



그거 맞는 통계예요?

ONE SIZE FITS ALL STATISTICS

최무이 소프트웨어정책연구소 가상융합연구실 책임연구원 | muyi@spri.kr

1 조사통계의 문제

데이터, 근거를 기반으로 정부의 정책이나 기업의 전략 수립을 지원하기 위해 다양한 통계들이 작성되고 있다. 산업이 세분화되고 고도화되면서 새롭게 정의되는 신산업들에 대한 통계 수요나, 인력 수급 차이, 연구개발 활동 등 특정 정책 목적 등에 따라 다양한 통계의 작성이 요구된다.

국가 전체 산업, 기업, 인구 등에 대한 기초통계들이 작성되고, 각 수요에 의해 구체적 분야의 전문통계들이 작성되다 보면 통계의 영역이 중첩되는 지점들이 있고, 비교가능성이 있는 유관 통계들이 존재하게 된다. 예를 들어 정보통신산업 전체에 대한 기초통계를 작성하는 ICT산업실태조사가 있고, ICT라는 큰 우산의 줄기에 SW산업, 인공지능산업, 클라우드산업 등 각 산업별 전문통계조사가 존재한다. 또한 한국은행에서 전체 산업을 대상으로 작성하는 기업경영분석통계, 서비스산업에 대한 서비스무역세분류통계 등도 각각의 산업 영역과 중첩되는 부분들이 있어 각 조사 간 비교 가능한 통계 항목들이 존재한다. 이 경우 각 통계가 리포트하고 있는 수치 간에 차이가 존재할 수밖에 없는데, 통계 사용자들의 입장에서는 도대체 어떤 게 맞고 어떤 게 틀린 지를 따져 묻게 된다.

통계 간 비교를 시작하게 되면, 주로 거론되는 것들이 조사를 통해 작성되는 통계에 대한 의구심과 회의 섞인 의견들이다. 인공지능 시대에 구시대적으로 일일이 기업들에게 물어봐서 자기 응답식으로 기록하게 하는 통계조사 결과를 신뢰할 수 있느냐와 같은 코멘트와 더불어 조사 신뢰성은 차치하더라도 조사 기반의 통계는 시의성이 너무 떨어진다는 것이 큰 약점으로 지적된다. 실시간으로 빠르게 변화하는 사회경제 환경에 신속하게 대응할 수 없는 통계는 정책 결정을 지원하는 데 한계가 있다는 의견이다.

따라서 통계의 신뢰성이나 시의성을 제고하기 위해서 행정자료를 적극 활용한 가공통계(또는 보고통계)¹를 좀 더 활용하자는 의견이 자연스럽게 나오게 된다. 행정자료를 활용하는 통계로는 국세청 법인세 신고첨부자료인 각 법인의 재무제표를 활용하는 한국은행의 기업경영분석통계, 관세청 통계 자료를 DB로 활용하여 작성하는 상품무역통계, 외국환거래신고서 DB를 기반으로 작성하는 서비스무역세분류통계, 지식재산권통계 등을 들 수 있다. 이러한 행정자료를 활용한 통계들을 각 산업에서 필요한 형식으로 재분류하고 가공하여 조사 기반의 통계들을 보완하자는 노력들이 진행되고 있다.

한국은행도 작년부터 콘텐츠, ICT, 보건의료 등 서비스 분야의 무역통계를 고도화하기 위해서 행정자료 기반으로 관련 통계를 개발하고 있다.(머니투데이, 2023.6) 이러한 노력들은 매우 중요한 가치를 가지고 있고, 지속적으로 추진되어야 할 필요가 있으나 행정자료 기반 통계작성에도 여러 가지 제약들은 존재한다.

¹ 통계청은 행정자료 활용 통계를 다음과 같이 정의함: '가공통계'란 한 종류 이상의 통계와 추가로 수집한 통계자료 또는 행정자료를 이용하여 작성한 통계, '보고통계'란 신고, 보고, 신청, 인·허가 등과 같이 다른 행정업무에 수반하여 수집된 자료를 이용하여 작성한 통계(통계승인업무 처리지침, 2023)

2 행정자료 활용 통계의 제약

행정자료는 공공기관이 직무상 작성·취득하여 관리하고 있는 문서·대장 및 도면과 데이터베이스(DB) 등 전산자료(통계법 제3조)를 말한다. 그런데 이러한 DB는 각 원시데이터의 특성에 따라 무역 통관기준 또는 외국환 거래기준 등으로 구축되어 있어 산업적 관점에서 분석하고 정책 목적에 활용하기 위해서는 해당 산업에 해당하는 기업들을 식별한 후, 이 기업들의 행정 데이터를 산업통계 목적에 맞게 재분류하는 작업이 필요하다.

각각의 행정자료는 자료의 특성에 따른 개별적 분류체계에 따라 분류되어 있겠지만, 이러한 분류체계들을 산업통계의 세세분류에 맞게 일대일로 모두 매칭시킬 수는 없기 때문에 실질적인 데이터의 성격에 맞게 일일이 데이터를 확인하고 재분류해야 한다. 일반적으로 행정자료를 가공하여 작성하는 통계는 조사비용이 소요되지 않기 때문에 상대적으로 매우 저렴한 비용으로 작성할 수 있다고 인식되지만, 실제로 이러한 데이터 정제와 재분류 작업에 소요되는 시간과 자원의 투입 또한 상당한 것이 현실이다. 또한 행정자료는 고유 수집 목적에 해당되는 주요항목 외의 정보는 데이터의 신뢰성이 낮을 수 있고, 국세청 신고자료 등 원자료의 신뢰성이 매우 높다 하더라도 이러한 자료 처리 과정의 품질에 따라 가공 작성되는 통계의 신뢰성도 크게 좌우될 수 있다는 제약도 존재한다.

행정자료 활용 통계의 활용성과 입수 가능성에 대한 제약도 중요한 이슈이다. 각각의 산업통계들은 그 통계작성의 기준이 되는 특수분류체계에 기반하여 통계를 작성하고 활용한다. 그러나 행정자료가 정리되어 있는 제반 행정단위들과 특수분류체계 등 통계단위 간의 근본적인 차이로 인해, 아무리 세밀한 재분류 작업을 거친다고 하더라도, 변환할 수 있는 통계단위에는 한계가 있기 마련이다. 예를 들어 SW산업의 경우 패키지SW>응용SW>산업특화SW>건설용 SW 등의 세세분류까지 통계분류체계가 정립되어 있는데, 한국은행의 기업경영분석 등이 활용하는 국세청 신고DB로는 이러한 세세분류 단위까지로 통계를 가공하여 집계하기 곤란하다. 기업의 재무자료 등 신고자료가 이러한 세부적인 품목 매출정보 등을 포함하지 않고 있기 때문이다. 이 경우 통계작성의 기초적 목적인 산업의 전체적인 규모를 추계하겠다는 목적은 달성할 수 있겠지만, 갈수록 디테일해지고 있는 정부의 정책 개발 지원을 위한 목적으로 활용하기에는 제약이 있을 수밖에 없다.

더 근본적인 문제는, 행정자료를 입수하기 위해서는 법률적·정책적·기술적·조직적 측면 등 다양한 영역에서 합의에 도달해야 한다(제주 청년통계, 2019.10)는 것이다. 법·제도적 한계로 개별 기업 식별 데이터를 받기는 요원할 것이고, 관련 부처 간 긴밀한 협조를 통해 원시데이터를 마스킹하여 받는다고 하더라도 조사 기반 데이터와의 연결자가 없이는 활용성이 떨어질 것이라고 생각된다. 몇 년 전 통계 관련 포럼에 참석하였을 때 한 공공기관 소속 데이터 담당 본부장께서 행정통계를 협조받아서 조사통계를 개선하는 작업을 3년째 열심히 추진하고 있는데 여전히 데이터를 협조받지 못하고 있다는 언급이 떠오른다.

3 통계 고도화를 위한 균형 찾기

행정자료 등을 활용한 통계의 작성 및 이에 기반한 조사통계의 고도화는 중장기적 관점으로 지속적 추진이 필요하다. 그러나 위와 같은 점들을 고려할 때 단기적으로는 현행 추진하고 있는 조사 기반 통계를 실질적으로 개선하는 작업을 병행하는 것이 더욱 중요하다. ICT산업 같은 경우 상위기업들의 실적이 전체 산업 규모의 대부분을 차지하는 쏠림현상이 있는데, 사실상 그 상위기업들 중 응답을 거부하는 기업도 있어서 조사 데이터 신뢰성과 활용성에 제약이 있을 수 있다. 표본조사라 하더라도 산업통계에 주요한 영향을 미치는 이러한 핵심기업에 대해서는 패널형태의 차별화된 관리와 2차 자료를 통한 수집 데이터의 검증을 한층 강화할 필요가 있다.

또한 SW, 콘텐츠 산업의 경우 마켓 플레이스, OTT와 같은 플랫폼 등을 통해 제품·서비스의 거래 유형이 복잡화되고 있다. 이러한 추세를 분석하고 구조화하여 설문 문항 고도화를 추진하고, 갈수록 중요성이 높아져 가는 플랫폼 사업자 등에 대한 신뢰성 있는 데이터 확보 방안 등을 면밀히 강구하는 것도 중요한 일이라 생각된다.

한편 더욱 중요한 것은 통계에 대한 이해 확대와 인식의 전환이다. 양립하는 통계·데이터에 대해서는 일반적으로 어떤 게 맞는 것인지에 천착하는 경우가 많은데, 그러다 보면 오히려 통계 신뢰성, 고도화 등 중요한 일에 집중하지 못하고 다른 통계와의 차이를 밝히는데 더 시간을 쏟게 된다. 정책담당자들은 기본적으로 담당 분야에 대한 높은 수준의 정책역량을 보유하고 있지만, 근거·통계에 대해서는 세부적인 지식이나 정보를 확보하는 데 한계가 있을 수 있다. 증거기반정책이 성공하기 위한 기반역량으로, 증거를 생산하는 행위자의 높은 역량도 중요하지만, 증거를 활용하는 정책담당자 역시 적정 수준의 통계 역량과 이해를 확보할 필요가 있다(윤건, 2019).

모든 현상을 하나의 통계체계로 제시해 줄 수 있는 ‘One size fits all’한 통계는 없다. 어떤 통계는 맞고 다른 건 틀리다는 이분법적인 시각에서 벗어나야 한다. 시의성이 높은 행정자료 기반 통계·데이터는 최신 동향 정보로 활용하고, 조사 기반 통계를 통해서는 디테일한 산업의 구조를 파악, 이슈를 발굴하여 정책 개발에 활용하는 등의 균형된 시각과 이해를 확산 정착시키는 것이 필요하다.

◎ 참고문헌 및 기사

통계청, 통계승인업무 처리지침(2023)

한국은행, 기업경영분석(2023), 서비스무역세분류 통계(2023)

머니투데이방송, “한국은행, 정부와 ‘서비스 무역통계’ 고도화 추진”(2023.06.05.)

제주특별자치도, 통계청 제주사무소, 행정자료를 활용한 행정통계작성 지침 - 제주 청년통계-(2019.10)

윤건, 증거기반정책: 이론과 실제(2019.9.)