

제3장 농가경제조사 표본개편에 따른 복합 이중추출 추정량 개발¹⁾

제3장



진 영 · 박중서

제1절 머리말

1. 연구개요

특정 집단의 어떤 특성을 알고 싶을 때, 전체에 대한 조사를 할 수도 있지만 시간, 비용, 신속성 등을 고려하여 특정 집단을 잘 대표할 수 있는 표본을 추출하여 표본조사를 실시하는 경우가 더욱 많다.

우리나라에서는 매 5년마다 실시되는 농림어업총조사를 목표모집단으로 다양한 농어업통계 표본조사를 실시한다. 농림어업총조사 조사시기와 각종 농어업통계 표본설계 실시시기는 약 2~3년 정도 차이가 있는데, 우리나라의 농촌 환경의 급변화로 표본설계 당시 농가인 가구가 실제 조사시점에서는 농가가 아닌 일반가구로 변경되어 표본교체가 많이 이루어진다. 또한, 표본설계 당시 농가의 주요 특성인 전겸업과 영농형태 등이 조사시점에서는 변동되는 경우가 다수 발생하여 농업통계의 단층현상이 심해질 수 있고, 이에 따른 농업동향 분석이 실제와 다르게 파악되거나 분석되어 현실과 괴리현상이 나타날 수 있다. 따라서, 최근 농촌 환경변화를 반영한 목표모집단을 확보하여 농어업

1) 2012년 농업통계 표본개편과 관련된 「복합 이중추출방법에 따른 추정량 연구」에 대한 통합보고서에서 두 번째로 진행되었던 연구로서 2012년 상반기(1~6월)에 연구했던 내용을 정리한 연구보고서임

조사 목적에 맞게 목표모집단을 분류하고 허용오차에 맞는 표본규모를 결정하여 대표성 높은 표본 농가를 추출하는 표본설계 및 추정방법의 연구가 필요하다.

2012년 농업통계의 표본설계 개편방향은 농업조사 자료를 주 표본(master sample)으로 하여 4종 관련 농업조사(농가경제조사, 농축산물생산비조사, 가축동향조사, 양곡소비량조사)의 표본 농가를 이중추출하여 2013년부터 각각의 통계조사를 실시할 예정이다.

이중추출방법은 중복추출의 한 방법으로 모집단의 사전정보가 부족하거나 정확한 표본추출틀이 없을 때 실시하는 표본설계 방법이다. 2012년부터 개편될 농업통계 표본 개편은 급격하게 변하는 농촌환경을 빠르게 파악하기 위하여 5종 농업통계의 교차분석을 할 수 있으며, 최근의 농촌 정보로 표본설계를 할 수 있는 주 표본에 의한 이중추출방법(two-phase sampling)을 고려하게 되었다. 따라서, 이중추출 표본설계와 추정방법에 대해 단계별로 체계적인 정리를 겸한 연구가 필요하다.

2. 연구목표 및 범위

본 연구의 궁극적인 목표는 2012년부터 점차적으로 개편될 농업통계 표본설계의 기본방향인 이중추출방법의 표본설계 효과를 측정하는 것으로서 복합 이중추출에 의한 추정량 모의 산출하여 추정량의 안정성과 효율성 등을 사전에 분석해 보는 것이 연구의 핵심이다. 본 연구에서는 농업통계의 핵심인 농가경제조사 자료를 이용하여 이중추출방법의 표본설계 효과를 연구하였다.

과거에는 0·5년도 농업총조사 자료를 모집단으로 관련 농업통계의 표본설계를 하였지만, 이번에는 매년 12월에 실시하는 농업조사를 주표본(마스터모집단)으로 농가경제조사, 농축산물생산비조사, 가축동향조사, 양곡소비량조사를 위한 표본농가를 추출하는 이중추출 표본설계를 할 것이다. 여기서 주표본으로 이용되는 농업조사는 2010년 농업총조사 자료를 모집단으로 2011년 하반기에 표본설계 되었으며, 2011년 12월에 통계조사를 실시했다.

이중추출방법은 통계청에서 농업통계에 처음 적용하는 표본설계 방법이므로 차근차근 단계별로 체계적인 정리를 겸한 연구를 하는 것이 중요하다고 외부 전문위원들이 권고하였다. 이런 권고사항을 고려하여 이중추출 표본설계 효과 분석 연구는 크게 3단계로 이루어질 것이다.



•1단계: 이중추출방법의 선행연구로 기초문헌연구와 국내외 연구동향을 파악하는 것이다. (완료) 이중추출방법이란 어떤 방법인지를 구체적으로 설명하고, 주요 특성과 추정방법 등을 요약하고 정리한다. 또한, 국내·외의 주요 연구동향과 적용사례를 검토하였다.

•2단계: 농가경제조사 표본개편(안)을 참조한 층화집락-층화를 위한 복합 이중추출 추정량을 검토하여 1차적으로 제시할 것이다. 이중추출 표본설계 효과를 분석하기 위해서 농가경제조사 주요 변수들의 모집단이 필요하지만 현실적으로 구할 수 없으므로 유사한 표본모집단을 구축하여 추정량을 산출하여야 한다. 이중추출방법은 1차 표본인 농업조사 표본에서 2차 표본인 농가경제조사 표본농가를 추출하는 것으로서 2차 표본은 1차 표본의 집합표본이다. 하지만, 현재 사용가능한 농업조사와 농가경제조사 자료는 2005년 농업총조사를 모집단으로 각각 표본설계하여 중복부분이 없으므로 이용가능한 농업조사 자료를 활용하여 복합 이중추출을 적용할 농가경제조사 유사 모집단 자료를 생성해야 하며, 이에 따른 연구가 필요하다.

•3단계: 2단계에서 생성된 농가경제조사 유사 표본모집단에서 제시된 복합 이중추출 추정량을 산출하여 이중추출 효과를 측정한다. 제시한 추정량과 반복 분산추정량 등의 안정성과 효율성 등을 비교하여 농가경제조사 추정량으로 적합한지 검토할 것이다.

본 연구는 제시된 3단계 연구과정 중 2단계에 해당하는 연구로서 복합 이중추출 추정량을 제시하고, 농가경제조사 유사 표본모집단을 구축하는 방안을 연구하였다. 제2절에서 2010년 기준 농가경제조사 표본개편(안)을 참조하여 층화집락-층화 복합 이중추출 추정량을 제시하였다. 제3절에서는 농가경제조사와 관련된 이용가능 자료들의 주요 특성을 감안한 자료연계 방법과 이에 대한 검토 및 분석 등을 통해 농가경제조사 유사 표본모집단을 구축하였고, 구축된 유사 표본모집단의 주요 특성을 분석하여 모집단으로 활용가능한지 검토하였다. 제4절에서 연구결과를 요약하고, 향후 진행될 연구사항과 연구방향에 대해 언급하면서 이번 연구를 마무리하였다.

본 연구의 범위와 관련해서 간단히 언급해 둘 것이 있다. 이중추출방법은 두 단계에 걸쳐 동일표본에 대해 간략조사와 심층조사를 하여 2차 표본들의 결과를 추정할 때 1차의 간략조사 정보를 보조정보로 이용할 수 있어서 비추정 또는 회귀추정에 매우 유익한 것으로 알려져 있다. 하지만, 본 연구에서는 복합 이중추출 표본설계에 따른 추정에 대해서만 연구하였고 보조정보를 활용한 추정량 연구는 향후에 다시 연구할 계획이다.

표본개편에 따른 복합 이중추출 추정효과를 사전에 측정하여야 하기 때문에 2005년 기준 농가경제조사 집계자료를 복합 이중추출 추정결과로 간주하여 연구를 진행하였고,

이와 관련된 농업조사 관련 자료는 2010년 기준을 사용하였으므로 이를 감안하여 추정량에 대한 검토와 분석을 해야 한다.

향후에 복합 이중추출 표본농가를 대상으로 조사한 결과자료를 기초자료로 이용한다면 더 나은 연구결과가 나올 것으로 여겨진다.

제2절 복합 이중추출에 따른 추정량 개발

1. 개요

농가경제조사는 매 5년마다 표본개편을 실시한다. 2008년에 개편된 ‘2005년 기준 농가경제조사’는 ‘2005년 농업총조사’를 표본추출틀로 하여 표본농가를 층화2단 확률비례 계통추출 하였는데, 2012년에 개편되는 ‘2010년 기준 농가경제조사’ 표본농가는 복합 이중추출할 계획이다.

‘2010년 농업총조사’를 표본추출틀로 하여 표본설계한 농업조사를 1차 표본이면서 2차 표본추출틀로 하고, 2차 표본인 농가경제조사의 표본농가를 추출한다. 1차 표본은 층화집락 추출하고 2차 표본은 층화추출을 적용하므로 층화집락-층화에 따른 복합 이중추출방법을 고려하여 추정량을 검토하고 제안하였다.

2. 농가경제조사 표본개편(안)

농업통계는 2010년 농어업총조사 결과를 토대로 2012년 하반기에 표본개편을 완료할 예정으로 농가경제조사 표본설계를 현재 추진하는 농업통계 표본설계 개편(안)을 중심으로 개략적으로 기술하였다. 상세한 표본개편 내용은 2012년 하반기에 발간될 『농업부문 표본설계 개편보고서』를 참고하기 바란다.

이번 농업통계 표본개편의 가장 큰 변화는 이중표본설계(double sampling)와 리스트추출법(list sampling)으로 표본농가를 추출한 것이다. 1차에서 대규모 연간조사인 농업조사를 2차 표본의 표본추출틀(주표본, master sampling frame)로 활용하여 2차에서 리스트 추출법을 적용하여 표본가구를 직접 추출하였다. 조사결과의 연속성 유지 및 정도 제고를 위해 축종 규모별로 층화계통추출방법을 적용하였다. 이런 방법을 이번 연구보고서에서는 ‘복합 이중추출’로 기술하였다.

이중추출에서 ‘2010년 기준 농가경제조사’의 2차 표본을 추출하기 위해 ‘2010년 농업총조사’에서 층화집락 추출하여 2011년 12월에 조사된 ‘2011년 농업조사’ 표본농가와 가축동향조사를 조사모집단으로 이용하여 2차 표본을 위한 이중모집단(dual population)을



구축하였다. 이렇게 구축된 이중모집단을 영농형태별로 층화하여 각 층에서 2차 표본 농가를 층화추출하였다.

1차 추출²⁾에서는 농업조사의 주요변수인 농가수, 농가인구수, 농업종사자수 등과 상관관계가 높은 농가수를 층화변수로 선정하였고, 층의 개수는 분산 감소폭을 고려하여 시군별로 3~4개 층으로 구분하였으며, 총계분산이 가장 작은 $W_h\sigma_h$ 균등화(Dalenius & Gurney) 층화 방법으로 층간 경계점을 결정하였다.

표본규모는 각 시군별 목표상대표준오차(CV) 및 업무량을 고려하여 결정하였으며, 네이만 배분법을 이용하여 시군별 층별로 표본을 배분하였다. 따라서, 1차 추출 설계가 중치(Design Weight)는 층별 자체가중(self-weight) 설계로 단위가 표본에 포함될 확률의 역수이다.

2차 표본의 표본추출틀은 2011년 농업조사 결과 가축사육농가 리스트 및 2012년 1사분기 가축동향조사 결과 가축사육농가 리스트이다. 2차 표본의 층화변수는 표본추출틀에서 농축산물판매액과 관련이 높고 현 조사결과와의 일관성 유지를 위해 영농형태 및 전겸업별 특성을 층화변수로 결정하였다. 층화변수로 결정된 ‘영농형태 및 전겸업별 특성’에 따라 논벼, 식량, 채소, 특약용및버섯, 과수, 화훼, 기타, 축산, 2중겸업과 같이 9개로 층화되고, 리스트 추출방법의 적용을 통해 층화효율의 증대를 기대할 수 있다.

2차 표본규모는 농업조사의 표본조사구 내에서 리스트 샘플링으로 전환하여 조사구간 분산, 조사구 내 분산, 1차 추출단위(조사구당) 현지 조사에 필요한 비용과 면접비용, 조사표류 등 2차 추출단위(가구)당 조사에 필요한 비용 등 현지 조사 및 이동비용을 고려하여 최종적으로 2,600여 가구를 추출하였다.

부차모집단인 9개도 지역의 통계생산을 위해 특광역시는 인근 도지역에 포함시켜 2,600개 표본을 여러 가지 배분방법과 업무량 및 예산 등을 고려하여 지역별로 배분하였다.

3. 복합 이중추출 추정량

가. 개요

이번 연구에서는 ‘2010년 기준 농가경제조사’의 2차 표본추출틀로 이용되는 1차 표본은 ‘2010년 농업총조사’에서 층화집락 추출하여 2011년 12월에 조사된 ‘2011년 농업조사’의 표본농가로 한정하였다. 또한, 층화집락 추출된 1차 표본농가를 영농형태별로 층화하여 각 층에서 2차 표본농가를 층화추출할 것이다.

2) 1차 추출의 표본은 다목적 표본이다. 첫째는 농가, 농업인구 등을 파악하기 위한 농업조사를 목적으로 하며, 두 번째는 농가경제조사와 농축산물생산비조사 표본을 추출하기 위한 이중추출(Double Sampling)의 주표본(Master Sample)으로 사용한다.

1차 표본농가는 조사구 단위로 추출하고 집계되므로 각 조사구의 농가는 동일한 조사구 가중값을 가지고 있으며, 1차 표본농가를 영농형태별로 층화하여 각 층에서 2차 표본농가를 추출할 것이므로 2차 표본농가는 각각 다른 조사구 가중값과 층 가중값을 가진다.

2차 표본추출틀은 층화집락 추출된 1차 표본농가 전부를 층화하여 구축하는데, 2차 표본추출틀에서 층화된 각 층은 서로 독립이라는 가정하에 추정량을 검토하였다.

나. 2차 표본추출의 층화 기준

현재 표본개편중인 '2010년 기준 농가경제조사'의 최종 표본인 2차 표본농가는 표본추출틀인 '2011년 농업조사'를 층화추출할 것인데, 기본적으로 9개 영농형태로 층화하고 다른 농가 특성을 반영하여 추가로 층화할 것이다. 본 연구에서는 9개 영농형태로 층화하여 2차 표본농가를 추출한 것으로 추정량 연구³⁾를 진행하였다.

집락-층화 이중추출은 임의추출-층화의 이중추출에 비해 분산 추정량의 전개과정이 상당히 복잡하였다. 복합 이중추출 추정량을 연구할 때 최대한 농가경제조사의 표본 설계 가정을 반영하였지만, 실제 농가경제조사의 분산 추정 시에는 다시 한번 가정과 수식을 상세히 검토하여 적용하여야 할 것이다. 여기에서는 상세한 전개과정은 생략하고, 간략하게 추정량 연구 결과만 수록하였다.

다. 농가경제조사의 복합 이중추출 추정량

1) 전국 가구당 평균추정량

2차 표본농가의 전국 가구당 평균 추정량은 (3.1)과 같다.

$$\begin{aligned}
 \widehat{Y_M^{(2)}} &= \widehat{\overline{Y^{(2)}}} = \frac{\widehat{Y^{(2)}}}{M} = \frac{1}{M} \sum_{j=1}^{n_k^{(2)}} w_j^{(1)} Y_j^{*(2)} \\
 &= \frac{1}{M} \sum_{h=1}^L \sum_{j=1}^{n_h^{(2)}} \sum_{k=1}^{m_{hj}^{(2)}} w_j^{(1)} \frac{M_j^*}{m_j^*} \frac{m_h^{(1)}}{m_h^{(2)}} Y_{hjk}^{(2)} \\
 &= \frac{1}{M} \sum_{h=1}^L \sum_{j=1}^{n_h^{(2)}} \sum_{k=1}^{m_{hj}^{(2)}} w_j^{(1)} w_h^{(2)} Y_{hjk}^{(2)} \tag{3.1}
 \end{aligned}$$

3) 집락-층화 이중추출인 복합 이중추출의 추정량 연구는 한국외국어대학교 신기일 교수와 경기대학교 이상은 교수와 공동으로 연구를 진행하였다.



여기에서

$N(n)$ 조사구 수, $M(m)$ 농가 수

Y 분석변수, h 층, j 집락, k 농가자료, * 집락

(¹)은 1차 표본, (²)는 2차 표본

$w_j^{(1)} = 1$ 차 표본의 조사구당 가중값, $w_h^{(2)} = m_h^{(1)} / m_h^{(2)}$

(3.1)에서 층화집락 추출한 1차 추출 표본농가 전부를 2차 추출에서는 9개 영농형태별로 층화했으므로 $\frac{M_j^*}{m_j^*} = 1$ 이 된다.

2차 표본을 이용한 가구당 평균은 불편추정량이다.

$$E(\widehat{Y_M^{(2)}}) = \bar{Y}_M$$

2) 전국 가구당 평균의 분산추정량

조건부 분산식인 $V(\widehat{Y}) = V_1 E_2(\widehat{Y}) + E_1 V_2(\widehat{Y})$ 을 이용하여 분산 추정량을 구한다. 분산의 불편추정량은 (3.2)와 같다.

$$\begin{aligned} \widehat{Var}(\widehat{Y_M^{(2)}}) &= v(\widehat{Y_M^{(2)}}) \\ &= \frac{N^2}{M^2} \frac{1-f_1}{n} \frac{\sum_{j=1}^n (\widehat{Y}_j^* - \widehat{Y}_N)^2}{n-1} + \frac{1}{M^2} \frac{N}{n} \sum_{j=1}^n M_j^{*2} \frac{s_{2j}^2}{m_j^{*(1)}} (1-f_{2j}) \quad (3.2) \\ &+ \frac{1}{M^2} \frac{N^2}{n^2} \sum_{h=1}^L \left\{ \frac{1-f_1^{(2)}}{n_h^{(2)}} \frac{\sum_{j=1}^{n_h^{(1)}} (\widehat{M}_{hj}^{*(2)})^2 (\widehat{Y}_{hj}^{*(2)} - \widehat{Y}_h)^2}{n_h^{(1)} - 1} + \frac{1}{n_h^{(1)} n_h^{(2)}} \sum_{j=1}^{n_h^{(1)}} (\widehat{M}_{hj}^{*(2)})^2 \frac{s_{2hj}^{(2)}}{m_{hj}^{(2)}} (1-f_{2hj}^{(2)}) \right\} \end{aligned}$$

여기에서

$$\widehat{Y}_j^* = \frac{M_j^*}{m_j^*} \sum_{h=1}^L \sum_{k=1}^{m_{hj}^{(1)}} Y_{hjk}^{(1)} = \sum_{h=1}^L \sum_{k=1}^{m_{hj}^{(1)}} Y_{hjk}^{(1)}$$

$$\widehat{Y}_N = \frac{\widehat{Y}}{N} = \frac{1}{N} \sum_{j=1}^n w_j^{(1)} \widehat{Y}_j^*$$

$$\widehat{Y}_{hj}^{*(2)} = \frac{\sum_{k=1}^{m_{hj}^{(2)}} y_{hjk}^{(2)}}{m_{hj}^{(2)}}, \quad \widehat{Y}_h = \frac{1}{n_h^{(2)}} \sum_{k=1}^{m_h^{(2)}} y_{hjk}^{(2)}$$

$$f_1 = \frac{n}{N}, \quad f_{2j} = \frac{m_j^{*(1)}}{M_j^*}, \quad f_1^{(2)} = \frac{n_h^{(2)}}{n_h^{(1)}}$$

$$\widehat{M}_{hj}^{*(2)} = \frac{M_j^*}{m_j^*} \frac{m_h^{(1)}}{m_h^{(2)}} m_{hj}^{(2)} = \frac{m_h^{(1)}}{m_h^{(2)}} m_{hj}^{(2)} \quad (\because M_j^* = m_j^*)$$

$$s_{2hj}^{2(2)} = \frac{\sum_{k=1}^{m_j^{*(2)}} (y_{jk}^{(2)} - \bar{y}_j^{*(2)})^2}{m_j^{*(2)} - 1}, \quad f_{2hj}^{(2)} = \frac{m_{hj}^{(2)}}{m_{hj}^{*(1)}}$$

$$\bar{y}_j^{*(2)} = \frac{1}{m_j^{*(2)}} \sum_{k=1}^{m_j^{*(2)}} y_{jk}^{(2)}$$

2차 표본을 이용한 가구당 평균의 분산추정량은 불편추정량이다.

제3절 농가경제조사 유사 표본모집단 구축

1. 개요

복합 이중추출 표본설계 효과를 분석하기 위하여 농가경제조사 표본모집단 자료가 필요한데, 현실적으로 자산, 부채, 소득 등이 포함된 ‘2010년 기준 농가경제조사 표본모집단’ 자료는 없으므로 유사한 표본모집단을 임시로 생성해야 한다.

2차 표본추출틀인 농업조사는 2011년 12월에 조사되어 2012년 4월에 발표될 예정이고, 복합 이중추출로 추출된 표본농가에 대한 농가경제조사 결과는 2014년에 발표될 예정이므로 현재 농가경제조사 관련 이용가능한 자료들을 활용하여 농가경제조사 유사 표본모집단을 구축하는 방법을 연구하였다.

2. 모집단 구축을 위한 자료 연계 및 특성분석

현재 사용할 수 있는 농가경제조사 관련 자료중 가장 유용한 자료들을 활용하여 ‘2010년 기준 농가경제조사 유사 표본모집단’을 구축하였다.



가. 이용가능 자료

‘2010년 기준 농가경제조사’의 표본조사가 완료되었다는 가정하에 추정값을 산출하여야 하는데, 현재는 표본설계 중이므로 ‘2005년 기준 농가경제조사 자료’를 이용하여야 한다. 표본모집단으로는 ‘2010년 기준 농업총조사’ 자료를 이용할 수 있으나 두 자료간의 기준이 다르므로 이를 보완해주기 위해 ‘2010년 농가경제조사’의 표본모집단 역할을 할 수 있는 ‘농가명부조사’ 자료를 이용하였다.

현재 이용가능한 자료는 모두 4가지 종류이며, 각 조사의 간략한 설명 및 사용기간은 <표 3-1>과 같은데, 분석변수가 포함된 농가경제조사 집계자료와 표본추출틀인 농업총조사, 농업조사의 기준연도가 서로 차이가 있으므로 추정결과의 검토 또는 분석 시 이를 유의해야 한다.

본 연구에서는 ‘2011년 농업조사 집계결과’가 4월 중순에 공표된 관계로 ‘2010년 농가명부조사 자료’, ‘2010년 농가경제조사 집계자료’, ‘2011년 농업조사 표본농가 명부자료’를 이용하였다. 추정량 검토 후에 ‘2011년 농업조사 집계자료’를 이용하여 다시 추정값을 산출할 예정이다.

<표 3-1> 농가경제조사 유사 표본추출틀 구축 시 이용가능 자료

이용 여부	해당 년도	자료	기준 년도	조사시기	농가수	비고
이용	2010년	농가명부조사 자료	2005년	2010. 3.	15,000농가	농가경제조사 농가유고 시 교체목적으로 매년 조사
이용	2010년	농가경제조사 집계자료	2005년	2010.1.~12.	2,649농가	월간조사에서 연간집계
이용 안함	2010년	농업총조사 집계자료	2010년	2010.12.	117만농가	매 5년 조사
이용	2011년	농업조사 표본농가 명부자료	2010년	2010.12.	75,000농가	농업총조사에서 추출
분석 예정	2011년	농업조사 집계자료	2010년	2011.12.	75,000농가	매년 조사

나. 자료연계방법

1) 요약

유사 표본모집단을 구축할 때 가장 먼저 고려해야 될 것은 분석자료인 ‘2010년 농가경제조사’의 모집단 정보를 알아야 되는데, 집계결과만 있으므로 2010년 모집단 정보를 얻기 위해서 농가명부 자료를 연계하였다. 두 조사는 2005년 기준으로 조사되었으므로 동일한 조사시점의 연결이 매우 중요하다.

우선, 표본모집단 정보를 얻을 수 있는 ‘2010년 농가명부 자료’와 ‘2010년 농가경제조사 자료’를 일대일로 연계하여 2,649농가의 모집단 정보를 구축하였으며, 이를 ‘2010년 농가경제조사 연계자료’라고 한다.

모집단 정보를 포함한 ‘2010년 농가경제조사 연계자료’와 ‘2011년 농업조사 표본농가명부자료’의 농가특성 중 전겸업, 영농형태를 공통변수로 자료연계하여 ‘2010년 기준 농가경제조사 유사 표본모집단’ 74,349농가를 구축하였다.

유사 표본모집단에서 자산, 부채, 농가소득, 농업소득, 농업총수입, 겸업소득 등을 분석변수로 이용하였다.

2) 주요 특성 분석

농가경제조사는 일계부와 농가원부 연간 실적을 충분히 고려하여 전겸업, 영농형태, 주부업 형태 등 농가특성을 구분하므로, 매년 조사되는 동일 표본농가의 전겸업, 영농형태, 주부업 형태 등이 변경될 수 있고, 이에 따라 소득·자산·부채 등에 영향을 미칠 수 있으므로 영향이 큰 농가들의 집중관리가 필요하다. ‘2010년 농가경제조사 집계자료’에서 자산, 부채, 농가소득, 농업소득, 농업총소득, 겸업소득 등을 분석변수로 이용하였다.

주요 농가특성을 간략하게 살펴보면 다음과 같다.

2003년에서 2010년 사이에 전업농가는 크게 감소하고, 1종·2종겸업 농가는 증가하였으며, 논벼·축산 농가는 감소하였고, 과수·특작·기타 농가는 소폭 증가하였다.

일반적으로 우리나라는 1농가에서 2가지 이상의 영농형태가 나타나는데, 주 수입원에 따라 영농형태가 변화하므로 농업정책·농촌환경 변화에 주의를 기울여야 한다. 화훼·축산·특작 농가 변화는 농가부채에 미치는 영향이 크다. 또한, 영농형태의 유지·변경과 농가소득은 매우 밀접하며, 전겸업·영농형태가 동일한 농가는 소득이 높고 변동된 농가는 소득이 낮다.

농촌사회의 소가족화와 경영주의 고령화 현상이 심화되었는데, 60세 미만 경영자는 동일 영농형태를 유지하려는 경향이 있고, 60세 이상 경영자는 영농형태를 변경하려는 경향이 있으며, 경영주 연령이 높을수록 영농형태 변동이 잦으며 농업부채가 많다.



3) 자료연계방법

2차 표본모집단으로 사용될 '2011년 농업조사 집계자료'는 농가 가구주 또는 가구원 한테 전겸업·영농형태·주부업 형태 등을 면접조사하여 구분하였으며, 농업총조사에서 추출된 '2011년 농업조사 표본농가 명부자료'도 면접조사하여 농가특성을 구분하였다. 따라서, 면접조사한 자료들과 연간 실적을 고려한 '2010년 농가경제조사 집계자료'는 농가특성이 다르게 분포되므로 '2010년 농가경제조사 집계자료'의 표본모집단 농가특성 정보가 필요하다.

가) 1차 자료연계

'2010년 농가경제조사 집계자료'의 표본모집단 정보를 얻기 위해 '2010년 농가명부조사 자료'와 '2010년 농가경제조사 집계자료'를 이용하였다. 농가명부조사는 매년 3월에 농가경제조사의 유고 표본농가를 교체하기 위해 조사하며, 조사 시 농가특성을 면접조사하므로 농가경제조사의 표본모집단 농가특성으로 생각할 수 있다. '2010년 농가경제조사'의 집계농가와 동일한 표본농가의 '2010년 농가명부조사 자료'를 1:1 연계하였다.

농가명부 15,000농가자료 중 농가경제조사 2,649집계농가에 해당하는 농가자료만을 일대일로 연계하여 '2010년 농가경제조사 연계자료'를 생성하였다. 자료연계 자료에서 두 조사의 농가특성의 분포와 일치성을 알아보았다.

<표 3-2>는 '2010년 농가경제조사 연계자료'에서 '2010년 농가명부조사 자료'와 '2010년 농가경제조사 집계자료'의 농가특성을 비교한 통계표이다.

<표 3-2>에서 농가명부조사 자료와 농가경제조사 집계자료간의 전겸업, 영농형태, 가경영주 연령 및 가구원수 등 주요 농가특성에서 차이가 있음을 알 수 있다.

전겸업 특성이 가장 큰 차이를 보였는데, 2649농가 중 31.6%인 837농가가 변동이 있었고 전업·1종겸업에서 변동이 많았다.(①, ①-1 참조)

농가명부조사 자료의 영농형태에서는 2종겸업이 구분되지 않아서 먼저 전겸업의 2종겸업을 영농형태별 2종겸업으로 구분한 다음, 남은 농가를 8개 영농형태로 재구분하여 비교하였다([그림 3-1] 참조). 영농형태는 2,649농가 중 29.0%인 770농가가 변동하였으며, 논벼·채소·축산·2종겸업에서 변동이 많았다(②, ②-1 참조).

경영주 연령은 전체 2,649농가 중 8.3%인 221농가, 가구수원수는 15.4%인 409농가가 변동하였다. 이런 차이가 발생하는 주요 원인은 농업명부조사는 면접조사로 농가특성을 구분한 반면에 농가경제조사는 일 년간의 농사 결과를 반영하여 농가특성을 구분하므로 구분방법의 차이가 주요 원인이라고 할 수 있다.

'2010년 농가경제조사 연계자료'에서 농가명부의 전겸업, 영농형태 등 주요 농가특성을

모집단 특성으로 간주하였고, 자산, 부채, 농가소득, 농업소득, 농업총수입, 겸업소득 등 분석변수는 농가경제조사 자료를 이용하였다.

<표 3-2> '2010년 농가경제조사 연계자료'의 농가특성 비교

① 전겸업별 분포

	전체	농가명부조사(a)			농가경제조사(b)			차이(a-b)		
		전업	1종겸업	2종겸업	전업	1종겸업	2종겸업	전업	1종겸업	2종겸업
농가수	2649	685	1093	871	1118	770	761	-433	323	110
(%)	100.0	25.9	41.3	32.9	42.2	29.1	28.7	-16.3	12.2	4.2

①-1 두 자료의 전겸업 변동 농가 분포

		농가명부조사			
		전체	전업	1종겸업	2종겸업
농가 경제 조사	전체	837	92	535	210
	전업	525	.	443	82
	1종겸업	212	84	.	128
	2종겸업	100	8	92	.

② 영농형태별 분포

		논벼	과수	채소	특작	화훼	전작	축산	기타	2종겸업
농가명부 조사(a)	농가수	968	320	585	123	72	86	163	332	
	(%)	36.5	12.1	22.1	4.6	2.7	3.2	6.2	12.5	
농가명부 조사(a')	농가수	700	270	448	93	61	63	129	14	871
	(%)	26.4	10.2	16.9	3.5	2.3	2.4	4.9	0.5	32.9
농가경제 조사(b)	농가수	500	286	616	90	50	81	227	38	761
	(%)	18.9	10.8	23.3	3.4	1.9	3.1	8.6	1.4	28.7
차이 (a'-b)	농가수	200	-16	-168	3	11	-18	-98	-24	110
	(%)	7.5	-0.6	-6.3	0.1	0.4	-0.7	-3.7	-0.9	4.2

* a' : 재구분한 영농형태별 분포



②-1 두 자료의 영농형태 변동 농가 분포

		농가명부조사									
		전체	논벼	과수	채소	특작	화훼	전작	축산	기타	2종겸업
농가 경제 조사	전체	770	295	39	88	36	23	38	31	10	210
	논벼	95	.	10	11	6	3	3	10	2	50
	과수	55	16	.	6	1	1	1	2	.	28
	채소	256	126	9	.	15	6	25	8	2	65
	특작	33	8	2	8	.	.	1	2	.	12
	화훼	12	2	1	1	.	.	.	1	3	4
	전작	56	14	2	20	5	.	.	2	.	13
	축산	129	76	3	12	1	.	2	.	.	35
	기타	34	13	1	8	1	6	2	.	.	3
	2종겸업	100	40	11	22	7	7	4	6	3	.

③ 경영주 연령별 분포

		2(20대)	3(30대)	4(40대)	5(50대)	6(60대 이상)
농가명부 조사(a)	농가수	21	175	555	846	1052
	(%)	0.8	6.6	21.0	31.9	39.7
농가경제 조사(b)	농가수	21	193	575	870	990
	(%)	0.8	7.3	21.7	32.8	37.4
차이 (a-b)	농가수	0	-18	-20	-24	62
	(%)	0.0	-0.7	-0.7	-0.9	2.3

③-1 두 자료의 경영주 연령별 변동 농가 분포

		농가명부					
		전체	2 (20대)	3 (30대)	4 (40대)	5 (50대)	6 (60대 이상)
농가 경제 조사	전체	221	3	10	42	77	89
	2(20대)	3	.	2	.	.	1
	3(30대)	28	2	.	25	1	.
	4(40대)	62	1	8	.	50	3
	5(50대)	101	.	.	16	.	85
	6(60대 이상)	27	.	.	1	26	.

④ 가구원수별 분포

		1명	2명	3명	4명	5명 이상
농가명부 조사(a)	농가수	1638	490	269	153	99
	(%)	61.8	18.5	10.2	5.8	3.7
농가경제 조사(b)	농가수	1690	504	239	135	81
	(%)	63.8	19.0	9.0	5.1	3.1
차이 (a-b)	농가수	-52	-14	30	18	18
	(%)	-2.0	-0.5	1.2	0.7	0.6

④-1 두 자료의 가구원수 변동 농가 분포

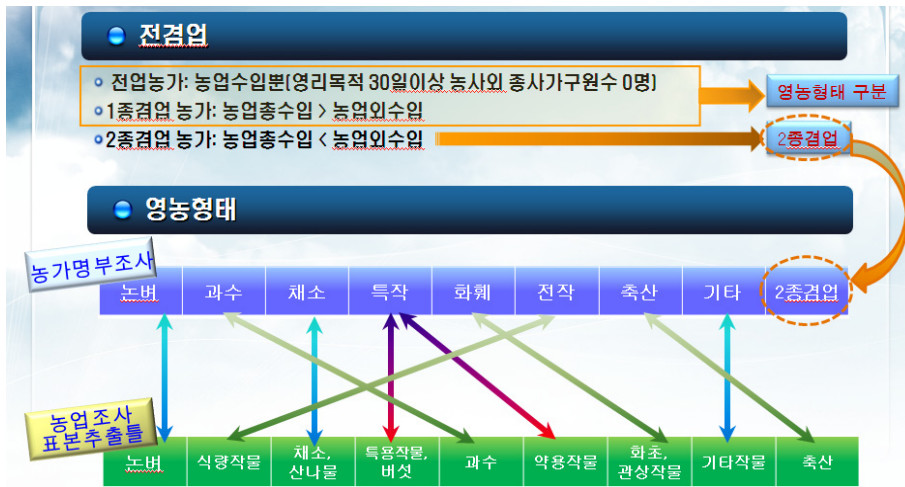
		농가명부조사					
		전체	1명	2명	3명	4명	5명 이상
농 가 경 제 조 사	전체	409	75	134	104	63	33
	1명	127		105	16	2	4
	2명	148	64		65	14	5
	3명	74	7	25		36	6
	4명	45	3	3	21		18
	5명 이상	15	1	1	2	11	

나) 2차 자료연계(2차 표본모집단 생성과정)

‘2010년 농가경제조사 연계자료’와 ‘2011년 농업조사 표본농가 명부자료’를 지역별, 전겸업별, 영농형태별로 구분하여 ‘2010년 농가경제조사 연계자료’가 조사된 지역, 전겸업, 영농형태에 대해서만 ‘농업조사 표본농가 명부자료’에서 가져와 ‘2010년 기준 농가경제조사 유사 표본모집단 자료’를 생성하였다.

‘2010년 기준 농가경제조사 유사 표본모집단’을 구축하는 과정은 다음과 같다.

농업조사 표본농가 자료는 75,353농가이고, 영농형태 변수인 경영형태는 논벼, 식량작물, 채소·산나물, 특용작물·버섯, 과수, 약용작물, 화초·관상작물, 기타작물, 축산으로 농가경제조사 영농형태와 다르게 구분되어 있다. 영농형태를 동일하게 구분해 주기 위해, 농업조사의 전겸업 농가 중 2종겸업 농가를 영농형태의 2종겸업으로 구분한 다음, 논벼, 과수, 채소·산나물, 특작은 특용작물·버섯과 약용작물, 화훼는 화초·관상작물, 전작은 식량작물, 축산, 기타 순으로 구분하였다. 농업조사의 경영형태를 농가경제조사 영농형태와 같이 구분하는 방법은 [그림 3-1]을 참고한다.



[그림 3-1] 농업조사 경영형태를 농가경제조사 영농형태로 구분한 방법

다. 자료비교

앞에서와 같이 구분된 영농형태를 적용하여 생성된 ‘2010년 농가경제조사 연계자료’와 ‘2011년 농업조사 표본농가 명부자료’의 전겸업별, 영농형태별 분포를 각각 비교해 보았다. 비교 결과는 <표 3-3>과 같다.

<표 3-3>에서 ①, ②는 각각 ‘농가경제조사 연계자료’와 농업조사의 층화기준인 영농형태를 전겸업별로 구분하여 농가수와 전체에 대한 구성비를 비교한 결과이다.

전겸업별 구성비를 비교해 보면, 농가경제조사는 ‘1종겸업→2종겸업→전업’ 순이고, 농업조사는 ‘전업→2종겸업→1종겸업’ 순으로 ‘전업’과 ‘1종겸업’의 차이가 바뀌어 나타난다.

또한, 영농형태별 구성비를 비교해 보면, 농가경제조사는 ‘2종겸업→논벼→채소→과수→축산→특작→전작·화훼→기타’ 순이고, 농업조사는 ‘논벼→2종겸업→채소→과수→축산·전작→특작→화훼→기타’ 순으로 나타나며, ‘2종겸업’과 ‘논벼’, ‘전작’, ‘축산’ 등의 구성비에서 차이가 많은 것을 알 수 있다.

<표 3-3> 농가경제조사 연계자료와 농업조사 표본농가 명부자료 비교

① 2010년 농가경제조사 연계자료의 전겸업별×영농형태별 분포

		영농형태									
		전체	논벼	과수	채소	특작	화훼	전작	축산	기타	2중겸업
전겸업	전체	2,649	700	270	448	93	61	63	129	14	871
		100	26.43	10.19	16.91	3.51	2.30	2.38	4.87	0.53	32.88
	전업	685	275	117	163	30	17	30	50	3	0
		25.86	10.38	4.42	6.15	1.13	0.64	1.13	1.89	0.11	0
	1종 겸업	1093	425	153	285	63	44	33	79	11	0
		41.26	16.04	5.78	10.76	2.38	1.66	1.25	2.98	0.42	0
	2종 겸업	871	0	0	0	0	0	0	0	0	871
		32.88	0	0	0	0	0	0	0	0	32.88

* 이탤릭체는 전체에 대한 구성비임

② 2011년 농업조사 표본농가 명부자료의 전겸업별×영농형태별 분포

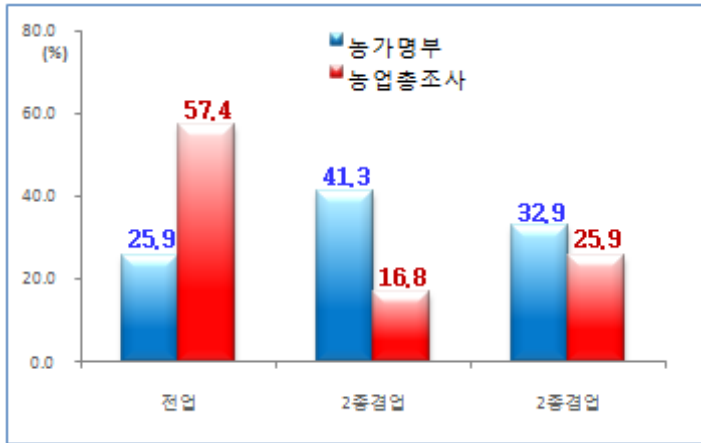
		영농형태									
		전체	논벼	과수	채소	특작	화훼	전작	축산	기타	2중겸업
전겸업	전체	75,353	23,729	7,804	11,736	1,842	935	4,577	4,766	448	19,516
		100	31.49	10.36	15.57	2.44	1.24	6.07	6.32	0.59	25.9
	전업	43,216	18,457	5,873	9,090	1,412	686	3,969	3,390	339	-
		57.35	24.49	7.79	12.06	1.87	0.91	5.27	4.5	0.45	-
	1종 겸업	12,621	5,272	1,931	2,646	430	249	608	1,376	109	-
		16.75	7.00	2.56	3.51	0.57	0.33	0.81	1.83	0.14	-
	2종 겸업	19,516	-	-	-	-	-	-	-	-	19,516
		25.9	-	-	-	-	-	-	-	-	25.9

* 이탤릭체는 전체에 대한 구성비임

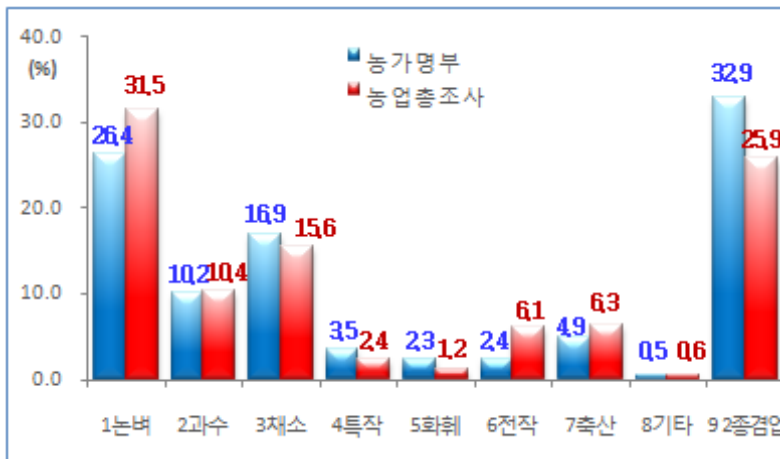
[그림 3-2], [그림 3-3]은 <표 3-3>의 두 자료들에 대해 각각 전겸업과 영농형태별로 구성비만을 비교한 것이다. 두 자료는 전겸업별 모든 구분에서 차이가 크게 났는데, 전업이 31.5%p 차이로 가장 크고 2중겸업의 차이가 가장 적었다. 영농형태에서는 2중겸업과 논벼, 전작, 축산, 채소, 특작, 화훼 순으로 차이가 발생했으며 기타, 과수는 매우 적게 차이가 발생했다.

두 조사의 전겸업별, 영농형태별 구성비 차이가 크므로 어느 자료를 농가경제조사 표본모집단 정보로 활용할 지를 결정해야 한다. 농가경제조사는 농가명부조사의 농가특성 자료이고, 농업조사 표본농가 명부자료는 2010년 기준 농업총조사의 농가특성이다. 2차

표본농가는 ‘2011년 농업조사 집계자료’에서 표본추출 할 것이므로 이와 가장 비슷한 ‘2011년 농업조사 표본농가 명부자료’의 농가특성을 ‘2010년 기준 농가경제조사 표본 모집단’의 농가특성으로 간주하였다.



[그림 3-2] 농가경제조사와 농업조사 표본명부 자료의 전겸업별 구성비 비교



[그림 3-3] 농가경제조사와 농업조사 표본명부 자료의 영농형태별 구성비 비교

3. 유사 표본모집단 구축 및 특성비교

가. 개요

두 자료에서 농가경제조사 유사 표본모집단을 생성하기 위해서 두 자료의 공통변수로 어떤 특성을 사용할지를 결정해야 하는데, 지역별로 분포가 골고루 나올 수 있는 농가 특성이어야 한다. 지역별로 가장 중요한 농가특성인 전겸업과 영농형태로 구분해보니, 1농가가 다수여서 더 이상 변수를 추가하지 않고 지역별×전겸업별×영농형태별로 구분하기로 결정하였다. 지역별×전겸업별×영농형태별 농가 분포는 <표 3-4>를 참고한다.

‘2010년 농가경제조사 연계자료’와 ‘2011년 농업조사 표본농가 명부자료’를 지역별×전겸업별×영농형태별로 구분한 다음, ‘2010년 농가경제조사 연계자료’의 지역별×전겸업별×영농형태별 자료가 있는 셀에 대해서 ‘2011년 농업조사 표본농가 명부자료’ 농가수 만큼 반복해서 임의추출하였다. 자산, 부채, 소득 등의 분석변수 또한 해당 셀의 ‘2011년 농업조사 표본농가 명부자료’ 수만큼 반복해서 임의추출하여 ‘2010년 기준 농가경제조사 유사 표본모집단’을 구축하였다.이용가능자료 연계 및 유사 표본모집단 구축과정을 간략하게 [그림 3-4]로 표현하였다.



[그림 3-4] 이용가능자료 연계 및 유사 표본모집단 구축과정

나. 유사 표본모집단 구축

농가경제조사 연계자료와 농업조사의 ‘지역별×전겸업별×영농형태별’ 농가 분포는 <표 3-4>의 ①, ②와 같다.

<표 3-4>에서 ① ‘2010년 농가경제조사 연계자료’의 14지역은 과수, 채소, 전작, 축산, 2종겸업만 조사되어 있으므로, ‘2011년 농업조사 표본농가 자료’에서 동일한 14지역의 과수, 채소, 전작, 축산, 2종겸업 셀을 농가수만큼 반복 임의추출하여 ‘2010년 기준 농가경제조사 유사 표본모집단’에 포함하였다.

<표 3-4>에서 ① ‘2010년 농가경제조사 연계자료’보다 ② ‘2011년 농업조사 표본농가 자료’가 부족한 경우가 있었는데, <표 3-4>의 ①, ②에서 음영부분 14지역의 특작, 화훼와 같은 경우이다. 이런 경우에는 부족한 ‘2011년 농업조사 표본농가 자료’ 농가수만큼 반복해서 임의추출하여 유사 표본모집단에 포함하였다. 농가경제조사 2차 표본추출 시에는 해당 그룹의 2011년 농업조사 표본농가수로 축소하여 최종적으로 2,646농가를 표본추출하였다.

위와 같은 방법으로 구축된 ‘2010년 기준 농가경제조사 유사 표본모집단’ 농가수는 74,349농가이며, ‘2011년 농가경제조사 연계자료’의 특성변수인 전겸업, 영농형태 등은 해당 셀에서 임의추출하였다. ‘2010년 기준 농가경제조사 유사 표본모집단’ 자료의 ‘지역별×전겸업별×영농형태별’ 분포는 <표 3-5>와 같다.

농가경제조사의 집계자료의 분석변수는 한정되어 있어 동일한 값이 너무 많아 임의추출 시 오차를 주어 임의추출하여 비교해보았다. 동일 값을 반복추출한 분석변수와 오차를 고려한 분석변수는 범위, 평균값, 분산 등에서 차이가 조금 있었지만 변동계수와 왜도는 동일하여 추정 시 큰 차이가 없었다. <표 3-6>를 참고한다.



② 2011년 농업조사 표본추출틀의 지역별×전겸업별×영농형태별 분포

지역	전겸업	영농형태									
		전체	논벼	과수	채소	특작	화훼	전작	축산	기타	2종겸업
전국	전체	75,353	23,729	7,804	11,736	1,842	935	4,577	4,766	448	19,516
	전업	43,216	18,457	5,873	9,090	1,412	686	3,969	3,390	339	.
	1종	12,621	5,272	1,931	2,646	430	249	608	1,376	109	.
	2종	19,516	19,516
14	전체	1211	.	475	253	9	10	98	31	.	325
	전업	504	0	277	129	0	0	0	21	0	.
	1종	382	0	198	124	1	2	39	0	0	.
	2종	325	325
20	전체	6,818	1,398	181	1,249	273	26	1,236	427	16	2,012
	전업	3,674	1,065	124	945	211	21	1,000	294	0	.
	1종	1,132	333	57	304	62	5	236	133	2	.
	2종	2,012	2,012
22	전체	5,734	1,580	724	719	310	30	471	332	21	1,547
	전업	3,295	1,295	558	523	230	26	416	230	0	.
	1종	892	285	166	196	80	4	55	102	4	.
	2종	1,547	1,547
26	전체	7,245	3,180	515	954	195	106	382	407	.	1,469
	전업	4,291	2,333	367	707	160	75	333	286	0	.
	1종	1,485	847	148	247	35	31	0	121	0	.
	2종	1,469	1,469
91	전체	10,764	2,891	689	1,055	240	377	532	686	18	4,276
	전업	4,871	2,115	474	831	0	257	0	498	0	.
	1종	1,617	776	215	224	49	120	37	188	8	.
	2종	4,276	4,276
92	전체	10,243	3,002	1,069	1,977	71	169	352	775	87	2,741
	전업	6,009	2,501	836	1,547	55	127	314	562	67	.
	1종	1,493	501	233	430	16	42	38	213	20	.
	2종	2,741	2,741
93	전체	12,466	3,107	2,808	2,194	319	27	629	857	30	2,495
	전업	8,091	2,573	2,237	1,766	258	22	586	625	24	.
	1종	1,880	534	571	428	61	0	43	232	6	.
	2종	2,495	2,495
94	전체	9,114	3,917	492	1,068	290	78	350	672	.	2,156
	전업	5,275	2,944	361	826	207	68	313	485	0	.
	1종	1,683	973	131	242	83	10	37	187	0	.
	2종	2,156	2,156
95	전체	11,758	4,654	851	2,267	135	112	527	579	138	2,495
	전업	7,206	3,631	639	1,816	92	82	453	389	104	.
	1종	2,057	1,023	212	451	43	30	74	190	34	.
	2종	2,495	2,495

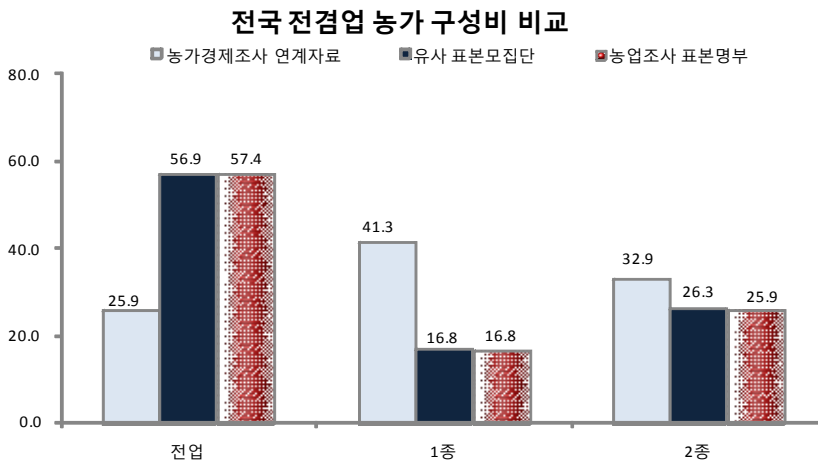


다. 유사 표본모집단의 특성비교

최종 연계한 ‘2010년 기준 농가경제조사 유사 표본모집단’의 주요 특성이 ‘2011년 농업조사 표본농가 명부자료’ 특성과 얼마나 유사한지를 비교하여 표본모집단으로서의 유효성을 검토하였다.

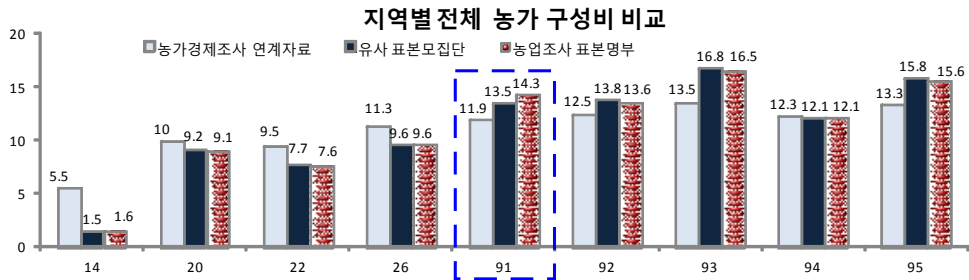
‘2010년 농가경제조사 연계자료’, ‘2010년 농가경제조사 유사 표본모집단’, ‘2011년 농업조사 표본농가 명부자료’의 전겸업별, 영농형태별, 지역별×전겸업별×영농형태별 분포를 살펴보면, ‘2010년 기준 농가경제조사 유사 표본모집단’과 ‘2010년 농업조사 표본농가 명부자료’의 주요 특성이 비슷하게 되어 있음을 알 수 있었다.

[그림 3-5]와 같이 전국의 전겸업 농가 구성비를 비교해보면, 농가경제조사 연계자료와는 차이가 있지만 농가경제조사 유사 표본모집단과 농업조사 표본농가 명부자료는 구성비가 거의 비슷한 것을 알 수 있다.

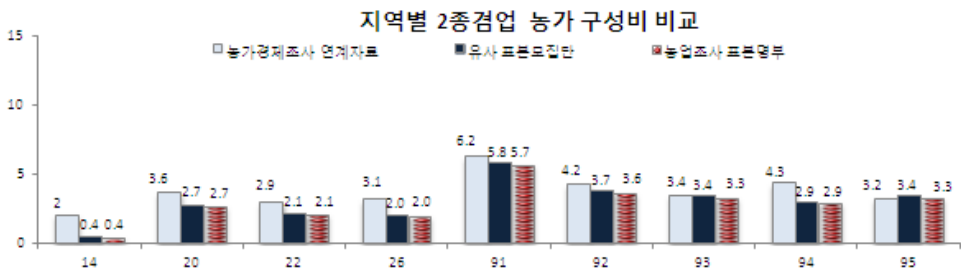
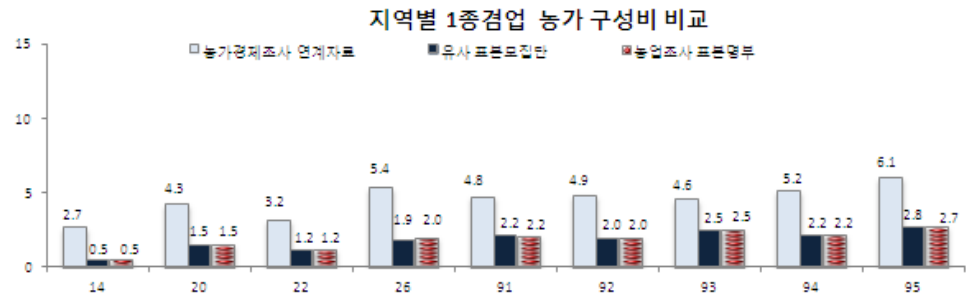
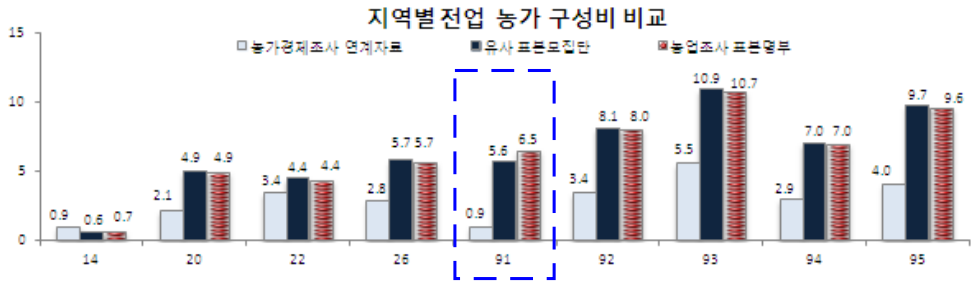


[그림 3-5] 전국 전겸업별 농가 구성비 비교

[그림 3-6], [그림 3-7]에서 지역별로 전겸업 전체와 전업, 1종겸업 및 2종겸업의 분포를 살펴보았는데, 농가경제조사 연계자료와는 차이가 있지만 농가경제조사 유사 표본모집단과 농업조사 표본농가 명부자료의 구성비가 거의 비슷하게 되어 있음을 알 수 있었다. 다른 지역에 비해 91지역에서 전업의 0.9%p 차이로 인해 전겸업 전체에서 0.8%p 차이를 보였다.

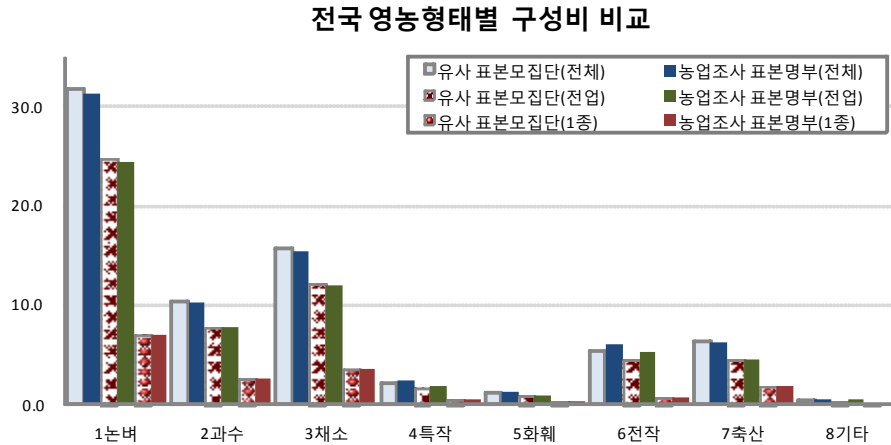


[그림 3-6] 지역별 전체 농가 구성비 비교



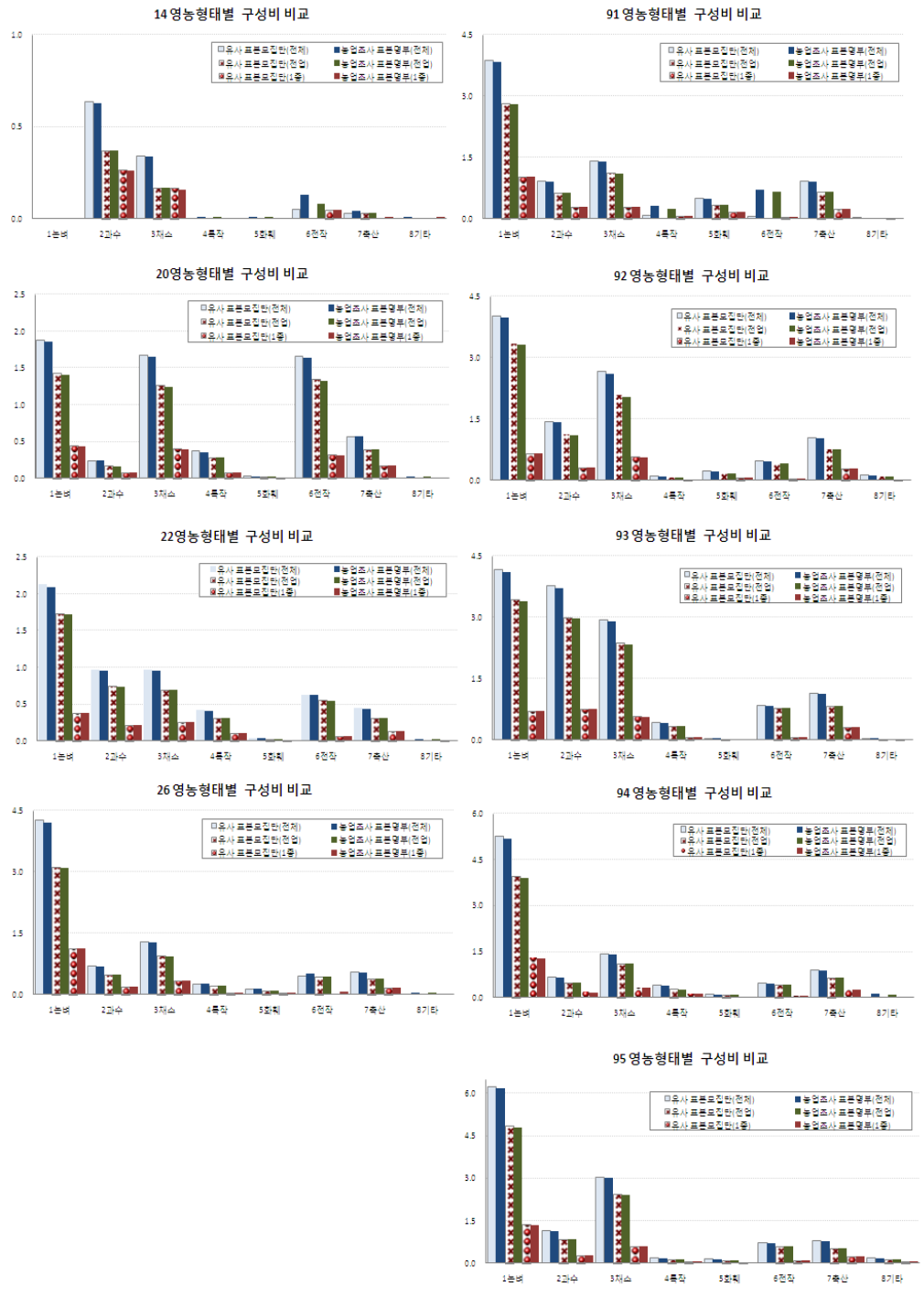
[그림 3-7] 지역별 전겸업별 농가 구성비 비교

[그림 3-8]은 전국의 영농형태별 농가 구성비를 비교한 것으로 농가경제조사 유사 표본모집단과 농업조사 표본농가 명부자료의 구성비가 거의 비슷하게 되어 있음을 알 수 있다.



[그림 3-8] 전국 영농형태별 농가 구성비 비교

[그림 3-9]는 지역별로 영농형태별 농가 구성비를 나타낸 것이다. 농가경제조사 유사 표본모집단과 농업조사 표본농가 명부자료의 영농형태별 구성비가 거의 비슷하게 되어 있음을 알 수 있다.



[그림 3-9] 지역별 영농형태별 농가 구성비 비교



<표 3-6>을 보면, ‘2010년 농가경제조사 유사 표본모집단’의 분석변수의 분포는 ‘2010년 농가경제조사 연계자료’의 분석변수와 거의 비슷한 분포임을 알 수 있다.

<표 3-6>의 ② ‘2010년 기준 농가경제조사 유사 표본모집단’의 주요 통계량에서 동일한 값을 임의추출한 농가소득과 오차를 고려한 농가소득(정규), 농가소득(일양)을 비교하면, CV, 왜도 등은 같고, 평균과 표준편차 등에서 매우 적은 차이가 있음을 알 수 있다. 평균추정값 산출 시에도 3개 자료의 차이가 거의 없었다.

<표 3-6> ‘2010년 농가경제조사 연계자료’와 ‘2010년 농가경제조사 유사 표본모집단’의 자산, 부채, 소득의 주요 통계량

① ‘2010년 농가경제조사 연계자료’의 주요 통계량

(단위: 농가, 천원)

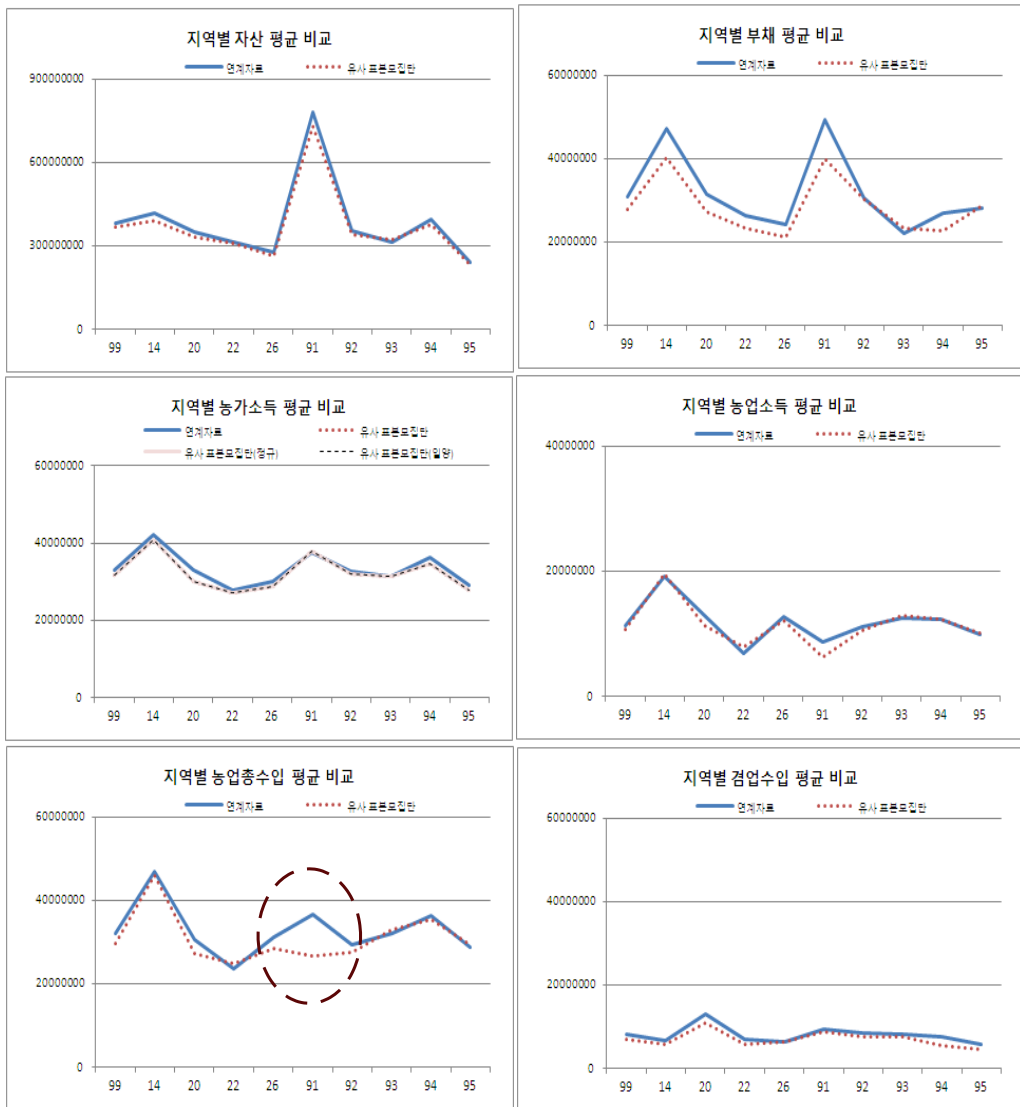
지역	전국	14	20	22	26	91	92	93	94	95
농가수	2,649	147	265	252	299	316	332	358	327	353
<평균> 자산	379,326	417,196	349,132	311,158	278,762	779,174	352,635	314,270	393,531	240,048
부채	30,816	47,304	31,542	26,439	24,150	49,457	30,633	22,032	26,914	28,184
농가소득	32,848	42,055	33,137	27,693	30,226	37,436	32,688	31,283	36,420	29,018
농업소득	11,381	19,105	12,994	6,880	12,620	8,602	11,043	12,519	12,261	9,952
농업총수입	32,084	47,022	30,542	23,517	31,048	36,749	29,361	32,023	36,201	28,643
겸업수입	8,067	6,739	13,028	6,997	6,321	9,501	8,559	8,233	7,518	5,731
<표준편차> 자산	435,608	308,226	362,276	270,215	218,496	856,959	383,195	276,857	349,551	172,168
부채	77,463	64,106	69,797	63,824	45,379	148,416	73,214	47,864	70,319	53,864
농가소득	36,445	35,972	51,136	25,474	28,525	45,300	27,687	29,710	47,955	25,529
농업소득	31,165	28,527	49,084	18,663	22,979	34,631	19,548	24,714	46,861	18,077
농업총수입	71,959	65,991	86,265	35,907	50,792	98,148	66,649	64,208	102,361	39,710
겸업수입	47,751	18,699	120,250	30,674	27,047	32,591	32,614	32,737	39,333	19,287
<왜도> 자산	5.75	1.8	4.07	2.06	2.36	3.99	3.18	2.57	2.86	2.43
부채	8.79	2.55	4.48	4.93	3.45	7.37	6.53	4.44	6.63	4.56
농가소득	6.16	1.12	12.12	2.01	2.45	1.78	2.06	3.16	5.69	3.29
농업소득	9.33	1.3	13.51	3.83	3.21	6.8	3.07	5.13	5.2	2.87
농업총수입	9.94	4.74	12.89	4.03	5.7	6.91	10.56	8.42	9	4.72
겸업수입	26.03	4.76	14.98	8.08	8.19	6.76	6.6	7.51	12.64	6.89
<CV> 자산	114.84	73.88	103.76	86.84	78.38	109.98	108.67	88.1	88.82	71.72
부채	251.37	135.52	221.28	241.4	187.9	300.09	239	217.25	261.27	191.11
농가소득	110.95	85.54	154.32	91.99	94.37	121	84.7	94.97	131.67	87.98
농업소득	273.83	149.32	377.75	271.25	182.08	402.57	177.01	197.41	382.19	181.64
농업총수입	224.28	140.34	282.44	152.68	163.59	267.07	226.99	200.5	282.76	138.64
겸업수입	591.92	277.45	922.96	438.37	427.88	343.02	381.02	397.63	523.16	336.53

② ‘2010년 기준 농가경제조사 유사 표본모집단’의 주요 통계량

(단위: 농가, 천원)

지역	전국	14	20	22	26	91	92	93	94	95
농가수	74,349	1116	6804	5717	7159	10068	10243	12461	9023	11758
<평균> 자산	368,299	39,2404	332,982	309,126	265,140	732,220	342,346	323,776	376,658	229,797
부채	27,729	40,307	27,389	23,389	21,291	39,827	30,209	23,281	22,720	28,801
농가소득	31,624	40,957	30,026	27,290	28,769	37,935	31,967	31,483	34,612	27,664
농가소득(정규)	31,624	40,957	30,026	27,290	28,769	37,935	31,967	31,483	34,612	27,664
농가소득(일양)	31,624	40,957	30,026	27,290	28,769	37,935	31,967	31,483	34,612	27,664
농업소득	10,636	19,550	11,329	7,804	12,079	6,269	10,487	12,892	12,336	10,063
농업총수입	29,720	46,061	27,172	24,874	28,490	26,688	27,654	32,995	35,438	29,289
겸업수입	6,996	5,705	10,769	5,790	6,368	8,778	7,582	7,448	5,553	4,495
<표준편차> 자산	409,739	321,187	310,796	272,329	200,597	748,102	368,711	296,652	314,078	156,641
부채	71,692	56,936	53,155	58,348	43,989	126,118	73,140	49,190	63,341	59,853
농가소득	31,754	35,103	29,958	25,443	29,541	38,395	25,455	30,463	42,838	24,394
농가소득(정규)	31,754	35,103	29,958	25,443	29,541	38,395	25,455	30,463	42,838	24,394
농가소득(일양)	31,754	35,103	29,958	25,443	29,541	38,395	25,455	30,463	42,838	24,394
농업소득	25,881	27,738	26,603	19,757	22,821	27,123	17,437	24,878	40,324	20,269
농업총수입	61,448	64,307	48,794	37,958	43,284	68,637	61,231	63,479	91,987	45,972
겸업수입	41,079	17,621	107,221	23,489	30,939	29,224	27,280	29,613	26,490	16,447
<왜도> 자산	5.77	2.16	3.75	2.04	2.44	4.48	3.01	2.74	2.98	2.33
부채	9.34	2.72	3.74	4.74	3.79	8.92	6.25	4.28	6.09	5.26
농가소득	4.24	1.35	12.18	2.52	3.24	1.31	1.81	2.95	5.92	2.41
농가소득(정규)	4.24	1.35	12.18	2.52	3.24	1.31	1.81	2.95	5.92	2.41
농가소득(일양)	4.24	1.35	12.18	2.52	3.24	1.31	1.81	2.95	5.92	2.41
농업소득	7.47	1.36	16.94	4.23	3.57	8.32	2.90	4.68	6.57	2.82
농업총수입	9.99	5.03	14.46	3.74	5.77	8.53	11.13	7.97	9.83	4.89
겸업수입	30.10	5.49	17.04	7.88	7.74	7.12	7.37	7.65	17.67	7.40
<CV> 자산	111.25	81.85	93.34	88.10	75.66	102.17	107.70	91.62	83.39	68.16
부채	258.54	141.26	194.07	249.46	206.61	316.66	242.11	211.29	278.79	207.82
농가소득	100.41	85.71	99.77	93.23	102.68	101.21	79.63	96.76	123.77	88.18
농가소득(정규)	100.41	85.71	99.77	93.23	102.68	101.21	79.63	96.76	123.77	88.18
농가소득(일양)	100.41	85.71	99.77	93.23	102.68	101.21	79.63	96.76	123.77	88.18
농업소득	243.33	141.88	234.81	253.14	188.92	432.65	166.27	192.98	326.88	201.42
농업총수입	206.75	139.61	179.58	152.60	151.93	257.18	221.42	192.39	259.57	156.96
겸업수입	587.14	308.86	995.56	405.66	485.85	332.90	359.76	397.57	477.00	365.88

[그림 3-10]은 두 자료의 지역별 자산, 부채, 소득의 평균을 그래프로 표현한 것이다. [그림 3-10]에서 부채평균의 유사 표본모집단 평균이 적게 나타났으며, 농업총수입에서 91지역 유사 표본모집단 평균이 적게 나타났으나 크지 않아 대체로 두 자료의 지역별 차이가 거의 없다고 할 수 있다.



[그림 3-10] 농가경제조사 연계자료와 유사 표본모집단 자료의 자산, 부채, 소득 평균 비교

제4절 맺음말

2010년 기준 농가경제조사 표본개편이 완전히 이루어지지 않았지만 농가경제조사의 표본개편(안)을 토대로 증화집락-증화 복합 이중추출 추정량을 검토하고 제시해 보았다. 농업조사에 이중추출방법을 처음으로 적용하는 것이므로 추정량의 효율성을 사전에 검토할 필요가 있다.

복합 이중추출 효과를 검토하기 위해서는 농가특성과 부채, 자산, 소득 등의 모집단 자료가 필요하지만, 현실적으로 존재하지 않으므로 ‘2010년 기준 농가경제조사 유사 표본모집단’을 구축하여 추정량을 모의산출하여 효율을 비교해 보고자 한다. 따라서, 현재 이용가능한 최근 자료인 ‘2010년 농가명부조사 자료’, ‘2010년 농가경제조사 집계자료’, ‘2011년 농업조사 표본농가 명부자료’를 자료연계하여 유사 표본모집단을 구축하였다.

연계과정을 간략하게 설명하면, ‘2010년 농가명부조사 자료’와 ‘2010년 농가경제조사 집계자료’를 연계하여 농가경제조사 집계농가의 모집단 특성을 부여하고, ‘2010년 농가경제조사 연계자료’와 ‘2011년 농업조사 표본농가 명부자료’를 지역별×전겸업별×영농형태별로 구분하고 자료연계하여 ‘2010년 기준 농가경제조사 유사 표본모집단’을 구축하였다. 지역별×전겸업별×영농형태별로 구분된 ‘2011년 농업조사 표본농가 명부자료’의 해당 농가수만큼 ‘2010년 농가경제조사 연계자료’에서 반복 임의추출하여 유사 표본모집단을 구축하였다.

향후 연구에서는 본 연구에서 구축한 ‘2010년 기준 농가경제조사 유사 표본모집단’에서 복합 이중추출로 2차 표본농가를 증화추출하고, 제2절에서 제시된 평균 추정량을 적용하여 평균 추정량의 대표성을 검토하고, 더 나아가 제시된 분산추정량과 반복표본에 의한 분산추정량의 안정성과 효율성 등을 비교하고 검토할 계획이다.

참고문헌

- 박재수 (1996). 표본조사론. 박영사.
박홍래 (1997). 통계조사론. 영지문화사.
신민웅 외 (2001). 표본조사를 위한 표본설계. 교우사
Cochran, W. G. (1977). Sampling Techniques (3rd edition). New York: John Wiley & Sons, Inc.
Wolter, K. M. (2003). Introduction to Variance Estimation. Springer Verlag.

