사항 (4개)	1. 지난 4주간 구직 여부 2. 지난주 직장(일)이 주어졌다면 일을 시작할 수 있었는지 여부 3. 구직 경로 4. 구직 기간
기타 활동에 관한 사항 (5개)	1. 지난주 직장(일)을 원하였는지 여부 2. 지난주 직장(일)이 주어졌다면 시작할 수 있었는지 여부 3. 지난 4주 내 비구직 이유 4. 지난 2년간 구직 여부 5. 지난주 활동상태
이전 직장(일)에 관한 사항 (2개)	1. 과거 취업 여부 2. 이전 직장(일)을 그만둔 이유

2) 북한이탈주민 사회통합실태조사 조사항목

2021년 기준 북한이탈주민실태조사 사회통합 관련 문항은 57개 문항으로 노동(7문항), 사회활동(8문항), 복지(6문항), 보건 및 건강(11문항), 교육(6문항), 가족(11문항), 사회적 관계망(4문항), 가구 경제상태(4문항)으로 구성되어 있다.

<표 5-2> 북한이탈주민 사회통합실태조사 조사항목

조사부문	조사항목
노동 (7개)	1. 직업 선택 고려 요인 2. 일하고 싶은 직장 및 업종 3. 여성 취업에 관한 견해 3-1. 여성 취업 시기에 관한 견해 4. 여성 취업을 어렵게 하는 요인 5. 일과 가정생활 우선순위 6. 실직·전직에 대한 불안함 7. 근로여건 평가
사회활동 (8개)	1. 주로 하는 여가활동(주중/주말) 2. 여가활동 같이 하는 사람(주중/주말) 3. 여가활동 만족도 3-1. 여가활동 불만족 이유 4. 희망하는 여가활동 5. 사회단체 참여 여부 5-1. 참여한 사회단체 유형 6. 북한이탈주민 운영 단체 가입 여부 6-1. 북한이탈주민 운영 단체 참여 정도 7. 자원봉사 참여 여부 7-1. 참여한 자원봉사활동 유형 8. 향후 자원봉사활동 참여 의향
복지 (6개)	1. 생활여건 만족도 2. 경제적 어려움 경험 여부 3. 노후 준비 여부 3-1. 노후 준비 방법 3-2. 준비하지 않는 이유 4. 노후를 위한 사회적 관심사

	5. 노후를 보내고 싶은 방법 6. (만 60세 이상) 생활비 마련 방법 6-1. 구체적 방법
보건 및 건강 (11개)	1. 평소 건강상태 2. 건강관리 실천 여부 3. 흡연 여부 3-1. 흡연 정도 4. 음주 여부 5-1. 음주 빈도 5. 의료기관 이용 여부 5-1. 주로 이용한 의료기관 6. 의료기관 서비스 만족도 6-1. 의료기관 서비스 불만족 이유 7. 질병 및 사고 경험 여부 8. 치료 및 약 복용 경험 여부 8-1. 치료 및 약 복용 일수 9. 의료 서비스 받지 못한 경험 여부 9-1. 의료 서비스 받지 못한 이유 10. 스트레스 정도 11. 자살 충동 경험 11-1. 자살 충동 이유
교육 (6개)	1. 대학 휴학·중퇴 경험 2. 대학 휴학·중퇴 이유 3. 희망 교육 수준 3-1. 대학 이상 교육 기대 목적 4. 학생 자녀 유무 4-1. 자녀 거주 지역 5. 자녀 기대 교육 수준 5-1. 자녀의 대학 이상 교육 기대 목적 6. 자녀 교육비 부담 정도 6-1. 가장 부담되는 항목
가족 (11개)	1. 가구주와 관계 2. 현재 생계를 같이 하는 가구원수, 만 65세 이상·만 15세 이상· 취업자 가구원 수 3. 배우자 유무 3-1. 배우자 출신 지역 3-2. 배우자 거주 지역 4. 함께 사는 가족 5. 부모 노후 부양에 대한 견해 5-1. 가족 중 부모 부양해야 하는 구성원 6. 결혼에 대한 견해

	7. 이혼에 대한 견해 8. 재혼에 대한 견해 9. 가사분담에 대한 견해 10. 가사분담 실태 11. 가족관계 만족도
사회적 관계망 (4개)	1. 인간관계 만족도 2. 상황별 도움을 받을 수 있는 사람 유무 및 인원 3. 어려운 일이 닥쳤을 때 가장 의지가 되는 사람 4. 차별·무시당한 경험 4-1. 차별·무시당한 이유
가구 경제상태 (4개)	1. 월평균 가구 소비지출액 2. 가구 생활비 중 가장 부담이 되는 항목 3. 가구 소득 금액 4. 가구 자산 및 부채 현황

북한이탈주민실태조사의 조사목적에 부합하는 조사를 위해 무엇보다 중요한 것은 설문지에 포함된 조사항목들이 현실을 잘 파악할 수 있는가, 이러한 조사항목들을 실제 활용할 수 있는가이다. 그러나 공개된 자료의 한계와 낮은 활용성 문제가 통계 이용자들에게 지적되고 있다. 북한이탈주민실태조사는 이미 조사내용이 다른 조사에 비해 많아 추가하는 것은 어려울 것으로 보이나 조사목적 및 이용자들의 요구와 정책적 요구를 반영하여 조사항목을 검토 및 개선할 필요가 있다.

북한이탈주민실태조사의 경우 정착실태와 관련한 내용(생활안정, 입국 전 사항, 직업교육훈련, 일에 관한 사항, 구직에 관한 사항, 기타 활동에 관한 사항, 이전 직장(일)에 관한 사항) 뿐 아니라, 사회통합과 관련한 내용(노동, 사회활동, 복지, 보건 및 건강, 교육, 가족, 사회적 관계망, 가구 경제상태)을 조사하고 있어, 조사영역이 매우 광범위하다. 조사문항이 지나치게 길 경우 후반부로 갈수록 응답의 신뢰성을 담보할 수 없으므로, 응답자들이 집중하여 응답할 수 있는 시간에 조사가 종료될 수 있도록 불필요한 문항은 삭제하는 등 조정 작업이 필요하다.

제 2 절 조사항목 개선안

1. 북한이탈주민 정착실태조사 조사항목 개선안

2021년 기준 북한이탈주민실태조사에서 정착실태 관련 문항 55개 중 일에 관한 사항(26문항)에서 자영업자를 대상으로 하는 5개의 문항(현재 사업의 사업자금 조달방법, 현재 사업 처음 시작시 사업자금 규모, 현재 사업 처음 시작하기까지 준비기간, 현재 사업 시작 동기, 현재 사업 시작시 가장 어려웠던 점)은 매년 조사하는 것 보다는 필요시 부가조사로 하는 것이 바람직해 보인다.

따라서 2022년 기준 북한이탈주민실태조사부터는 정착실태 관련 문항이 50개 문항으로 입국 전 사항(4문항), 생활안정(10문항), 직업교육훈련(4문항), 일에 관한 사항(21문항), 구직에 관한 사항(4문항), 기타 활동에 관한 사항(5문항), 이전 직장(일)에 관한 사항(2개)으로 구성하는 다음과 같은 안을 제안하고자 한다.

<표 5-3> 북한이탈주민 정착실태조사 조사항목 개선안

조사부문	조사항목
입국 전 사항 (4개)	1. 탈북동기 2. 탈북시기 3. 북한에서의 직업 4. 제3국에서의 직업
생활안정 (10개)	1. 현재 거주 주택 2. 남한 생활 만족도 3. 남한 생활 만족 이유 4. 남한 생활 불만족 이유 5. 북한에서의 사회경제적 지위 6. 남한에서의 사회경제적 지위 7. 사회·경제적 성취 만족도 8. 사회·경제적 지위 향상 가능성 9. 자식세대 사회·경제적 지위 향상 가능성 10. 더 나은 남한생활을 위해 필요한 지원

직업교육훈련 (4개)	1. 직업교육훈련 경험 여부 2. 직업교육훈련 경험 분야 3. 직업교육훈련 수료 여부 4. (직업교육훈련 수료한 경우)동일분야 근무 여부
일에 관한 사항 (21개)	**1. 지난주 수입을 목적으로 1시간 이상 경제활동을 하였는지 여부 2. 지난주 함께 사는 가족의 사업체 또는 농장에서 보수를 받지 않고 일을 하였는지 여부 3. 지난주 일을 하지는 않았지만 직장(일)이 있었는지 여부 4. 지난주 주된 일 이외에 다른 일을 하였는지 여부 5. 평소 1주일간 총 근무 시간 6. 평소 1주일간 36시간 미만 근무 이유 7. 지난주 실제 총 근무 시간 8. 지난주 36시간 미만 근무 이유 9. 지난주 더 많은 시간 근무하기 원하였는지 여부 10. 지난주 더 많은 시간 근무할 수 있는 기회가 있었다면 할 수 있었는지 여부 11. 지난 4주간 더 많은 시간 일하기 위해 현재 하는 일이나 다른 일을 찾아 본 적이 있었는지 여부 12. 지난주 주된 일을 한 직장(사업체)의 사업체 유형 13. 지난주 주된 일을 한 직장(사업체)의 직원 수 14. 지난주 주된 일을 한 직장(사업체)에서의 본인 업무 유형 15. 지난주 주된 일을 한 직장(사업체) 근무 시작 년월 16. 지난주 주된 일을 한 직장(사업체)에서의 종사상 지위 17. 지난주 주된 일을 한 직장(사업체)와 고용계약기간 정하였는지 여부 및 고용계약기간 18. 지난 3개월간 세금공제 전 월평균 임금 19. 지난주 사회보험 가입여부 20. 지난주 일을 한 직장(사업체) 구직 경로 21. 지난주 일을 한 직장(사업체) 구직 기간**
구직에 관한 사항 (4개)	1. 지난 4주간 구직 여부 2. 지난주 직장(일)이 주어졌다면 일을 시작할 수 있었는지 여부 3. 구직 경로 4. 구직 기간
기타 활동에 관한 사항 (5개)	1. 지난주 직장(일)을 원하였는지 여부 2. 지난주 직장(일)이 주어졌다면 시작할 수 있었는지 여부 3. 지난 4주 내 비구직 이유 4. 지난 2년간 구직 여부 5. 지난주 활동상태
이전 직장(일)에 관한 사항 (2개)	1. 과거 취업 여부 2. 이전 직장(일)을 그만둔 이유

또한 북한이탈주민실태조사는 재단 소속 전문상담사가 북한이탈주민을 만나 대인면접조사를 실시하고 있으므로, 북한이탈주민들의 어려움이나 요구사항을 구체적으로 파악할 수 있다. 이처럼 실제조사가 일반 조사원이 아닌 전문상담사들에 이루어지고 있으므로 이러한 장점을 충분히 살려 북한이탈주민들의 요구사항을 묻는 내용을 정확히 기록하도록 하고, 연구진은 이를 범주화하여 분석하는 작업이 필요해 보인다. 이러한 분석 결과는 북한이탈주민들의 지원정책 개발에 활용될 수 있을 것이다.

2. 북한이탈주민 사회통합실태조사 조사항목 개선안

2021년 기준 북한이탈주민실태조사에서 사회통합 관련 문항은 57개*로 매년 정착실태(50 문항)와 함께 조사할 경우 응답자의 응답부담이 상당할 것으로 판단된다.

* 노동(7문항), 사회활동(8문항), 복지(6문항), 보건 및 건강(11문항), 교육(6문항), 가족(11문항), 사회적 관계망(4문항), 가구 경제상태(4문항)

이에 2022년 기준 북한이탈주민실태조사부터는 사회통합조사 Ⅰ과 Ⅱ로 구분하여 격년제로 조사는 방안을 제안하고자 한다. 이때 사회통합조사 Ⅰ과 Ⅱ로 분류하는 기준은 조사내용을 고려하여 통계청의 사회조사와 유사하도록 구성하였다.

현재 사회조사는 짝수 연도에는 보건·교육·안전·가족·환경 부문을, 홀수 연도에는 복지·사회참여·문화와 여가·소득과 소비·노동 부문을 조사하고 있으므로 북한이탈주민실태조사도 짝수 연도에는 보건 및 건강, 교육 등에 대한 사회통합조사 Ⅰ을 실시하고, 홀수 연도에는 노동, 사회활동, 복지 등에 대한 사회통합조사 Ⅱ를 실시하는 것이 적절할 것으로 판단된다.

1) 북한이탈주민 사회통합조사 Ⅰ 조사항목 개선안

사회통합조사 Ⅰ는 28개 문항이며, 보건 및 건강(11문항), 교육(6문항), 가족(11문항)으로 구성한 다음과 같은 안을 제안한다.

<표 5-5> 북한이탈주민 사회통합조사 Ⅰ 조사항목 개선안

조사부문	조사항목
보건 및 건강 (11개)	1. 평소 건강상태 2. 건강관리 실천 여부 3. 흡연 여부 3-1. 흡연 정도 4. 음주 여부 4-1. 음주 빈도 5. 의료기관 이용 여부 5-1. 주로 이용한 의료기관 6. 의료기관 서비스 만족도 6-1. 의료기관 서비스 불만족 이유 7. 질병 및 사고 경험 여부 8. 치료 및 약 복용 경험 여부 8-1. 치료 및 약 복용 일수 9. 의료 서비스 받지 못한 경험 여부 9-1. 의료 서비스 받지 못한 이유 10. 스트레스 정도 11. 자살 충동 경험 11-1. 자살 충동 이유
교육 (6개)	1. 대학 휴학·중퇴 경험 2. 대학 휴학·중퇴 이유 3. 희망 교육 수준 3-1. 대학 이상 교육 기대 목적 4. 학생 자녀 유무 4-1. 자녀 거주 지역 5. 자녀 기대 교육 수준 5-1. 자녀의 대학 이상 교육 기대 목적 6. 자녀 교육비 부담 정도 6-1. 가장 부담되는 항목
가족 (11개)	1. 가구주와 관계 2. 현재 생계를 같이 하는 가구원수, 만 65세 이상·만 15세 이상·취업자 가구원 수 3. 배우자 유무 3-1. 배우자 출신 지역

	3-2. 배우자 거주 지역 4. 함께 사는 가족 5. 부모 노후 부양에 대한 견해 5-1. 가족 중 부모 부양해야 하는 구성원 6. 결혼에 대한 견해 7. 이혼에 대한 견해 8. 재혼에 대한 견해 9. 가사분담에 대한 견해 10. 가사분담 실태 11. 가족관계 만족도

2) 북한이탈주민 사회통합조사 Ⅱ 조사항목 개선안

사회통합조사 Ⅱ은 29개 문항이며, 노동(7문항), 사회활동(8문항), 복지(6문항), 사회적 관계망(4문항), 가구경제상태(4문항)로 구성한 다음과 같은 안을 제안한다.

<표 5-4> 북한이탈주민 사회통합조사 II 조사항목 개선안

조사부문	조사항목
노동 (7개)	1. 직업 선택 고려 요인 2. 일하고 싶은 직장 및 업종 3. 여성 취업에 관한 견해 3-1. 여성 취업 시기에 관한 견해 4. 여성 취업을 어렵게 하는 요인 5. 일과 가정생활 우선순위 6. 실직·전직에 대한 불안함 7. 근로여건 평가
사회활동 (8개)	1. 주로 하는 여가활동(주중/주말) 2. 여가활동 같이 하는 사람(주중/주말) 3. 여가활동 만족도 3-1. 여가활동 불만족 이유 4. 희망하는 여가활동 5. 사회단체 참여 여부 5-1. 참여한 사회단체 유형 6. 북한이탈주민 운영 단체 가입 여부 6-1. 북한이탈주민 운영 단체 참여 정도 7. 자원봉사 참여 여부 7-1. 참여한 자원봉사활동 유형 8. 향후 자원봉사활동 참여 의향

복지 (6개)	1. 생활여건 만족도 2. 경제적 어려움 경험 여부 3. 노후 준비 여부 3-1. 노후 준비 방법 3-2. 준비하지 않는 이유 4. 노후를 위한 사회적 관심사 5. 노후를 보내고 싶은 방법 6. (만 60세 이상) 생활비 마련 방법 6-1. 구체적 방법
사회적 관계망 (4개)	1. 인간관계 만족도 2. 상황별 도움을 받을 수 있는 사람 유무 및 인원 3. 어려운 일이 닥쳤을 때 가장 의지가 되는 사람 4. 차별·무시당한 경험 4-1. 차별·무시당한 이유
가구 경제상태 (4개)	1. 월평균 가구 소비지출액 2. 가구 생활비 중 가장 부담이 되는 항목 3. 가구 소득 금액 4. 가구 자산 및 부채 현황

3) 사회통합조사 Ⅱ의 문항 5의 문구 수정

북한이탈주민실태조사(사회통합조사, 노동) 5번 문항(일과 가정생활 우선순위)의 선택문항에서 '① 일을 우선시함'과 '② 일을 우선시하는 편임'의 구분이 명확하지 않으므로 다음과 같이 ①번 선택 문항을 '① 일을 매우 우선시함'으로 수정하고, 같은 맥락으로 ⑤번 선택 문항을 '⑤ 가정을 매우 우선시함'으로 수정하는 것이 적절해 보인다.

5. 귀하는 <u>현재</u> 직장 일과 가정생활 중 어느 것을 우선시하고 있습니까?

① 일을 매우 우선시함 ② 일을 우선시하는 편임

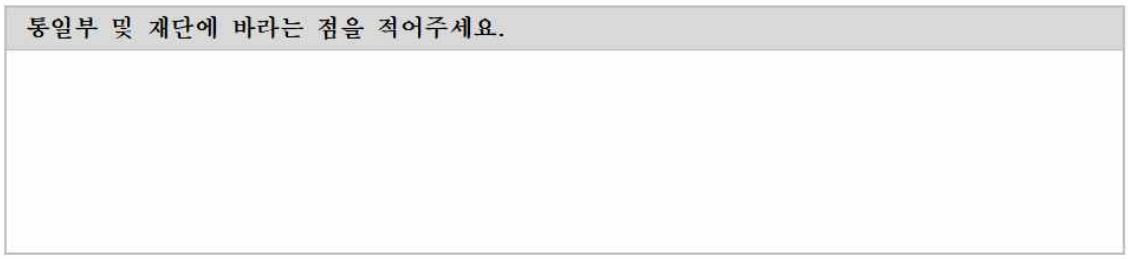
③ 일과 가정생활 둘 다 비슷함 ④ 가정생활을 우선시하는 편임

⑤ 가정생활을 매우 우선시함

4) 서술형 항목 분석

북한이탈주민실태조사의 조사항목은 대부분 선택형(폐쇄형)으로 구성되어 있어 응답자들에게 응답 편의를 제공하고 있어 항목무응답률을 줄이는 효과가 있다. 하지만 선택형으로만 구성된 조사항목은 새로운 사실을 발견하기 어려운 문제점이 있다.

현재 북한이탈주민실태조사의 조사표의 가장 마지막 부분에



라는 서술형 항목이 있지만, 이 항목에 대한 분석이 되지 않고 있다.

따라서 전문상담사들로 하여금 위 항목에 대하여 응답자들의 응답을 적극적으로 받는다면 북한이탈주민들의 욕구를 정확히 파악할 수 있을 것이다.

파악된 북한이탈주민들의 요구사항을 조사항목의 개발이나 정책에 연계할 수 있다면 북한이탈주민실태조사의 활용성을 높이는 계기가 될 수 있을 것으로 생각된다.

제 6 장 결론 및 제언

북한이탈주민실태조사는 북한이탈주민의 남한정착정도를 파악하고, 북한이탈주민에 대한 지원정책 수립 등에 필요한 기초자료를 제공하기 위해 실시하고 있다. 이러한 작성 목적에 맞는 통계 품질 개선을 위해 2020년 북한이탈주민실태조사 정기통계품질진단 결과 및 표본설계를 검토한 결과, '패널조사로의 전환 검토', '표본설계 개선방안 마련', '조사표 검토' 등이 필요한 것으로 파악되었다.

따라서 본 연구에서는 2020년 북한이탈주민실태조사 결과와 2020년 정기통계품질진단 결과를 토대로 새로운 표본설계 방안(표본배분, 가중치 작성 및 모수추정, 패널관리방안 등)과 조사항목 등의 개선안을 마련하여 북한이탈주민실태조사의 품질을 개선하고자 한다.

본 연구의 주요 결과를 정리하면 다음과 같다.

첫째, 중복 조사되는 표본의 결과를 적절히 활용할 수 있도록 조사방식을 고정패널이나 분할패널설계로 전환하는 문제를 검토하였고, 패널조사로 전환시 필요한 패널 이탈 방지를 위한 패널관리방안을 제시하였다.

둘째, 통일부에서 작성 및 관리하고 있는 북한이탈주민명부인 최신의 표본추출틀을 이용하여 새로운 표본설계 방안을 마련하였다. 지역(2개 층), 성별(2개 층), 연령(5개 층), 남한거주기간(2개 층)을 층화기준으로 사용하여 조사모집단에 대한 충분한 분석을 통해 공표수준별 허용오차 및 조사환경을 고려하여 표본크기를 2,500명 수준으로 줄이는 방안을 제안하였다. 그리고 모집단의 특성을 잘 반영한 효과적인 표본배분방법으로 제곱근비례배분 방법을 제안하였다. 또한 정확한 통계산출을 위해서는 표본설계에 따라 설계가중치와 조사과정에서 발생하는 무응답 조정을 위한 무응답조정가중치 등을 적용한 가중치를 부여하는 과정을 구체적으로 기술하였고, 모수추정식과 상대표준오차식을 제시하는 한편 현재 층화기준과 공표수준의 범위가 달라 세분화된 공표수준별 추정이 불가능하므로, 공표수준을 층화기준에 일치시킬 필요가 있다.

셋째, 조사항목 수가 너무 많아 모든 항목을 표본으로 선정된 응답자들에 대해 매해 조사하기보다 조사영역을 나누어 정착실태와 같은 주요한 분야는 매년 생산하고, 사회통합과 관련된 조사항목은 사회통합조사 Ⅰ(보건 및 건강, 교육, 가족)과 Ⅱ(노동, 사회활동, 복지, 사회적 관계망, 가구 경제상태)로 영역을 구분하여 짝수년에는 사회통합조사 Ⅰ을, 홀수년에는 사회통합조사 Ⅱ를 조사하는 방안을 제안하였다.

다음으로 북한이탈주민실태조사의 조사방법 및 조사항목 개선에 대한 제언을 하면 다음과 같다.

북한이탈주민실태조사를 패널화 하는 문제는 좀 더 심도있는 논의가 필요해 보이며, 현재 조사방법을 유지하면서 표본설계 개선방안을 반영하는 것이 바람직 할 것으로 생각된다. 공표수준별 허용오차 및 조사환경을 고려했을 때 표본크기를 2,500명 수준으로 유지하면서 심층면접방안을 병행하는 방향으로 개선할 필요가 있다.

북한이탈주민실태조사의 조사항목에 대한 통계이용자들의 요구는 다양하게 발생할 수 있지만, 조사항목의 개선은 보다 장기적인 관점에서 진행되어야 한다. 한편 조사항목 중 서술형 항목에 대해서는 상담원들의 교육을 강화하여 심층적인 응답을 이끌어내고 기록하고 관리 및 분석할 필요가 있다.

더불어 북한이탈주민지원재단에서는 현재의 조사내용이 갖는 한계를 보다 체계적이고 지속적으로 검토하여 반영 여부를 결정할 수 있어야 할 것이다. 다만 현재에도 조사내용이 많아 기존의 조사내용을 유지하면서 조사항목을 추가하는 것은 조사의 정확성과 효율성이 떨어질 수 있다. 따라서 현재 광범위하게 조사되고 있는 각 조사내용 및 항목의 활용도 및 중요도에 따라 조사항목을 검토하여 선택과 집중을 할 필요가 있다. 학계 및 정책 수립 분야 등 이용자를 대상으로 조사문항에 대한 요구를 조사하고, 광범위하게 포괄하고

있는 조사영역부터 중요도를 고려하여 재점검할 필요가 있다. 이와 더불어 현실과 괴리가 있거나 응답빈도가 낮은 조사항목을 삭제하고 정책결정에 필요한 문항들을 추가하는 작업이 필요하다. 이와 같은 조사항목 개선 작업은 북한이탈주민실태조사의 현실성 및 신뢰도를 제고하고 이를 통한 활용도 증가에 기여할 것으로 보인다.

북한이탈관련 전문가의 내용 관련 정성적 검토 및 통계관련 전문가의 정량적 검토 등 여러 전문가 의견수렴을 통해 조사항목들을 검토하고 이용자에게 필요한 정보를 제공할 수 있는 통계가 생산된다면 보다 가치있게 활용될 수 있을 것이다.

참고문헌

남북하나재단(2021). 북한이탈주민 사회통합조사

남북하나재단(2021). 북한이탈주민 정착실태조사

북한이탈주민지원재단(2020). 북한이탈주민실태조사 통계정보보고서

샘플링아카데미(2013). 표본조사입문-제3판-, 자유아카데미

손창균(2016). 통계적 관점에서 여성가족패널 설계의 발전 방향. 한국여성정책연구원 패널조사 학술대회, 93-119.

통계청(2007). 통계청 조사지침서, 통계청

통계청(2017). 국민노후보장패널조사 품질개선 컨설팅 최종결과보고서

통계청(2020). 북한이탈주민실태조사 정기통계품질진단 결과보고서

한국보건사회연구원(2006). 표본추출 및 관리 매뉴얼

한국여성정책연구원(2015). 한국여성패널조사 표본보완을 위한 설계연구

Beaumont, J-F., and Rivest, L-P.(2007). A Weight Smoothing Method for Dealing with Stratum Jumpers in Business Surveys. *Proceedings of the Survey Methods Section of SSC Annual Meeting.*

Bernier, J., Nobrega, K.(1998). Outlier Detection in Asymmetric Samples: A Comparison of an Inter-quartile Range Method and a Variation of a Sigma Gap Method, *Proceedings of the Survey Research Methods Section*, *American Statistical Association*, 137-141.

Bethlehem, J.(2009). *Applied Survey Methods*, Wiley.

Biemer, P. P., and Lyberg, L. E.(2003). *Introduction to Survey Quality*, John Wiley and Sons.

Cochran. W. G.(1977). *Sampling Techniques*, John Wiley & Sons, Inc.

Groves, R. M., Dillman, D. A. Eltinge, J. L., and Little, R. J. A.(2002). *Survey Nonresponse*, John Wiley and Sons.

Groves, R. M., Fowler, Jr., F. J., Couper, M. P., Lepkowski, J. M., Singer, E., and Tourangean, R.(2004). *Survey Methodology*, John Wiley and Sons.

Hansen, M. H., and Hurwitz, W. N.(1946). The Problem of Non-response in Sample Surveys, *Journal of the American Statistical Association*, 41, 517-529.

Hidiroglou, M. A., and Berthelot, J. M.(1986). Statistical Edit and Imputation for Periodic Surveys, *Survey Methodology*, 12, 73-83.

Hulliger, B.(1999). Simple and Robust Estimators for Sampling, *Proceedings of the Survey Research Methods Section, American Statistical Association*, 54-63.

Lavalee, P., and Hidiroglou, M. A.(1988). On the Stratification of Skewed Populations. *Survey Methodology*, 14(1), 33-43.

Rivest, L.-P.(1999). Stratum Jumpers: Can We Avoid Them? *Proceedings of the Survey Research Methods Section, American Statistical Association*, 64-72.

Rouseeuw and Zomeren(1990). Unmasking Multivariate Outliers and Leverage Points, *Journal of the American Statistical Association*, 85, 633-639.

SAS/STAT User's Guide, Version 8, SAS Publishing.

Srinath, K. P., and Carpenter, R. M.(1995). Sampling Methods for Repeated Business Surveys, *Business Survey Methods,* Eds. Cox, B.G. et al., Wiley.

Thompson, K.(2007). Investigation of Macro Editing Techniques for Outlier Detection in Survey Data, *Proceedings of the Third International Conference on Establishment Surveys, American Statistical Association*.

Thompson, K. J.(2001). Ratio Edit Tolerance Development Using Variations of Exploratory Data Analysis(EDA) Resistant Fences Methods, *Statistical Policy Working Paper 29, a Federal Committee on Statistical Methodology Conference Paper.*

부록 : 조사표

승인번호
제 437001 호

통계법 제 33조(비밀의 보호 등)

① 통계작성과정에서 알려진 사항으로서 개인 또는 법인이나 단체의 비밀에 속하는 사항은 보호되어야 한다.

② 통계작성을 위하여 수집된 개인 또는 법인이나 단체의 비밀에 속하는 기초자료는 통계작성의 목적 외에 사용해서는 아니 된다.

2021년 북한이탈주민 실태조사

안녕하십니까?

북한이탈주민지원재단에서는 북한이탈주민들에 대한 지원정책 개발을 위해 '**2021년 북한이탈주민 실태조사**'를 실시하고 있습니다. 본 조사는 「북한이탈주민의 보호 및 정착지원에 관한 법률」에 근거해 북한이탈주민을 대상으로 실시하고자 합니다. 바쁘시더라도 시간을 내어 다음의 질문에 응답하여 주시면 대단히 감사하겠습니다.

귀하의 응답내용은 「통계법」 제33조에 의해 비밀이 보장되며, 귀하의 개인정보는 「개인정보 보호법」 제3조에 의거 철저하게 보호됨을 알려드립니다. 귀하가 응답한 모든 설문결과는 북한이탈주민 지원정책 수립을 위한 목적으로만 사용할 것이며, 개인 정보가 외부로 알려지는 일은 없을 것입니다.

감사합니다.

2021년 5월

조사대상	15세 이상의 북한이탈주민 (2006년 5월 1일 이전 출생자)

북한이탈주민지원재단
North Korean Refugees Foundation

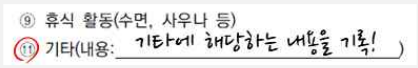
☎담당자: 북한이탈주민지원재단 조사개발팀 장인숙 팀장 02-3215-5727 cool@nkrf.or.kr
북한이탈주민지원재단 조사개발팀 윤설아 주임 02-3215-5854 seola4021@nkrf.or.kr
㈜ 현대리서치연구소 홍종호 이사 02-3218-9651 hong@hri4u.com
㈜ 현대리서치연구소 이희동 과장 02-3218-9630 lhd@hri4u.com

☞ 조사하신 전문상담사 선생님께서 직접 작성해 주시면 됩니다.

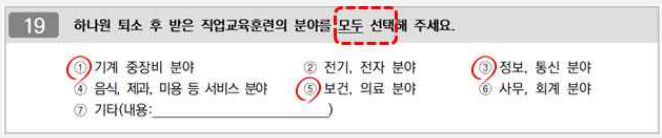
일련번호 (PID)		조사방법	① 직접 만나서 조사 ② 기타(_____)
조사장소	① 하나센터 ② 응답자 자택 ③ 그 외 장소(_____)		
조사 완료일자	2021년 _____월 _____일	조사 소요시간	시작 : ___시 ___분 완료 : ___시 ___분
상담자 성명	(인)	비고	

01 모든 질문에 대한 응답은 **보기번호(① ② ③ ④)에 동그라미(O)로** 해주세요.
표시한 응답을 바꾸고 싶을 때는 **가위표(X)를 하고, 바꾼 보기에 다시 동그라미로 표시** 해주세요.

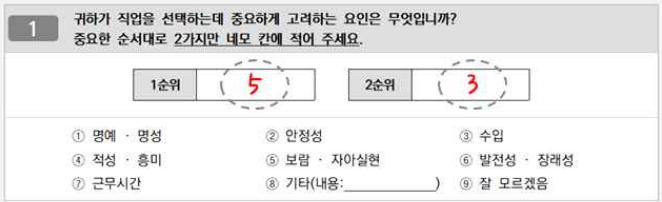
02 보기 중 '기타'에 응답하신 경우에는 **반드시 해당 내용을 괄호 안에 적어주세요.**



03 질문 중 '**모두 선택**' 또는 '**(복수응답)**'이 있는 경우에만, 해당하는 항목을 여러 개 선택할 수 있습니다.



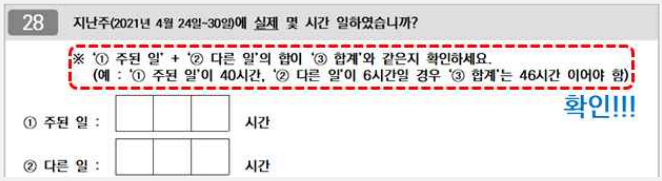
04 질문 중 **1순위, 2순위**를 묻는 질문이 있으면,
중요한 순서대로 2가지만 선택해서 **네모 칸**에 번호를 적어주세요.



05 응답을 선택한 후 **화살표**에 따라 반드시 **해당번호로** 이동해서 응답하세요.



06 밑줄, **당구장 표시(※)또는 점선 박스(⋯⋯) 부분은** 응답하기 전에, **꼭 읽어주세요.**



2021년

북한이탈주민 정착실태조사



인적 정보

1 성별

① 남성 ② 여성

2 생년월

□□□□년 □□월

3 귀하의 남한에서 최종 학력은 무엇입니까?

※ 남한에서 졸업한 학교 기준으로 응답, 북한 학력 X
※ 중등 검정고시 합격 = ② 중학교 졸업
※ 현재 대학생 = ③ 고등학교 졸업

① 초등학교 졸업
② 중학교 졸업
③ 고등학교 졸업
④ 대학(교)(4년제 미만) 졸업
⑤ 대학교(4년제 이상) 졸업
⑥ 대학원(석사) 졸업
⑦ 대학원(박사) 졸업
⑧ 남한에서 학교를 졸업한 적 없음

I 입국 전 사항

4 다음 중 탈북을 결심하게 된 주된 이유를 중요한 순서대로 2가지만 네모 칸에 적어주세요.

※ 여러 번 탈북하신 경우에는 최초에 탈북하셨을 때를 기준으로 응답해 주세요.

1순위		2순위	

① 가족(자녀 등)에게 더 나은 생활환경을 주려고
② 돈을 더 많이 벌고 싶어서
③ 식량이 부족해서
④ 신변의 위협을 느껴서
⑤ 먼저 탈북한 가족을 찾거나 함께 살기 위해
⑥ 먼저 탈북한 사람(가족 제외)이 권유해서
⑦ 북한에 함께 있던 가족을 따라서
⑧ 주변 사람(가족 제외)의 권유로
⑨ 북한 체제의 감시·통제가 싫어서(자유를 찾아서)
⑩ 기타(내용 :_____)

5 (최종)탈북 시기는 언제입니까?

□□□□년 □□월

6 북한에 있을 당시 가졌던 주된 직업을 1가지만 선택해 주세요.

※ 북한에서 주로 했던 일에 대해 네모 칸 안에 자세히 적고, 아래의 보기 중 해당하는 번호를 하나 선택해 주세요.

내가 북한에서 주로 했던 일(자세히) :

① 유급 당일군 또는 행정일군
② 군인
③ 노동자, 봉사원
④ 사무원
⑤ 교원, 연구사
⑥ 농장원(협동농장/국영농장)
⑦ 해외파견 일군
⑧ 기사(농업기사, 공업기사, 광업기사)
⑨ 의사, 약사
⑩ 사법검찰, 보위·안전원
⑪ 문화예술 체육인
⑫ 장사
⑬ 연로보장(은퇴자), 사회보장
⑭ 주부(부양)
⑮ 학생
⑯ 미취학아동
⑰ 기타(내용 :_____)

7 제3국(중국, 러시아, 베트남 등)에서 거주하는 동안 가졌던 주된 직업 1가지만 선택해 주세요.

① 제3국에서 거주한 적 없음
② 제3국에서 거주했지만, 직업을 가진 적은 없음
③ 서비스직(식당일, 가사도우미 등)
④ 농업·축산 관련 노동자(농장, 목장 등에서 일함)
⑤ 노동자(생산직, 기술직, 벌목업, 건축업 등)
⑥ 사무원(가이드·통역 포함)
⑦ 장사
⑧ 종교관련 직업
⑨ 기타(내용 :_____)

Ⅱ 생활안정

8 **현재 귀하는 주로 어디에서 살고 있습니까?**

① 본인 소유 집 (자기 소유)
② 하나원에서 배정받은 집 또는 임대아파트
③ 타인 소유 집 (전·월세)
④ 무연고 시설, 그룹홈, 쉼터 등 (주택 미배정자)
⑤ 기숙사, 고시원, 하숙집 등
⑥ 일정한 거주지 없이(여관 등) 자주 옮겨 다님
⑦ 친척 또는 친구집에서 함께 살고 있음
⑧ 기타(내용 :_____)

9 **귀하는 남한에서의 생활에 대하여 어느 정도 만족하십니까?**

① 매우 불만족 ② 불만족하는 편임 ③ 보통 ④ 만족하는 편임 ⑤ 매우 만족

10 **남한에서의 생활에 만족하는 주된 이유를 중요한 순서대로 2가지만 네모 칸에 적어 주세요.**

1순위		2순위	

① 북한보다 경제적 여유가 있어서
② 자유로운 삶을 살 수 있어서
③ 가족과 함께 생활할 수 있어서
④ 주변 사람들의 인정을 받아서
⑤ 내가 일한 만큼 소득을 얻을 수 있어서
⑥ 내가 하고 싶은 일을 할 수 있어서(자아실현)
⑦ 자녀에게 더 좋은 미래를 줄 수 있어서
⑧ 기타(내용 :_____)

11 **남한에서의 생활에 불만족하는 주된 이유를 중요한 순서대로 2가지만 네모 칸에 적어 주세요.**

1순위		2순위	

① 경제적으로 어려워서
② 남한 사회에 적응이 어려워서
③ 가족과 떨어져 살아야 해서(가족이 북한, 제3국에 있음)
④ 주변 사람(가족, 이웃 등)과 갈등 때문에
⑤ 북한이탈주민에 대한 남한사회의 차별/편견 때문에
⑥ 내가 하고 싶은 일을 하지 못해서
⑦ 경쟁이 너무 치열해서
⑧ 기타(내용 :_____)

12 **북한에서(재북시) 귀하의 사회경제적 지위(소득, 직업, 교육, 재산 등을 고려)는 어느 정도였다고 생각하십니까?**

① 최상층 ② 상층 ③ 중상층 ④ 중하층 ⑤ 하층 ⑥ 최하층

13 **현재 남한에서 귀하의 사회경제적 지위(소득, 직업, 교육, 재산 등을 고려)는 어느 정도라고 생각하십니까?**

① 최상층 ② 상층 ③ 중상층 ④ 중하층 ⑤ 하층 ⑥ 최하층

14 **귀하는 지금까지 인생에서 사회·경제적으로 성취한 것에 대해 어느 정도 만족하십니까?**

① 매우 불만족 ② 불만족하는 편임 ③ 보통 ④ 만족하는 편임 ⑤ 매우 만족

15 **남한 사회에서 노력한다면 개인의 사회·경제적 지위가 높아질 가능성은 어느 정도라고 생각하십니까?**

① 매우 낮음 ② 비교적 낮음 ③ 비교적 높음 ④ 매우 높음 ⑤ 잘 모르겠음

16 **남한 사회에서 현재 자신의 세대 보다 자식 세대의 사회·경제적 지위가 높아질 가능성은 어느 정도라고 생각하십니까?**

① 매우 낮음 ② 비교적 낮음 ③ 비교적 높음 ④ 매우 높음 ⑤ 잘 모르겠음

17 **귀하의 더 나은 남한생활을 위해 필요한 지원이 무엇인지 중요한 순서대로 2가지만 네모 칸에 적어주세요.**

1순위		2순위	

① 취·창업지원(취업알선, 취업교육 등)
② 교육지원(본인 또는 자녀의 학습관련 등)
③ 법률지원(가족관계, 형사사건 등)
④ 심리 상담 지원
⑤ 의료지원(병원비, 치료비 지원)
⑥ 요양시설 이용 지원(간병인 파견, 요양원 이용 등)
⑦ 주택문제(배정된 주택 상태, 동네 등) 관련 지원
⑧ (자녀 등) 보육 관련 지원
⑨ 남한주민과 교류 확대 지원
⑩ 북한이탈주민간 교류 확대 지원
⑪ 소득 지원(장려금, 생활비 보조 등)
⑫ 기타(내용 :_____)

Ⅲ 직업교육훈련

18 **하나원 퇴소 후에 학원, 기관 등에서 직업교육훈련을 받은 적이 있거나, 또는 현재 받고 있습니까?**

※ 제2하나원(화천)에서 받은 직업교육훈련도 포함해 주세요.

① 받은 적 있음(또는 현재 받고 있음) → 19번으로 ② 받은 적 없음 → 22번으로

19 **하나원 퇴소 후 받은 직업교육훈련의 분야를 모두 선택해 주세요.**

① 기계, 중장비 분야 ② 전기, 전자 분야 ③ 정보, 통신 분야
④ 음식, 제과, 미용 등 서비스 분야 ⑤ 보건, 의료 분야 ⑥ 사무, 회계 분야
⑦ 기타(내용:_____)

20 **하나원 퇴소 후 받은 직업교육훈련을 모두 수료하였습니까?**

① 직업교육훈련을 마쳤고 모두 수료하였음 → 21번으로
② 현재 처음으로 직업교육훈련을 받고 있음 → 22번으로
③ 직업교육훈련을 수료하지 못 한 적 있음

20-1 **수료하지 못한 이유를 모두 선택해 주세요.**

① 출석일수를 채우지 못해서(질병, 결석 등) ② 내용이 너무 어려워서
③ 적성에 맞지 않아서 ④ 취업에 도움이 될 것 같지 않아서
⑤ 기타(내용 :_____)

→ 21번으로

21 **[직업교육훈련을 하나라도 수료한 경우에만 응답]**
직업교육훈련을 수료한 분야에서 일하신 적이 있거나, 또는 현재 일하고 계십니까?

① 수료한 분야에서 일한 적 있음(또는 일하고 있음) → 22번으로
② 수료한 분야에서 일한 적 없음

21-1 **직업교육훈련을 수료하였음에도 해당 분야에서 일한 적이 없는 이유는 무엇입니까?**

※ '북한이탈주민이라 차별받아서' 등의 이유는 제외하고 응답해 주세요.

① 훈련받은 분야의 일자리가 없어서(일자리를 구하기 어려워서)
② 임금, 복지, 출퇴근거리 등 조건에 맞는 회사가 없어서
③ 다른 분야의 일을 하고 싶어서
④ 수료한 직업교육훈련의 내용이 취업에 도움이 되지 않아서
⑤ 직업훈련 장려금 수급을 목적으로 교육을 받아서
⑥ 기타(내용 :_____)

→ 22번으로

Ⅳ 일에 관한 사항

※ 문22번부터 16쪽 문47번까지의 응답 기준시점은 2021년 5월 1일입니다.

※ '지난주'는 5월 1일을 기준으로 지난주인 '2021년 4월 24일부터 4월 30일까지' 7일간을 의미합니다.

22 지난주(2021년 4월 24일~30일)에 '수입을 목적으로' 1시간 이상 일하였습니까? (일당·시간당 아르바이트 포함)

① 일하였음 → 25번으로 ② 일하지 않았음 → 23번으로

23 지난주(2021년 4월 24일~30일)에 함께 사는 가족의 사업체 또는 농장에서 보수를 받지 않고 1시간 이상 일하였습니까?

① 일한 적 있음 → 25번으로 ② 일한 적 없음 → 24번으로

24 지난주(2021년 4월 24일~30일)에 일을 하지는 않았지만 직장(일)을 가지고 있었습니까?

① 직장(일)을 갖고 있었음 → **24-1** 지난주에 일을 하지 않은 이유는 무엇입니까? (1개만 선택)

② 직장(일)이 없었음 → 17쪽 48번으로

① 일시적 병, 사고
② 휴가·연차(연가)
③ 교육·훈련
④ 육아
⑤ 가족적 이유(집안의 경조사, 가족간병)
⑥ 노사분규(파업 등)
⑦ 사업부진·조업중단(코로나19 여파 등)
→ 11쪽 26번으로

⑧ 기타(내용 : _____)
→ 17쪽 48번으로

25 지난주(2021년 4월 24일~30일)에 주된 일 이외에 다른 일(아르바이트 포함)을 한 적이 있습니까?

※ 여러 가지 일을 하는 경우, 적어도 한 개 이상 임금근로자로서 일을 하는 경우에만 주된 일(주업)과 다른 일(부업)을 구분함

주된 일 : 주된 직업, 본업, 조사대상 주간의 취업시간을 기준으로 가장 많이 한 일
- 매일 일하는 장소와 고용주가 바뀌는 경우는 모든 일을 포함하여 주된 일 1가지만을 하는 것으로 봄 (가사서비스 종사자, 건설현장 일용근로자 등)

다른 일 : 주된 일 외에 따로 하는 일
- 낮에는 기업체에서 사무직으로 일하고(주된 일), 퇴근 후 다단계 판매나 편의점 아르바이트 등(다른 일)을 함

① 있었음 ② 없었음

26 평소 1주일 동안 총 몇 시간 일합니까?

※ 총 시간은 '주된 일'과 '주된 일 이외에 다른 일'의 합계임
※ 일주일 평균 3일 근무, 하루 평균 8시간 일한다고 하면, 3일 X 8시간 = 24시간

총 [ ][ ] 시간

→ ① 0 ~ 35시간 → 27번으로
② 36시간 이상 → 28번으로

27 평소 1주일에 36시간 보다 적게 일하는 주된 이유는 무엇입니까?

① 정규 근무시간이 36시간 미만이라서
② 건강 때문에
③ 육아 때문에
④ 가사(집안일) 때문에
⑤ 학교나 학원을 다니기 때문에
⑥ 내가 원해서
⑦ 평소에 일거리가 없어서
⑧ 기타(내용 :_____)

28 지난주(2021년 4월 24일~30일)에 실제 몇 시간 일하였습니까?

※ '① 주된 일' + '② 다른 일'의 합이 '③ 합계'와 같은지 확인하세요.
(예 : '① 주된 일'이 40시간, '② 다른 일'이 6시간일 경우 '③ 합계'는 46시간이어야 함)

① 주된 일 : [ ][ ][ ] 시간
② 다른 일 : [ ][ ][ ] 시간
③ 합계 : [ ][ ][ ] 시간

→ 1) 1~17시간 일한 무급가족종사자 → 17쪽 48번으로
2) 0~35시간 → 29번으로
3) 36시간 이상 → 13쪽 33번으로

29 지난주(2021년 4월 24일~30일)에 36시간 보다 적게 일했던 주된 이유는 무엇입니까?

① 평소 1주일에 36시간 보다 적게 일하므로
② 일시적 병, 사고
③ 날씨가 좋지 않아서
④ 휴가·연가, 공휴일
⑤ 교육·훈련
⑥ 육아
⑦ 가족적 이유(집안의 경조사, 가족간병)
⑧ 노사분규(파업 등)
⑨ 일시적으로 일거리가 없어서
⑩ 사업부진·조업중단
⑪ 기타(내용 :_____)

30 지난주(2021년 4월 24일~30일)에 더 많은 시간 일하기를 원하였습니까?

① 하고 있는 일의 시간을 늘리고 싶었음
② 하고 있는 일 이외에 다른 일도 하고 싶었음
③ 더 많은 시간 일할 수 있는 일(직장)로 바꾸고 싶었음
→ 31번으로

④ 계속 그대로 일하고 싶었음 → 13쪽 33번으로

31 만약 지난주(2021년 4월 24일~30일)에 일이 더 주어졌거나, 더 많은 시간 일할 수 있는 다른 일(직장)이 주어졌다면, 할 수 있었습니까?

① 있었음 ② 없었음

32 지난 4주(2021년 4월 1일~30일)동안, 더 많은 시간 일하기 위해 현재 하는 일을 더 찾아보거나 다른 일(직장)을 찾아 본 적이 있습니까?

① 있었음 ② 없었음

※ 문33 ~ 47번은 지난주(2021년 4월 24일 ~ 4월 30일까지 7일간)의 '주된 일'을 기준으로 작성합니다.
주된 일 : 수입을 목적으로 주로 하는 일. 조사대상기간에 취업시간이 가장 많은 일

33 지난주(2021년 4월 24일~30일)에 주된 일을 한 직장(사업체)에 대해 말씀해 주세요.

사업체(직장) 명 :

※ 보기카드 참고하여 코드 기입
산업분류코드

사업체(직장)가 주로 하는 일 :

※ 본인의 업무가 아닌 회사의 사업 유형을 적어주세요. (아웃소싱, 휴대폰 액정 생산, 중국 음식점, 볼링장 운영 등)

34 지난주(2021년 4월 24일~30일)에 주된 일을 한 직장(사업체)의 직원 수는 얼마나 됩니까?

① 1 ~ 4명 ② 5 ~ 9명 ③ 10 ~ 29명 ④ 30 ~ 99명
⑤ 100 ~ 299명 ⑥ 300 ~ 499명 ⑦ 500명 이상

35 지난주(2021년 4월 24일~30일)에 주된 일을 한 직장(사업체)에서 귀하는 주로 무슨 일을 하였습니까?

내가 회사에서 주로 하는 일 :

※ 보기카드 참고하여 코드 기입
직업분류코드

나의 직명(직위) :

36 지난주(2021년 4월 24일~30일)에 주된 일을 한 직장(사업체)은 언제부터 다니기 시작했습니까?

				년			월부터 다니기 시작함

37 **지난주(2021년 4월 24일~30일)에 주된 일을 한 직장(사업체)에서 귀하의 종사상 지위(근로형태)는 무엇입니까?**

① 상용근로자(정규직, 무기계약직, 계약기간 1년 이상)
② 임시근로자(계약기간 1개월 이상 1년 미만)
③ 일용근로자
(계약기간 1개월 미만 또는 일일단위로 고용되어 일급 또는 일당제 급여를 받는 사람)

→ 14쪽 38번으로

④ 고용원이 있는 자영업자
(유급 고용원을 1인 이상 고용하고 있는 자영업자)
⑤ 고용원이 없는 자영업자
(유급 고용원 없이 자기 혼자 또는 무급가족종사자와 함께 일하는 사람)

→ 15쪽 43번으로

⑥ 무급가족종사자
(자영업자의 가족, 친인척으로 임금을 받지 않고 1주일에 18시간 이상 일하는 사람)

→ 23쪽 사회통합조사로 이동

38 지난주(2021년 4월 24일~30일)에 주된 일을 한 직장(사업체)에서 고용계약기간을 정하였습니까?

① 정하였음 → **38-1** 고용계약기간은 얼마입니까?

① 1개월 미만 ② 1개월 이상 ~ 6개월 미만
③ 6개월 이상 ~ 1년 미만 ④ 1년 이상 ~ 2년 미만
⑤ 2년 이상 ~ 3년 미만 ⑥ 3년 이상

② 정하지 않았음(정년제 포함) → 39번으로

39 지난 3개월(2021년 2월 ~ 4월)동안 세금을 내기 전의 귀하의 월평균 임금은 얼마입니까?

※ '반드시' 숫자로 기록하세요. (예 : 150)

귀하의 세전 월평균 임금 [ ] 만원

40 지난주(2021년 4월 24일~30일)에 다음의 사회보험에 가입되어 있었습니까?

40-1 국민연금 또는 특수직역연금

① 직장가입자 ② 지역가입자 ③ 가입하지 않음(수급권자 포함)

40-2 건강보험

① 직장가입자 ② 지역가입자 ③ 의료수급권자
④ 직장가입피부양자 ⑤ 가입하지 않음

40-3 고용보험(공무원, 사립학교 교직원, 별정우체국 직원은 응답하지 않고 문41번으로)

① 가입함 ② 가입하지 않음

41 **지난주(2021년 4월 24일~30일)에 주된 일을 한 직장(사업체)을 구하기 위해 정보를 알아볼 때, 어떤 경로로 구하였습니까? 최대 세 가지만 응답해 주세요.**

① 남북하나재단과 하나센터(전문상담사)
② 고용노동부 고용센터(워크넷, 고용복지플러스센터 등)
③ 여성새로일하기센터
④ 시청·군청·구청의 일자리 센터
⑤ 민간 취업알선기관 (알바천국, 잡코리아, 직업소개소 등)
⑥ 대중매체(인터넷 검색, TV, 신문, 잡지, 광고 등)
⑦ 학교, 학원의 추천, 소개, 의뢰
⑧ 북한이탈주민 친척, 친구, 동료
⑨ 북한이탈주민이 아닌 친척, 친구, 동료
⑩ 기타(내용 :_____)

42 **지난주(2021년 4월 24일~30일)에 주된 일을 한 직장(사업체)을 구하기 위해 구직(직장을 구하는) 활동을 얼마동안 하셨습니까?**

[ ] 년 [ ][ ] 개월

→ 23쪽 사회통합조사로 이동

※ 문43~47번은 '자영업자' 대상 문항입니다. (문37번 이동 로직 참고)

43 **현재의 사업을 처음 시작하였을 때, 필요한 사업자금의 조달방법은 주로 어떤 것이었습니까? 조달 자금의 크기 순으로 응답해 주세요.**

※ 조달방법이 2가지 이상인 경우 2순위까지 응답해 주세요.

1순위		2순위	

① 본인 또는 가족이 마련한 목돈(적금, 퇴직금, 명퇴금 등)
② 친지 또는 동업자의 자금
③ 친지 또는 동업자 이외 타인에게 빌려서 마련
④ 은행, 보험회사, 상호신용금고 등
⑤ 사채, 대부회사 등
⑥ 정부의 보조 또는 지원 등
⑦ 별도의 자본이 필요 없었음(사업승계, 무자본창업 등)
⑧ 기타(내용 :_____)

44 **현재의 사업을 처음 시작했을 때 사업자금의 규모는 어느 정도였습니까?**

① 5백만원 미만
② 5백만원 ~ 2천만원 미만
③ 2천만원 ~ 5천만원 미만
④ 5천만원 ~ 1억원 미만
⑤ 1억원 ~ 3억원 미만
⑥ 3억원 이상

45 **현재의 사업을 처음 시작하기까지 걸린 준비기간은 어느 정도였습니까?**

※ 사업 시작을 결심하면서부터 실제로 사업하기까지 소요된 시간

① 1개월 이상 ~ 3개월 미만
② 3개월 이상 ~ 6개월 미만
③ 6개월 이상 ~ 1년 미만
④ 1년 이상

46 **현재 사업을 시작하게 된 주된 동기는 무엇이었습니까?**

① 임금근로자로 취업이 어려워서
② 자신만의 사업을 직접 경영하고 싶어서
③ 기타(내용 :_____)

47 **현재 사업을 시작할 때, 가장 어려웠던 점은 무엇이었습니까?**

① 사업관련 인·허가, 지원기관 협조 등
② 사업 자금의 조달
③ 입지 선정, 기술개발 및 시설 확보
④ 필요인력의 확보
⑤ 사업정보 경영노하우 습득
⑥ 판매처 확보 및 홍보
⑦ 기타(내용 :_____)

→ 23쪽 사회통합조사로 이동

V 구직에 관한 사항

※ 문48번부터 19쪽 문58번까지의 응답 기준시점은 2021년 5월 1일입니다.

※ '지난주'는 5월 1일을 기준으로 지난주인 '2021년 4월 24일부터 4월 30일까지' 7일간을 의미합니다.

※ '지난 4주'는 5월 1일을 기준으로 4주 전인 '2021년 4월 1일부터 4월 30일까지' 30일간을 의미합니다.

48 지난 4주(2021년 4월 1일~30일) 내에 직장(일)을 구해 보았습니까?

※ 취업예정자(1개월 안으로 취업이 예정된 사람)는 '구해 보았음'에 해당

① 구해 보았음 → 49번으로 ② 구해 보지 않았음 → 52번으로

49 지난주(2021년 4월 24일~30일)에 직장(일)이 주어졌다면 일을 시작할 수 있었습니까?

① 있었음 → 50번으로 ② 없었음 → 19쪽 56번으로

50 직장(사업체)을 구하기 위해 정보를 알아볼 때, 어떤 경로로 구하였습니까?
최대 세 가지만 응답해 주세요.

① 남북하나재단과 하나센터(전문상담사)
② 고용노동부 고용센터(워크넷, 고용복지플러스센터 등)
③ 여성새로일하기센터
④ 시청·군청·구청의 일자리 센터
⑤ 민간 취업알선기관(알바천국, 잡코리아, 직업소개소 등)
⑥ 대중매체(인터넷 검색, TV, 신문, 잡지, 광고 등)
⑦ 학교, 학원의 추천, 소개, 의뢰
⑧ 북한이탈주민 친척, 친구, 동료
⑨ 북한이탈주민이 아닌 친척, 친구, 동료
⑩ 기타(내용 :_____)

51 직장(일)을 구하기 위해 구직(직장을 구하는)활동을 얼마 동안 계속 해왔습니까?

[ ][ ] 개월

→ 19쪽 57번(이전 직장(일)에 관한 사항)으로

Ⅵ 기타 활동에 관한 사항

※ '지난주'는 5월 1일을 기준으로 지난주인 '2021년 4월 24일부터 4월 30일까지' 7일간을 의미합니다.

※ '지난 4주'는 5월 1일을 기준으로 4주 전인 '2021년 4월 1일부터 4월 30일까지' 30일간을 의미합니다.

52 지난주(2021년 4월 24일~30일)에 직장(일)을 원하였습니까?

① 원하였음 → 53번으로 ② 원하지 않았음 → 19쪽 56번으로

53 지난주(2021년 4월 24일~30일)에 직장(일)이 주어졌다면 일을 시작할 수 있었습니까?

① 있었음 ② 없었음

54 지난 4주(2021년 4월 1일~30일) 내에 직장(일)을 구하지 않은 주된 이유는 무엇입니까?

① 전공이나 경력에 맞는 일거리가 없을 것 같아서
② 원하는 임금수준이나 근로조건이 맞는 일거리가 없을 것 같아서
③ 근처(주변)에 일거리가 없을 것 같아서
④ 교육, 기술, 경험이 부족해서
⑤ 나이가 어리거나 많다고 고용주가 생각할 것 같아서
⑥ 이전에 찾아보았지만 일거리가 없었기 때문에
⑦ 육아 때문에
⑧ 가사(집안일) 때문에
⑨ 학교나 학원을 다니기 때문에
⑩ 심신장애(건강 문제) 때문에
⑪ 기타(내용 :_____)

55 지난 1년 동안(2020. 5. 1 ~ 2021. 4. 30) 구직(직장을 구하는)활동을 해본 적이 있습니까?

① 있었음 ② 없었음 → 19쪽 56번으로

55-1 지난 1년 동안(2020. 5. 1 ~ 2021. 4. 30) 마지막으로 구직(직장을 구하는)활동을 한 시기는 언제입니까?

2	0	2		년			월

→ 19쪽 56번으로

56 지난주(2021년 4월 24일~30일)에 주로 무엇을 하였습니까?

① 육아
② 가사(집안일)
③ 정규교육기관 통학
④ 입시학원 통학
⑤ 취업을 위한 학원·기관 통학(고시학원, 직업훈련기관 등)
⑥ 취업준비
⑦ 진학준비
⑧ 연로
⑨ 심신장애(건강문제)
⑩ 군 입대 대기
⑪ 쉬었음
⑫ 기타(내용 :_____)

→ 57번으로

Ⅶ 이전 직장(일)에 관한 사항

57 이전에 수입을 목적으로 일한 적이 있습니까?

① 있었음 → ② 없었음 → 23쪽 사회통합조사로 이동

57-1 그 일을 그만둔 지는 얼마나 되었습니까?

① 1년 미만 → 그만 둔 시기 : (2020.5.1.~2021.4.30) 2020 년 [ ][ ] 월 → 58번으로

② 1년 이상 → 23쪽 사회통합조사로 이동

58 이전 직장(일)을 그만둔 주된 이유는 무엇입니까?

① 개인·가족적 이유
② 육아
③ 가사(집안일)
④ 심신장애(건강 문제)
⑤ 정년퇴직, 연로
⑥ 작업여건(시간·보수 등) 불만족
⑦ 직장의 휴업·폐업
⑧ 명예·조기퇴직, 정리해고
⑨ 임시 또는 계절적 일의 완료
⑩ 일거리가 없어서 또는 사업부진
⑪ 기타(내용 :_____)

→ 23쪽 사회통합조사로 이동

2021년
북한이탈주민
사회통합조사



Ⅰ 노동

1 귀하가 직업을 선택하는데 중요하게 고려하는 요인은 무엇입니까?
중요한 순서대로 2가지만 네모 칸에 적어 주세요.

1순위		2순위	

① 명예 · 명성　② 안정성　③ 수입
④ 적성 · 흥미　⑤ 보람 · 자아실현　⑥ 발전성 · 장래성
⑦ 근무시간　⑧ 기타(내용:_____)　⑨ 잘 모르겠음

2 귀하가 일하고 싶은 직장 혹은 업종은 무엇입니까?

① 국가기관·공기업　② 대기업　③ 중소기업
④ 외국계 기업　⑤ 해외 취업　⑥ 자영업(창업 등)
⑦ 영농업　⑧ 기타(내용:_____)

3 귀하는 여성 취업에 대하여 어떻게 생각하십니까?

① 직업을 가지는 것이 좋음 → **3-1**

② 취업보다 가정 일에 전념하는 것이 더 중요함
③ 잘 모르겠음 → 4번으로

3-1 어느 시기에 취업하는 것이 좋다고 생각하십니까?

① 결혼 전까지만
② 첫 자녀 출산 전까지만
③ 자녀가 성장한 후
④ 출산 전과 자녀 성장 후
⑤ 가정 일에 관계없이 계속 취업

4 귀하는 여성 취업을 어렵게 하는 가장 큰 요인이 무엇이라고 생각하십니까?
순서대로 2가지만 네모 칸에 적어 주세요.

1순위		2순위	

① 여성에 대한 사회적 편견 및 차별적 관행 ② 불평등한 근로여건(채용, 임금 등)
③ 구인 정보 부족 ④ 육아 부담
⑤ 가족 돌봄(육아 제외) ⑥ 가사 부담
⑦ 기타(내용:_____) ⑧ 잘 모르겠음

※ 문5~문7은 현재 직장이 있는 경우만 응답합니다. 현재 직장이 없으신 분은 문8번부터 응답해 주세요.

5 귀하는 현재 직장 일과 가정생활 중 어느 것을 우선시하고 있습니까?

① 일을 우선시함 ② 일을 우선시하는 편임
③ 일과 가정생활 둘 다 비슷함 ④ 가정생활을 우선시하는 편임
⑤ 가정생활을 우선시함

6 귀하는 평소 가까운 미래에 직장(직업)을 잃거나, 직장(직업)을 바꾸어야 한다는 불안함을 느끼십니까?

① 전혀 느끼지 않음 ② 느끼지 않는 편임 ③ 느끼는 편임 ④ 매우 많이 느낌

7 귀하가 다니는 직장의 근로여건에 어느 정도 만족하고 계십니까?
다음 각 항목마다 각각 해당되는 것을 선택해 주세요.

	매우 불만족	불만족하는 편임	보통	만족하는 편임	매우 만족	해당 없음
1) 하는 일(전반적인 일)	①	②	③	④	⑤	
2) 인사관리(승진, 배치, 이동)	①	②	③	④	⑤	⑥
3) 임금(일에 비해 받는 정도)	①	②	③	④	⑤	
4) 복리후생(기념일, 수당 등 복지)	①	②	③	④	⑤	⑥
5) 직장에서 나의 장래성	①	②	③	④	⑤	⑥
6) 근무환경(안전, 위생상태 등)	①	②	③	④	⑤	
7) 인간관계(상하, 동료 간)	①	②	③	④	⑤	⑥
8) 근로시간	①	②	③	④	⑤	
9) 일·가정 양립 직장 문화	①	②	③	④	⑤	⑥
10) 직장 내 폭력(언어폭력, 성희롱 등) 방지 노력	①	②	③	④	⑤	⑥
11) 직장 내 교육훈련의 기회	①	②	③	④	⑤	⑥
12) 북한이탈주민에 대한 동등한 대우	①	②	③	④	⑤	⑥
13) 전반적인 일자리 만족도	①	②	③	④	⑤	

Ⅱ 사회활동

8 귀하는 주로 무엇을 하며 여가를 활용하고 계십니까?
(주중과 주말(휴일 포함)로 구분하여 모두 응답해 주세요.)

※ 여가란 직업상의 일이나 필수적인 가사 활동 외에 소비하는 시간을 의미합니다.
의무적인 활동시간(식사, 수면, 출퇴근, 수업, 숙제, 집안일, 생필품구입)을 제외한 자유시간의 활동을 고려하여 응답해 주세요.

1) 주중	1순위		2) 주말 (휴일 포함)	1순위	
	2순위			2순위	

① TV시청(유선방송, IPTV, DVD(비디오) 시청 등)
② 문화 예술 관람(연극, 영화, 음악연주회, 미술전시회 등)
③ 문화 예술 참여(독서 토론, 미술 활동, 악기 연주, 사진 촬영 등)
④ 스포츠 관람
⑤ 스포츠 활동(축구, 테니스, 골프, 당구, 체조, 승마 등)
⑥ 관광 활동(국내, 해외 여행, 캠핑 등)
⑦ 컴퓨터 게임, 인터넷 검색 등
⑧ 취미, 자기개발 활동(요리, 독서, 기술, 자격증 취득 공부 등)
⑨ 휴식 활동(수면, 사우나 등)
⑩ 사회 및 기타 활동(봉사 활동, 종교 활동, 사교 활동 등)
⑪ 기타(내용:_____)

9 귀하는 주로 누구와 여가 활동을 하십니까?
(주중과 주말(휴일 포함)로 구분하여 각각 하나씩 응답해 주세요.)

※ 명수는 본인을 제외하고 응답합니다. (가족이 나 포함 4명이면, 3명으로 기록)

항목		명수 (본인 제외)	
1) 주중	① 가족	➡ 총 [ ] 명	➡ 이 중에 북한이탈주민 [ ] 명임
	② 친구(연인 포함)	➡ 총 [ ] 명	➡ 이 중에 북한이탈주민 [ ] 명임
	③ 동호회(종교단체 등 포함) 활동	➡ 총 [ ] 명	➡ 이 중에 북한이탈주민 [ ] 명임
	④ 직장 동료	➡ 총 [ ] 명	➡ 이 중에 북한이탈주민 [ ] 명임
	⑤ 혼자서		
	⑥ 기타(내용:_____)	➡ 총 [ ] 명	➡ 이 중에 북한이탈주민 [ ] 명임
2) 주말	① 가족	➡ 총 [ ] 명	➡ 이 중에 북한이탈주민 [ ] 명임
	② 친구(연인 포함)	➡ 총 [ ] 명	➡ 이 중에 북한이탈주민 [ ] 명임
	③ 동호회(종교단체 등 포함) 활동	➡ 총 [ ] 명	➡ 이 중에 북한이탈주민 [ ] 명임
	④ 직장 동료	➡ 총 [ ] 명	➡ 이 중에 북한이탈주민 [ ] 명임
	⑤ 혼자서		
	⑥ 기타(내용:_____)	➡ 총 [ ] 명	➡ 이 중에 북한이탈주민 [ ] 명임

10 귀하는 평소 여가생활에 대해 전반적으로 얼마나 만족하십니까?

① 매우 불만족
② 불만족하는 편임
③ 보통
④ 만족하는 편임
⑤ 매우 만족

→ 11번으로

10-1 만족하시지 못한다면 그 주된 이유는 무엇입니까?

① 경제적 부담 때문에
② 시간이 부족하여
③ 교통 혼잡 및 이용 교통수단이 불편하여
④ 여가 시설이 부족하므로
⑤ 여가 정보 및 프로그램이 부족하여
⑥ 적당한 취미가 없어서
⑦ 체력이나 건강이 좋지 않아서
⑧ 여가를 함께 즐길 사람이 없어서
⑨ 기타(내용:_____)

11 앞으로 시간적·경제적 여유가 생긴다면, 귀하는 여가 시간에 무엇을 하고 싶으십니까? 순서대로 2가지만 네모 칸에 적어주세요.

1순위		2순위	

① TV시청(유선방송, IPTV, DVD(비디오) 시청 등)
② 문화 예술 관람(연극, 영화, 음악연주회, 미술전시회 등)
③ 문화 예술 참여(독서 토론, 미술 활동, 악기 연주, 사진 촬영 등)
④ 스포츠 관람
⑤ 스포츠 활동(축구, 테니스, 골프, 당구, 체조, 승마 등)
⑥ 관광 활동(국내, 해외 여행, 캠핑 등)
⑦ 컴퓨터 게임, 인터넷 검색 등
⑧ 취미, 자기개발 활동(요리, 독서, 기술, 자격증 취득 공부 등)
⑨ 휴식 활동(수면, 사우나 등)
⑩ 사회 및 기타 활동(봉사 활동, 종교 활동, 사교 활동 등)
⑪ 기타(내용:_____)

12 귀하는 지난 1년 동안2020.05.01.~2021.04.30.) 사회단체에 참여하여 활동한 적이 있습니까?

① 있음
② 없음 → 13번으로

12-1 어떤 단체에 참여하셨습니까?
적극적으로 참여하신 순서대로 2가지만 네모 칸에 적어 주세요.

1순위		2순위	

① 친목 및 사교 단체(계, 동창회 등)
② 종교 단체(자비원, 선교원, 교리 연구회 등)
③ 취미, 스포츠 및 레저 단체
④ 시민 사회단체(환경 단체, 봉사 단체 등)
⑤ 학술 단체
⑥ 이익 단체(노조, 의사회, 약사회 등)
⑦ 정치 단체
⑧ 지역사회 모임(반상회, 입주자 대표회 등)
⑨ 기타(내용:_____)

13 귀하는 북한이탈주민이 운영하는 단체에 가입되어 있습니까?

① 가입되어 있음 → ② 가입되어 있지 않음 → 14번으로

13-1 활동에 실제로 어느 정도 참여하셨습니까?

① 적극 참여 ② 소극 참여 ③ 전혀 활동하지 않음

14 귀하는 지난 1년 동안(2020.05.01.~2021.04.30.) 자원봉사활동에 참여하신 적이 있습니까?

※ 사람들이 보통 보수(수입)라고 생각하지 않을 정도의 점심값, 교통비 등 실제 비용만 받고 봉사하는 경우도 포함됩니다.

① 있음 → ② 없음 → 15번으로

14-1 있다면 각 활동분야의 참여여부를 모두 표시해 주세요.(지난 1년(2020.05.01.~2021.04.30.) 기준)

활동분야	참여 경험 있음	참여 경험 없음
1) 아동, 청소년, 노인, 장애인, 재소자 등과 관련 (복지시설, 병원, 개인 가정 봉사, 방과후 교실 등)	①	②
2) 환경 보전, 범죄 예방 등과 관련(마을 청소, 방범 활동 등)	①	②
3) 자녀 교육 등과 관련(일일 교사, 교통 지도 등)	①	②
4) 국가 및 지역 행사 등과 관련 (스포츠, 문화 행사 등에서의 안내, 통역 등)	①	②
5) 재해 지역 주민 돕기 및 시설 복구 등과 관련 (해외 구호 포함)	①	②
6) 기타 일반인을 위한 봉사 관련 (관공서 업무 지원, 무료 상담 등)	①	②

15 귀하는 향후 1년 이내에 자원봉사 활동을 하실 의사가 있으십니까?

① 있음 ② 없음

Ⅲ 복지

16 귀하는 남한 사회의 생활 여건에 대해 어떻게 생각하십니까?

복지 분야	매우 불만족	불만족하는 편임	보통	만족하는 편임	매우 만족
1) 보건 의료 서비스	①	②	③	④	⑤
2) 사회보장제도 (건강보험, 사회복지서비스, 기초생활보장제도 등)	①	②	③	④	⑤
3) 문화·여가 생활 향유 여건	①	②	③	④	⑤

17 귀하는 지난 1달 동안(2021.04.01.~2021.04.30.) 다음의 경제적 어려움을 경험한 적이 있습니까?

※ 지난 1달 동안 남한에서 경험하신 것만 기준으로 응답해 주세요. (북한 및 제3국에서의 경험은 제외)

복지 분야	경험 있음	경험 없음
1) 돈이 없어서 끼니를 거른 적이 있다	①	②
2) 병원비가 부담되어 진료를 받지 못한 적이 있다	①	②
3) 건강보험료, 전기/전화/수도요금 등의 공과금을 내야하는 날짜까지 내지 못한 적이 있다	①	②

18 귀하는 노후를 위하여 준비를 하고 계십니까?

① 준비하고(되어) 있음 → 18-1

② 준비하고(되어) 있지 않음 → 18-2

18-1 준비하고 있다면 어떤 준비를 하고 계십니까? 순서대로 2가지만 네모 칸에 적어 주세요.

1순위		2순위	

① 국민연금
② 기타 공적 연금(공무원, 군인, 교원)
③ 사적 연금(은행, 보험회사 등을 통한 개인연금)
④ 퇴직급여(퇴직금, 퇴직연금)
⑤ 예금, 적금, 저축성 보험
⑥ 부동산 운용
⑦ 주식, 채권 등
⑧ 기타(내용:_____)

→ 19번으로

18-2 준비하지 않는 주된 이유는 무엇입니까?

① 아직 생각하고 있지 않음
② 앞으로 준비할 계획임
③ 준비할 능력이 없음
④ 자녀에게 의지함
⑤ 기타(내용:_____)

→ 19번으로

19 귀하는 노후 생활을 위해 사회적으로 주로 관심을 두어야 하는 것은 무엇이라고 생각합니까?

① 노후 소득 지원
② 노후 취업 지원
③ 의료 및 요양 보호 서비스
④ 문화, 여가 복지 서비스
⑤ 세대 간 이해 증진을 위한 교육 및 홍보
⑥ 교통 및 주거 환경 조성
⑨ 기타(내용:_____)

20 귀하는 노후를 주로 어떻게 보내고 싶으십니까?

① 취미 활동
② 학습 및 자아 개발 활동
③ 자원봉사 활동
④ 종교 활동
⑤ 소득 창출 활동(수입을 얻을 수 있는 활동)
⑥ 기타(내용:_____)

※ 만 60세가 안 되신 분(주민등록증에 양력 1961.5.1. 보다 생년월일이 늦거나 같으신 분)은 30쪽의 문22번으로 이동하세요.

※ 만 60세 이상(주민등록증에 양력 1961.5.1. 보다 생년월일이 빠르신 분)만 응답합니다.

21 귀하의 생활비는 주로 어떻게 마련하고 있습니까?

① 본인 및 배우자 부담 → 21-1번으로
② 자녀 또는 친척 지원
③ 정부 및 사회단체(정부 보조 등) → 22번으로
④ 기타(내용:_____)

21-1 생활비를 마련하는 구체적인 방법은 무엇입니까?

① 근로소득 및 사업소득
② 재산소득(이자, 임대소득, 배당소득 등)
③ 연금(개인연금 포함), 퇴직급여
④ 예금(적금)
⑤ 기타(내용:_____)

→ 22번으로

Ⅳ 보건 및 건강

22 귀하는 평소 자신의 건강상태에 대해 어떻게 생각하십니까?

① 매우 나쁨 ② 나쁜 편임 ③ 보통 ④ 좋은 편임 ⑤ 매우 좋음

23 귀하는 평소 다음 각 사항을 실천합니까?

	실천하지 않음	실천함
1) 아침 식사하기	①	②
2) 적정 수면(6~8시간)	①	②
3) 규칙적 운동	①	②
4) 정기 건강검진	①	②

24 귀하는 현재 담배를 피우십니까?

① 피움 ② 안 피움

25 귀하는 지난 1년 동안(2020.05.01.~2021.04.30.) 술을 한 잔 이상 마신 적이 있습니까?

① 마심 → **25-1** 얼마나 자주 드셨습니까?

① 월 1회 이하 ② 월 2~3회
③ 주 1~2회 ④ 주 3~4회
⑤ 거의 매일

② 안 마심

26 귀하는 지난 1년 동안(2020.05.01.~2021.04.30.) 의료기관을 이용한 적이 있습니까?

① 있음 → ② 없음 → 28번으로

26-1 아팠을 때 주로 이용한 의료기관은 어디입니까? (한 곳만 선택해 주세요) → 27번으로

① 종합병원(공공의료원 제외)
② 공공의료원
③ 병원(의원)
④ 치과 병원(의원)
⑤ 한의원(한방병원)
⑥ 보건소
⑦ 약국(한약국)

27 귀하는 이용하신 의료기관의 서비스에 대해서 어느 정도 만족하십니까?

① 매우 불만족
② 불만족하는 편임
③ 보통
④ 만족하는 편임
⑤ 매우 만족

→ 28번으로

27-1 의료 서비스에 대해 만족하지 못한 이유가 무엇입니까? 순서대로 2가지만 네모 칸에 적어 주세요.

1순위		2순위	

① 불친절함
② 의료비(약값 포함)가 높음
③ 치료 결과가 미흡함
④ 진료가 불성실함
⑤ 진료 대기 및 입원 대기 시간이 김
⑥ 의료 시설이 낙후되거나 미비함
⑦ 필요 이상으로 진료(검사 등)를 많이 함
⑧ 전문 의료 인력이 부족함
⑨ 집에서 의료 기관까지 거리가 멂
⑩ 기타(내용: _____)

28 귀하는 지난 2주 동안(2021.04.17.~04.30.) 질병이나 사고로 아팠던 적이 있습니까?

※ 고혈압, 당뇨 등 만성질환으로 계속 약을 복용한 경우도 '① 있음'에 표시해 주세요.

① 있음 → 29번으로 ② 없음 → 30번으로

29 귀하는 지난 2주 동안(2021.04.17.~04.30.) 의료기관에서 치료를 받거나 약을 복용한 적이 있습니까?

① 있음 → ② 없음

29-1 어느 기관에서 치료(또는 복용)를 받으셨습니까? 각각 며칠 치료(또는 복용) 받으셨습니까?

※ 응답 가능한 일 수는 최소 1일, 최대 14일까지입니다.
※ 병원 진료를 1일 받고, 약국에서 3일치의 약을 받아 복용한 경우 : ③ 병(의)원 1일, ⑦ 약국(한약국) 3일 응답
※ 병원이나 약국을 방문하지 않았지만, 약은 2주 동안 복용한 경우 : ⑦ 약국(한약국) 14일 응답

① 종합병원(공공의료원 제외) ····▶ [ ] 일
② 공공의료원 ····▶ [ ] 일
③ 병원(의원) ····▶ [ ] 일
④ 치과 병원(의원) ····▶ [ ] 일
⑤ 한의원(한방병원) ····▶ [ ] 일
⑥ 보건소 ····▶ [ ] 일
⑦ 약국(한약국) ····▶ [ ] 일

30 귀하는 지난 1년 동안(2020.05.01.~2021.04.30.) 병의원 진료 또는 검사를 받아 볼 필요가 있었으나 받지 못한 적이 한 번이라도 있었습니까? (※치과치료, 치과검사 제외)

① 예, 받지 못한 적이 한 번이라도 있었음 → ② 아니오, 받지 못한 적이 한 번도 없었음 → 31번으로

30-1 지난 1년 동안(2020.05.01.~2021.04.30.) 필요한 병의원 진료나 검사를 받지 못한 가장 중요한 이유는 무엇입니까?

① 경제적 이유(치료비용이 너무 많이 들어서)
② 병의원 등에 예약을 하기가 힘들어서
③ 교통편이 불편해서
④ 내가 갈 수 있는 시간에 병의원 등이 문을 열지 않아서
⑤ 증상이 가벼워서 필요성을 느끼지 못함
⑥ 어디로 가야할 지 잘 몰라서(정보 부족)
⑦ 기타(내용: _____)

31 귀하는 지난 2주 동안(2021.04.17.~04.30.) 일상생활에서 전반적으로 스트레스를 어느 정도 느끼셨습니까?

	전혀 느끼지 않음	느끼지 않는 편임	느끼는 편임	매우 많이 느낌	해당 없음
1) 가정생활	①	②	③	④	
2) 직장생활	①	②	③	④	⑤
3) 학교생활	①	②	③	④	⑤
4) 전반적인 일상생활	①	②	③	④	

32 귀하는 지난 1년 동안(2020.05.01.~2021.04.30.) 한 번이라도 자살하고 싶다는 생각을 해 본 적이 있습니까?

※ 단순히 죽고 싶다고 생각한 경우에는 자살에 대한 충동을 느낀 적이 '② 없음'에 응답하고, 구체적으로 죽을 수 있는 방법을 생각해 본 경우는 '① 있음'에 응답합니다.

① 있음 ② 없음 → 33번으로

32-1 자살하고 싶다는 생각을 했던 주된 이유는 무엇입니까?

① 경제적 어려움 때문에
② 이성 문제가 원만치 않아서(실연, 파혼 등)
③ 신체적·정신적 질환, 장애 때문에
④ 직장 문제 때문에(실직, 미취업 등)
⑤ 외로움, 고독 때문에
⑥ 가정불화로 인해
⑦ 학교 성적, 진학 문제 때문에
⑧ 친구나 동료들과의 불화 및 따돌림 때문에
⑨ 기타(내용:_____)

V 교육

33 귀하는 남한에서 대학(교)를 휴학 혹은 중퇴한 경험이 있습니까?

※ 남한에서 다닌 학교를 기준으로 응답합니다.
※ 대학(교)는 일반대학(4년제 이상), 전문대학, 사이버대학, 평생교육시설 등을 모두 포함합니다.

① 경험 있음 → 34번으로 ② 경험 없음 → 35번으로

34 귀하가 남한에서 대학(교)를 휴학 혹은 중퇴한 이유를 중요한 순서대로 2가지만 네모 칸에 적어주세요.

1순위		2순위	

① 학업(공부내용)을 따라가기 어려워서
② 학업(공부내용)이 적성에 안 맞아서
③ 취업준비를 해야 해서
④ 다른 학교로 편입·유학하기 위해서
⑤ 학비를 마련하기 어려워서
⑥ 생활비를 마련하기 위해서
⑦ 질병 또는 신체적 장애 때문에
⑧ 가사와 육아를 해야 해서
⑨ 북한(또는 제3국)의 가족, 친척에게 경제적 도움 또는 남한으로 데려오기 위한 비용 마련을 위하여
⑩ 기타(내용:_____)

35 귀하는 어느 정도까지 교육 받기를 원하십니까?

① 고등학교 이하 → 36번으로
② 대학(교)(4년제 미만)
③ 대학(교)(4년제 이상)
④ 대학원(석사)
⑤ 대학원(박사)
⑥ 없음 → 36번으로

35-1 귀하가 대학 이상 교육을 받고자 하는 주된 목적은 무엇입니까?

① 좋은 직업을 갖기 위해서
② 자신의 능력과 소질을 개발하기 위해서
③ 주위(부모 등)의 기대 때문에
④ 결혼, 친구관계 등 사회적으로 유리하기 때문에
⑤ 인격이나 교양을 쌓기 위해
⑥ 기타(내용:_____)

36 귀하는 현재 학생(휴학생·재수생 포함) 자녀가 있습니까?

※ 북한, 제3국에 거주하는 자녀, 재혼 배우자 자녀를 포함해서 응답해 주세요.

① 있음 → 36-1번으로

② 없음 → 39번으로

36-1 자녀는 어느 지역에 거주하고 있습니까?

※ 자녀가 여러 명인 경우, 모두 응답해 주세요.
※ 복수응답 중에 "① 남한"이 있으면 문37번으로 이동해 주세요.

① 남한 → 37번으로

② 북한 → 39번으로

③ 중국 → 39번으로

④ 기타(내용: _____)

37 귀하는 자녀를 어느 정도까지 교육시키기를 원하십니까?

※ 남한에 함께 있는 자녀를 기준으로 응답해 주세요.

① 고등학교 이하 → 39번으로

② 대학(교)(4년제 미만) → 37-1번으로

③ 대학(교)(4년제 이상) → 37-1번으로

④ 대학원(석사) → 37-1번으로

⑤ 대학원(박사) → 37-1번으로

⑥ 없음 → 39번으로

37-1 귀하의 자녀에게 대학 이상 교육을 시키고자 하는 주된 목적은 무엇입니까?

① 인격이나 교양을 쌓게 하기 위해서
② 좋은 직업을 갖게 하기 위해서
③ 자녀의 능력과 소질을 개발하기 위해서
④ 결혼, 친구관계 등 사회적으로 유리하기 때문에
⑤ 내가 교육을 받지 못한 것을 자녀를 통해 보상받고 싶어서
⑥ 기타(내용: _____)

38 귀댁의 자녀 교육비는 가구의 소득에 비하여 얼마나 부담되십니까?

※ 남한에 함께 있는 자녀를 기준으로 응답해 주세요.
※ 교육비는 학교납입금, 보충 교육비, 학원비, 교재비, 기숙사비 등 교육을 위해 지출되는 비용을 의미합니다.

① 전혀 부담스럽지 않음 → 39번으로

② 별로 부담스럽지 않음 → 39번으로

③ 보통 → 39번으로

④ 약간 부담스러움 → 38-1번으로

⑤ 매우 부담스러움 → 38-1번으로

38-1 어느 부분이 가장 부담스럽습니까?

① 학교 납입금(수업료, 학교 운영 지원비, 학교 보충 교육비 등)
② 보충 교육비(학원비, 개인 및 그룹 과외비 등)
③ 하숙·자취·기숙사비
④ 기타(내용: _____)

Ⅵ 가족

39 귀하는 가구주와 어떤 관계입니까?

※ 가구주는 세대주와는 관계없이 가계를 실제로 책임지고 있는 생계 책임자를 의미합니다.

① 가구주 본인
② 배우자
③ 미혼 자녀
④ 기혼 자녀 및 그 배우자
⑤ 손자·손녀 및 그 배우자
⑥ 부모(배우자 쪽 포함)
⑦ 조부모(배우자 쪽 포함)
⑧ 미혼 형제자매(배우자 쪽 포함)
⑨ 기타 친인척
⑩ 기타 동거인

40 아래 가구 관련 질문에 답변해 주세요.

항목	명수
1) 현재 남한에서 생계를 같이 하는 가구원 수는 본인을 포함하여 총 몇 명입니까?	□□ 명
2) 위에 응답한 가구원 중 만 65세 이상인 가구원은 몇 명입니까? (1956.04.30. 이전 출생자)	□□ 명
3) 위에 응답한 가구원 중 만 15세 이하인 가구원은 몇 명입니까? (2006.04.30. 이후 출생자)	□□ 명
4) 귀하의 가구원 중 취업자(일하고 있는 사람) 수는 몇 명입니까?	□□ 명

41 귀하는 현재 배우자가 있습니까?

※ 북한에 배우자가 있거나 사실혼 관계의 경우에도 '배우자 있음'에 해당합니다.

① 있음 ↓
② 없음 → 42번으로

41-1 배우자는 어느 지역 출신입니까?

① 남한 ② 북한 ③ 중국 ④ 기타(내용: _____)

41-2 배우자는 어느 지역에 거주하고 있습니까?

① 남한 ② 북한 ③ 중국 ④ 기타(내용: _____)

42 **현재 같이 살고 있는 가족을 모두 선택해 주세요.**

① 배우자
② 미혼 자녀
③ 기혼 자녀 및 그 배우자
④ 손자·손녀 및 그 배우자
⑤ 부모(배우자 쪽 포함)
⑥ 조부모(배우자 쪽 포함)
⑦ 미혼 형제자매(배우자 쪽 포함)
⑧ 기타 친인척
⑨ 기타(내용:_____)
⑩ 없음

※ 문43~47은 사회 구성원으로서 해당 주제에 대한 일반적인 생각을 표시해 주세요.

43 **부모님의 노후는 주로 누가 돌보아야 한다고 생각하십니까?**

① 스스로 해결
② 가족
③ 가족과 정부·사회
④ 정부·사회
⑤ 기타(내용:_____)

44 **귀하는 결혼에 대해 어떻게 생각하십니까?**

① 반드시 해야 함
② 하는 것이 좋음
③ 해도 좋고, 하지 않아도 좋음
④ 하지 않는 것이 좋음
⑤ 하지 말아야 함
⑥ 잘 모르겠음

45 **귀하는 이혼에 대해 어떻게 생각하십니까?**

① 어떤 이유라도 이혼해서는 안 됨
② 이유가 있더라도 가급적 이혼해서는 안 됨
③ 경우에 따라 이혼할 수도 있고, 하지 않을 수도 있음
④ 이유가 있으면 이혼을 하는 것이 좋음
⑤ 잘 모르겠음

46 **귀하는 재혼하는 것에 대해 어떻게 생각하십니까?**

① 반드시 해야 함
② 하는 것이 좋음
③ 해도 좋고, 하지 않아도 좋음
④ 하지 않는 것이 좋음
⑤ 하지 말아야 함
⑥ 잘 모르겠음

47 가사 분담에 대해 어떻게 생각하십니까?

① 아내가 전적으로 책임져야 함
② 아내가 주로 하고 남편도 분담해야 함
③ 공평하게 분담해야 함
④ 남편이 주로 하고 아내도 분담해야 함
⑤ 남편이 전적으로 책임져야 함

※ 부부가 함께 사는 경우만 응답합니다. (주말부부 포함)

48 귀댁에서는 <u>평소</u> 가사분담을 어떻게 하고 계십니까?

① 아내가 전적으로 책임짐
② 아내가 주로 하고 남편도 분담함
③ 공평하게 분담함
④ 남편이 주로 하고 아내도 분담함
⑤ 남편이 전적으로 책임짐

49 귀하는 가족 관계에서 다음 각 항목에 대하여 어느 정도 만족하고 계십니까?
다음 항목마다 해당되는 곳에 ○표를 해 주세요.

	매우 불만족	불만족하는 편임	보통	만족하는 편임	매우 만족	해당 없음
1) 배우자와의 관계	①	②	③	④	⑤	⑥
2) 자녀와의 관계	①	②	③	④	⑤	⑥
3) 자기 부모와의 관계	①	②	③	④	⑤	⑥
4) 배우자 부모와의 관계	①	②	③	④	⑤	⑥
5) 자기 형제자매와의 관계	①	②	③	④	⑤	⑥
6) 배우자의 형제자매와의 관계	①	②	③	④	⑤	⑥
7) 전반적인 가족 관계	①	②	③	④	⑤	⑥

Ⅶ 사회적 관계망

50 귀하는 가족, 친척, 친구, 이웃, 직장 동료 등 전반적인 개인적 인간관계에 얼마나 만족하십니까?

① 매우 불만족
② 불만족하는 편임
③ 보통
④ 만족하는 편임
⑤ 매우 만족

51 귀하는 다음과 같은 상황이 발생했을 때,
도움을 받을 수 있는 사람(가족, 친척, 친구, 이웃, 직장 동료 등)이 몇 명 있습니까?

※ "도움을 받을 수 있는 사람"에는 가족, 친척, 친구, 이웃, 직장 동료 등이 포함됩니다.
단, 파출부 등 돈을 주고 고용한 사람, 사회복지단체에서 지원받은 경우는 포함되지 않습니다.

항목	유무	명수
1) 몸이 아파 집안일을 부탁할 경우	① 있음 →	총 [ ] 명 ↓ 이 중에 북한이탈주민 [ ] 명임
	② 없음	
2) 갑자기 많은 돈을 빌릴 일이 생길 경우	① 있음 →	총 [ ] 명 ↓ 이 중에 북한이탈주민 [ ] 명임
	② 없음	
3) 낙심하거나 우울해서 이야기 상대가 필요한 경우	① 있음 →	총 [ ] 명 ↓ 이 중에 북한이탈주민 [ ] 명임
	② 없음	

52 귀하는 평소에 대면, 인터넷(SNS), 전화 등의 방식으로 교류하는 사람이 있습니까?
있다면 하루 평균 몇 명의 사람과 교류하십니까?

※ 교류는 전화, 우편, 대면 등의 방식으로 다른 사람과 개인적으로 이야기를 나누는/의사소통하는 것을 말합니다.

항목	유무	명수
1) 가족 또는 친척 (같이 살고 있는 사람들은 제외)	① 있음 →	하루 평균 [ ] 명
	② 없음	
2) 가족 또는 친척 이외의 사람 (업무상 접촉은 제외)	① 있음 →	하루 평균 [ ] 명
	② 없음	

53 귀하는 지난 1년 동안(2020.05.01.~2021.04.30.) 북한 출신이라는 이유로 차별이나 무시를 당한 적이 있습니까?

※ 언어나 행동으로 구체적인 차별이나 무시를 당한 경우만 있다고 응답해 주세요.

① 있음 → ② 없음

53-1 귀하가 생각하기에 어떤 이유로 차별이나 무시를 당했다고 생각합니까?
순서대로 2가지만 네모 칸에 적어주세요.

1순위		2순위	

① 전문적 지식과 기술 등에 있어 남한사람에 비해 능력이 부족하다고 생각되어서
② 말투, 생활방식, 태도 등 문화적 소통방식이 다르다는 점에서
③ 남한사회에서 경제적 소득수준이 낮은 계층이라서
④ 북한의 호전적인 도발(예: 미사일, 핵실험, 군사교전 등)의 영향으로
⑤ 언론에서 북한체제와 북한이탈주민들에 대한 부정적 보도의 영향으로
⑥ 남한사람이 북한이탈주민의 존재에 대한 부정적 인식(세금부담 증가 등)을 가지고 있어서
⑦ 기타(내용: _____)

Ⅷ 가구경제상태 관련

※ 문54번부터 57번까지는 본인이 아닌, '가구 기준'으로 응답해 주세요.

54 **귀 가구의 지난 1년 동안(2020.05.01.~2021.04.30.) 월 평균 소비지출액을 응답해 주세요.**

천만	백만	십만	만

만원

※ 월 평균 소비 지출액이란?
1) 가구 내에서 1달 동안의 생활을 위해 사용되는 모든 지출의 평균치를 의미합니다.
2) 지출액 예시 : 주거비(월세 등), 식료품비(외식 포함), 교통/통신비, 교양오락비(영화관람 등의 문화생활)
3) 지출액 예시 : 보건의료비(병원진료, 약값 등), 피복비(옷, 신발 구입), 가구집기/가사용품(서랍장 구매 등)

55 **귀 가구에서 지출하는 생활비 중 가장 부담이 되는 항목은 무엇입니까? 1개만 응답해주세요.**

① 식비(주식비와 부식비)
② 교육비
③ 월세
④ 주거 관련비(주택관리비, 냉난방비, 수도비 등)
⑤ 보건의료비(건강보험료 제외)
⑥ 가구·집기·가사용품
⑦ 피복비(옷, 신발 등 구입비)
⑧ 교양오락비
⑨ 교통비(차량 구입비 제외)
⑩ 통신비(전화, 인터넷 사용료 등)
⑪ 경조사비(축의금, 조의금 등)
⑫ 부채상환(대출이자·원금 갚기)
⑬ 기타(내용:_____)

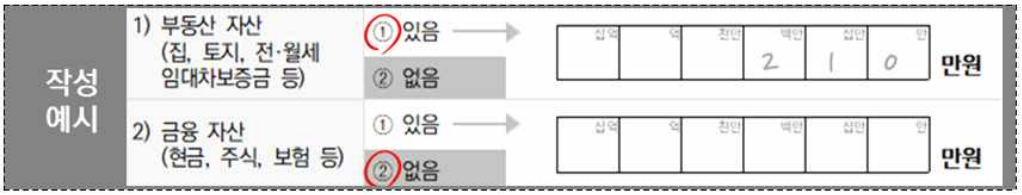
56 **귀 가구의 지난 1년 동안(2020.05.01.~2021.04.30.) 소득 금액(세금 및 사회보험료 공제 후)을 아래에 적어 주세요.**

항목	금액
1) 가구 내 총 소득(수입)	십억 / 억 / 천만 / 백만 / 십만 / 만 : 만원

※ 가구 내 총 소득(수입)에 포함되는 항목
1) 가구 내에서 직장(일) 또는 사업을 통해 벌어들이는 수입 : 직장급여(연봉), 사업수익(이윤) 등
2) 가구 내 보유하고 있는 자산(재산)으로부터 발생하는 소득 : 예금 이자, 임대료 등
3) 가구원, 친지, 지인 또는 후원자 등으로부터 정기적으로 지원받는 금품(수익) : 현금지원, 후원금 등
4) 정부 및 기관에서 지원하는 각종 수당 및 연금 : 정착지원금, 양육수당, 실업급여, 기초생활수급 등

※ 가구 내 총 소득(수입) 작성 방법
1) 본인 뿐만 아닌 가구 내 모든 구성원의 소득을 합산하여 적어 주십시오.
2) 지난 1년(2020년 5월부터 2021년 4월까지)의 소득분을 적어 주십시오.

57 현재 귀 가구의 자산(재산) 및 부채(빚, 대출) 현황을 아래의 항목별로 적어 주세요.



항목	자산 유무	금액
1) 부동산 자산 (집, 토지, 전·월세 임대차보증금 등)	① 있음 → / ② 없음	십억 억 천만 백만 십만 만 만원
2) 금융 자산 (현금, 주식, 보험 등)	① 있음 → / ② 없음	십억 억 천만 백만 십만 만 만원
3) 기타 자산(그외 자산) (빌려준 돈 등)	① 있음 → / ② 없음	십억 억 천만 백만 십만 만 만원
4) 부채(빚, 대출 등)	① 있음 → / ② 없음	십억 억 천만 백만 십만 만 만원

통일부 및 재단에 바라는 점을 적어주세요.

조사에 참여해 주셔서 대단히 감사합니다.

2021년 북한이탈주민실태조사
품질개선 컨설팅 최종결과보고서

발 행 일 2021년 6월
발 행 인 통계청장 류근관
발 행 처 통계청
대전시 서구 청사로 189



안 내

1. 이 보고서는 통계청에서 수행한 국가통계 품질개선 컨설팅 연구용역 결과보고서로 보고서의 내용을 발표 또는 인용할 때에는 반드시 올바른 인용 및 출처 표시 방법을 준수해야 합니다.

2. 이 보고서에 대한 지식재산권은 통계청에 있습니다.