캐나다의 노동패널조사

(The Survey of Labour and Income Dynamics in Canada)

1998. 1

한국노동연구원

고용보험연구센터 금재호

目 次

Ⅰ. 캐나다 노동패널조사(SLID)의 배경과 목적
1. 패널조사의 의의
2. SLID의 배경
3. SLID의 목적 및 주요 연구주제

Ⅱ. 패널설계와 표본추출방법
1. 패널설계
2. 표본선정방법
3. 캐나다 노동력조사(LFS)의 표본설계

Ⅲ. 설문조사방법과 패널관리
1. 조사방법
2. 표본추적
3. 응답자 및 면접원 관리

Ⅳ. 자료의 관리
1. 가중치의 결정
2. 자료의 편집과 무응답자 처리

3. 개인 비밀보호

Ⅴ. 설문내용

1. 개인적 특성

2. 교육

3. 소득과 자산

4. 노동시장활동

Ⅵ. 한국노동패널연구(KLPS)에 대한 시사점

參考文獻

Ⅰ. 캐나다 노동패널조사(SLID)의 배경과 목적

1. 패널조사의 의의

경제분석에 가장 많이 사용되는 데이터로 시계열자료(time-series data)와 횡단면 자료(cross-section data)의 두 가지가 있다. 시계열자료는 국민총생산이나 이자율과 같이 시간에 따라 변동하는 자료로 국민경제의 거시적 분석에 주로 사용된다. 이에 반하여 횡단면자료는 가계소득의 분포나 지역별 실직자수와 같이 어떤 일정 시점을 기준으로 각각의 표본(sample)들이 어떤 값을 지니는가를 보여주는 자료로 주로 미시적 분석에 사용된다.

패널자료는 시계열자료와 횡단면자료의 성격을 모두 지니는 데이터로 특히 (a) 동일한 개인 및 가구들을 표본(sample)으로 하여, (b) 이들을 지속적으로 추적하고, (c) 일정 기간(주로 1년)마다, (d) 동일한 질문을 하여 얻어지는 데이터를 종단면자료(longitudinal data)라 한다. 이러한 패널자료는 노동시장분석에 주로 사용되는 횡단면자료의 장점을 갖추고 있으며, 동시에 횡단면자료로는 얻기 어려운 추가적인 장점을 갖추고 있다. 즉, 사회에 진출한 젊은이들이 안정적인 직장을 얻을 때까지의 이동과정이나 육아장려금이 주부의 취업에 미치는 영향, 그리고 실업급여의 변화가 실업률에 미치는 영향 등과 같은 분석에 종전의 횡단면자료는 근본적인 한계를 가지고 있으며, 이에 패널데이터가 크게 기여할 수 있다.

경제현상의 분석 및 정책결정에 미치는 패널데이터의 유용성과 중요성이 널리 알려진 반면, 실제로 패널조사가 이루어지는 경우는 많지 않다. 이의 가장 큰 이유는 패널조사에 상당한 비용이 소요되기 때문인데, 특히 지속적으로 추적하여야 하는 조사대상가구들의 관리 및 유지에 많은 비용이 소요된다. 실사에서도 사전

적으로 지정된 가구만을 대상으로 소득, 자산 등 민감한 부문을 질문하기 때문에 면접원 비용 등 실사에 소요되는 비용이 다른 조사에 비하여 훨씬 많이 들게 된다. 따라서, 미국, 영국, 독일, 캐나다와 같은 선진국들을 중심으로 패널조사가 실시되고 있다. 우리나라에서는 1993년부터 대우경제연구소가 한국가구패널조사(KHPS; Korea Household Panel Survey)를 실시하고 있는데 소득, 소비, 자산, 그리고 고용 등 가구 및 개인의 경제생활 전반에 관한 데이터를 제공하고 있다.

2. SLID(The Study of Labour and Income Dynamics)의 배경

1980년경부터 캐나다 통계청은 개인 및 가족의 노동시장활동에 관한 패널데이터를 수집하기 시작하였는데 특히, 1980년대 후반의 Labor Market Activity Survey(LMAS)는 캐나다 노동패널조사의 기원을 이루고 있다. LMAS는 1986~87년까지의 제1패널과 1988~1990년까지의 제2패널의 둘로 구성되어 있었으며, 이를 발전시킨 것이 현재의 SLID이다.

SLID는 1993년 1월에 제1차 조사를 시작하였는데 1차년도에는 응답자의 부담을 줄이고, 면접조사를 용이하게 하기 위하여 인구학적·개인적 배경에 관한 질문들이 주된 내용을 이루었다. SLID는 국민들의 경제적 풍요를 측정하기 위하여 종전의 LMAS에서는 묻지 않던 소득에 관한 항목들을 추가하고 있고, 자산(wealth)에 관한 질문들도 일부 던지고 있다. 이외에도 SLID는 개인 및 가족들의 노동시장활동과 소득에 영향을 미치는 다양한 내용들을 질문하고 있다.

SLID는 비록 연륜은 짧으나 미국의 National Longitudinal Survey of Youth(NLSY)와 함께 대표적인 노동패널조사로의 위치를 공고히 하고 있다. 미국의 NLSY가 1979년부터 14~22세까지의 청소년을 대상으로 이들을 계속 추적하는 개인조사인 반면, SLID는 연령과 상관없이 패널로 선정된 가구(패널가구)에 속한 16세 이상의 모든 가구원을 조사하는 가구조사라는 점에 가장 큰 차이점이 있다. 14~22세의 동일 연령층을 대상으로 하는 NLSY는 동시태생효과(cohort effect)를 정확하게 파악할 수 있는 장점을 지니고 있다. 그러나, 그 조사결과는 전국의 모집단을 대표하지 못하고 시간이 지날수록 응답자들이 노화되어 은퇴

또는 사망하게 됨에 따라 응답자수가 급격히 줄어든다는 문제점이 있다.

SLID 방식의 패널조사는 동시태생효과 및 동류집단효과를 정확하게 분석하기 위해 표본수가 매우 커야한다는 한계점을 가지고 있지만 분석결과가 전국을 대표하는 대표성을 지니고 있고, 결혼, 출산 등의 과정을 통하여 새로운 표본이 계속 충원된다는 특징이 있다.

노동관련 통계가 아직 충분치 못한 우리나라의 현실에서는 NLSY와 같이 특정 연령이나 계층만을 대상으로 하는 패널조사는 시기상조로 여겨지며, 이러한 측면에서 캐나다의 SLID는 우리나라의 노동패널조사(KLPS)의 설계와 발전방향에 많은 시사점을 제공하고 있다.

3. SLID의 목적 및 주요 연구주제

SLID는 개인과 가족들의 경제적인 풍요(well-being)와 이러한 풍요에 영향을 미치는 요인들에 대한 연구를 기본 목적으로 하고 있으며, 보다 구체적으로 다음과 같은 목적들을 들 수 있다.

① 소득의 동태적 분석(income dynamics) : 개인 및 가족의 소득에 영향을 미치는 결정요인의 분석, 빈곤의 동태적 요인분석
② 노동시장의 동태분석(labor market dynamics) : 학교에서 직장으로의 전이 등 노동시장에서 라이프사이클 움직임에 대한 분석
③ 가족의 동태분석(family dynamics) : 특히, 경제적 풍요에 관한 분석과 이해

이처럼, SLID는 소득, 노동시장, 가족 등의 다방면에 걸쳐 그 동태적 변화의 양상을 이해하고 분석하는데 기본목적을 두고 있다. 또한, 실업, 임금, 직장생활 등에 관한 다양한 횡단면 데이터를 제공하고 있다. 조사표본의 기본단위가 개인이긴 하지만 가족 구성원들은 공통의 경제·사회적 환경에 부딪치고 있으며, 가족의 부나 소득을 함께 영위하기 때문에 가구의 특성 및 그 동태적 변화에 대한 조사와 분석이 중요한 과제의 하나로 자리잡고 있다.

SLID의 데이터를 이용하여 분석, 연구할 수 있는 분야는 매우 광범위하다. 그 중에서도 다음과 같은 과제들이 주된 연구분야로 관심의 대상이 되고 있다.

① 고용 및 실업의 동태적 변화
② 노동시장에서의 이동과정
③ 직장생활
④ 가구경제의 동태적 변화
⑤ 저소득층 가구의 동태적 변화
⑥ 가족사
⑦ 학교와 직장의 연계

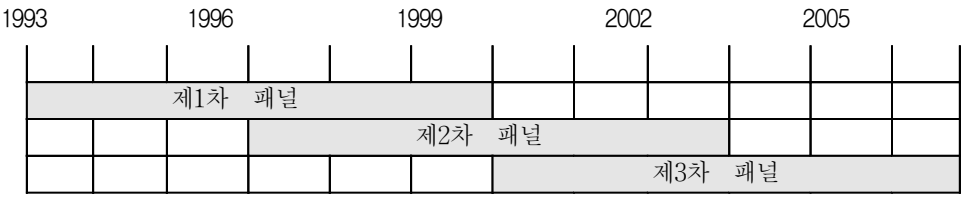
Ⅱ. 패널설계와 표본추출방법

1. 패널설계(Panel Design)

일단 패널에 포함된 가구 또는 개인은 패널이 지속되는 한 표본에 계속 머물러 있는 것이 일반적이다. 그러나, 캐나다의 SLID의 경우 표본으로 선정된 가구는 6년동안만 패널에 머물러 있게 된다.

1993년에 최초로 시작된 SLID는 노동력조사(LFS : Labor Force Survey)의 표본에서 추출한 15,000가구로 제1차 패널을 구성하였는데 1995년까지의 초기 3년 동안은 표본수의 변화없이 제1차 패널만으로 패널조사를 실시하였다. 그러나, 1996년 LFS에서 추출한 제2차 패널이 기존의 제1차 패널에 추가됨에 따라 1996년부터 제1차 패널이 완료되는 1998년까지의 3년 동안은 제1차 및 제2차의 2개 패널이 동시에 유지된다. 또한, 1998년을 마지막으로 6년 동안 유지되던 제1차 패널이 탈락하게 되고 1999년에 제3차 패널이 새로 참가하게 됨에 따라 1999년부터 SLID는 제2차 및 제3차의 두 패널이 구성될 예정이다.(<표 1> 참조)

<표 1> SLID의 표본설계



이처럼, 1996년부터 SLID는 2개의 패널로 이루어지고 매 3년 마다 패널구성이 변동하게 되는데 이와 같이 표본설계한 주된 이유는 표본의 대표성을 보장하고,

조사내용의 유연성을 도모하기 위한 것이다. 또한, 응답 가구의 부담을 감안하여 패널이 교체되는 시기를 6년으로 정하였다.

2. 표본선정(Sample Design)방법

SLID의 모집단은 캐나다 노동력조사(Canadian Labor Force Survey)의 모집단과 거의 일치하며, 추가적으로 아동들에 대한 정보를 수집하고 있다. SLID는 연령과 상관없이 캐나다에 거주하고 있는 모든 사람들을 포함하고 있으며, 지역별로는 아래의 지역을 제외한 모든 지역을 포괄한다.

- Yukon과 Northwest Territories
- 사회보호시설의 6개월 이상 장기수용자[1)]
- 인디안 보호지역
- 군부대의 영내거주자

SLID의 표본은 캐나다 노동력조사(LFS: Labor Force Survey)의 표본에서 추출하였다. LFS는 순환그룹(rotation group)으로 불리는 6개의 표본그룹으로 구성된다. 매달 조사를 실시하는 LFS의 표본으로 선정된 가구는 6개월 동안 표본에 머물러 있게 되며, 매달 전체 표본의 1/6이 새로운 표본으로 대체된다.

SLID의 제1차 패널을 구성하기 위해 LFS의 6개 순환그룹 중에서 2개의 그룹을 먼저 선정하였다. 이 2개의 그룹은 약 20,000가구로 구성되어 있으며, 이들의 88% 정도인 17,000가구가 SLID의 표본이 될 것을 승낙하였다. 예산제약으로 인하여 이들 중 2,000가구를 제외한 나머지 15,000가구로 제1차 패널을 구성하였다. 이들 15,000가구에 속한 15세 이상의 남·녀는 모두 31,000명에 달하였다.

3. 캐나다 노동력조사(LFS)의 표본설계

1) 사회보호시설에는 고아원 등 아동보호시설, 장기요양소, 양로원, 병원, 정신병원, 장애인 수용시설, 교도소, 소년원, 유치장 등이 포함되며 학교기숙사, 수도원, 수녀원 등은 제외된다.

LFS는 지역을 기준으로한 다단계층화추출법을 통하여 표본을 추출하고 있다. 표본선정에 10년마다 실시하는 인구조사(census)의 결과를 반영하기 위하여 매 10년마다 표본을 재설계한다. 이에 따라, SLID의 제1차 패널은 1981년도의 인구조사를 토대로 선정되었고, 제2차 패널은 1991년도의 인구조사에 기초하고 있다.

구체적인 표본의 설계방법은 먼저 각 지방(province)을 여러개의 경제지역(economic regions)으로 구분하고, 다시 각 경제지역(economic region)을 한 개 이상의 SRUs(self representing areas)와 한 개의 NSRU(non-self representing area), 그리고 여러개의 특별구역(special areas)으로 나눈다. SRU는 해당 구역내 주민들의 노동시장 특성이 같은 인구 15,000명 이상의 도시구역이 대부분을 차지한다. NSRU는 SRU가 아닌 economic region의 모든 지역을 의미하고, 특별지역(special area)는 병원이나 학교내 주거지역, 군부대내 민간인 거주지역, 또는 산간벽지 등 특수한 지역으로 전체인구의 1% 정도를 차지하고 있다.

각각의 SRU는 실업률, 취업률, 산업구조, 인구의 연령별 분포, 자가(自家)보유율 등 다양한 기준에 따라 동일한(homogeneous) 사회·경제적 특성을 지니는 집단으로 다시 층화(statification)된다. 층화된 각각의 집단은 다시 거주지역에 따라 구별되는데 이러한 세부 집단(cluster)을 기본표본단위(PSUs: primary sampling units)라고 한다.

이러한 층화방법은 NRSU에도 마찬가지로 적용되는데 지역의 규모가 큰 NRSU의 경우에는 먼저 지역을 도시지역과 비도시지역으로 층화한 다음, 다시 동일한 사회·경제적 특성을 지니는 집단으로 층화(stratification)한다. 특별지역은 지역의 특수성에 따라 세부적으로 나누어지고, 이렇게 나누어진 세부 집단은 SRU의 기본표본단위와 같은 역할을 하게 된다.

SRU, NSRU, 그리고 특별지역에서 층화된 세부 집단(cluster)들 중 일부를 먼저 추출하고, 추출된 세부 집단(cluster)에 속한 가구들 중에서 임의로 조사대상 가구들을 선정하게 된다. 이러한 다단계층화추출법은 전국적 규모의 실태조사들이 가장 많이 사용하고 있는 방법으로 우리나라의 한국가구패널조사(KHPS)나 미국의 대표적 패널조사인 PSID(Panel Survey of Income Dynamics)도 이러한

표본추출 방법을 채택하고 있다.

Ⅲ. 설문조사방법과 패널관리

1. 설문조사방법(collection method)

가. 일반적 조사방법(general method)

(1) SLID의 설문조사는 매년 2회에 걸쳐 인터뷰를 실시한다. 먼저 매년 1월에 전년도의 노동시장활동을 조사하고, 전년도의 세금보고(tax report)가 끝난 5월에 전년도 소득에 대한 소득조사를 실시한다.

(2) 또한, 각각의 패널에 대하여 포함된 가구 및 개인들의 사회·인구적 배경과 특성을 파악하기 위한 예비조사(preliminary survey)를 1회 실시한다. 1993년도의 첫 번째 패널에 대하여는 1993년 1월의 노동력조사(LFS)에 추가하여 패널 본조사와는 독립적으로 예비조사를 실시하였다. 그러나, 1996년에 시작되는 두 번째 패널부터는 비용절감을 목적으로 예비조사를 따로 실시하지 않고 첫 번째의 노동조사와 함께 실시하도록 하고 있다.

(3) 따라서, 한 번 패널에 포함된 가구 및 개인들은 패널에 머무르는 6년 동안 6차례의 노동조사와 6차례의 소득조사에 대하여 응답하게 되며, 첫 번째 노동조사시 한 차례의 예비조사에 대하여 함께 응답한다.

(4) 가능한 동일한 조사원이 매년 같은 응답자와 접촉하도록 하고 있으며 노동력 조사(LFS)의 조사원이 패널조사도 겸임하도록 유도하고 있다.

(5) 조사방법은 컴퓨터를 이용한 전화조사(computer assisted interviewing: CAI)가 주로 사용되고 있으며, 수집된 응답결과는 일반적으로 모뎀(modem)을 통하여 근처의 통계청 사무소로 전송된다. 이 때, 보안상의 이유로 데이터는 암호화되어 보내진다.

(6) 조사는 일반적으로 한 가구당 한 사람씩과 행하여지기 때문에 그 사람이 다른 가구원에 관련된 내용도 응답하는 대리응답(proxy response)이 허용된다. 만약, 응답자가 다른 가구원에 대한 충분한 정보를 갖고 있지 못한 경우에는 조사원은 해당 가구원과 직접 접촉하여 조사를 시도한다.

(7) 매 조사기간 전에 'respondent newsletter'가 표본가구에 보내어지는데 이는 조사의 실시를 통고함과 동시에 조사에 대한 협조를 당부하고, 조사시의 유의사항을 전달하려는 목적을 가지고 있다. 또한, 조사가 끝난 이후인 매년 7월 감사편지를 발송한다. 더불어, 응답거절자에게도 다음 인터뷰의 참여를 유도하는 편지가 때때로 보내진다.

나. 소득조사의 방법(special methods for income data)

소득에 관한 자료의 질을 향상시키기 위하여 SLID에서는 다양한 방법을 사용하고 있다. 소득에 관한 질문은 응답자들이 전년도의 소득보고(tax report)를 마쳐 전년도의 소득에 관하여 정확하게 기억하고 있는 5월에 노동조사와는 독립적으로 실시된다. 응답자들은 소득조사가 실시되기 2주전 소득인터뷰에 관한 정보가 담긴 편지를 받게 된다. 이는 응답자들이 자신 및 가구구성원의 소득에 관한 자료와 정보를 검토할 시간적 여유를 주기 위한 것이다.

소득에 관한 데이터는 직접 인터뷰를 통하는 방법과 국세청의 세금자료를 얻어 사용하는 방법의 두 가지가 있는데, SLID는 응답자들이 어떤 방법을 택할지 스스로 선택하게 하고 있다. 국세청에 제출하는 세금보고서의 열람을 허락하는 방법은 다음해에도 소득에 관한 인터뷰를 반복적으로 하지 않아도 되고, 또한 조사비용의 절감, 응답자의 조사에 대한 부담의 경감, 그리고 양질의 자료를 획득할 수 있게 하는 장점을 가지고 있다. 1995년의 경우 과반수를 넘는 63%의 응답자들이 자신의 세금보고서 내용을 통계청에서 열람할 수 있도록 동의하고 있었다.

2. 표본추적

가. 조사대상(population coverage)

패널이 선정되어 최초로 예비조사를 실사하였을 당시 표본가구에 거주하고 있던 개인들은 모두 조사대상이 되며, 이들을 종단적응답자(LR: longitudinal respondent)라고 한다. SLID는 패널이 유지되는 6년 동안 별거, 분가, 이혼, 이사 등의 변동과 상관없이 지속적으로 이들을 추적하여 조사한다. 패널구성 당시에는 가구구성원이 아니여서 종단적응답자는 아니지만 이후의 표본조사에서 종단적응답자와 함께 거주하는 개인들을 cohabitant라고 하며, 종단적응답자와 함께 거주하는 동안은 이들도 조사대상에 포함된다. 이후, 어떤 이유로든지 종단적응답자와 떨어져 거주하게 되는 경우에는 이들 cohabitant는 조사대상에서 탈락된다. 조사대상에서 벗어난 cohabitant에 대하여는 비록 설문조사를 실시하지는 않지만 추후 다시 종단적응답자와 함께 거주하여 조사대상에 포함될 가능성에 대비하여 그 소재지를 가능한 범위내에서 추적을 계속한다.

cohabitant들을 조사대상에 포함시키는 이유는 첫째 종단면응답자(longitudinal respondent)의 가족 및 가구에 대한 자료의 충실도를 높이고, 둘째 매년 실시되는 조사의 횡단면적(cross-sectional) 대표성을 유지하기 위한 것이다. 또한, 15세 미만의 아이들은 조사하지 않지만 부모를 통하여 지속적으로 그들의 신상에 관한 정보를 축적한다.

나. 주거이동의 추적(following movers)

종단적응답자(LR: longitudinal respondent)가 이사한 경우에도 일반적으로 계속 추적조사를 한다. 특히, Yukon, Northwest territories, 또는 미국으로 이주하거나 아니면 군부대나 인디안보호구역내로 이주한 경우에도 비록 표본선정지역이 아닌 장소로 이주하였지만 가능한 범위내에서 지속적으로 조사를 실시한다. 그러나, 미국이 아닌 다른 해외로 이주하였거나 각종 사회보호시설내로 이주한

경우에는 거주지에 대한 추적만 하고, 조사는 실시하지 않는다.

사망한 경우를 제외하고 모든 종단적응답자(longitudinal respondent)를 추적하며 비록 추적이 끊기거나 조사범위 밖의 지역으로 이동하더라도 표본에서 탈락시키지는 않는다. 추후, 이들이 다시 조사범위 대상지역으로 돌아왔을 때 다시 조사를 실시한다.[2)]

3. 응답자 및 면접원 관리

가. 응답자 관리

캐나다 통계청의 정책은 자료수집을 위해서 응답자에게 대가를 지불하지 않는 것이다. 통계청에서는 대가의 지불이 응답률을 높인다는 결정적인 증거가 없고 응답자에 대한 대가의 지불에 많은 비용을 지출하는 것에 대한 정당한 이유를 발견하지 못하고 있다.

따라서, 응답률을 유지하기 위해 SLID에서는 대가를 지불하는 대신 응답자와의 역동적(dynamic) 관계를 추구하거나 응답자가 6년 동안의 조사에 참여하도록 유도하고 면접원의 사기를 진작시키는 방법의 개발에 많은 투자를 하고 있다. 응답률을 유지하도록 도와주는 두 가지 요인으로 SLID에서는 ① 흥미롭고 재미있으며 정보를 주는 동기유발 자료와 ② 응답자와 면접원 사이의 인간적 관계를 중요시한다.

응답자의 동기유발을 위한 자료는 응답자들의 관심을 끌 수 있으며 동시에 정부예산의 낭비라는 비난으로부터 자유로울 수 있어야 한다. 이러한 입장에서 자료는 두 가지의 범주로 나누어질 수 있다.

---

2) SLID에서는 조사대상 가구원의 연령을 16세로 정하였지만 첫해년도의 조사에 한해 15세 이상의 모든 가구원을 조사하였다. 이러한 이유는 15세 청소년들이 다음해부터는 조사대상으로 포함되기 때문에 이들에 대한 인구,사회학적 기초자료를 수집하기 위한 것이였다.

○ 응답유도를 위한 자료(promotional material): 안내책자, 편지 등
○ 수집된 자료의 분석 결과 및 활용에 관한 자료

응답자의 관리를 위해 SLID에서는 채택하고 있는 방법은 다음과 같이 정리될 수 있다.

1) 1월의 노동시장활동조사

조사시작 전

○ 첫 해년도에는 조사대상자에게 조사의 목적 및 시기를 알리는 편지나 안내책자를 발송하였다.

○ 1994년 1월에 SLID의 조사대상자들은 편지와 안내책자, 그리고 회신용 응답카드를 제공받았다. 응답자들이 SLID에 대하여 더 많은 정보를 알고 싶어하는 경우 회신용 응답카드를 중앙사무소에 보내도록 요청하였고, 이에 대해 650개 이상의 카드가 돌아왔는데 이는 전체 표본가구주의 4% 이상에 해당하였다. 더 많은 정보를 원하는 응답자에게는 통신문(newsletter)이 우송되었다.

○ 2차년도부터는 편지와 응답자 통신문이 발송되었는데 통신문에서는 조사결과에 대한 분석결과들이 설명되어 있다. 이러한 분석결과들은 응답자들이 이해하기 쉽게끔 표와 그림을 주로 사용한다.

○ 이와 함께 주소변경카드가 보내어 진다.

2) 5월의 소득조사

4월경

○ 조사의 목적과 조사시점을 알리는 사전 예보 편지가 발송된다.
○ 1월의 노동조사를 거절한 사람들에게 설득 편지를 보낸다
○ 전화면접조사 이전 응답자들이 세금보고서나 다른 관련 자료를 응답에 이용하도록 질문지를 보낸다.
○ 응답자의 질문에 대답할 뿐만이 아니라 응답자들이 질문에 정확하게 대답하도록 안내한다.

3) 조사완료 이후 6월

○ 감사편지나 카드를 응답자에게 우송한다.
○ 주소변경카드를 제공한다.

나. 면접원 관리

SLID에 대한 가능한 최대한의 정보를 면접원에게 제공하는 것은 중요한 것으로 면접원이 응답자들의 질문에 충실히 답변할 수 있도록 SLID의 내용을 잘 인지하는 것이 응답률의 유지에 긍정적 효과를 준다. 면접원이 SLID의 내용을 잘 인지할수록 응답을 꺼리는 조사대상자를 보다 잘 설득할 수 있었다.

또한, 면접원의 동기유발을 위해 팀의 한 부분이라고 느끼게 하는 것이 중요하다. 이를 위해서 SLID에서는 다음의 방법이 제안되었다.

○ 면접원 통신문을 통해서 SLID에 대해서 알리고, 통신문에 개재할 현장 경험을 요청한다. 이러한 경험을 면접원 서로에게 알려 그들이 팀의 일원임을 느끼도록 한다.

○ 면접원에게 책임자의 서명이 된 감사편지를 보내고 통신문에 그들이 직무를 잘 수행하는데 대한 감사의 글을 싣는다.

Ⅳ. 자료의 관리

1. 가중치의 결정

SLID는 다음과 같은 4개의 가중치를 가지고 있다.

(1) 종단면가중치(longitudinal weight)

종단면가중치(longitudinal weight)는 표본선정 당시의 인구분포를 반영하고 있으며, 시간에 걸쳐 자료를 비교분석하는 시계열 분석에 주로 적용된다. 종단면가중치는 매년 무응답자와 탈락자(attrition)를 반영하도록 다시 계산된다. 제1차 패널의 경우 1992년 12월말 인구분포의 가중치가 종단면가중치(longitudinal weight)이다. SLID가 LFS로부터 표본을 추출하였기 때문에 LFS의 가중치를 이용하여 종단면가중치(longitudinal weight)를 계산한다.

(2) 부분표본가중치(sub-sampling weight)

부분표본가중치(sub-sampling weight)는 소득, 부 등 민감한 질문사항에 대하여 비밀보장 등의 이유로 표본의 일부만이 응답한 경우 적용되는 가중치로 일반적으로 종단면가중치(longitudinal weight)를 대신하여 사용된다.

(3) 횡단면가중치(cross-sectional weight)

횡단면가중치(cross-sectional weight)는 매년 조사된 데이터의 횡단면 분석에 사용되는 가중치로 패널이 구성된 첫해를 제외하고는 매년 다시 계산하여야 한다. 패널이 구성된 첫해에는 종단면가중치(longitudinal weight)와 횡단면가중치가 동일하지만 이후의 연도부터는 새로운 표본의 진입과 이탈로 인하여 두 가중치

는 서로 달라지게 된다. 예를 들어, Cohabitant는 종단면가중치의 계산에는 포함되지 않지만 횡단면가중치의 계산에는 포함된다.

(4) 횡단면노동가중치(cross-sectional labor weight)

횡단면노동가중치(cross-sectional labor weight)는 매년 1월에 실시하는 노동조사에 응답하지 않는 개인들이 있을 때, 횡단면가중치를 대체하여 사용하는 가중치이다. 응답을 하지 않는 경우는 조사에 응하였으나 특정 질문에 대하여 무응답인 경우와 노동조사 자체에 대하여 응답하지 않는 경우의 둘로 구별된다. 따라서, 분석에 횡단면가중치를 적용하는 것은 분석결과를 왜곡시킨다. 이를 수정하기 위하여 횡단면노동가중치를 사용한다. 노동조사의 분석에서는 이 가중치를 이용하는 것이 일반적으로 직업의 특성, 결근율, 실직기간, 교육활동 등의 인구학적인 자료와 소득 자료가 필요하지 않은 분석의 경우 주로 사용한다.

2. 자료의 편집과 무응답자 처리

가. 자료의 가공단계(processing steps)

설문조사에 의하여 수집된 결과는 다음과 같은 몇 가지의 단계를 거쳐 분석가능한 데이터로 전환된다.

(1) 사전가공(preprocessing) : 일련의 수작업 단계로 데이터의 정확성을 위해 면접자가 설문조사의 내용을 다시 검토하는 과정이다. 중복질문을 발견하여 제거하고, 가구 구성원의 변화, 이사한 날 등 인구학적 조사내용의 일관성과 정확성을 확인한다.

(2) 로딩(loading) : 수집한 자료를 변수(variable)의 값으로 변환하는 과정으로 데이터베이스화하는 단계이다.

(3) 코딩(coding) : 문서화된 자료를 분석에 사용되는 수적인 부호로 전환하는 단계이다.

(4) 편집(editing) : 전통적인 조사에 비교해서 컴퓨터로 수집된 인터뷰는 수집

단계에서 많은 편집이 이루어 진다.

(5) 전가(imputation) : 응답하지 않은 항목에 과거의 응답결과나 또는 다른 응답자들의 경우를 참작하여 알맞는 값을 임의로 부여하는 과정이다.

나. 편집(editing)

편집단계의 가장 중요한 점은 전체 패널의 자료수집이 끝나기 전에는 모든 편집이 예비적이라는 사실이다. 즉, 더 많은 자료가 수집됨에 따라 편집내용의 변화가 발생하게 된다.

더불어, SLID에서는 컴퓨터를 이용한 전화조사(CAI : computer assisted interviewing)를 통하여 응답내용을 수집하기 때문에 on-line 편집이 가능하다. 이에따라, 이전 응답으로부터 질문의 순서나 흐름을 조절하고 일관성을 검사할 수 있다. 즉, 컴퓨터가 현재의 질문내용뿐만 아니라 이전의 조사들에서 응답자가 유사한 질문에 어떻게 대답하였는지를 즉각적으로 알려주기 때문에 현재의 응답이 이전의 응답들과 일치하는 지를 쉽게 판단할 수 있다.

또 하나의 문제는 1월의 노동조사에서 응답된 내용과 5월의 소득조사에서 응답한 내용들과의 상호일치성 문제이다. 노동조사에서 질문한 직업에 대한 정보와 소득에 대한 정보를 비교하여 자료에 모순되는 점이 있는지 확인해야 한다.

다. 응답률(response rates)

응답률은 응답자의 기준을 어디에 두는가에 따라 횡단적응답률(cross-sectional response rate)과 종단적응답률(longitudinal response rate)의 둘로 나누어 진다.

횡단적인 응답률은 개인보다는 가구를 단위로 계산하는데 종단적응답자(longitudinal respondent)가 현재 거주하고 있는 가구들을 대상으로 그 가구에 속한 가구원들 중 최소한 한 명이라도 조사에 응하였으면 그 가구는 조사에 응한 것으로 간주한다. 이 때 조사에 응한 응답자가 종단적응답자가 아닌 cohabitant일지라도 응답가구로 간주한다.

종단적응답률(longitudinal response rate)의 계산에는 단지 종단적응답자(longitudinal respondent)만이 포함된다. 즉, 조사에 응한 가구원들 중 최소한 한 명이 종단적응답자이여야 그 가구는 응답가구로 간주된다. 따라서, 일반적으로 종단적 응답률은 횡단적 응답률보다 낮은 것이 보통이며, SLID는 두 가지의 응답률 중에서 불필요한 혼동을 피하기 위하여 횡단적 응답률만을 공표하고 있다.

응답률에 대한 자료를 보면 조사가 반복될수록 횡단적 응답률(cross-sectional response rate)은 증가하며, 반대로 종단적 응답률(longitudinal response rate)은 감소하는 추이를 발견할 수 있다. 한편, 소득조사의 경우 첫 번째 조사에서는 76%의 가구만이 응답을 했으나 다음해에는 88.5%의 응답율을 보였고 점점 응답률이 증가하여 결국 95%가 넘는 응답률을 보였다. 컴퓨터를 이용한 전화조사(CAI)는 종이와 연필에 의한 직접조사에 비하여 응답률을 높이는데 커다란 기여를 한 것으로 나타나고 있다.

라. 전가(imputation)

설문항목들의 일부분은 응답자가 대답하지 않았을 때 전가(imputation)할 수 있는 것이 있다. 전가(imputation)는 무응답된 항목에 허위이지만 합리적인 값을 부여하는 방법이다. 만약 응답자가 어떤 항목에 대하여 응답하지 않았다 해도 관련된 정보를 통하여 간접적으로 알 수 있는데 예를 들어, 거주지역, 나이, 성별 등의 변수들은 과거의 조사결과를 통하여 간접적으로 파악할 수 있다. 특히, SLID의 표본이 노동력조사(LFS)를 바탕으로 하고 있기 때문에 노동력조사(LFS)의 결과를 통하여 무응답된 항목의 값들을 알 수 있다. 소득에 관한 인터뷰 정보의 경우 전가(imputation)는 가장 근접한 이웃의 경우를 미루어 판단하는 방법을 사용하였다. 즉, 가장 유사한 특성을 지닌 개인 또는 가구를 찾아 그들의 소득정보를 대신 사용하는 방법을 채택하고 있다.

3. 개인 비밀보호(confidentiality)

SLID는 개인의 비밀을 보호하기 위해 다음과 같은 자료들은 공포(公布)된 데이터에서 제외하고 있다.

(1) 개인의 신원을 확인할 수 있는 이름, 전화번호 등의 모든 변수는 파일에서 삭제했다.

(2) 지방(province) 수준 이하의 모든 지역에 관한 세부자료는 파일에서 삭제하였다.

(3) 생년월일과 같은 민감한 변수는 그 값을 범주화하였다.

(4) 범주화된 값이라도 특정 개인 및 지역을 식별할 수 있는 경우에는 범주를 통합하였다.

(5) 아주 높거나 아주 낮은 값을 가지는 경우는 드물게 발생하기 때문에 사전에 결정한 상하한의 값으로 대체하여 발표한다.

(6) 소득액 등과 같은 몇가지의 값에 대하여는 어림수로 대체하였다.

(7) 종교, 직업, 가족수, 아이수 등 몇가지의 값을 함께 종합적으로 분석할 때 응답개인이나 가구가 누구인지를 찾아낼 수 있는 위험성이 있어 이러한 변수들의 일부를 제거하였다.

Ⅴ. 설문내용

SLID에서는 개인의 인구·사회적 배경, 가정환경, 교육, 소득, 자산과 부채, 취업, 직장생활 등 16개의 주제에 걸쳐 다양한 질문을 하고 있는데 이러한 질문내용들은 크게 ① 인구·사회적 특성, ② 교육, ③ 소득, 그리고 ④ 노동시장활동의 네 분야로 구분될 수 있다.(<표 2> 참조)

1. 개인적 특성(Personal Characteristics)

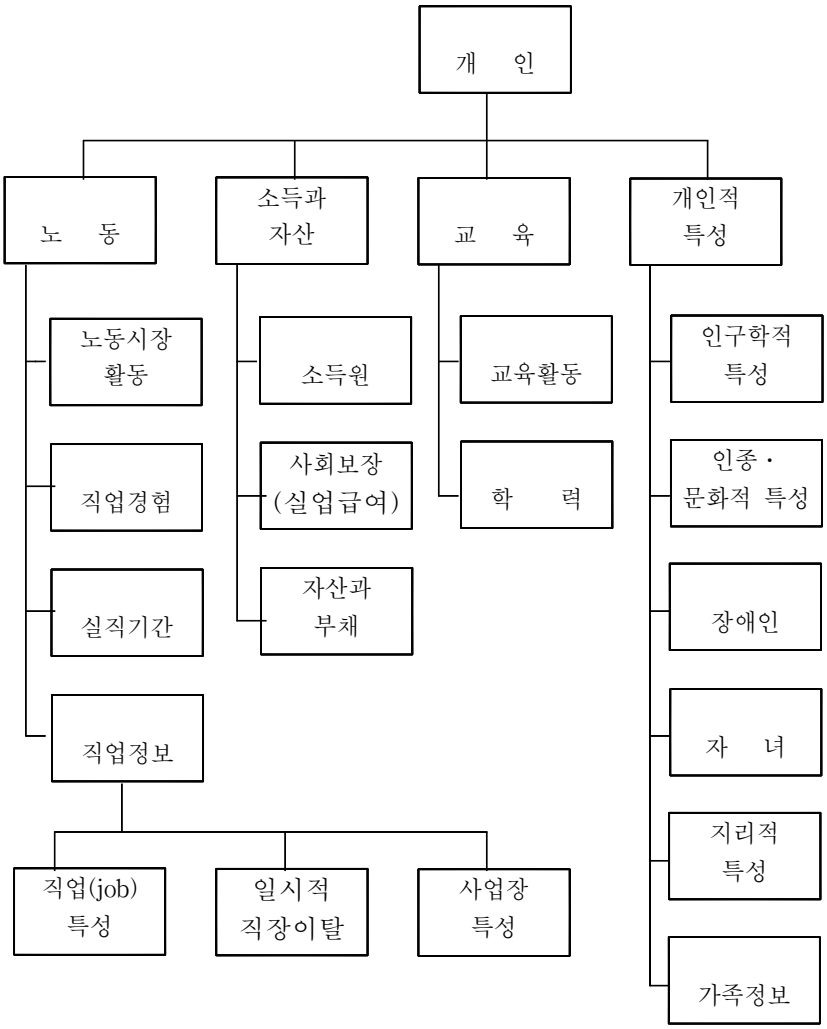
가. 인구학적 특성(Demographics)

패널표본에 속한 가구 및 개인의 인구학적 특성과 식별을 위하여 다음과 같은 변수들이 조사되었다.

○ ID(Identifiers):
식별을 위한 목적으로 모든 가족과 개인은 고유한 ID를 부여받으며 이 번호는 패널에 남아있는 한 언제나 동일하다. 외부로 공개되는 데이터파일에서 개인의 ID는 그 가구에 속한 다른 가구구성원들과 함께 묶을 수 없도록 무작위로 선정된 ID를 가지게 된다.

○ 출생일과 성(sex):
출생일은 패널에 들어온 첫해에 한 번 조사하며 외부로 공개되는 데이터에는 5년마다 한 번씩 출생일에 대한 정보가 제공된다.

<표 2> 설문내용의 구성



○ 혼인상황(Marital status):

혼인상황은 기혼, 동거, 별거, 이혼, 사별, 미혼, 동거남(녀)와의 별거의 7가지로 분류된다. 또한, 첫 번째의 면접과정에서 당시의 혼인상태를 질문하며 생애 처음으로 결혼한 날짜와 결혼횟수, 그리고 결혼을 시작한 시기와 지속

기간 등을 질문하여 혼인에 대한 완벽한 사건사를 구성할 수 있도록 하였다.

나. 인종문화적 특성(Ethnocultural Characteristics)

미국과 같이 다민족국가이고 이민이 많은 캐나다의 특성에 따라 인종과 문화적 배경의 파악은 언제나 매우 중요한 주제로 간주되고 있다. SLID에서도 응답자의 모국어(mother tongue), 출생국, 인종(race)과 같은 인종 및 문화적 배경에 대한 여러 가지의 질문을 던지고 있다. 특히, 이민온 경우 이민온 시기와 국적취득여부, 캐나다에서의 거주기간 등에 대해 묻고 있다. 더불어 세대간 변동과정(intergenerational mobility)을 파악하기 위해 부모의 교육수준에 대한 정보를 수집하고, 응답자가 소수민족이나 아메리카 원주민 등 고용평등(Employment Equity)의 대상인지를 알아보기 위한 설문도 제시한다.

다. 주거이동

응답자가 거주지에 대한 조사는 조사 기준 년도의 연말을 기준으로 답변하도록 하고 있다. 그러나, 외부에 공개되는 데이터에는 응답자의 거주지역을 정확하게 추적할 수 없도록 제한적인 주거정보만이 포함된다. 조사에는 응답자의 지리적 이동성(mobility)과 가족의 離合集散과정을 파악하기 위한 몇 가지의 추가적인 문항들이 질의된다. 특히, 기준 년도 기간 중 개인이나 가족이 이사하였는지의 여부와 함께 이사하였다면 그 날짜와 이사사유를 묻고 있다. 가족구성원의 동태적 벼 변화도 SLID의 주요한 관심사로 일 년 동안 가족구성원의 변동상황에 대하여 구체적으로 질문한다.

라. 자녀에 대한 정보

자녀 양육의 문제는 여성의 경우 경제활동 참여여부를 결정하는 중요한 변수

로 작용한다. SLID에서는 양육하고 있는 자녀의 수 및 생일에 대한 질문을 던지고 있으며, 이와 함께 자녀와 부모 사이의 혈연관계에 대하여 조사한다. 즉, 양육하고 있는 자녀가 자신의 피를 이어받은지 아니면 입양하였는지 등 혈연관계에 대한 정보를 수집한다.

마. 장애(Disability)

정신적, 신체적 장애자에 대한 질문은 장애자들의 현황을 파악하려는 목적보다는 장애가 노동시장에서의 활동에 미치는 영향을 분석하기 위한 목적을 가지고 있다. 장애의 종류와 장애가 나타난 시기 및 기간과 더불어 장애로 인하여 직장이나 직업을 가지지 못하고 있는지의 여부를 질문한다. 직장이나 직업을 가지고 있는 장애자의 경우에는 장애가 직장생활에 어떤 제약을 주는지를 조사하며, 또한 현재 원하는 만큼 직장에서 일하고 있는지의 만족도에 대한 문항도 응답자에게 제시된다. 만약, 비자발적인 이유로 희망하는 것보다 적은 시간을 일하는 경우 이것이 장애 기인한 것인지를 추가적으로 질문하고 있다.

바. 가족 및 가구특성(Family and Household Characteristics)

SLID에서 사용하는 가족(family)의 개념은 항상 경제적 가족(economic family)을 의미하고 한다. 즉, 가족은 혈연, 결혼, 동거(common-law) 또는 입양으로 맺어져 함께 거주하고 있는 개인들로 정의된다. 이러한 가족의 개념은 일반적으로 광범위하게 사용되는 정의로 가족사항에 대한 응답은 기준 년도말의 시점을 기준으로 답하도록 요구하고 있다. 가족 및 가구의 특성에 관해 구체적으로 다음과 같은 사항들이 질문되고 있다.

- ○ 가족수와 가구주와의 관계
- ○ 함께 거주하는 경제적 가족의 수
- ○ 함께 거주하는 비가족원(other household member)에 대한 사항:

가족과의 관계, 성별, 생일, 소득, 교육수준, 경제활동상태(labour force status) 등

- ○ 가족의 소득원, 가족 구성원별 소득액과 소득세
- ○ 가족 구성원별의 경제활동상태(취업자수, 실업자수, 비경제활동자수)
- ○ 교육 및 인종
- ○ 정신적, 신체적 장애 여부
- ○ 실업급여, 사회부조(social assistance), worker's compensation benefits 등의 수급여부 및 기간

2. 교육(Education)

가. 교육활동(Educational Activity)

교육활동에 대한 조사는 응답자가 받고 있는 정식 교육(formal education)의 내용을 질문한다. 구체적으로 학교에 다닌 기간(months), 학교에 full-time으로 다니는지 아니면 part-time으로 다니는지의 여부, 재학하고 있는 학교의 종류, 고등학교 졸업여부, 학위취득여부 등에 대한 문항들이 설문지에 포함되어 있다.

나. 교육수준(Educational Attainment)

패널에 참여한 첫 번째 면접에서 SLID는 응답자의 교육수준에 대한 정보를 수집하는데 15세 미만의 아동은 재학하고 있는 학교와 학년을 조사한다. 인구의 동태적 이동과정을 파악하기 위한 목적으로 응답자가 어느 지역에서 초등교육과 중등교육을 이수하였는지를 질문하고 있다.

이외에도 출생후 현재까지 받은 교육기간, 고등학교 졸업여부, 고등교육의 경험유무, 학위나 졸업증명서를 수여 받은 나이와 연도에 대한 정보도 수집하고 있다. 고등교육의 경험이 응답자에 대하여는 교육기관의 종류, 전공분야, 교육받은 나이와 연도, 이수 때까지의 소요기간 등에 대한 추가적인 질문을 던진다.

3. 소득과 자산

가. 소득원(Income Sources)

소득에 관한 조사는 조사의 정확성을 기하기 위하여 세금보고(tax report)가 끝난 직후인 5월에 실시하는데 응답자의 동의를 전제로 면접조사대신 세금보고서로 소득조사를 대체하기도 한다. SLID의 소득질문지와 세금보고서 사이에는 약간의 차이가 존재하는데 이러한 차이는 데이터의 가공과정 중에 조정된다. 소득에 대한 질문문항의 주요 내용은 다음과 같다.

○ 가족소득: 가족구성원의 소득합계
○ 주 소득원(main income sources) 및 소득액
- 임금과 봉급
- 농가 자영업소득(farm self-employment income)
- 비농가 자영업소득
- 투자소득(investment income): 이자, 배당소득 등
- 정부의 이전소득: Child Tax Benefit, Old Age Security, Guaranteed Income Supplement and Spouse's Allowance, Canada Pension Plan, Unemployment Insurance Benefits, Social Assistance, Workers' Compensation, GST credits
- Taxable Capital Gains
- 연금소득
- 이혼위자료 및 자녀양육비(Alimony and Child Support)

나. 실업급여, 사회부조 및 근로자보상

정부로부터의 이전소득 중에서 실업급여(unemployment insurance benefits), 사

회부조(social assistance), 근로자보상(workers' compensation)의 세 가지 항목에 대하여는 매월 수령한 금액을 조사하고 있다. 이 데이터는 고용과 실업 등 고용구조의 변화과정을 이해하는 기초자료로 활용되고 있다.

4. 노동시장활동

가. 노동시장활동 형태

지난 1년 동안 노동시장에서의 활동을 종합적으로 이해하기 위한 여러 가지의 질문을 응답자에게 던지고 있다. 구체적으로 다음의 <표3>과 같은 내용들이 조사되고 있다.

<표 3> 노동시장활동에 관한 질문

조사내용	비고
1. 매주 가지고 있던 일자리수	1년 동안 매주 단위로 조사
2. 일년 동안 가지고 있던 일자리수	
3. 근로한 개월수(months)	
4. 연간 총근로시간	모든 일자리에서의 전체 근로시간
5. 연간 총유급근로시간	앞에서 무급휴가 등을 제외
6. 경제활동상태(취업, 실업, 비경제활동)	
7. 매 주간 경제활동상태	53주 동안 매 주의 경제활동상태
8. 연간 고용된 주(週)의 숫자	
9. 미취업기간(無職)(jobless spell)	연간 무직 횟수
10. absences from work	무급휴가, lay-off 등
11. 연간 총비근로시간 및 주(週)	
12. 임금	시간당 임금수준
13. 직장탐색(job search)	
14. 직장경험(work experience)	

나. 미취업 기간중의 활동(Jobless Periods)

직장이 없는 기간동안에 대하여 SLID는 미취업의 시기 및 기간(dates and

durations), 미취업기간 중 구직활동여부 등에 대한 문항을 질문하고 있으며, 구직활동을 하고 있지 않으면 그 이유는 무엇 때문인지 그리고 직장을 원하는지의 여부 등을 조사한다.

다. 직업의 특성(Job Characteristics)

응답자가 일년 동안 하나 이상의 직장에 근무한 경우는 많지만 3개 이상의 다른 직장에서 근무한 경우는 거의 없는 것으로 나타나 SLID에서는 3개까지의 직장에 대하여 근무기간, 근무시작일과 종료일, 직장을 그만 둔 이직사유 등 그 특성을 질문하고 있다. 응답자가 하나 이상의 직장을 가졌을 경우 수집된 자료를 이용하여 직장간 임금, 근로기간, 주당 평균근로시간 등에 대하여 상호 비교할 수 있다. 응답자의 직장에 대한 질문에 대하여는 구체적으로 다음과 같은 내용들이 다루어지고 있다.

- ㅇ 직장을 시작한 때와 그만 둔 때
- ㅇ 이직사유
- ㅇ 종사상의 지위
- ㅇ 재택근무여부
- ㅇ 직장에 근무한 개월수
- ㅇ 주당 평균근무시간
- ㅇ 근무형태 및 근무시간대
- ㅇ 시간당 임금
- ㅇ 직장에서 얻은 연간 총소득
- ㅇ 직종
- ㅇ 직위 또는 직책
- ㅇ 노동조합에 관련된 사항
- ㅇ 연금 등 사회보험 가입여부

Ⅵ. 한국노동패널연구(KLPS)에 대한 시사점

이상에서 설명한 것과 같이 SLID는 캐나다의 통계청이 과거 AWPS(Annual Work Patterns Survey)와 LMAS(Labor Market Activity Survey)를 근간으로 발전시킨 종단면조사(longitudinal Survey)로 PSID(The Panel Study of Income Dynamics)나 NLSY(National Longitudinal Survey of Youth) 등 다른 종단면조사와는 여러 가지의 측면에서 특징을 보이고 있다. SLID가 지니는 특징들과 그것이 한국노동패널연구에 던져주는 시사점을 정리하면 다음과 같다.

먼저, SLID는 표본규모를 15,000가구로 시작하여 PSID의 7,190가구(1992년), NLSY의 12,686명, 영국 BHPS의 5,000가구, 독일 The Socio-Economic Panel의 6,819가구 등 다른 패널조사의 표본수에 비해 상당히 큰 편이다. 특히, 1996년부터는 두 개의 패널이 동시에 운용됨에 따라 표본규모는 거의 30,000가구에 달하게 되며, 이는 캐나다의 인구가 미국, 영국, 독일 등 다른 국가들보다 훨씬 적은 3천만명 정도라는 점을 감안할 때 이러한 규모의 표본규모는 패널조사의 대표성과 신뢰도 확보에 매우 긍정적으로 작용할 것이다.

표본의 규모는 기본적으로 예산에 의하여 제약받을 수밖에 없다. 제한된 예산으로 무리하게 표본수를 늘이는 것은 패널유지관리에 사용되는 비용이나 면접조사시 면접원 수당 등을 과소책정하게 되어 직접적으로 수집된 데이터의 정확도나 신뢰도를 저하시키고 장기적으로는 패널유지관리의 실패에 따른 표본탈락률의 증가를 초래할 가능성이 높다.[3] 특히, 실사를 외부 민간조사기관에 용역을 주어 위탁하는 경우 영리를 추구하는 기업의 성격상 충실한 실사결과를 기대하기 어려운 상황이다. 이러한 측면에서 예산에 맞게 표본수를 적정한 수준으로 조절

3) 이러한 현상은 한국가구패널조사(KHPS)에서 발견되어질 수 있다.

하는 것이 매우 중요하며, 동시에 예산의 변동에 따라 표본을 추가하거나 줄이는 것이 쉽도록 표본의 유연한 관리체계를 개발하는 것이 필요하다.

둘째로 다른 패널조사에서는 패널이 유지되는 한 표본가구 또는 개인을 지속적으로 추적, 조사하는 것이 일반적인 데 이에 반하여 SLID에서는 일단 표본으로 선정된 가구는 향후 6년 동안만 패널에 머무른다. 따라서, 앞의 <표 1>에서 보는 것처럼 1996년부터는 매 3년마다 조사대상가구의 절반 정도가 교체되도록 패널이 설계되었다. SLID에서 이처럼 패널을 설계한 이유는 표본을 주기적으로 교체하여 시간의 경과에 따라 표본일부가 탈락하기 때문에 발생하게 되는 대표성의 문제를 해결하고 조사내용의 유연성을 도모하기 위한 목적을 지니고 있다. 또한, 이는 응답가구의 조사부담을 줄여 조사의 효율성을 증대시키려는 목적도 함께 가지고 있다.

셋째로 SLID에서는 조사가 1월의 일반조사와 5월의 소득조사의 둘로 나누어진다는 점이다. 전년도 1년간의 노동시장활동 및 가족의 변동사항, 교육 등 관심분야에 대한 정확한 정보를 얻기 위하여 1월중에 조사를 실시하며, 이 때 응답자들이 질문에 대답하는 기준시점은 전년도말로 규정하고 있다. 또한, 소득에 관한 질문에 관하여는 세금보고(tax report)가 끝난 직후인 5월에 조사를 하여 응답자가 지난해의 소득에 대해 가장 정확한 기억을 가지고 있는 시점을 택하고 있다. 특히, 소득조사는 응답자가 세금당국에 제출한 세금보고서를 응답자의 사전동의가 있을 때 통계청이 열람하여 소득조사를 대체하도록 하고 있다. 세금보고서의 열람에 동의한 응답자는 전체의 63%에 이르고 있으며, 이는 별도의 소득조사에 따른 비용과 시간을 절약할 뿐만 아니라 조사의 정확성 및 다른 데이터와 일관성에 커다란 기여를 하고 있다.

조사시점을 어떻게 결정하는가는 조사의 정확성에 상당한 영향을 미치게 된다. 일반적으로 패널조사는 1년에 한 번씩 기준기간(일반적으로 지난해) 동안의 활동 및 변화상황을 조사하게 되는데 조사기간이 기준기간에서 멀어질수록 기억의 감퇴 및 각종 자료의 상실에 따라 응답의 정확도가 떨어지게 된다. 따라서, 가능한

범위 내에서 기준기간과 면접시점과의 간격을 좁히는 것이 필요하다.

캐나다의 SLID는 기준기간을 전년도 1월 1일에서 12월 31일까지의 1년으로 하고 대부분의 응답에서 기준시점은 전년도말(12월 31일)로 정하고 있다. 이와 같은 기준기간의 설정방식은 대부분의 패널조사에서 채택하고 있는데 현실적으로 조사기간이 다음 해 중반 이후로 늦추어질 경우에 이는 기준기간과 조사 기간의 차이가 많이 발생하여 조사의 정확성이 떨어지는 원인으로 작용한다.

따라서, 하나의 대안은 기준년도를 변경하는 방안이다. 예를 들어 4월부터 조사를 실시할 경우 기준 기간을 지난해 4월 1일에서 금년 3월 31일로 조정하여 3월 31일을 기준으로 응답하도록 하는 것이다. 이는 현재 한국가구패널조사(KHPS)에서 사용하고 있는 방식이다. 그렇지만 이러한 방식은 응답자가 기준기간 및 기준시점에 관한 정확한 인식을 가지지 못한 경우 응답자로 하여금 질문에 대답하기 어렵게 만들거나 부정확한 답변을 유도할 소지가 많다. 즉, 응답자가 기준시점을 지난해 말이나 아니면 올해 초로 잘못 인식하여 답변할 가능성이 높다. 이는 또 다른 측면에서 조사의 정확도를 저해하는 요인으로 작용하게 된다. 따라서, 노동패널조사에서는 기준기간을 지난해 1년 동안으로 하고 조사시점이 늦추어지지 않도록 노력할 필요성이 있다.

네 번째로는 표본추출에서 기존의 노동력조사(Labor Force Survey)의 표본을 이용하고 있다는 점이다. SLID는 LFS의 6개 순환표본그룹 중에서 2개를 선정하여 이를 모표본으로 하고 있다. 따라서, LFS의 조사결과와의 비교를 통하여 조사결과의 정확성 및 대표성에 관한 분석을 효율적으로 수행할 수 있는 장점이 있음과 동시에 사전적으로 조사대상가구에 대한 기본적인 성격 및 정보를 가질 수 있었다. LFS는 지역을 기준으로 한 다단계층화추출법에 의하여 표본을 추출하고 있는데 여기에서 질문하는 내용의 대부분이 SLID 질문내용의 기초를 이루고 있어 사전적인 예비조사를 실시한 것과 유사한 효과를 SLID에 가져온다.

다섯 번째로 SLID가 개인의 노동시장에서의 동태적 변화와 소득의 동태적 분석을 기본 목적으로 하고 있지만 미국의 NLSY처럼 개인을 표본단위로 하여 선

정하기보다는 가구를 표본단위로 하고 있다.

동일한 연령계층을 대상으로 조사하는 NLSY는 동시태생효과(cohort effect)를 정확하게 파악할 수 있다는 장점을 가지고 있지만 그 조사결과가 전국의 모집단을 대표하지 못하고 시간이 지날수록 응답자들의 나이가 들어 장기적으로는 응답자수가 급격하게 줄어드는 문제점이 있다. SLID방식의 가구단위조사는 동시태생효과와 동류집단효과를 정확하게 파악하기 위해 표본수가 매우 커야한다는 한계점을 가지고 있지만 분석결과가 전국을 대표하고 결혼, 출산 등의 과정을 통하여 새로운 표본이 계속 충원된다는 장점이 있다.

아직 초보적인 수준에 머무르고 있는 우리 나라 노동관련 개인통계 수준을 감안할 때 NLSY와 같은 성격의 패널조사는 시기상조로 여겨지며 다목적 성격을 지닌 가구단위의 패널조사를 실시하는 것이 현실감을 준다. 패널조사의 성격을 다목적화하는 것은 특정계층의 특정문제에 대한 분석을 할 경우, 유효표본의 부족현상을 일으켜 패널조사의 의의를 반감시키는 결과가 발생할 우려가 있다. 이러한 문제에 대하여는 가능한 전체 표본수를 늘이고, 매 조사 때마다 특정계층 또는 특정문제에 대한 조사를 집중적으로 실시하는 등 패널조사의 유연한 운영을 통하여 해결하도록 하는 것이 바람직하다.

다음으로 SLID에서도 응답자 및 면접원 관리에 많은 관심을 기울이고 있지만 응답자에 대한 물질적 보상은 캐나다 통계청의 원칙에 따라 제공하고 있지 않다. 응답자와의 원만한 관계의 유지는 표본의 탈락률을 낮추고, 응답 내용의 신뢰도 제고에 결정적으로 중요한 역할을 하는데 특히, 직접 응답자를 만나고 질문하는 면접원의 사기와 의욕이 매우 중요하다. 많은 경우 면접원과 응답자와의 장기적인 접촉은 둘 사이의 인간적 관계를 형성하여 면접의 정확성과 응답률의 유지에 긍정적으로 기여한다. 이러한 측면에서 한국노동패널연구에서도 응답자 관리에 많은 관심과 노력을 기울여야 할 필요성이 있으며, 특히 면접원의 사기와 의욕촉진을 위한 면접원의 관리의 중요성이 부각되고 있다.[4)]

4) 응답자 및 면접원 관리의 중요성 및 구체적 방안에 대하여는 패널조사의 응답자 관리(금재호, 1998)을 참조할 수 있다.

이와 함께 SLID에서는 개인의 비밀보장을 위해 개인의 ID를 임의로 결정하고 개인 또는 가족을 식별할 수 있는 변수들을 삭제 또는 변경하는 등 다양한 노력을 기울이고 있으며, 더불어 면접과정에 CAI(Computer Assisted Interviewing)를 도입하여 응답의 일관성을 점검하고, 면접시간을 단축시키며, 자료관리의 체계화와 신속화를 기하고 있다. 한국노동패널연구에서도 자료의 외부공개시 개인의 비밀보장에 보다 많은 신경을 쓰고 장기적으로는 CAI시스템을 이용한 설문조사 기법을 채택하여 응답자의 응답부담을 줄이고, 자료의 체계적 관리 및 데이터베이스화에 기여할 수 있도록 할 필요성이 있다.

參 考 文 獻

Fahy, T. W., *The National Longitudinal Surveys of Labor Market Experience An Annotated Bibliography of Research, the Ohio State University, 1995.*

CHRR, *The NLS Handbook, the Ohio State University,* 1995.

Statistics Canada, *Dealing with the Seam Problem for the Survey of Labour and Income Dynamics,* 1992.

Statistics Canada, *Objectives and Content of the Preliminary Interview,* 1992.

Statistics Canada, *Respondent Relations in Longitudinal Surveys,* 1994.

Statistics Canada, *SLID Coding Structure: Mother Tongue, Country of Birth and Ethnic Background,* 1995.

Statistics Canada, *SLID Derived Variables: Demographic, Cultural and Geographic,* 1994.

Statistics Canada, *SLID: Microdata User's Guide,* 1994.

Statistics Canada, *SLID Following Rules: Who to Trace and Who to Interview,* 1994.