

강원 영동지역 관광경기가 지역경제에 미치는 영향과 시사점: 고빈도 데이터에 기반한 관광경기지수 개발을 중심으로*

강규호** · 배주원*** · 이준영****

[국문요약]

본 연구는 신용카드 사용액, 영동지역 고속도로 톨게이트 차량 통행량, 내비게이션 영동 지역 목적지 검색 건수, KT 이동통신 데이터 기준 영동지역 외지인 방문자수, 영동지역 시군 별 네이버 검색어 트렌드 등 고빈도 자료를 이용하여 강원 영동지역 관광경기지수를 산정한다. 산정된 관광경기지수와 지역 관광업 실물경기 간 통계적으로 높은 상관관계가 추정되었는데, 이는 본 고의 고빈도 자료 기반 관광경기지수가 현재 관광경기 상황에 대한 높은 속보성을 지닐 뿐만 아니라, 관광경기의 실물효과를 적절히 반영하고 있다는 것을 의미한다. 더불어 대규모 구조 VAR 분석결과, 영동지역 관광경기 충격은 관광업 고용과 신용카드 소비를 상승시키며, 영동지역 전체 고용에도 긍정적인 효과를 주는 것으로 추정되었다. 본 연구 성과는 향후 영동지역 관광경기 실시간 모니터링과 관광업 진흥 및 파급효과 확대를 위한 정책개발에 유용하게 활용될 것으로 기대한다.

핵심주제어: 고빈도자료, 주성분 분석, 대규모 구조벡터자기회귀
경제학문헌목록 주제 분류: G12, C11, E43

I. 연구배경

1. 강원 영동지역 관광업 현황

강원 영동지역은 전체 산업 중 서비스업의 비중이 높은 지역으로 특히 관광업은 지

* 유익한 논평을 해주신 한국은행 강릉본부 중간세미나 토론자 및 참가자분들께 감사드립니다.
본 연구는 한국은행의 재정지원을 받아 작성되었습니다. 본 연구내용은 집필자들의 개인의견이며 한국은행의 공식견해와는 무관합니다.

** 교신저자, 고려대학교 경제학과 교수, Tel: +82-2-3290-5132, E-mail: kyuho@korea.ac.kr

*** 공동저자, 한국은행 강릉본부 조사역, Tel: +82-33-640-0157, E-mail: jw.bae@bok.or.kr

**** 공동저자, 한국은행 강릉본부 조사역, Tel: +82-33-640-0161, E-mail: junyoung.lee@bok.or.kr

역의 주력산업 중 하나이다.¹⁾ 강원 영동지역 관광업이 2019년 중 총부가가치에서 차지하는 비중은 11.3%로 전국에서 가장 높은 수준이며 전체 고용에서 차지하는 비중도 높은 편이다. 강원 영동지역 관광업(숙박·음식점업, 예술·스포츠·여가 관련 서비스업)의 고용 규모는 2020년 하반기 기준 5.3만명으로 전체 산업(31.8만명) 대비 16.5%, 서비스업(25.3만명) 대비 20.8%로 높은 비중을 차지한다. 역내 GRDP 및 고용에서 관광업의 비중이 높은 만큼 지역경제는 관광업 동향에 큰 영향을 받는 것으로도 나타났다.

강원 영동지역의 여행 수요는 2010년 이후 국내 여행 수요가 늘어나는 추세를 보이는 가운데²⁾, 고속도로, 고속철도(KTX), 항공 등 광역 교통망의 확충으로 수도권 지역과의 접근성이 크게 개선됨에 따라 양적으로 증가하였다. 아울러 평창동계올림픽 개최(2018년 2월)를 전후로 국내외 인지도가 상승하고 숙박 및 레포츠 등 관광인프라 투자가 활발히 진행되면서 관광업이 활성화되었다.³⁾ 2020년 코로나19 팬데믹으로 강원 영동지역 방문객 수가 일시적으로 감소하였으나 해외여행 제한의 풍선효과 등으로 빠른 회복세를 보이면서 지역경제에 긍정적인 영향을 미치고 있는 것으로 나타났다.

다만 <표 1>에서 볼 수 있듯이 장기적인 시계에서 관광업의 성장률이 높은 변동성을 보이면서 지역경제의 성장률에 안정적인 양의 효과를 미치지 못하는 것으로 나타난다. 또한 <그림 1>에서 보듯이 코로나 2020년에는 일부 시군을 제외한 영동지역 대부분에서 관광업 취업자 수⁴⁾가 상당폭 감소세로 전환되었으며 시군별로도 장기적인 성장세가 상이하게 나타내는 등 고용 또한 안정적인 모습을 보여주지 못하고 있다.

<표 1> 강원 영동지역 총부가가치 대비 관광업 비중 및 성장률

(단위 : %)

	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
비 중	11.4	11.2	11.3	11.2	10.6	10.4	11.1	10.4	11.3	11.3
성장률	-	-4.9	0.0	0.0	-2.6	-2.9	5.9	-2.5	6.8	4.7

자료: 통계청

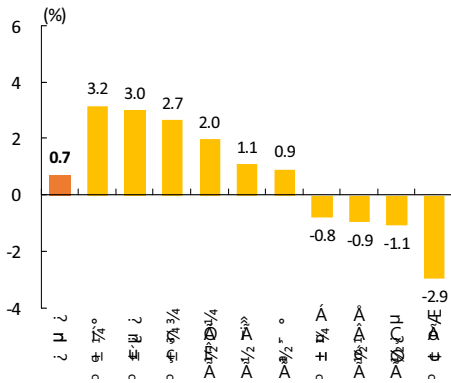
1) 2019년 기준 강원 영동지역 총부가가치 대비 서비스업의 비중은 69.7%로 제조업(8.1%), 건설업(10.9%), 농업·임업 및 어업(4.6%), 광업(3.1%) 등 다른 산업에 비해 높은 수준이다.

2) 한국문화관광연구원의 국민여행조사에 따르면 국내 관광여행 경험률은 2010년 59.7%에서 2019년 85.0%, 2020년 75.5%로 상승하였다.

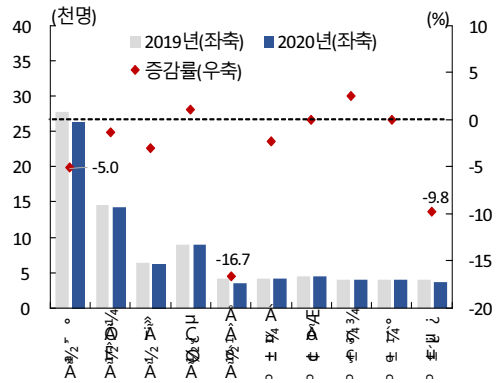
3) 관광업은 관련 법규, 통계자료, 분석 방법 등에 따라 포괄범위를 다양하게 정의할 수 있으며, 본 연구는 통계청 GRDP 분류기준에 따라 관광업을 “숙박 및 음식점업”과 “문화 및 기타서비스업”으로 한정하여 정의한다.

4) 통계청에서 공표하는 “시군구/산업별 취업자” 자료의 경우 도소매·숙박음식점업과 사업·개인·공공서비스 및 기타 등으로 산업분류를 제공하며, 앞서 “시도/산업별 취업자” 자료에서 관광업 중 숙박 및 음식점업의 취업자 수 비중이 평균적으로 80%를 초과하므로 본 단락에서는 시군별 관광업의 취업자수를 사업·개인·공공서비스 및 기타 업종을 제외하고 도소매·숙박음식점업 취업자 수로 분석하였다.

<그림 1> 영동지역의 시군별 관광업 취업자 수 증가율

(a) 시군별 관광업 취업자 수 증가율¹⁾
(2014~20년 연평균 기준)

주: 1) 도소매·숙박음식점업의 취업자 수 증가율
자료: 통계청

(b) 시군별 관광업 취업자 수 및 증가율¹⁾
(2019-2020년)

주: 1) 도소매·숙박음식점업의 취업자 수 증가율
자료: 통계청

2. 연구의 필요성

강원 영동지역의 경제구조는 지리적 여건, 인구구조, 산업 등 여러 특성상 양질의 일자리 창출에 취약한 상황이다. 특히, 탄광산업의 사양화 이후 지역 내 근간산업이 부재한 상황인 데다 높은 산맥 등으로 불리한 지리적 입지, 전·후방 파급효과가 큰 제조업 성장 기반 취약 등으로 젊은 연령층의 역외유출 현상 및 급속한 고령화가 지속되고 있다. 이는 적은 인구 수에 따른 소비시장 제한 등 지역경제 위축의 악순환을 야기시키고 있다. 이에 따라 현재 영동지역 경제는 제조업 비중이 낮고 농림어업, 비금속광물, 도소매 및 음식숙박업, 음식료품 등 전통산업과 경기민감 업종 비중이 높은 현상이 지속되고 있다.

강원 영동지역의 이러한 산업구조를 고려해 볼 때, 지역경제 활성화에 있어 관광업의 지속적 발전은 매우 중요하다고 할 수 있다. 관광업은 숙박·음식점업, 예술·스포츠·여가 관련 서비스업 외에도 건설토목업, 운송업, 유통업, 소매업 등 지역 내 다양한 세부산업과 밀접하게 관련되어 있어 경제적 유발효과가 크고 그 범위도 상대적으로 넓기 때문이다.

최근 영동지역의 관광업은 평창동계올림픽 이후의 관광인프라 투자, 코로나19로 인한 해외여행 제한의 풍선효과 등으로 높은 성장세를 나타내고 있다. 다만 향후 해외여행 수요 회복시 성장률이 하락하면서 변동성이 확대될 가능성도 상존하는 만큼 관광업의 지속적이고 안정적인 성장을 도모하는 것이 중요하다. 따라서 관광업 발전을 통해 지역경제 활성화 효과를 극대화하기 위해서는 관광경기 결정요인을 정확히 파악하여 올바른 정책 방향을 설정할 필요가 있다. 이에 본 연구에서는 관광경기에 대한 실시간 모니터링

시스템 구축하는 한편 관광경기가 실물경제 전반에 미치는 파급효과 등을 고빈도 데이터를 사용하여 정량적으로 연구해 보고자 한다.

II. 연구내용

1. 연구목표

본 연구의 목표는 세 가지이다. 첫째, 영동지역 관광업 동향을 면밀하게 실시간으로 파악할 수 있도록 빅데이터를 이용한 “영동지역 관광경기지수”를 개발하는 것이다. 이와 비슷한 지표로 현재 문화체육관광부가 산출하는 지역관광발전지수가 있는데, 이 지수는 17개 광역시·도와 152개 시·군을 대상으로 산출되며 지자체별로 6개 등급으로 분류된다. 하지만 이 지수는 격년별로 산출되는 등 속보성이 낮아 관광경기를 실시간으로 모니터링하는 데 한계가 있다. 둘째, 개발된 관광경기지수를 활용하여 관광경기가 관광업 고용에 미치는 영향을 분석하고 나아가 지역 실물경제 전반에 대한 파급경로를 분석한다. 마지막 목표는 분석결과를 바탕으로 영동지역 관광업 활성화를 위한 정책적 시사점을 제시하는 것이다.

구체적으로 관광경기지수 산정을 위해 한국관광공사 빅데이터 플랫폼에서 제공하는 이동통신 기반 외부방문객 수, 내비게이션 검색 건수 데이터와 신용카드 사용액, 코레일과 도로공사에서 제공하는 KTX 이용객 수 및 고속도로 차량 통행량, 네이버 검색 트렌드 등 속보성이 높은 각종 통계자료를 면밀히 검토하여 강원 영동지역의 관광경기지수를 산출한다. 그런 다음, 지역 실물경제 지표와의 시차상관관계 분석을 통해 유용성과 정보력을 검증한다.

산정된 관광경기지수를 활용하여 관광경기 변화가 고용을 포함한 지역 실물경제 전반에 미치는 영향과 파급경로를 분석하기 위해 구조 VAR 분석을 한다. 이를 위해 관광경기지수, 지역 실물경제 변수 및 국내 거시변수들을 사용하여 내생변수 벡터를 구성한다. 관광경기 파급효과 분석 시 변수 개수가 많고 자료 주기가 변수별로 상이할 뿐만 아니라, 일부 변수는 표본기간이 길지 않아 전통적인 분석기법으로는 통계적으로 유의미한 결과를 도출하기가 쉽지 않다. 이러한 통계분석 기법상 어려움을 해결하기 위해 최근에 개발된 베이지안 Large scale 구조 VAR모형과 추정방법을 도입한다. 파급효과 등 추정결과를 바탕으로 영동지역 관광업 활성화를 위한 정책적 시사점을 도출한다.

2. 기존 문헌 검토

강원지역 관광업과 실물경기 간 연관성에 관한 최근 연구로는 강승호·하세호(2015)와 류시영·유선옥(2017)이 있다. 우선 강승호·하세호(2015)는 2010년과 2012년 문화·관광업 업종별 종사자 수 데이터와 지역성장률 시차분석방법을 통해 문화·관광업이 지역성장률과 일자리 창출에 미치는 영향을 평가하였다. 추정결과, 강원지역의 관광산업은 지역경제 성장에 매우 긍정적인 요인으로 작용하였으나 문화콘텐츠산업은 기여도가 작은 것으로 나타났다. 류시영·유선옥(2017)은 소셜미디어 빅데이터 분석을 통해 관광목적지로서 강원도에 대한 인식을 조사한 연구로, 강원도의 거점 도시 각각에 대한 관심 이슈와 연관 키워드들을 세부적으로 고찰하고 강원지역의 관광 마케팅 전략과 각 지역의 관광자원화 방향을 제시했다.

이 연구는 여러 측면에서 기존 연구와 차별화된다. 강승호·하세호(2015)의 연구는 문화·관광업 업종별 종사자 수를 관광경기로 측정하였으나, 본 연구는 보다 광범위하고 종합적인 자료를 활용하여 관광경기지수를 산정한다. 또한 강승호·하세호(2015)가 정태적 분석방법을 사용한 것과 달리 관광경기충격이 지역실물경제 전반에 미치는 동태적 효과를 추정해 관광경기지수의 결정요인을 분석한다. 류시영·유선옥(2017)은 데이터의 범위를 블로그(네이버블로그, 다음블로그), 주요 커뮤니티 사이트 게시판(다음 아고라, 디시인사이드, 네이트판 등), 국내 트위터의 게시글 및 댓글로 한정하였으나 본 연구는 이동통신, 신용카드, 내비게이션, KTX 이용객 수 및 차량 통행량 등 관광경기과 보다 밀접하게 관련되고 정형화된 통계자료를 사용하여 관광경기를 측정한다. 더 나아가 본 연구는 관광경기가 지역실물경제에 미치는 영향을 분석한다는 점에서 류시영·유선옥(2017)의 연구와 차별화된다.

3. 연구 구성

본 연구는 다음과 같이 구성된다. III장에서는 관광경기지수 산정과정과 유효성 평가에 대해 설명한다. IV장은 산출된 관광경기지수의 결정요인 및 실물경제에 미치는 영향을 수치화하고 파급경로를 분석한다. V장은 분석결과를 요약하고 이를 바탕으로 정책점 시사점을 도출한다.

III. 관광경기지수 개발

영동지역 관광경기지수는 세 가지 조건을 만족해야 한다. **(조건 1)** 첫 번째 조건은

다양한 영동지역 관광경기 관련 지표에 공통적으로 내포된 정보를 반영하는 것이고, **(조건 2)** 두 번째 조건은 영동지역 관광경기를 실시간으로 모니터링하여 정책적 대응 방안을 신속하게 마련하기 위해 속보성을 갖추는 것이다. **(조건 3)** 세 번째 조건은 관광경기 지수가 관광업 취업자 수와 같은 영동지역 관광실물 지표와 통계적으로 유의한 당기 및 시차 상관관계를 갖는 것이다.

1. 통계자료 소개

(조건 1)과 (조건 2)를 충족하기 위해서는 짧은 주기로 수집할 수 있으면서도 관광경기기에 대한 정보를 반영하는 통계지표를 선정하는 것이다. 본 연구는 이러한 조건을 만족하는 통계자료로

- 신용카드 사용액(숙박·음식점업, 예술·스포츠·여가 관련 서비스업)⁵⁾
- 영동지역 고속도로 톨게이트 차량 통행량(1·6종)
- T-map 내비게이션 영동지역 목적지 검색 건수
- KT 이동통신 데이터 기준 영동지역 외지인 방문자 수
- 영동지역 시군별 관광 관련 네이버 검색어 트렌드

를 사용한다.⁶⁾ 표본기간은 모든 통계지표가 동시에 가용한 기간인 2018년 1월부터 2021년 6월이다.⁷⁾

<표 2>은 관광경지지수 산정뿐만 아니라 관광경지지수의 유효성 평가, 관광경기가 지역실물경제에 미치는 동태적 영향분석에 사용하는 모든 통계자료를 요약한 것이다.

5) 신용카드 사용액은 현지인과 외지인의 사용액이 합산된 것이다. 분석의 목적상 외지인의 신용카드 사용액만 사용하는 것이 바람직하지만 아쉽게도 자료가 가용하지 않다. 다만 관광경지지수 산정에 사용된 여타 통계지표와 신용카드 사용액이 비슷한 움직임을 보이는 것을 봤을 때, 현지인의 신용카드 사용액이 외지인의 사용액에 비해 안정적이거나 동조성이 높은 것으로 짐작된다.

6) 참고로 KTX 강릉선·동해선 하차객 수는 KTX 도입 초기 단계에 발생한 누적수요, 코로나19에 따른 발매 수 제한 및 여타 관광통계지표와의 낮은 상관관계 등을 고려할 때 데이터의 신뢰성과 안정성이 낮아 관광경지지수 산정에서 제외하였다.

7) 통계지표생성에 고려된 자료 중 네이버 검색어 트렌드를 제외한 나머지 자료들이 모두 2017년 1월부터 월별로 가용하다. 2017년 중 자료의 동태성과 2018년 이후의 동태성이 지나치게 이질적이어서 2018년 이후의 자료만을 사용하였다. 아마도 2017년은 통계자료수집 초기 단계여서 수집기준이나 방법이 안정화되지 않았기 때문으로 짐작된다.

<표 2> 통계자료 목록

자료	출처	비고
신용카드 사용액(내구재)	신한카드 하나카드	신한·하나카드 사용액 중 가맹지가 강원 영동지역으로, 소매재화가 내구재로 분류된 결제금액
신용카드 사용액(관광서비스)	신한카드 하나카드	신한·하나카드 사용액 중 가맹지가 강원 영동지역으로, 업종이 숙박·음식점업/예술·스포츠·여가 관련 서비스업으로 분류된 결제금액
고속도로 톨게이트 1·6종 입출입	한국도로공사	강원 영동지역 고속도로 톨게이트를 통과한 소형 승용·승용·승합차 대수
비제조업 BSI	한국은행 강릉본부	강원 영동지역 사업체 대상 설문조사를 통해 집계한 비제조업 업체들의 당월 업황 지수
내비게이션 목적지 검색건수	한국관광공사	강원 영동지역을 자동차로 방문한 인원의 지역내 목적지 내비게이션(T-map) 검색건수
이동통신 데이터 기준 외지인 방문자수	한국관광공사	이동통신(KT) 가입자 중 강원 영동지역을 방문한 외부방문객 수
관광업 고용보험 피보험자수	한국고용정보원	강원 영동지역에 등록된 고용보험 피보험자 중 업종이 숙박·음식점업/예술·스포츠·여가 관련 서비스업으로 분류된 인원수
고용보험 피보험자수	한국고용정보원	강원 영동지역에 등록된 고용보험 피보험자수
강원지역 소비자물가지수	통계청	총지수, 15년 1월=100
전국 숙박 및 음식점업 생산지수	통계청	서비스업생산지수의 숙박·음식점업 관련 하위 항목
전국 여행비지출전망 CSI	한국은행	소비자들의 6개월 후 여행비 지출에 대한 전망을 집계한 지수
네이버 검색어 트렌드(강릉)	네이버	강원 영동지역 시군별 관광 관련 검색건수 추이(강릉 홍수(18.8월), 속초 산불(19.4월) 등 극단값은 제외하고 직전월과 직후월의 평균으로 대체하였으며 가장 많이 검색된 시점(강릉 21.7월)을 100으로 설정)
네이버 검색어 트렌드(속초)		
네이버 검색어 트렌드(양양)		
네이버 검색어 트렌드(고성)		
네이버 검색어 트렌드(영월)		
네이버 검색어 트렌드(동해)		
네이버 검색어 트렌드(삼척)		
네이버 검색어 트렌드(평창)		
네이버 검색어 트렌드(태백)		
네이버 검색어 트렌드(정선)		

2. 지수산정 방법

본 연구는 관광경기 관련 통계지표에 주성분 분석(Principal Component Analysis)을 적용하여 지수를 산정한다. 주성분 분석은 다수의 시계열 자료를 대표하는 하나의 지수를 산정하는 목적으로 가장 널리 사용되는 차원축소 방법이다. 주성분 분석으로 계산되는 주성분의 수는 관측변수와 동일하다. 각 주성분은 관측 시계열의 선형조합으로 표현

되며, 각 시계열의 중요도에 따라서 가중치가 부여된다. 주성분들은 원분포에 대한 설명력을 기준으로 정렬할 수 있으며, 통상적으로 설명력이 가장 높은 주성분이 지수로 활용된다. 따라서 주성분 분석을 통해 다양한 관광경기 관련 지표들에 설명력이 가장 큰 시계열을 추출하고 이를 관광경기지수로 활용한다.

주성분 분석 이외에 널리 사용되는 지수산정 방법은 회귀분석기반 지수산정법이 있다. 기본적으로 주성분 분석은 실물경기와 무관하게 관광경기 관련 지표들만으로 추정된다는 단점이 있다. 반면 회귀분석기반 지수산정법은 관광경기지수가 개별 관광경기 관련 지표의 선형결합으로 구성되고 단기 경기예측에 활용된다는 성질을 감안하여, 분포 시차(distributed lag) 회귀분석 기반으로 관광경기지수를 산정하는 방식이다. 종속변수는 정책목표에 해당하는 고용이나 소비 등 실물경제 변수이며 설명변수는 관광경기 관련 통계지표들이다. 하지만 회귀분석 기반 지수산정법은 어떤 종속변수를 쓰느냐에 따라서 결과가 민감하게 달라진다는 한계가 있다. 게다가 종속변수가 관광경기 통계지표와 동일한 빈도로 관측되고 동일한 수준의 속보성을 지닌다는 조건도 만족해야 한다. 본 연구에서 사용하는 종속변수는 관광업 분야 고용보험 피보험자 수인데, 이는 속보성이 상대적으로 현저하게 떨어진다. 또한 정책당국이 관광경기지수를 사용하는 목적이 취업자 수에 국한되는 것이 아니기 때문에 회귀분석에 기반하여 산정된 지수는 정책적 유용성이 낮다는 한계가 있다. 이러한 이유로 본 연구는 회귀분석기반 지수산정법 대신 주성분 분석을 사용하고자 한다.

참고로, 주성분 분석과 회귀분석을 결합하여 지수를 산정할 수도 있다. 회귀분석 기반 지수 산정법은 다수의 관광경기 관련 지표 중 기계학습을 통한 차원축소 과정에서 극소수의 변수만이 선택될 가능성이 있다. 이러한 경우를 방지하기 위해 주성분 분석과 회귀분석의 결합방식은 관광경기 관련 지표 관측치 대신 동일한 수의 주성분을 설명변수로 하여 분포 시차 회귀분석을 실시하는 방식인데, 이 방식은 회귀분석기반 지수산정법과 마찬가지로 속보성과 범용성 측면에서 동일한 문제점이 있다.

3. 지수산정 과정

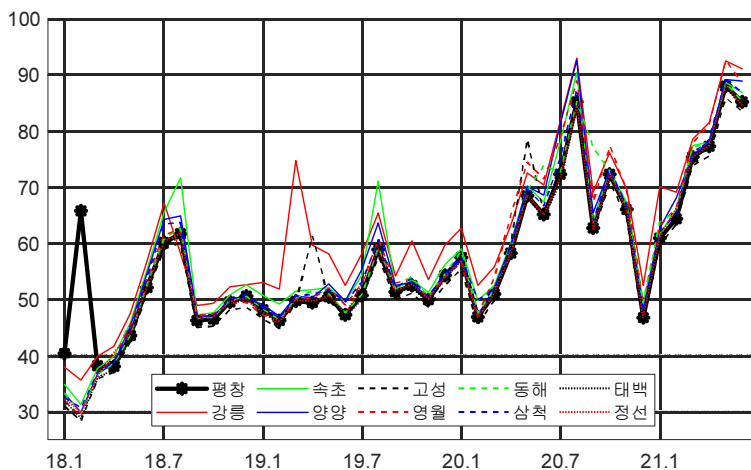
앞서 선정된 관광경기 관련 통계지표를 활용하여 주성분 분석을 통해 영동지역 관광경기지수를 산정하며, 이는 총 2단계에 걸쳐서 진행한다.

[단계 1] 네이버 검색 지수 산정

본 연구의 영동지역 관광경기 통계지표 중에는 영동지역 10개 시군 네이버 검색어 트렌드(이후 네이버 트렌드)가 포함된다. <그림 2>은 시군별 네이버 트렌드의 시계열을 나타낸 것인데, 이 표를 보면 10개 시군의 네이버 트렌드가 대체로 유사하다는 것과 평

창과 강릉의 동태성이 여타 시군과 다소 차이가 있다는 것을 볼 수 있다. 우선 평창의 경우 2018년 1월 네이버 트렌드 수치가 다른 시군보다 높게 나오는 데 이는 거의 전적으로 동계올림픽의 영향으로 보인다. 다음으로 2019년 4월 강릉의 네이버 트렌드 급등은 4월 초에 발생한 강릉·고성 지역 산불의 영향으로 판단된다. 여기서 동계올림픽은 관광경기 상승으로 인한 고용증가와 밀접한 관련이 있는 반면 산불은 그렇지 않다. 동계올림픽의 효과를 관광경기지수에 일부 반영하기 위해 평창의 네이버 트렌드를 독자적인 관광경기 통계지표로 분류하는 대신, 강릉을 포함한 나머지 9개 시군의 네이버 트렌드 자료에 주성분 분석을 적용하여 1차 주성분을 추출하였다.⁸⁾ 가성회귀 가능성을 배제하기 위해 모든 자료에서 시간 추세를 제거하였다. 주성분 분석 결과, 1차 주성분이 총 변동(total variation)을 설명하는 비중은 0.926으로 추정되므로 1차 주성분이 대부분의 총 변동을 설명하며 9개 네이버 트렌드 자료의 공통요인인 것으로 판단된다. <그림 3>의 패널 (a)는 평창의 네이버 트렌드를, 패널 (b)는 나머지 9개 시군의 네이버 트렌드 주성분을 나타낸 것이다.

<그림 2> 네이버 트렌드



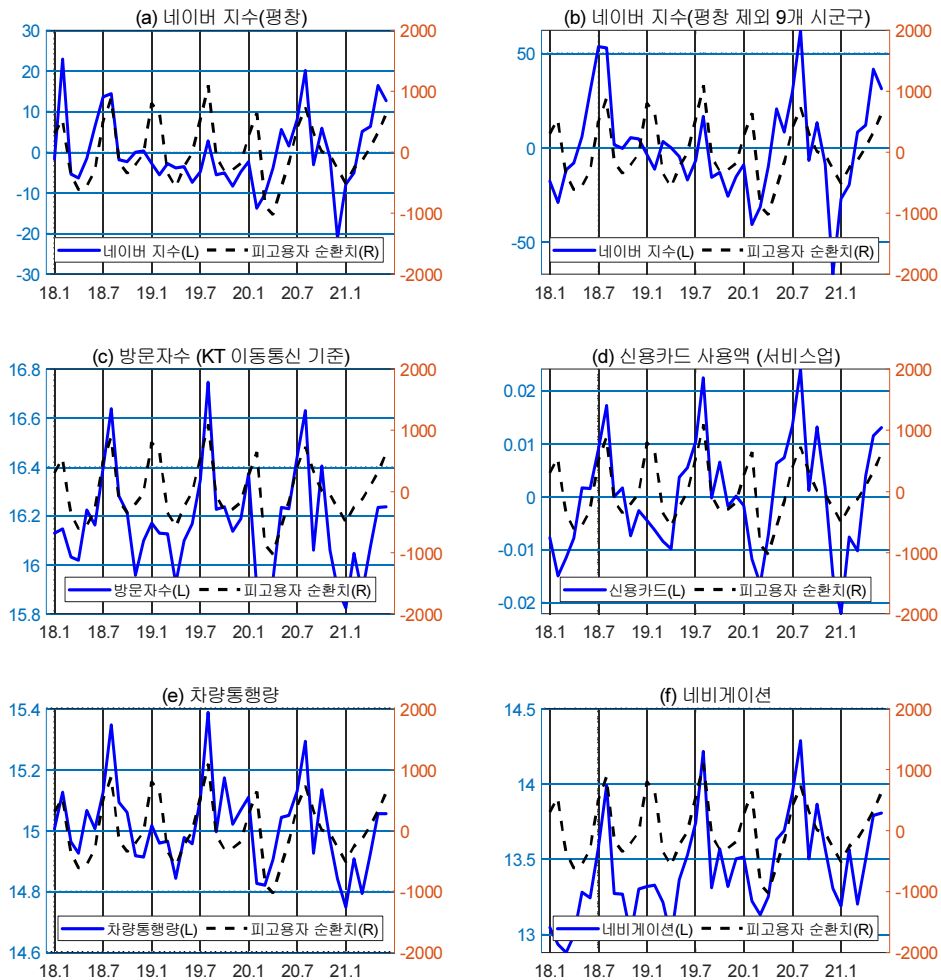
[단계 2] 관광경기지수 산정

두 번째 단계에서는 <그림 3>에 나타난 총 6개의 영동지역 관광 관련 통계지표(평창의 네이버 트렌드, 여타 9개 시군 네이버 트렌드 주성분, KT 이동통신 데이터 기준 외지인 방문자수, 신용카드 사용액(숙박·음식점업, 예술·스포츠·여가 관련 서비스업), 고

8) 평창의 네이버 트렌드를 따로 분류하지 않고 10개 시군구의 네이버 트렌드를 대상으로 주성분 분석을 하더라도 본 연구의 전체적인 통계분석결과에는 유의미한 차이가 발생하지 않았다.

속도로 톨게이트 차량 통행량(1·6종), 내비게이션 목적지 검색건수)에 주성분 분석을 실시한다. 기본적으로 주성분 분석은 변수들간의 공통 요인을 잡아내는 것이므로 변수간 상관관계가 높을수록 주성분 분석 적용이 적절하다고 사전에 판단할 수 있다. <그림 3>의 점선은 관광업 분야 고용보험 피보험자의 순환치를 나타낸 것인데, 모든 통계지표가 비슷한 시계열적 행태를 나타낸다는 것을 시각적으로 충분히 짐작할 수 있다.⁹⁾ 실제 변수들의 상관관계 행렬을 추정한 결과, 모든 상관관계들이 높은 양의 값을 갖는 것으로 추정된다.(<표 3> 참조)

<그림 3> 관광경기 통계지표와 고용보험 피보험자 순환치



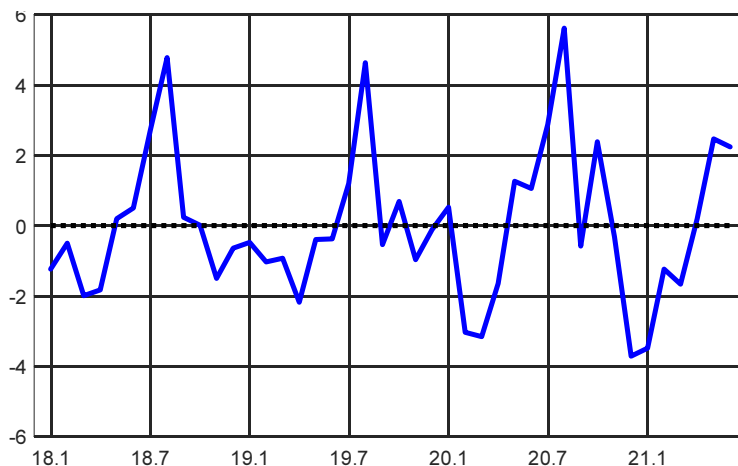
9) 고용보험 피보험자수 순환치 추정과정은 III-(4)절에서 논의한다.

<표 3> 관광경기 통계지표 간 상관관계

	신용카드	네이버 (평균)	네이버 (평균제외)	방문자수 (KT)	차량 통행량	내비게이션
신용카드	1.000					
네이버 (평균)	0.555	1.000				
네이버 (평균제외)	0.749	0.799	1.000			
방문자수 (KT)	0.872	0.558	0.655	1.000		
차량 통행량	0.813	0.575	0.585	0.958	1.000	
내비게이션	0.887	0.446	0.593	0.795	0.701	1.000

6개의 영동지역 관광경기 통계지표로부터 추출된 1차 주성분을 <그림 4>에 나타내었으며, 이 값을 영동지역 관광경기지수로 사용한다. 이때 1차 주성분이 통계지표들의 총변동을 75% 이상 설명하는 등 대표성이 충분하고 볼 수 있다. 한편 1차 주성분에 포함된 신용카드, 네이버(평균), 네이버(평균제외), 방문자수(KT), 차량 통행량, 내비게이션 통계자료에 대한 가중치(특성벡터)는 0.35~0.44 내외로 산정된 관광경기지수가 개별 관광 관련 통계지표들을 대체로 균일하게 반영한다. 참고로 1차와 2차 주성분의 총분산 설명력은 각각 75.6%와 12.3%이고 3차 주성분은 6.5%에 불과하다. 2차 주성분의 경우에는 네이버 트렌드에 대한 가중치가 다른 변수에 대한 가중치보다 절대적으로 높기 때문에 2차 주성분은 네이버 트렌드의 정보를 반영하는 것으로 볼 수 있다.

<그림 4> 관광경기지수 추이



4. 유효성 평가

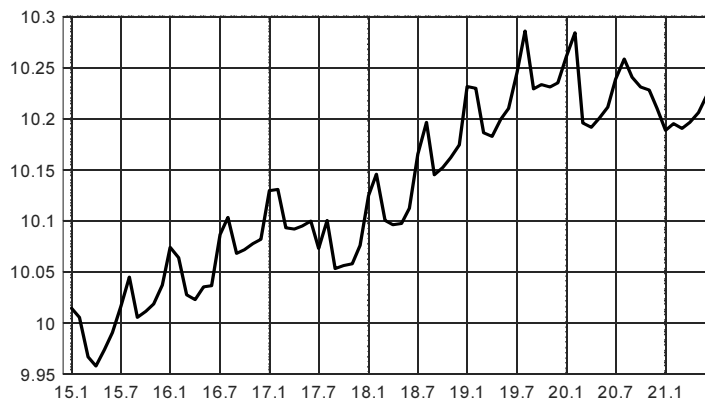
영동지역 관광경기지수가 유효하기 위한 (조건 3)은 영동지역 관광실물경기과 통계적으로 유의한 당기 또는 시차 상관관계를 가지는 것이다. 앞서 주성분 분석으로부터 추정된 관광경기지수가 이 조건을 만족하는지를 판단하기 위해 본 연구는 영동지역 관광실물경기의 대리변수(proxy variable)를 제시한 다음, 관광경기지수와 관광실물경기 간 당기 및 시차 상관관계의 통계적 유의성을 검증한다.

1) 관광실물경기의 측정

본 연구는 영동지역 관광실물경기를 영동지역 관광업 고용보험 피보험자수로 측정한다. 그 이유는 고용보험 피보험자수가 관광업의 산출량에 비해 지역사회후생에 밀접히 관련될 뿐만 아니라 월별자료(2015년 1월~)가 가용하기 때문이다. 또한 고용보험 피보험자수는 고용보험 가입자를 대상으로 하기 때문에 고용의 질을 일부 반영한다는 장점이 있다. 한편 <그림 5>에서 보는 바와 같이 로그 관광업 피고용자수는 불안정계열이다. 반면 관광경기지수는 안정계열이기 때문에 두 변수간 (시차) 상관관계를 편의없이 추정하기 위해 피보험자수의 추세를 제거한다.

통상적으로 거시변수의 추세제거에는 Hodrick-Prescott (HP) 필터가 주로 사용된다. 월별자료에 주로 사용되는 파라미터를 설정하고 HP 필터를 실시하였으나, 추세가 지나치게 선형으로 추정되어 순환부분이 I(1)에 가깝게 산출되었다. 이는 아마도 시계열이 짧고 임의보행 과정에 속한 상수항의 시변성이 강하기 때문인 것으로 보인다. 다양한 파라미터들을 시도해보았으나 결과는 별로 개선되지 않았으며, 이런 이유로 불가피하게 본고는 관광업 로그 피고용자수 추세제거의 적합한 추세제거 모형을 따로 설계하여 추정한다.

<그림 5> 관광업 고용보험 피보험자수(로그) 추이



피보험자수는 결정적 시간 추세(deterministic time trend) 보다는 확률적 추세(stochastic trend)를 가지기 때문에 은닉인자 모형(unobserved component model)을 설정하여 추세를 추정한다. 우선 Emp_t 는 t 월의 로그 피보험자수를 나타내고, Emp_t^{Trend} 과 Emp_t^{Cycle} 는 각각 확률적 추세와 순환변동을 나타낸다. 로그 피보험자수는 확률적 추세와 순환치의 합으로 결정된다.

$$Emp_t = Emp_t^{Trend} + Emp_t^{Cycle}$$

확률적 추세는 시변하는 증가율(μ_t)을 갖는 임의 보행(random-walk)을 가지며, 순환 변동분은 안정적인 AR(2) 과정을 따른다고 가정한다. BIC를 기준으로 순환부분의 자기 상관 시차를 2로 선택하였으며 모든 충격은 정규분포를 따른다고 가정한다.

$$\begin{aligned} Emp_t^{Trend} &= \mu_{t-1} + Emp_{t-1}^{Trend} + v_t^{Trend} \\ \mu_t &= \mu_{t-1} + v_t^{Drift} \\ Emp_t^{Cycle} &= \phi_1 Emp_{t-1}^{Cycle} + \phi_2 Emp_{t-2}^{Cycle} + v_t^{Cycle} \end{aligned}$$

추정을 위해 모형을 상태공간모형 형태로 나타내면 아래와 같다.

(추정식)

$$Emp_t = Emp_t^{Trend} + Emp_t^{Cycle} = (1 \ 1 \ 0 \ 0) \begin{pmatrix} Emp_t^{Trend} \\ Emp_t^{Cycle} \\ Emp_{t-1}^{Cycle} \\ \mu_t \end{pmatrix}$$

(전환식)

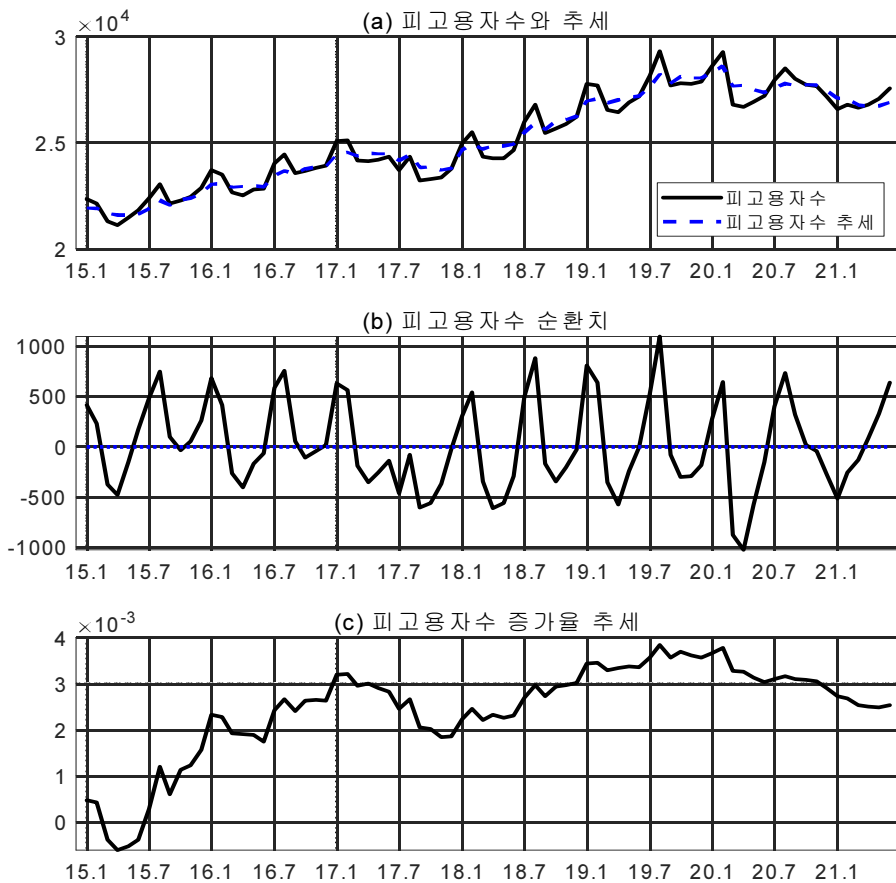
$$\begin{pmatrix} Emp_t^{Trend} \\ Emp_t^{Cycle} \\ Emp_{t-1}^{Cycle} \\ \mu_t \end{pmatrix} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & \phi_1 & \phi_2 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} Emp_{t-1}^{Trend} \\ Emp_{t-1}^{Cycle} \\ Emp_{t-2}^{Cycle} \\ \mu_{t-1} \end{pmatrix} + \begin{pmatrix} v_t^{Trend} \\ v_t^{Cycle} \\ 0 \\ v_t^{Drift} \end{pmatrix}$$

$$\begin{pmatrix} v_t^{Trend} \\ v_t^{Cycle} \\ v_t^{Drift} \end{pmatrix} \sim N \left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{Trend}^2 & & \\ & \sigma_{Cycle}^2 & \\ & & \sigma_{Drift}^2 \end{pmatrix} \right)$$

추정방법은 최우추정(maximum likelihood estimation)이며, pile-up 문제를 방지하기 위

해 median estimator를 사용한다. 추정에 사용된 자료는 통계자료가 수집되기 시작한 2015년 1월 이후 월별 관측치이다. 로그 피보험자수를 이용하여 추정하였으며, 로그 추세와 로그 순환치 추정값에 지수를 취하여 피보험자수로 복원하였다. 추정결과는 <그림 6>에 나타내었다.

<그림 6> 관광업 고용보험 피보험자수 추세-순환 분해 결과



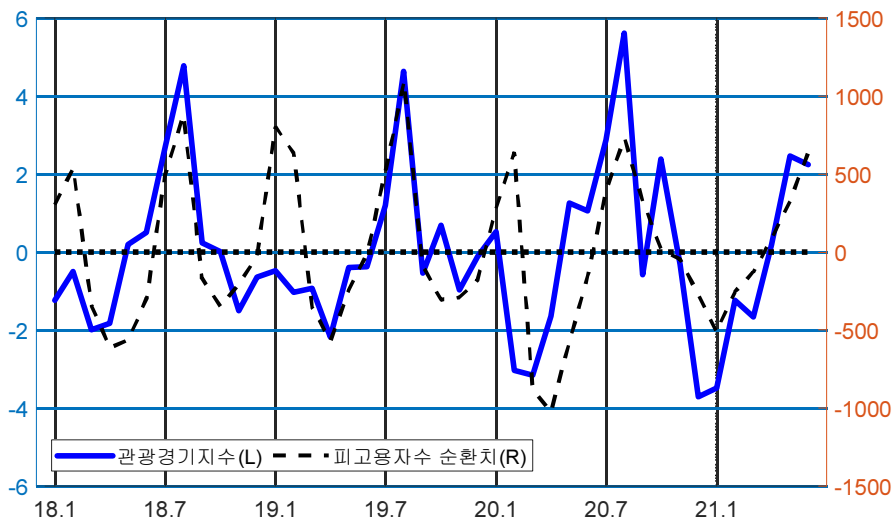
<그림 6>을 보면 피보험자수가 19년 말까지 추세적으로 상승하다가 코로나19 여파가 시작된 20년 1월부터 하향 추세로 전환되었음을 알 수 있다. 추정된 순환치는 전국경기, 지역경기, 계절성, 타지역과 해외 관광상품에 대한 지역관광상품의 상대적 경쟁력 변화를 반영한다. 매년 7, 8월은 여행성수기 여파로 순환치가 상승하는 반면, 매년 1, 2월은 연도별로 상이하다. 18년부터 20년까지는 순환치가 양수이지만 21년에는 순환치가 음수를 가진다. 이는 20년 1, 2월은 코로나19 상황의 초창기여서 여파가 크지 않아 이전

과 유사한 순환변동치를 나타내었던 반면, 21년 1, 2월은 코로나19 3차 유행의 영향으로 이전과 달리 순환치가 하락한 것으로 보인다.

2) 관광실물경기와 관광경기지수 간 당기 및 시차상관관계

피보험자수 순환치와 관광경기지수를 <그림 7>에 함께 나타내었다. 이 그림을 보면 관광경기지수와 고용변동이 18년 1월과 19년 1월경을 제외한 기간에서 전반적으로 상당히 유사함을 알 수 있다.

<그림 7> 관광경기지수와 관광업 고용보험 피보험자수 순환치



두 변수의 상관관계를 통계적으로 검증하기 위해 관광실물경기에 대한 $ADL(p)(p$ -month-ahead autoregressive distributed lag) 모델을 설정하고 추정한다.

- 당기상관관계 추정식($p = 0$)

$$Emp_t^{Cycle} = \sum_{d=1}^{12} M_{t,d} \delta_d + Emp_{t-1}^{Cycle} \phi + X_t' \gamma + TBCI_t \beta + error_t$$

- 시차상관관계 추정식($p = 1, 2, \dots$)

$$Emp_t^{Cycle} = \sum_{d=1}^{12} M_{t,d} \delta_d + Emp_{t-p}^{Cycle} \phi + X_{t-p}' \gamma + TBCI_{t-p} \beta + error_t$$

$M_{t,d}(d = 1, 2, \dots, 12)$ 는 월별 더미변수이고 Emp_{t-p}^{Cycle} 는 시차변수, X_{t-p} 는 시차통제 변

수의 벡터, $TBCI_{t-p}$ 는 관광경기지수이다. 본 연구는 강원 영동지역 관광경기가 전국 관광 수요의 영향을 받는다는 점을 감안하여 전국 여행비지출전망 CSI와 전국 숙박 및 음식점업 생산지수를 통제변수로 사용한다. $TBCI_{t-p}$ 의 계수인 β 의 통계적 유의성을 검증함으로써 관광경기지수의 관광실물경기에 대한 설명력 또는 예측력 여부를 판단할 수 있다.

추정결과는 <표 4>에 나타내었다. 추정결과를 보면 당기 상관관계와 1개월 시차상관관계가 통계적으로 유의한 양의 값으로 추정되었다. 또한 $TBCI_{t-p}$ 이 포함된 식과 포함되지 않은 식의 R^2 와 $\overline{R^2}$ 를 비교하면, $TBCI_{t-p}$ 가 포함된 식의 R^2 와 $\overline{R^2}$ 모두 상당히 크게 상승하였다. 당기 상관관계식의 경우, 관광경기지수가 설명변수로 포함되면 R^2 와 $\overline{R^2}$ 가 각각 0.792에서 0.825로, 0.680에서 0.721로 증가한다. 1개월 시차상관관계식의 경우에는 R^2 와 $\overline{R^2}$ 가 각각 0.790에서 0.846으로, 0.676에서 0.753으로 증가한다. 반면 2개월 이상의 시차에 대해서는 통계적으로 무의미한 것으로 나타났으며, $\overline{R^2}$ 의 개선

<표 4> 관광실물경기와 관광경기지수 간 당기 및 시차상관관계식 추정결과

시차	P=0			P=1			P=2	
	추정치	S.E.		추정치	S.E.		추정치	S.E.
더미(1월)	1.728*	0.709		2.078**	0.688		0.851	1.050
더미(2월)	1.644*	0.651		1.803**	0.611		1.665	1.042
더미(3월)	-2.186**	0.718		-1.938**	0.676		-2.173*	0.911
더미(4월)	-0.807	0.717		-0.558	0.688		-2.311*	1.009
더미(5월)	-0.583	0.734		0.265	0.694		-1.299	1.076
더미(6월)	0.111	0.660		-0.316	0.647		-0.222	1.104
더미(7월)	1.359	0.811		1.980**	0.669		1.671	1.123
더미(8월)	0.556	1.178		1.565*	0.768		3.509**	1.005
더미(9월)	-1.786*	0.910		-4.495**	1.185		-0.324	1.201
더미(10월)	-1.170	0.699		-0.441	0.675		-1.223	1.998
더미(11월)	0.123	0.714		-1.000	0.683		-0.607	1.028
더미(12월)	0.435	0.727		0.170	0.680		-0.848	1.018
시차변수	0.517**	0.148		0.350*	0.154		-0.051	0.238
전국여행비 지출전망CSI	0.005	0.029		0.010	0.031		-0.009	0.052
전국 숙박 및 음식점업 산지수	0.016	0.030		0.008	0.033		0.032	0.059
관광경기지수	0.429*	0.196		0.638**	0.212		0.132	0.353
R^2	0.825			0.846			0.669	
$\overline{R^2}$	0.721			0.753			0.462	

주: *는 $p<0.05$, **는 $p<0.01$

효과도 존재하지 않는다. 따라서 본 연구에서 산출된 관광경기지수는 관광실물경기에 대한 상황 판단 및 단기 선행지표로서 유효성이 클 것으로 추정된다.

IV. 관광경기가 실물경제에 미치는 영향

1. 분석 모형

VAR 분석시 내생성 문제를 해결하면서 동시에 통계적 추론에 충분한 정보를 제공하기 위해서는 많은 내생변수가 포함된 VAR system이 필요하다. 하지만 system의 차원이 큰 VAR의 경우, 시차 항을 하나만 늘려도 파라미터 수가 급격히 증가하여 차원의 저주(curse of dimensionality)가 발생하는 경향이 강하다. 차원의 저주가 발생하게 되면 추정의 불확실성이 증가하고 과적합 문제가 발생하여 예측 정확도가 떨어지게 된다. 또한 충격반응함수의 통계적 유의성이 크게 약화된다. 따라서 고차원(Large scale) VAR 추정 시 가장 중요한 이슈는 모형의 설명력을 높게 유지하면서 과적합 현상을 해결하는 것이다.

이를 위해서는 변수누락(missing variable) 문제를 일으키지 않으면서 동시에 불필요한 회귀계수에 페널티를 부여해야 한다. 일변수 선형회귀에 적용되는 대표적인 방식이 바로 Ridge 회귀분석이나 Lasso 회귀분석과 같은 shrinkage 기법이다. 수학적으로 사전 정규분포하에서 선형 모형을 추정했을 때 사후 분포 최빈값은 ridge 추정량과 같고, 사전 라플라스 분포 하에서 선형 모형을 추정했을 때 사후 분포 최빈값은 Lasso 추정량과 같다. 일변수 선형회귀와 달리 다변량 모형에서의 shrinkage 기법 역시 사전 분포 설정을 통해 적용하는데 그러한 방법 중 가장 대표적이고 널리 사용되는 것이 베이지안 VAR 모형에 미네소타 사전(Minnesota prior) 분포를 적용하는 것이다.

지금부터는 미네소타 사전 분포 하의 베이지안 VAR 모형을 구체적으로 설명하고자 한다. 우선 $y_t = (y_{1t} \ y_{2t} \ \dots \ y_{kt})'$ 는 k 개 변수로 이루어진 k -차원 벡터이다. 주어진 모수 하에서 y_t 는 조건부 VAR(p) 확률과정을 따른다고 가정한다.

$$y_t \mid \theta, Y_{t-1} \sim \text{Normal}(\Phi_1 y_{t-1} + \Phi_2 y_{t-2} + \dots + \Phi_p y_{t-p}, \Omega)$$

$Y_{t-1} = \{y_i\}_{i=1}^{t-1}$ 는 $(t-1)$ 시점까지의 관측치를 나타내고, $\Phi_i (i = 1, 2, \dots, p)$ 는 $k \times k$ VAR 계수행렬이고, Ω 는 축약형 오차항의 $k \times k$ 분산-공분산 행렬이다. θ 는 VAR 계수 행렬과 분산-공분산 행렬을 포함한 모형의 모든 모수집합이다. 위 VAR 계수행렬을 벡터로 표시하게 되면 아래와 같다.

$$\beta = \text{vec} \begin{pmatrix} \Phi_1' \\ \Phi_2' \\ \vdots \\ \Phi_p' \end{pmatrix} : k^2 p \times 1$$

위 VAR 모형은 일반화된 최소자승법을 적용할 수 있는 통상적인 선형회귀형태,

$$y_t = x_t' \beta + \text{Normal}(0, \Omega)$$

로 표현할 수 있다. 단, $x_t' = I_k \otimes (y_{t-1}' \ y_{t-2}' \ \cdots \ y_{t-k}'): k \times k^2 p$ 는 시점별 설명변수 행렬이다.

베이지안 모델링은 사전분포를 설정함으로써 완결된다. 모수의 사전분포는 깃스 샘플링이 가능하도록 켄레(conjugate)사전분포로 설정하고자 한다. 먼저 오차항의 분산-공분산 행렬은 역위사트 분포를 따른다고 가정한다.

$$\Omega \sim \text{Inverse Wishart}(\nu_0, R_0 = \Omega_0(\nu_0 - k - 1))$$

이때, Ω 의 사전 평균은 $E[\Omega] = \Omega_0 = \text{diag}([\omega_1^2, \omega_2^2, \dots, \omega_k^2])$ 이다. 사전 평균은 대각행렬로 가정하며, i 번째 대각요인, ω_i^2 은 i 번째 변수의 관측치에 대한 AR(p) 모형을 최소자승법으로 추정한 오차항의 분산 추정치이다.

다음으로 λ 과 γ 는 각각 역감마 분포,

$$\lambda \sim \text{Inverse Gamma}(\alpha_0, \delta_0)$$

$$\gamma \sim \text{Inverse Gamma}(\tau_0, \kappa_0)$$

를 따른다고 가정한다. λ 과 γ 가 주어지면 $\phi_{ij}^{(r)}$ 는 Φ_r 의 (i, j) 인자이고, 평균이 $b_{ij}^{(r)}$, 분산이 $v_{ij}^{(r)}$ 인 정규분포,

$$\phi_{ij}^{(r)} \mid \lambda, \gamma \sim \text{Normal}(b_{ij}^{(r)}, v_{ij}^{(r)})$$

을 따른다고 가정한다.

이때, 사전 평균은

$$b_{ij}^{(r)} = 0.5 \text{ if } r=1 \text{ and } i=j$$

$$b_{ij}^{(r)} = 0, \text{ otherwise}$$

와 같이 설정한다. 이러한 사전 평균은 거시변수의 지속성을 고려한 것인데, 종속변수와 동일한 시차변수의 영향은 평균적으로 0.5이고, 여타 시차변수의 영향은 평균적으로 0이라고 가정한다. 사전 분산은 shrinkage 효과를 반영하기 위해 모든 $r = 1, 2, \dots, p$ 에 대해서

$$v_{ij}^{(r)} = \frac{\lambda}{r^2} \text{ if } i=j$$

$$v_{ij}^{(r)} = \frac{\lambda}{r^2} \gamma \frac{\omega_i}{\omega_j} \text{ if } i \neq j$$

으로 설정된다. 여기서 λ 는 전반적인 shrinkage 정도를 잡아내고, $\gamma (< 1)$ 는 횡단-시차 계수(cross-lag coefficients)의 shrinkage 정도를 λ 에 추가한 것이다. 한편 ω_i/ω_j 은 변수들 간의 scale 차이를 조정하는 역할을 한다.

결과적으로 시차계수행렬의 사전분포는

$$\Phi_r \mid \lambda, \gamma \sim$$

$$Multi-Normal \left(\Phi_{r,0} = \begin{bmatrix} b_{11}^{(r)} & b_{12}^{(r)} & \dots & b_{1k}^{(r)} \\ b_{21}^{(r)} & b_{22}^{(r)} & & \\ \vdots & & \ddots & \\ b_{k1}^{(r)} & & & b_{kk}^{(r)} \end{bmatrix}, V_r = \begin{bmatrix} v_{11}^{(r)} & v_{12}^{(r)} & \dots & v_{1k}^{(r)} \\ v_{21}^{(r)} & v_{22}^{(r)} & & \\ \vdots & & \ddots & \\ v_{k1}^{(r)} & & & v_{kk}^{(r)} \end{bmatrix} \right)$$

이다. 따라서 β 의 조건부 사전분포는

$$\beta \mid \lambda, \gamma \sim Normal \left(\beta_0 = vec \left(\begin{bmatrix} \Phi'_{1,0} \\ \Phi'_{2,0} \\ \vdots \\ \Phi'_{p,0} \end{bmatrix} \right), B_0 = diag \left(vec \left(\begin{bmatrix} V'_1 \\ V'_2 \\ \vdots \\ V'_p \end{bmatrix} \right) \right) \right)$$

이 된다. $\beta_0: k^2 p \times 1$ 은 β 의 사전 평균이고 $B_0: k^2 p \times k^2 p$ 는 사전 분산-공분산이다.

Litterman(1980)은 VAR을 통해 변수를 예측하고자 할 때, 다른 변수들의 과거값보다는 예측하고자 하는 대상의 과거값에 더 가중치를 줘야 하고, 먼 과거의 관측치보다는 가까운 거리의 관측치에 더 가중치를 줘야 한다고 주장한다. VAR 예측에서의 가중치관 결국 계수 추정치를 의미하므로, 어느 변수의 가중치를 적게 준다는 것은 결국 그 변수에 대응하는 계수 추정치를 0에 가깝게 두겠다는 것과 동일하다. 기존의 미네소타 사전분포

는 여러 변형된 형태가 존재하지만 사전에 설정된 하이퍼-파라미터(Hyper-parameter)가 고정된다는 한계가 존재한다.

최근 Giannone, Lenza, and Primiceri (2015)은 하이퍼-파라미터 역시 추정해야 할 대상으로 보고 그 값들이 데이터에 의해 결정되도록 설계하는 방법을 고안하였다. 특히 기존 미네소타 사전분포에 더하여 계수 행렬에 대해 두 가지 사전분포를 추가로 도입하면 모형의 설명력이 향상될 수 있음을 보였다. 이러한 이유로 본 연구도 λ 과 γ 를 특정 값에 고정하지 않고 계층적 사전분포(hierarchical prior) 틀 안에서 추정해야 할 모수로 간주하고 사후분포를 시뮬레이션한다. $(\Omega_0, \lambda, \gamma)$ 에 대한 하이퍼-파라미터값은 <표 5>에서 볼 수 있다.

<표 5> 모수의 사전분포

하이퍼-모수	값		하이퍼-모수	값
α_0	100		ω_1^2	32.5198
δ_0	$0.16 \times \alpha_0$		ω_2^2	0.0585
τ_0	100		ω_3^2	0.1122
κ_0	$0.04 \times \tau_0$		ω_4^2	0.0047
ν_0	$k + 2$		ω_5^2	3.1515
			ω_6^2	0.9646
			ω_7^2	2.6617

2. 추정결과

추정에 사용된 자료는 앞서 산출된 영동지역 관광경기지수와 더불어 전국 관광경기를 반영하는 전국 여행비지출전망 CSI, 지역 가격변수로는 강원도 물가상승률(전품목)과 강원도 물가상승률(서비스제)을 사용한다. 전국 여행비지출전망 CSI는 전국 관광경기 효과를 통제하는 역할을 하며, 강원도 전품목 물가상승률과 서비스제 물가상승률은 지역관광경기가 지역의 전반적인 물가수준과 관광업 분야의 물가수준에 미치는 영향을 보기 위함이다. 다음으로 지역관광경기가 지역 소비 또는 실물경제에 미치는 영향을 살펴보기 위해 신용카드 사용액(내구재 및 비내구재)을 고려한다. 마지막으로 주요한 정책목표인 고용에 대한 관광경기의 파급효과를 분석하기 위해 고용보험 피보험자수(전산업) 전년동기대비 증가율과 고용보험 피보험자수 순환치(관광업)를 내생변수에 포함한다.

구조충격은 Recursive 제약으로 식별하고, 영동지역 관광경기지수의 장단기 결정요인과 관광경기충격이 고용을 포함한 실물경제에 미치는 영향에 대한 파급경로를 분석한다. Recursive 구조제약 부여시 외생성이 강한 순으로 내생변수 벡터를 구성해야 하는데, 우선 전국 변수인 전국 여행비지출전망 CSI를 가장 외생적인 변수로 설정한다. 물가는

산출량에 비해 경직적이기 때문에 강원도 물가상승률(전체)과 강원도 물가상승률(서비스재)은 두 번째와 세 번째로 외생적인 변수로 설정한다. 전국 관광경기의 영향을 받는 영동지역 관광경기지수가 그 다음에 위치한다. 지역 실물변수이면서 외지인의 관광소비와 밀접히 관련된 로그 신용카드 사용액이 다섯 번째 변수이다. 지역 전체 경기와 관광경기의 영향을 받는 고용보험 피보험자수(전산업) 전년동기대비 증가율과 고용보험 피보험자수 순환치(관광업)가 각각 여섯 번째와 일곱 번째 변수에 해당한다. 자료는 월별이며 표본 기간은 2018년 1월부터 2021년 6월이다.

1) 시차선택

충격반응함수의 추정결과는 시차에 크게 의존하기 때문에 시차의 결정을 대단히 엄밀하게 할 필요가 있다. 최소자승법에 기반한 기존 연구들은 대부분 주변우도의 근사치인 BIC 또는 AIC를 기준으로 시차를 결정하였는데, Large scale VAR을 최소자승법으로 추정하면 shrinkage가 적용되지 않아 과적합문제가 발생한다. 이로 인해 BIC나 AIC로 시차선택을 하면 대부분의 경우 시차가 1로 선택된다. 시차가 2가 되면 패널티 항이 급격히 증가하기 때문이다. 또한 미네소타 사전분포를 적용하여 베이지안으로 추정하더라도 BIC나 AIC는 근사 오차(approximation error)로 인해서 시차를 과소추정하는 경향이 있다. 이런 이유로 본 연구는 BIC나 AIC에 의존하지 않고 주변우도를 직접 추정하여 최적 시차를 결정한다.

시차별 주변우도 추정결과는 <표 6>에 제시하였다. 시차가 2일 때 로그 주변우도가 극대화되며, 시차가 1이거나 3 이상인 모형의 로그 주변우도보다 6 정도의 차이가 나므로 통계적으로도 유의하다고 볼 수 있다. 이 표의 결과를 좀 더 살펴보면, 시차가 커질수록 우도값이 지속적으로 증가하지만, 시차가 3 이상이 되면 모형복잡성에 대한 패널티 항(로그 사후밀도 - 로그 사전밀도)이 더 빠르게 증가하면서 로그 주변우도는 감소하는 것을 볼 수 있다. 마지막으로 이 표의 하단에 최소자승법으로 추정한 BIC 결과를 추가로 제공한다. 예상한 바대로 시차가 1일 때 BIC가 최소화되는 것을 볼 수 있다.

<표 6> 시차선택결과

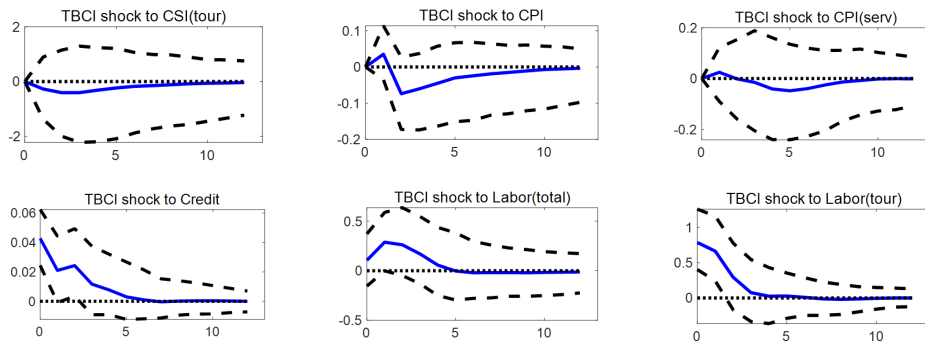
시차	1	2	3	4
로그 주변우도	-309.99	-299.78	-305.86	-335.62
로그 우도	-326.14	-300.33	-294.61	-261.56
로그 사전밀도	-5.87	25.88	57.53	54.60
로그 사후밀도	-22.02	25.34	68.79	128.66
BIC	812.72	858.28	949.06	920.24

2) 충격반응함수

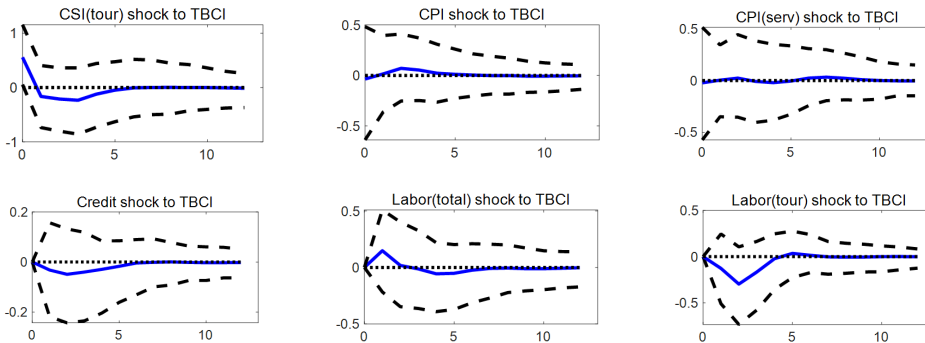
<그림 8>는 최적 시차 2에서의 충격반응함수 추정결과 중 일부를 나타낸다. 전체 추정결과와는 <그림 9>에서 확인할 수 있다. 충격반응함수의 추정결과를 크게 두 가지로 분류할 수 있는데, 하나는 영동지역 관광경기지수 충격이 고용 등 실물경제 변수에 미치는 동태적 영향이고, 다른 하나는 고용 등 실물경제 변수 충격이 영동지역 관광경기지수에 미치는 동태적 영향이다.

<그림 8> 충격반응함수

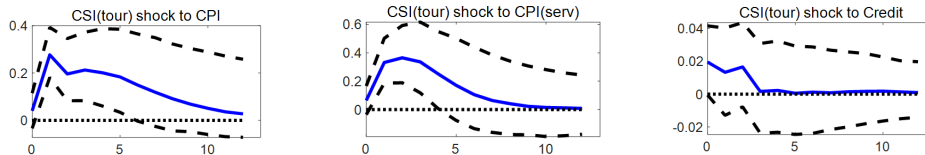
(a) TBCI 충격이 여타 변수에 미치는 영향



(b) 여타 구조충격이 TBCI에 미치는 영향



(c) CSI(tour) 충격이 물가 및 신용카드 사용액에 미치는 영향



변수순서는 CSI(tour)=전국 여행비지출전망 CSI, CPI=영동지역 물가상승률(전품목), CPI(serv)=영동지역 물가상승률(서비스재), TBCI=영동지역 관광경기지수, Credit=신용카드 사용액, Labor(total)=고용보험 피보험자수(전산업), Labor(tour)=고용보험 피보험자수 순환치(관광업)임. 실선은 사후 중간값이고 점선은 95% 사후 신용구간이다.

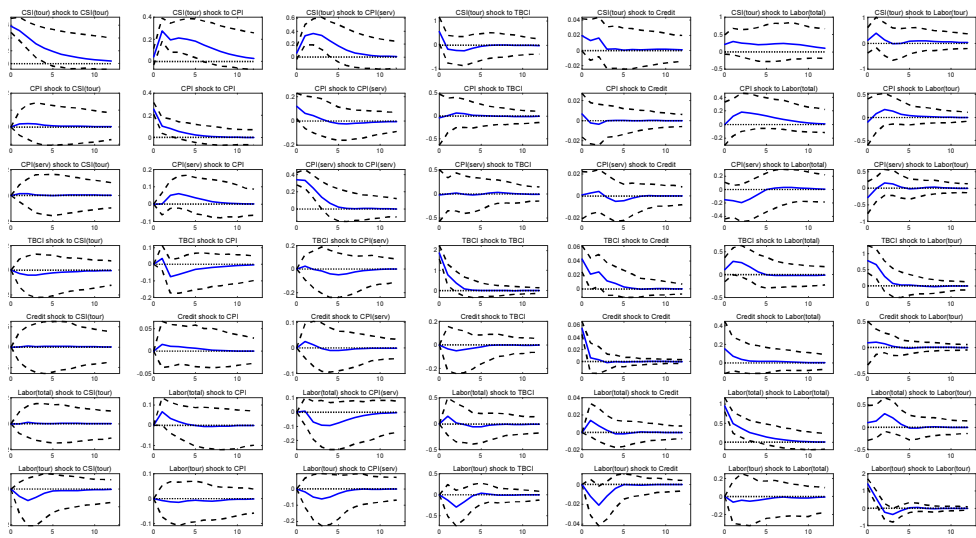
먼저 <그림 8>의 패널 (a)에 나타난 결과를 통해 영동지역 관광경기지수가 여타 실물경제 변수에 미치는 영향부터 살펴보도록 한다. 먼저 영동지역 관광경기지수 충격이 관광업 고용과 신용카드 소비에 통계적으로 유의하게 양의 효과를 미치는 것으로 추정되었으며, 파급효과가 2개월간 지속되는 것으로 나타났다. 성수기와 비성수기가 대체로 명확하게 구분되며, 성수기가 2~3개월 정도라는 점이 반영된 결과로 볼 수 있다. 관광경기지수 충격의 지속성이 길지 않다는 것은 관광경기지수 충격에 대한 관광경기지수의 충격반응함수를 통해서도 확인할 수 있다. 신용카드 소비 자료는 오프라인 결제기록만 취합한 것이고 외지인과 거주인을 구별하지 않는다. 따라서 관광경기 상승으로 인한 신용카드 소비 증가는 외지인의 직접적인 관광 소비 증가와 거주인의 간접적인 소비 증가가 함께 나타난 결과이다. 추가적으로 영동지역 관광경기지수가 영동지역 전체 고용에 미치는 영향 또한 양의 값으로 추정되었다. 다만 영동지역 전체 고용에서 관광업이 차지하는 절대적인 비중이 크지 않아서 충격반응함수의 통계적 유의성이 아주 강하지는 않다. 그럼에도 관광경기지수 충격이 2개월의 시차를 두고 지역 전체 고용에 상당한 영향을 준다는 점은 주목할 만하다. 즉, 관광경기지수 충격이 지역 실물경기에 파급되는 데 2개월의 시차가 소요된다는 것을 알 수 있다. 마지막으로, 소비자물가 상승률의 충격반응 추정결과를 봤을 때, 관광경기지수 충격이 지역 물가에 미치는 영향은 거의 없는 것으로 보인다.

다음으로 <그림 8>의 패널 (b)에 있는 실물경제 변수 충격에 대한 영동지역 관광경기지수의 충격반응을 살펴보도록 하자. 먼저 전국 관광경기 충격은 영동지역 관광경기지수를 즉각적으로 상승시키며, 충격의 효과는 해당 월에 그친 것으로 나타났다. 전국 관광경기에서 영동지역 관광경기가 차지하는 비중이 상당하고, 전국적으로 성수기 및 비성수기가 유사하여 전국 관광경기와 영동지역 관광경기 간 동조성이 강하기 때문에 나타난 결과로 해석된다. 특히 전국 관광경기 충격은 강원도 전체 소비자물가와 서비스업 물가 및 신용카드 소비를 통계적으로 유의미하게 상승시키는 등 영동지역 경기부양에 상당하게 기여한다는 것이 통계적으로 확인된다. (<그림 8>의 패널 (c) 참조)

추정결과의 강건성 검증을 위하여 변수순서를 달리하여 추정해보았다. 영동지역의 관광 산업 비중이 전국 관광 산업에서 큰 비중을 차지하고 있음을 고려하여 영동지역 관광경기지수를 가장 외생적인 변수로 가정하고 전국 여행비지출전망 CSI, 영동지역 물가상승률, 영동지역 물가상승률(서비스제), 로그 신용카드 사용액, 고용보험 피보험자수(전산업), 고용보험 피보험자수 순환치(관광업) 순으로 설정하였다. 또한 영동지역의 관광경기지수와 전국 여행비지출전망 CSI의 위치를 바꿔서도 추정해보았다. 전자의 추정결과는 <그림 10>이고 후자의 추정결과는 <그림 11>이다. 주요 추정결과, 즉 영동지역 관광경기지수 충격이 신용카드 사용액과 지역 고용을 상승시킨다는 것과 전국 여행비지출전망 CSI 충격이 지역 물가를 상승시키고 신용카드 사용액을 증가시킨다는 것은 변수순서와 무관하게 일치하였다. 다만, 전국 여행비지출전망 CSI와 영동지역 관광경기지수

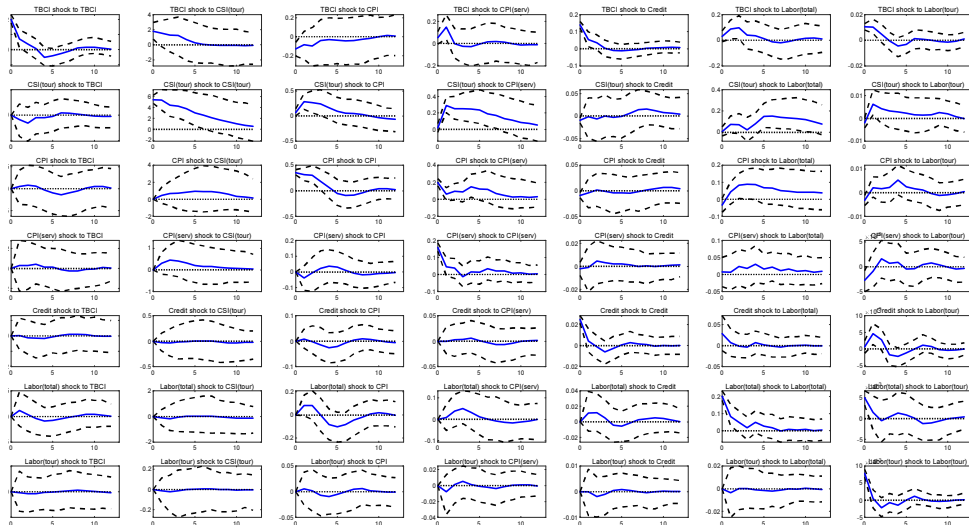
간 충격반응함수는 어떤 변수를 더 외생적인 변수로 두느냐에 따라서 달라지는 데, 영동 지역 관광경기가 전국 관광경기에서 비중의 대부분을 차지하는 것은 아니므로 전국 관광경기를 영동지역 관광경기보다 상대적으로 외생적인 변수로 보는 것이 합리적인 것이다. 이러한 가정하에서 전국 관광경기 충격이 영동지역 관광경기에 유의한 영향을 미치지만 영동지역 관광경기 충격은 전국경기에 유의한 영향을 미치지 않는 것으로 추정되었으며, 이 추정결과는 여타 변수의 순서와 무관하게 유지된다.

<그림 9> 충격반응함수(전체)



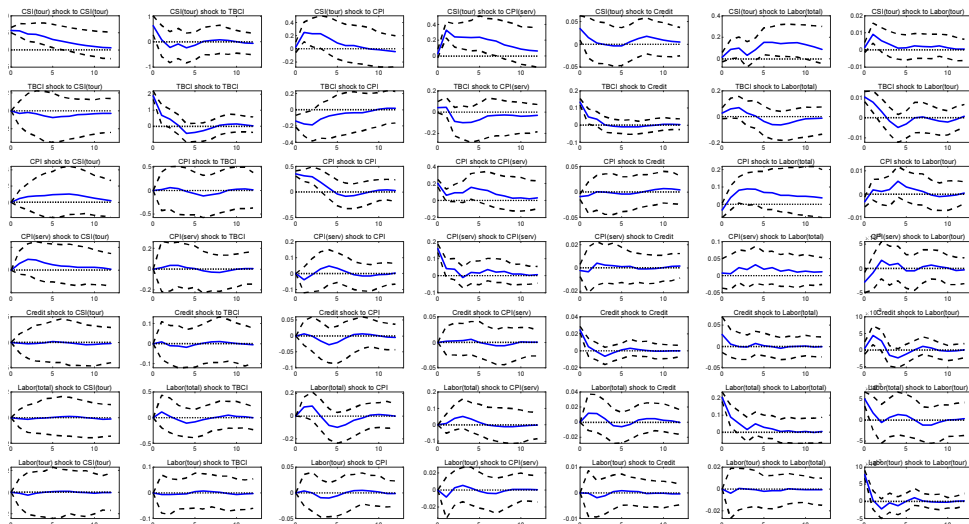
변수 순서는 CSI(tour)=전국 여행비지출전망 CSI, CPI=영동 지역 물가상승률(전품목), CPI(serv)=영동지역 물가상승률(서비스재), TBCI=영동지역 관광경기지수, Credit=로그 신용카드 사용액, Labor(total)=고용보험 피보험자수(전산업), Labor(tour)=고용보험 피보험자수 순환치(관광업)임. 실선은 사후 중간값이고 점선은 95% 사후 신용구간이다.

<그림 10> 충격반응함수(전체, 강건성 검증(a))



변수 순서는 TBCI=영동지역 관광경기지수, CSI(tour)=전국 여행비지출전망 CSI, CPI=영동지역 물가상승률(전품목), CPI(serv)=영동 지역 물가상승률(서비스제), Credit=로그 신용카드 사용액, Labor(total)=고용보험 피보험자수(전산업), Labor(tour)=고용보험 피보험자수 순환치(관광업)임. 실선은 사후 중간값이고 점선은 95% 사후 신용구간이다.

<그림 11> 충격반응함수(전체, 강건성 검증(b))



변수 순서는 CSI(tour)=전국 여행비지출전망 CSI, TBCI=영동지역 관광경기지수, CPI=영동지역 물가상승률(전품목), CPI(serv)=영동 지역 물가상승률(서비스제), Credit=로그 신용카드 사용액, Labor(total)=고용보험 피보험자수(전산업), Labor(tour)=고용보험 피보험자수 순환치(관광업)임. 실선은 사후 중간값이고 점선은 95% 사후 신용구간이다.

IV. 요약 및 시사점

1. 요약과 한계

본 연구에서는 신용카드 사용액, 영동지역 고속도로 톨게이트 차량 통행량, 내비게이션 영동지역 목적지 검색 건수, KT 이동통신 데이터 기준 영동지역 외국인 방문자수, 영동지역 시군별 네이버 검색어 트렌드 등 고빈도 자료를 이용하여 영동지역 관광경기 지수를 개발하였다. 이러한 지역 관광경기지수와 관광실물지표 간에 통계적으로 상관관계가 높게 나타났는데, 이는 고빈도 데이터에 기반한 관광경기지수가 관광업 동향을 판단하는 지표로서 속보성과 정보력이 우수함을 의미한다. 한편 관광경기지수와 관광업 고용보험 피보험자수를 통해 2018년 이후 영동지역의 관광실물경기를 측정해 본 결과 영동지역 관광경기는 성수기와 비수기가 확연히 구분되고 강한 계절성을 띠는 것을 확인할 수 있었다.

대규모 구조 VAR모형을 이용한 관광경기와 실물경제와의 관계분석에서는 영동지역 관광경기의 양(+)의 충격은 관광업 고용과 신용카드 소비를 증가시키며, 영동지역 전체 고용에도 긍정적인 효과가 있는 것으로 추정되었다. 다만 영동지역 관광경기 충격이 실물 부문으로 파급되는 효과의 지속성은 길지 않은 것으로 나타났다. 한편 전국 관광경기 충격이 영동지역 관광경기 및 신용카드 소비 증가에 상당 부분 영향을 주는 것으로 확인되었다.

한편 본 연구에서 산정된 관광경기지수는 계절성을 갖기 때문에 추정결과 해석에 상당한 주의를 요한다. 애초에 지수산정에 사용된 통계자료에 계절성이 강하기 때문에 계절성이 높은 관광경기지수가 산정되었다. 이로 인해 충격반응함수 추정결과가 외생적이고 임의의 충격이 아니라 주기적 현상에 의한 것이기 때문에 계절 변화가 구조충격에 포함되었을 가능성이 높다. 하지만 계절성을 제거하고자 하면 관측치의 변동성 중 계절성이 차지하는 부분이 대단히 클 뿐만 아니라 계절성을 정확하게 제거하는 행위 자체가 자료의 왜곡을 가져올 위험이 높다. 결과적으로 본 연구는 계절성을 제거하지 않았는데, 이는 본문에서도 언급한 바와 같이 코로나의 여파로 2020년과 2021년 겨울의 계절성이 이전에 비해 크게 약화되거나 아예 사라지는 현상이 자료에 드러났기 때문이다. 즉, 계절성의 시변성을 통제했을 때 잠재적으로 발생하는 편의가 너무 커서 주기적 효과를 압도할 것으로 판단했기 때문이다.

2. 시사점

영동지역 관광경기가 계절성이 강하고, 실물경제로의 파급효과가 단기적이며, 전국 관광경기에 영향을 받는다는 연구결과를 토대로 다음과 같은 정책적 시사점을 도출할

수 있다.

첫째, 관광경기의 성수기와 비성수기 간 격차를 완화하고 신규 관광수요를 지속적으로 창출하기 위해 다양한 관광자원을 개발해야 한다. 관광업의 특성상 성수기·비성수기 간 경기의 고저차 발생이 불가피하나 영동지역의 경우 주요 관광자원이 해안지역에 집중됨에 따라 관광경기의 계절성이 강한 것으로 나타난다. 계절성을 완화하기 위해서는 여름 휴가철 뿐만 아니라 비수기에도 관광객을 적극 유치할 수 있는 지역 특색을 살린 새로운 관광콘텐츠 개발이 필요하다.

대표적으로 올림픽 개최시설을 활용하여 동계레포츠 체험 프로그램을 제공하거나 국내·국제 스포츠 경기, 국제회의 유치를 확대하는 방안과 아울러 특색있는 지역을 묶어 테마관광 클러스터로 조성하는 방안도 고려해 볼 수 있을 것이다.

둘째, 영동지역의 관광경기가 전국 관광경기에 영향을 크게 받는다는 점을 고려할 때 정부의 단계적 일상회복 시행 등에 따른 향후 국내 관광수요 감소 위험에 대비해야 한다. 트래블버블 시행 등으로 해외여행에 대한 이연소비가 증가할 경우 영동지역 관광경기가 위축될 수 있다. 이에 대비하여 다양한 관광상품의 개발을 통해 지역관광 역량을 강화하는 한편 지역사회가 협심하여 외래객에 보다 친화적인 관광인프라를 조성할 필요가 있다.

향후 과제로 본 연구에서 개발한 관광경기지수를 제주도 등 관광업 비중이 높은 지역이나 세부 관광 업종별로 확장하여 산출할 경우 관광경기 모니터링 및 관광업 진흥을 위한 정책개발에 보다 유용하게 활용될 것으로 기대한다.

(투고일 : 2021.11.08. 수정 후 게재 : 2022.01.04. 게재확정일 : 2022.04.21.)

참고문헌

- 강승호 · 하세호(2015), “강원도 문화,관광업의 고용창출력 분석,” 『한국도시지리학 회지』 17(3), 141-154.
- 김보일 · 조강철(2012), “강원 및 제주지역 관광산업 분석을 통한 지역관광 발전방안에 관한 연구,” 한국은행 강원 · 제주본부.
- 김영국 · 민병기(2018), “2018 평창 동계올림픽 개최를 통한 강원도 관광의 공유가치 창출(creating shared value),” 한국은행 강원본부.
- 김영표(2021), “철도관광상품 연계를 통한 지역발전방안,” 한국은행 강릉본부 세미나 발표자료.
- 류시영 · 유선옥(2017), “소셜미디어에 나타난 강원도 관광에 대한 인식 연구: 빅데이터 분석을 중심으로,” 『관광연구저널』 31(2), 63-81.
- 문화체육관광부(2018), “지역 관광 활성화 방안 계획”.
- 박근화(2020), “코로나19로 인한 국민의 국내 여행 행태 변화,” 한국문화관광연구원
- 박정현 · 이은국(2020), “강원 영동지역 해안 관광도시를 중심으로 커피전문점 창업 지속,” 한국은행 지역경제보고서 2020.3월호
- 박재성 · 채희준(2018), “강원 영동지역 고용 동향 및 주요 특징,” 한국은행 강릉본부.
- 서정원 · 이준영(2021), “강원 영동지역의 관광인프라 투자 현황 및 시사점,” 한국은행 강릉본부.
- 송운강 · 박종필 · 윤승완(2016), “강원지역 동계레포츠산업 현황 및 발전방안,” 한국은행 강원 · 강릉본부.
- 송운강 · 양희원(2017), “계절에 따른 관광네트워크 차이 분석,” 강원대학교.
- 신동수(2021), “코로나19 충격의 파급경로 및 강원경제에 미친 영향,” 한국은행 강원본부.
- 양희원(2019), “2018 평창동계올림픽과 강원관광의 Shift-up,” 강원연구원.
- 유영선 · 김성은(2019), “강원도 관광업 현황 및 시사점,” 한국은행 강원본부.
- 윤승완 · 장하주(2018), “2018 평창 동계올림픽 기간 중 지역경제 동향 및 향후 전망,” 한국은행 강릉본부.
- 윤주(2020), “지역관광 활성화를 위한 주민참여확대 방안 연구,” 한국문화관광연구원.
- 이원희 · 박주영 · 조아라(2019), “관광 트렌드 분석 및 전망 : 2020-2024,” 한국문화관광연구원.
- 장이문 · 문동규(2020), “코로나19 발생 이후의 강원 영동 경제의 회복 현황 및 시사점,” 한국은행 강릉본부.
- 전현정(2019), “최근 강원지역 고용의 주요 특징 및 향후 과제,” 한국은행 강원본부.
- 한국관광공사(2021), “코로나19가 바꿔놓은 대한민국 관광지도”.

Giannone, D., Lenza, M., & Primiceri, G. E.(2015), “Prior selection for vector autoregressions,” *Review of Economics and Statistics*, 97(2), 436-451.

Litterman, R.,(1980), “A Bayesian Procedure for Forecasting with Vector Autoregression,” MIT, Department of Economics working paper.

[Abstract]

**Influence of Tourism in Yeongdong, Gangwon Province
Economy and Policy Implications: Development
of Regional Tourism Business Cycle
Index using Big data Analysis***

Kyu Ho Kang^{**} · Joo Won Bae^{***} · Jun Young Lee^{****}

This study develops a tourism business cycle index of Yeongdong area, and Naver search term trends by city and county in Yeongdong area, Gangwon-do using high-frequency data such as credit card usage amount, Yeongdong area highway tollgate vehicle traffic, navigation Yeongdong area destination searches, KT mobile communication data based on visitors to Yeongdong area. Calculate the index. A statistically high correlation was estimated between the calculated tourism economy index and the real economy of the local tourism industry. This means that the real-time tourism condition is properly reflected. In addition, as a result of large-scale structural VAR analysis, it was found that the shock of the tourism industry in Yeongdong increased employment and credit card consumption in the tourism industry, and had a positive effect on total employment. We expect that the results of this study will be used in real-time monitoring of tourism in the Yeongdong region and policy development to promote tourism and expand its real effects.

Key words: High frequency data, principal component analysis, large scale vector autoregressive
JEL Classification Number: G12, C11, E43

* The views expressed in this paper are those of the authors and do not necessarily reflect the position of the Bank of Korea. Kang acknowledges financial support from the Bank of Korea.

** Korea University, department of economics, Tel: +82-2-3290-5132, E-mail: kyuho@korea.ac.kr

*** Bank of Korea, Gangneung Branch Office, Junior Economist, Tel: +82-33-640-0157,
E-mail: jw.bae@bok.or.kr

**** Bank of Korea, Gangneung Branch Office, Junior Economist, Tel: +82-33-640-0161,
E-mail: junyoung.lee@bok.or.kr